Norman Schofield
Gonzalo Caballero  *Editors*

# Political Economy of Institutions, Democracy and Voting

Springer

# Political Economy of Institutions, Democracy and Voting

Norman Schofield · Gonzalo Caballero
Editors

# Political Economy of Institutions, Democracy and Voting

*Editors*
Prof. Norman Schofield
Washington University in St. Louis
Center in Political Economy
Brookings Drive 1
Campus Box 1027
Saint Louis
MO 63130
USA
schofield.norman@gmail.com

Assoc. Prof. Gonzalo Caballero
University of Vigo
Faculty of Economics
Lagoas Marcosende
36310 Vigo
Spain
gcaballero@uvigo.es

# Contents

# Contributors

**Carles Boix**   Department of Politics and Public Affairs, Princeton University, Princeton, NJ, USA, cboix@princeton.edu

**Gonzalo Caballero**   University of Vigo, Vigo, Spain; Faculty of Economics, University of Vigo, Campus as Lagoas, 36310 Vigo, Spain, gcaballero@uvigo.es

**Christopher Claassen**   Washington University in Saint Louis, St. Louis, MO, USA, classen@wustl.edu

**Bruno Codenotti**   Institute for Informatics and Telematics, Consiglio Nazionale delle Ricerche, Pisa, Italy

**Alfred W. Darnell**   Department of Political Science, Washington University in St. Louis, 1 Brookings Drive, St. Louis, MO 63130, USA, adarnell@wustl.edu

**Yadira González de Lara**   Universidad CEU-Cardenal Herrera, Elche, Spain; Universidad de Valencia, València, Spain, yadira.gonzalez@uch.ceu.es

**Desiree A. Desierto**   School of Economics, University of the Philippines, Diliman, Quezon City, 1101 Philippines, dadesierto@econ.upd.edu.ph

**Jon X. Eguia**   Department of Politics, New York University, 19 West 4th Street, 2nd Floor, New York, NY 10012, USA, eguia@nyu.edu

**Lawrence Ezrow**   Department of Government, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ UK, ezrow@essex.ac.uk

**Maria Gallego**   Washington University in Saint Louis, St. Louis, MO, USA; Department of Economics, Wilfrid Laurier University, Waterloo, ON, Canada, mgallego@wlu.ca

**Andrei Gomberg**   Centro de Investigación Económica, Instituto Tecnológico Autónomo de México, Av. Camino a Santa Teresa 930, DF 10700 México, México, gomberg@itam.mx

**Avner Greif**   Department of Economics, Stanford University, Stanford, CA 94305, USA, avner@stanford.edu

**JeeSeon Jeon**   Center in Political Economy, Washington University, 1 Brookings Drive, Saint Louis, MO 63130, USA, jsjeon@wustl.edu

**Christopher Kingston**   Department of Economics, Amherst College, Box 2201, Amherst, MA 01002, USA, cgkingston@amherst.edu

**Daniel Max Kselman**   Center for Advanced Studies in the Social Sciences, Juan March Institute, C/Castelló 77, 28006 Madrid, Spain, dkselman@march.es

**Carol Mershon**   Political Science Program, National Science Foundation, Arlington, VA, USA; Department of Politics, University of Virginia, Charlottesville, VA 22904 USA, cmershon@nsf.gov

**Marina Muskhelishvili**   Center in Political Economy, Washington University, 1 Brookings Drive, Saint Louis, MO 63130, USA; Apt. 27, 19 Chavchavadze ave., Tbilisi 0179, Georgia, marinmus@yahoo.com

**Antonio Nicolò**   Department of Economics "M. Fanno", Università degli Studi di Padova, Via del Santo 33, 35123 Padova, Italy

**John V.C. Nye**   Economics Department, George Mason University, Carow Hall MSN 1D3, 4400, University Drive, Fairfax, VA 22030, USA, jnye@gmu.edu

**Ugur Ozdemir**   Center in Political Economy, Washington University, 1 Brookings Drive, Saint Louis, MO 63130, USA, uozdemir@wustl.edu

**Sunita Parikh**   Department of Political Science, Washington University in St. Louis, 1 Brookings Drive, St. Louis, MO 63130, USA, saparikh@wustl.edu

**Giovanni Resta**   Institute for Informatics and Telematics, Consiglio Nazionale delle Ricerche, Pisa, Italy

**Norman Schofield**   Center in Political Economy, Washington University, 1 Brookings Drive, Saint Louis, MO 63130, USA, schofield.norman@gmail.com

**Olga Shvetsova**   Department of Political Science, Binghamton University, P.O. Box 6000, Binghamton, NY, 13902 USA, shvetsova@gmail.com

**Margit Tavits**   Center in Political Economy, Washington University, 1 Brookings Drive, Saint Louis, MO 63130, USA, tavits@wustl.edu

**Fernando Toboso**   University of Valencia, Valencia, Spain, Fernando.Toboso@uv.es

**Joshua A. Tucker**   Department of Politics, New York University, 19 West 4th Street, 2nd Fl., New York, NY 10012, USA, joshua.tucker@nyu.edu

**Ignacio Urquizu-Sancho**   Department of Sociology VI, Complutense University of Madrid, Madrid, Spain, iurquizu@march.es

# Introduction

**Norman Schofield and Gonzalo Caballero**

Modern Political Economy cannot be understood without considering the work of two pre-eminent scholars, Douglass C. North and William H. Riker.

The work by North[1] gave a boost to the New Institutional Economics (Coase 1984), and institutions have become an important research topic in political science and economics in recent years. The contributions by North have had increasing influence and this multi-disciplinary approach has propelled the New Institutional Social Sciences. Work by several institutionalist scholars, such as Williamson (1985), Libecap (1989), Eggertsson (1990), Ostrom (1990), Menard and Shirley (2005) and Greif (2006), as well as the recent book by North et al. (2009), have had a significant influence on current research in social sciences, as well as on policy making in both developed and developing countries.

Riker's work in positive political theory and federalism[2] had a major impact in political science itself, and has influenced the way scholars study democracy. The development of a theory of institutions, combined with the formal theory of elections, has engendered a new political economy involving political scientists, economists and economic historians.

Political economists have used insights about the role of ideas and institutions in an attempt to explain why rapid economic and social development occurred in Great Britain in the seventeenth and eighteenth centuries,[3] and spread to Europe and North America,[4] why the American colonies fought for independence,[5] why

---

[1]North (1961, 1981, 1990, 1994), North and Thomas (1973), North and Weingast (1989).

[2]Riker (1962, 1964, 1982, 1986, 1996), Riker and Ordeshook (1973).

[3]Acemoglu and Robinson (2000, 2006), Schofield (2006), Clark (2007), Mokyr (2010).

[4]Acemoglu et al. (2004, 2005).

[5]Schofield (2006).

N. Schofield (✉)
Washington University in St. Louis, St. Louis, MO, USA
e-mail: schofield.norman@gmail.com

G. Caballero
University of Vigo, Vigo, Spain
e-mail: gcaballero@uvigo.es

Latin America has seemed to fall behind North America,[6] why post-communist states move to democracy, and sometimes fall back to autocracy,[7] why autocracy can be stable,[8] and why economic and political development seems so difficult to implement in some countries, particularly in Africa.[9] Moreover, the research program on institutions has focused on the passage of time and the process of institutional change (North 2005; Greif 2006; Kingston and Caballero 2009). Other authors have examined the links between democracy and economic development, the so-called *modernization hypothesis* that development facilitates the transformation of the polity to democracy.[10] Recently, Jones and Romer (2010), in reviewing theories of economic growth, have suggested that the next major task is to build a theoretical apparatus that focuses on political and economic institutions, and on the difference between oligarchic and democratic societies.[11]

For this new research trajectory, studies of institutions, democracy and voting can provide the key to an understanding of societies both in the present and in the past.

The current volume includes contributions from authors of papers that were presented at conferences on the *Political Economy of Institutions, Democracy and Voting*, held at the Hoover Institution, Stanford in May 2009, at ECARES, Universite Libre de Bruxelles, August 2009, and at Baiona, Spain, June 2010, the latter under the auspices of the University of Vigo. The editors thank the Hoover Institution, ECARES and the University of Vigo for the support they provided.

Each chapter in this book went through a review process before publication. These chapters deal with theoretical and empirical issues over the behavior of institutions and the operation of democratic elections. Below we briefly sketch the topics discussed in these chapters.

## Part 1: Institutions

1. *Institutions: Rules or Equilibria?*
    by Avner Greif and Christopher Kingston
    Recent scholarship has demonstrated the power of the rational choice framework for advancing our understanding of institutions and institutional change. Stimulated by these developments, the conceptual frameworks employed by scholars studying institutions have also been evolving, as old frameworks have been adapted and new

---

[6]Acemoglu et al. (2001, 2002), Sokoloff and Engerman (2000), Przeworski and Curvale (2006).

[7]Schofield (2009), Bunce and Wolchik (2010).

[8]Epstein et al. (2006), Gallego and Pitchik (2004).

[9]Collier (2007, 2009), Easterly (2007).

[10]Persson and Tabellini (1999, 2003), Przeworski et al. (2000), Przeworski (1991, 2006), Boix (2003), Acemoglu et al. (2008, 2009).

[11]See also Acemoglu (2008).

frameworks have emerged to explore how institutions function, how they change, and how they affect economic behavior and outcomes. This involves two key questions: first, how institutions are selected and second, how people are motivated to follow institutionalized patterns of behavior. One strand of thought within the rational-choice approach to institutional analysis, the 'institutions-as-rules' approach, focuses on a theory of how the "rules of the game" in a society are selected. An emerging alternative approach instead emphasizes the importance of a theory of motivation and thereby endogenizes the "enforcement of the rules", by studying 'institutions-as-equilibria'. In this chapter, the authors survey these developments and highlight promising directions for future research. They argue that by endogenizing the issue of enforcement, the institutions-as-equilibria approach enables a more satisfactory treatment of several key issues, including promoting our understanding of processes of institutional change.

2. *War, Wealth and the Formation of States*
by Carles Boix, Bruno Codenotti and Giovanni Resta

Employing agent-based modelling techniques, the authors examine the evolution of a world with sovereign states that maximize power. They show that: (1) the size (number) of states increases (decreases) as war technologies become capital-intensive; (2) the number of states declines with development and population expansion; (3) capital-rich (capital-poor) economies lead to smaller (larger) economies (mainly because war is less frequent if capital is mobile); (4) world government may become possible in the future (given the evolution of military technology) yet only with a very low probability (given the distribution of economic activities throughout the globe); (5) the possibility of secession leads to a permanent increase in the number of countries if all effects when the countries involved in the split are democratic. These stylized findings fit well the historical evolution of Europe and most of the territorial dynamics of state formation over time, at least until the nineteenth century. The last point accommodates the explosion of the number of countries we have witnessed in the twentieth century.

3. *Why Do Weak States Prefer Prohibition to Taxation?*
by Desirée A. Desierto and John V.C. Nye

Why do weak states prefer prohibition to taxation? Desier to and Nye show that keeping an undesirable good illegal is more efficient than legalizing and taxing it, even if producers of the prohibited goods pay out large bribes to prohibition enforcers. If the bribes are recognized as revenues to the enforcers, this additional benefit keeps welfare losses small. This chapter further supports this finding with preliminary empirical evidence and graphical analyses of the likely net welfare losses from prohibition and taxation. It provides a positive rationale for the preference for prohibition in states prone to corruption and imperfect enforcement.

4. *Self-Enforcing, Public-Order Institutions for Contract Enforcement: Litigation, Regulation, and Limited Government in Venice, 1050–1350*
by Yadira González de Lara

The spectacular economic growth of Venice during the late medieval period (1050–1350) was based on the expansion of its trade along the Mediterranean and beyond. Crucial to this expansion was the mobilization of large amounts of capital

into risky investments. However, this mobilization required the development of institutions that protected creditors and shareholders from expropriation by controlling merchants. This chapter finds that legal and administrative institutions conjointly provided investor protection and explores the interactions between these public-order institutions for contract enforcement and the emergence of a limited government, a coercion-constraining institution that motivated judges and regulators to use their coercive power for protecting rather than abusing investor rights.

5. *Judicial Stability During Regime Change: Apex Courts in India 1937–1960* by Alfred Darnell and Sunita Parikh

In this chapter, Alfred Darnell and Sunita Parikh examine the conflictual relationship of two apex courts with the executive branches of India under British colonial rule and after Independence. One, the Federal Court of India, existed in the closing decades of British colonial rule, the other, the Supreme Court of India, replaced the Federal Court in independent India. Little changed between the two courts institutionally or organizationally. However, each court has been characterized quite differently: the former as weak and ineffectual, the latter as elitist and obstructionist. Why has this been the case? In order to answer this question the authors examine major rulings of each court that involved the executive branch and assess each court's decision according to two prevailing theories of judicial decision making: those that emphasize preferences over policy and those that emphasize decisions based on "black letter law". They find both explanations lacking because of evidence that both apex courts in India were concerned not only with issues of law and policy, but also with the stability and security of the institution of each Court.

6. *Institutional Arrangements Matter for Both Efficiency and Distribution: Contributions and Challenges of the New Institutional Economics* by Fernando Toboso

Are scholars in the New Institutional Economics tradition systematically disregarding distributive aspects when approaching policy issues as was the case during the 1970s and 1980s? Do economic and political agents usually care about distribution too? To provide an answer to these questions is the basic purpose of this chapter. The analysis carried out demonstrates that not all NIE oriented scholars disregard distributive issues. Some contributions are examined as examples, mainly in the so-called political economy branch of NIE. By means of a well-known graphical tool, the chapter also emphasizes that all of us clearly care about distribution, not just about efficiency, when participating in market transactions as well as in collective political decisions. The analysis also reveals very persuasively how institutional reforms affect participants' relative rights and capacities to act and bargain, not just the total amount of transaction costs experienced by them. Though unfamiliar to many new institutionalists, the author concludes that all this has been acknowledged by authors such as North (1990), Eggertsson (1990) and Libecap.

7. *Institutional Foundations, Committee System and Amateur Legislators in the Governance of the Spanish Congress: An Institutional Comparative Perspective (USA, Argentina, Spain)* by Gonzalo Caballero

Legislative organization matters for policy-making. Institutional rules determine the role of property rights, hierarchies, individual deputies, parliamentary groups, transactions and committees in the industrial organization of Congress. The New Institutional Economics and Transaction Cost Politics have given rise to a useful research program on legislative organization. This chapter analyses the institutional foundations of legislative organization of the Spanish Congress from an institutional and transactional comparative perspective. Electoral rules and Committee systems are institutional determinants of the political property rights of congressmen and the structure of governance of legislative organization. The industrial organization of the Spanish Congress is studied, and compared with the traditional model of the US Congress and the Argentine Congress.

8. *Coalition Governments and Electoral Behavior: Who is Accountable?*
by Ignacio Urquizu Sancho

Elections have been studied in political sciences from two different points of view: either by looking at the selection of 'good types' – prospective mechanism-or by studying the sanctions – retrospective mechanism. If we assume that elections are a question about sanctioning, it is widespread that citizens may not assign responsibilities to multiparty cabinets. Thus, scholars have concluded that economic voting does not work properly in the case of coalition governments. This argument has been coined as the hypothesis of 'clarity of responsibility'. However, if so, how do they explain the electoral results of coalition governments? What do voters consider when they evaluate a multiparty cabinet? In this chapter, Urquizu Sancho discusses some theoretical arguments that question that hypothesis. In fact, this research develops the causal mechanisms that explain how economic voting work for multiparty cabinets.

# Part 2: Democracy and Voting

9. *Empirical and Formal Models of the United States Presidential Elections in 2000 and 2004*
by Norman Schofield, Christopher Claassen, Maria Gallego, and Ugur Ozdemir

This chapter develops a general stochastic model of elections in which the electoral response is affected by the *valence* (or quality) of the candidates. In an attempt to explain non-convergence of candidate positions in the 2000 and 2004 Presidential elections, a formal spatial stochastic model, based on *intrinsic valence*, is presented. A pure spatial model of the election is constructed. It is shown that the equilibria, under vote maximization, do indeed lie at the electoral origin. Other work on Presidential elections in the United States has suggested that a superior empirical model should incorporate the electoral perceptions of the candidate character traits. The chapter then considers a joint model with sociodemographic valences as well as electoral perception of traits and shows by simulation that the vote maximizing equilibrium positions were close to, but not precisely at, the electoral origin. This model used electoral estimates of the candidates' positions. These

differed substantially from the estimated equilibria of the traits model. To account for this difference, a more general formal model is then considered where the valence differences between the candidates were due to resources that were contributed to the candidates by party activists. The trade off between activist and electoral support is given by a (first order) *balance condition* involving, called the *centrifugal marginal activist pull*. Survey information on party activists, who contributed resources to the candidates, was obtained. It is argued that the difference between the equilibrium obtained from the spatial model with traits, and the estimated candidate positions, is compatible with the location of these activists.

The final model is one where the activist resources are used by candidates to target individual voters or groups of voters. The balance condition in this case involves a complex constrained optimization problem, that captures the essence of modern electoral politics.

10. *Modelling Elections in Post-Communist Regimes: Voter Perceptions, Political Leaders and Activists*

by Norman Schofield, JeeSeon Jeon, Marina Muskhelishvili, Ugur Ozdemir and Margit Tavits

This chapter uses the stochastic electoral model to examine elections in Poland in 1997, 2001 and 2005, in Georgia in 2008, and in Azerbaijan in 2010. In contrast to the result for the U.S. elections presented in Schofield et al. (2011), it was found that in Poland the valence differences were sufficiently large to force low valence parties to adopt divergent positions. This implies a fundamental difference between an electoral system based on plurality rule in contrast to one based on proportional representation. In addition, in "anocracies" such as Georgia and Aizerbaijan, the limited access to the media by the parties in opposition to the president means that their support groups find it difficult to coalesce. As a consequence, they are unable to press successfully for greater democratization. In these countries, the presidential electoral system is highly majoritarian, and the President's party dominates the political arena, controlling political resources and the media. The chapter concludes by giving an overview of the empirical results that have been obtained so far for the three plurality democracies of the USA, Britain and Canada, three polities with proportional electoral systems, and the three anocracies of Georgia, Azerbaijan and Russia.

11. *Electoral Systems and Party Responsiveness*

by Lawrence Ezrow

Do political parties respond to shifts in the preferences of their supporters or to shifts in the mean voter position? Also, do electoral systems mediate these crucial citizen-party linkages? The central finding of this chapter is that electoral systems do condition these effects. Parties in proportional systems are systematically responsive to the mean voter position while parties in disproportional systems do not display the same tendency. Additionally, neither system induces parties to systematically respond to their supporters.

12. *Electoral Institutions and Political Corruption: Ballot Structure, Electoral Formula, and Graft*

by Daniel Max Kselman

Most research on the consequences of electoral institutions examines the distinction between majoritarian and proportional electoral formulae. Recent work has also examined the impact of a system's ballot structure, i.e. the formal rules governing how citizens vote, on a variety of political phenomena. This chapter develops a game theoretic model to study the interactive impact of formulae and ballot structures on political corruption. In contrast to received wisdom, the theoretical results suggest that Open-List voting systems should outperform both Closed-List voting systems and First-Past-The-Post systems in constraining corruption. Also in contrast to received wisdom, the results identify a set of conditions under which Closed-List systems might themselves outperform First-Past-The-Post systems. Analysis of cross-national data provides support for the chapter's theoretical model. Taken together, the chapter's formal and empirical results provide a strong counter-argument to the notion that majoritarian institutions generate better governance than their proportional representation counterparts.

13. *A Model of Party Entry in Parliamentary Systems with Proportional Representation*
by Daniel M. Kselman and Joshua A. Tucker

Spatial models with a party entry decision largely fall into one of two classes. The first of these preserves the Downsian assumptions that candidates are office-seeking and can announce policy positions anywhere in the policy space. A distinct class of models features what are now known as "citizen-candidates" who combine policy- and office-seeking incentives, and who cannot credibly commit to implementing any policy other than their own 'ideal point' as a platform in electoral campaigns. The chapter develops a game theoretic model of party entry which employs mechanisms from each of these classes of analyses, but departs from both bodies of literature in studying party entry in Parliamentary regimes with Proportional Representation. Preliminary analysis of Subgame Perfect Nash Equilibrium suggests that, when parties are exclusively concerned with policy, party entry should be somewhat more likely when status quo parties are well-dispersed around the median voter's ideal point than when they are both fairly centrist. However, as candidates' office-seeking incentives begin to outpace their policy-seeking incentives, the relationship between status quo dispersion and entry becomes more complicated, and depends crucially on the ideal point of the entering candidate.

14. *Moving in Time: Legislative Party Switching as Time-Contingent Choice*
by Carol Mershon and Olga Shvetsova.

Why would a sitting legislator leave the party on whose label she has won election and join another parliamentary party? The premise of this chapter is that a politician's calculus on party affiliation involves not only what she stands to gain or lose, but also when the potential gains or losses likely occur. The theoretical model demonstrates that an MP times a shift in party allegiance so as to minimize losses and maximize gains. The empirical illustrations bearing on our predictions afford variation on the key parameter of electoral laws and drive home the strategic importance of timing in systems. The theoretical and empirical findings on when incumbents switch party during a legislative term shed new light on why they switch.

15. *On the Distribution of Particularistic Goods*
by Jon X. Eguia and Antonio Nicolò

This chapter characterizes the set of equilibria in a model of distributive politics with inefficient local public goods. Candidates compete for office in three districts under a majoritarian rule. For each district there is a project that brings a benefit only to this district if implemented, but the aggregate cost for society of financing the project surpasses the localized benefit. Candidates can commit to implement the project surpasses the localized benefit. Candidates can commit to implement the projects in any number of districts. If projects are very inefficient, in equilibrium candidates commit not to implement any of them. However, if projects are inefficient but not too inefficient, in the unique equilibrium candidates randomize between financing projects in zero, one or two districts, so that in expectation 43% of projects are implemented.

16. *Vote Revelation: Empirical Content of Scoring Rules*
by Andrei Gomberg

In this chapter Gomberg considers choice correspondences defined on an extended domain: the decisions are assumed to be taken not by individuals, but by committees and, in addition to the budget sets, committee composition is observable and variable. In this setting, he establishes a restriction on the choice structures that is implied by the scoring decision-making by rational committee members.

# References

Acemoglu D (2008) Oligarchic versus democratic societies. J Eur Econ Assoc 6:1–44

Acemoglu D, Robinson J (2000) Why did the west extend the franchise? Growth, inequality and democracy in historical perspective. Quart J Econ 115:1167–1199

Acemoglu D, Robinson J (2006) Economic origins of dictatorship and democracy. Cambridge University Press, Cambridge

Acemoglu D, Johnson S, Robinson J (2001) The colonial origins of comparative development. Am Econ Rev 91:1369–1401

Acemoglu D, Johnson S, Robinson J (2002) Reversal of fortune. Quart J Econ 118:1231–1294

Acemoglu D, Johnson S, Robinson J (2004) Institutions as the fundamental cause of long run growth. NBER, Washington, DC

Acemoglu D, Johnson S, Robinson J (2005) The rise of Europe: Atlantic trade, institutional change, and economic growth. Am Econ Rev 95:546–579

Acemoglu D, Johnson S, Robinson J, Yared P (2008) Income and democracy. Am Econ Rev 98:808–842

Acemoglu D, Johnson S, Robinson J, Yared P (2009) Reevaluating the modernization hypothesis. J Monet Econ 56:1043–1058

Boix C (2003) Democracy and redistribution. Cambridge University Press, Cambridge

Bunce V, Wolchik S (2010) Democracy and authoritarianism in the post communist world. Cambridge University Press, Cambridge

Clark G (2007) A farewell to alms. Princeton University Press, Princeton, NJ

Coase RH (1984) The new institutional economics. J Inst Theor Econ 140:229–231

Collier P (2007) The bottom billion. Oxford University Press, Oxford

Collier P (2009) Wars. Guns and votes, Harper, NY

Easterly W (2007) Globalization, poverty and all that: factor endowment versus productivity views. In: Harrison A (ed) Globalization and poverty. Chicago University Press, Chicago, IL

Eggertsson T (1990) Economic behaviour and institutions. Cambridge University Press, Cambridge

Epstein D, Bates R, Goldstone J, Kristensen I, O'Halloran S (2006) Democratic transitions. Am J
    Polit Sci 50:551–568
Gallego M, Pitchik C (2004) An economic theory of leadership turnover. J Public Econ 88:
    2361–2382
Greif A (2006) Institutions and the path to modern economy. Cambridge University Press,
    Cambridge
Jones CI, Romer P (2010) The Kaldor facts: ideas, institutions, population and human capital. Am
    Econ J Macroecon 2:224–245
Kingston C, Caballero G (2009) Comparing theories of institutional change. J Inst Econ 5(2):
    151–180
Libecap GD (1989) Contracting for property rights. Cambridge University Press, Cambridge
Menard C, Shirley MM (eds) (2005) Handbook of new institutional economics. Springer, Berlin
Mokyr J (2010) The enlightened economy: an economic history of Britain 1700–1850. Yale
    University Press, New Haven, CT
North DC (1961) The economic growth of the United States. Norton, New York
North DC (1981) Structure and change in economic history. Norton, New York
North DC (1990) Institutions, institutional change and economic performance. Cambridge Uni-
    versity Press, Cambridge
North DC (1994) Economic performance through time. Am Econ Rev 84:359–368
North DC (2005) Understanding the process of economic change. Princeton University Press,
    Princeton, NJ
North DC, Thomas R (1973) The rise of the western world: a new economic history. Cambridge
    University Press, Cambridge
North DC, Weingast B (1989) Constitutions and commitment: the evolution of institutions
    governing public choice in seventeenth century England. J Econ Hist 49:803–832
North DC, Wallis JJ, Weingast BR (2009) Violence and social orders: a conceptual framework for
    interpreting recorded human history. Cambridge University Press, Cambridge
Ostrom E (1990) Governing the commons. Cambridge University Press, Cambridge
Persson T, Tabellini G (1999) Political economics: explaining economic policy. MIT, Cambridge,
    MA
Persson T, Tabellini G (2003) The economic effect of constitutions. MIT, Cambridge, MA
Przeworski A (1991) Democracy and the market: political and economic reforms in Eastern
    Europe and Latin America. Cambridge University Press, Cambridge
Przeworski A (2006) Democracy and economic development. In: Mansfield E, Sisson R (eds)
    Political science and the public interest. Ohio State University Press, Columbus, OH
Przeworski A, Curvale C (2006) Does politics explain the economic gap between the United States
    and Latin America? In: Fukuyama F (ed) La Brecha entre America Latina y los Estados
    Unidos. Fondo de Cultura Economica, Buenos Aires
Przeworski A, Alvarez ME, Cheibub JA, Limongi F (2000) Democracy and development: political
    institutions and well-being in the world, 1950–1990. Cambridge University Press, Cambridge
Riker WH (1962) The theory of political coalitions. Yale University Press, New Haven, CT
Riker WH (1964) Federalism: origin, operation, maintenance. Little Brown, Boston, MA
Riker WH (1982) Liberalism against populism: a confrontation between the theory of democracy
    and the theory of social choice. WH Freeman, San Francisco, CA
Riker WH (1986) The art of manipulation. Yale University Press, New Haven, CT
Riker WH (1996) The strategy of rhetoric: campaigning for the ratification of the constitution.
    Yale University Press, New Haven, CT
Riker WH, Ordeshook PC (1973) An introduction to positive political theory. Prentice-Hall,
    Englewood Cliffs, NJ
Schofield N (2006) Architects of political change: constitutional quandaries and social choice
    theory. Cambridge University Press, Cambridge
Schofield N (2009) The political economy of democracy and tyranny. Oldenbourg, Munich

Schofield N, Claassen C, Gallego M, Ozdemir U (2011) Empirical and formal models of the United States presidential elections in 2000 and 2004. Chapter 9 in this volume

Sokoloff KL, Engerman SL (2000) Institutions, factor endowments and the paths of development in the new world. J Econ Perspect 14:217–232

Williamson OE (1985) The economics institutions of capitalism: firms, markets, relational contracting. The Free Press, New York

# Part I
# Institutions

# Institutions: Rules or Equilibria?

**Avner Greif and Christopher Kingston**

## 1 Introduction

In recent years, the interest in rational choice analysis of institutions has received substantial impetus from an accumulating body of evidence demonstrating the importance of a society's institutions in determining its economic outcomes. Econometric studies have uncovered correlations between institutional variables such as the security of property rights, the rule of law, and trust, and economic and political outcomes including levels of production, saving, and corruption.[1] Historical studies have revealed the role that institutions played in long-run trajectories of industrial and commercial development.[2] Studies of the developing world and of countries transitioning from socialism have revealed the challenges involved in creating well-functioning institutions, the benefits that can be obtained when institutional change and economic reform are successful, and the dangers that ensue when they are not.[3] Stimulated by these developments, the conceptual frameworks employed by scholars studying institutions have also been evolving, as old frameworks have been adapted and new frameworks have emerged to explore old and new questions about how institutions function, how they change, and how they affect economic behavior and outcomes.

The rational-choice approach to institutional analysis does not require us to assume that people are always 'rational', or that institutions are chosen rationally. Rather, it holds that a rational-choice perspective enables us to generate a theory with empirically refutable predictions about the institutions that can prevail in a given situation.

---

[1]For example, La Porta et al. (2008), Keefer and Knack (1997), Easterly and Levine (2003).

[2]For example, Milgrom et al. (1990), North (1990), Greif (1989, 1994, 2006).

[3]For example, Roland (2000), Aoki (2001), Qian (2003).

A. Greif (✉)
Department of Economics, Stanford University, Stanford, CA 94305, USA
e-mail: avner@stanford.edu

C. Kingston
Department of Economics, Amherst College, Amherst MA 01002, USA
e-mail: cgkingston@amherst.edu

This involves two key questions: first, how institutions are selected and second, how people are motivated to follow institutionalized patterns of behavior. One strand of thought within the rational-choice approach to institutional analysis, the so-called 'institutions-as-rules' approach, emphasizes the importance of a theory of selection of institutions, while an emerging alternative approach, the 'institutions-as-equilibria' line of analysis emphasizes the importance of a theory of motivation.

The institutions-as-rules approach, following North (1990, p. 3), identifies institutions as "the rules of the game in a society", including both "formal" rules such as constitutions and laws enforced by the state, and "informal" constraints such as "codes of conduct, norms of behavior, and conventions", which are generally enforced by the members of the relevant group (North, 1990, p. 36). Many kinds of formal rules are selected through a centralized process of bargaining and political conflict between individuals and organizations who attempt to change the rules for their own benefit. In other cases, formal or informal rules may be selected in a decentralized way through evolutionary competition among alternative institutional forms. In either case, the institutions-as-rules view holds that institutions are ultimately best understood from a functionalist perspective that recognizes that they are responsive to the interests and needs of their creators (although there is no guarantee that the rules selected will be efficient).

Within the institutions-as-rules view, the enforcement of the rules is considered as a distinct issue from the formation and content of the rules themselves. Enforcing the rules involves "enforcement costs". The formal and informal rules, together with their "enforcement characteristics" constitute the institutional structure within which interactions occur. Thus, the institutions-as rules approach employs a rational-choice perspective to study the formation of institutions, but a theory of motivation – explaining *why* people follow particular rules of behavior – is not integrated into the analysis.

A growing body of recent research on institutions places a theory of motivation at the center of the analysis, and thereby endogenizes the "enforcement of the rules", by studying 'institutions-as-equilibria'. This perspective focuses on how interactions among purposeful agents create the structure that gives each of them the motivation to act in a manner perpetuating this structure. To give a simple example: in the United States, people (nearly always) drive on the right-hand side of the road. This regularity of behavior generates expectations that motivate the behavior itself: people drive on the right because they expect others to do so, and wish to avoid accidents. Of course, it is also a "rule" that one must drive on the right. However, many alternative technologically feasible rules (for example, women drive on the right and men on the left) would generate expectations which would fail to motivate a pattern behavior consistent with the rule: that is, such patterns of behavior are not equilibria, and even if they were formally specified as a "rule" we would not expect them to emerge as institutions, because the "rule" would not be self-enforcing. For everyone to drive on the right, however, is one of two potentially self-enforcing "rules" which could emerge (or be enacted) as an equilibrium.

The crucial point is that while a "rule" may serve as a coordination device, it is fundamentally the expected behavior of others, rather than the rule itself, which

motivates people's behavior. A similar logic can be used to examine economic, political, and social institutions even in situations involving specialized actors and more complex formal "rules". From the institutions-as-equilibria perspective, it is always ultimately expectations about the behavior of the other actors (including those in specialized enforcement roles such as police, judges, etc.) that create the institutional constraints which mold people's behavior, and all such behavior must therefore ultimately be explainable endogenously as part of the equilibrium.

Despite their differences, the institutions-as-rules and institutions-as-equilibria approaches have much in common and are best viewed as complements rather than substitutes. Both seek to advance a positive analysis of the non-technological determinants of order and regularities of human behavior. Recent advances in the literature combine elements of the two perspectives. This chapter surveys these developments and highlights promising directions for future research. As we will discuss, the institutions-as-rules framework has been fruitfully applied to shed light on the emergence and functioning of a variety of institutions, including communities, organizations, and political and legal institutions. However, we will argue that by endogenizing the issue of enforcement, the institutions-as-equilibria approach enables a more satisfactory treatment of several key issues, including promoting our understanding of processes of institutional change.[4]

## 2   Institutions as Rules: Conceptual Issues

As discussed above, the most commonly cited definition of institutions is that advanced by Douglass North: institutions "are the rules of the game in a society, or more formally, are the humanly devised constraints that shape human interaction" (North 1990, p. 3). Institutions include both formal rules, which are explicit, written rules such as laws and constitutions, and informal constraints such as conventions and norms. In North's theory, formal rules are created by the polity, whereas informal norms "are a part of the heritage that we call culture" (p. 37) and therefore impervious to deliberate human design. The focus of the analysis is therefore on formal rules, namely, rules that are explicitly and intentionally created.

To illustrate the institutions-as-rules approach, consider the framework developed by Ostrom (2005), who envisages a hierarchy with several levels of rules: "operational rules" which govern day-to-day interactions; "collective-choice rules", which are rules for choosing operational rules; "constitutional rules" (rules for choosing collective-choice rules); "meta constitutional rules" (rules for choosing constitutional rules); and at the highest level, the biophysical world (p. 58).[5] That is, each level in this

---

[4]For a recent discussion, see Greif (2006). Kingston and Caballero (2009) survey theories of institutional change.

[5]North (1990, p. 47) envisages a similar hierarchy with four levels of formal rules: constitutions, statute and common laws, specific bylaws, and individual contracts.

hierarchy of rules consists of rules that govern how rules at the lower level are created. For example, constitutional and collective-choice rules provide the structure that governs the choice of operational rules. Higher-level rules are also more difficult and costly to change.

When they perceive that existing rules governing their interactions at one level are unsatisfactory, individuals are driven to "shift levels" and try to change the rules. A political bargaining process ensues. Each individual calculates their expected costs and benefits from any proposed institutional change, and an institutional change can occur only if a "minimum coalition" necessary to effect change agrees to it. What constitutes a "minimum coalition" is determined by the higher-level rules; for example, in a democracy, a majority would constitute a minimum winning coalition; in a dictatorship the dictator alone might constitute a minimum coalition. Therefore, the set of rules that ultimately emerges will depend on the perceived interests of the actors involved in setting the rules, on the ability of various interest groups to act collectively to make their interests count (Olson 1982), and on the higher-level rules that determine how those individual interests are aggregated.

There is no guarantee that this process will lead to the selection of efficient rules. In many cases, those with political power may try to select rules to generate distributional benefits for themselves; that is, to maximize their welfare rather than that of society as a whole. To explain why societies "choose" inefficient institutions, however, it is not sufficient to note that the groups in power have interests that diverge from the rest of society. If an institutional change could increase efficiency and economic output, why cannot the beneficiaries of the change agree to redistribute the gains to compensate the losers? Acemoglu (2003) argues that the key problem is commitment: the powerful cannot credibly commit not to use their power for their own benefit as the opportunity arises, and other groups cannot credibly commit to compensate the powerful for giving up their power. As a result, the set of bargains which can be struck is restricted to those bargains which can be sustained as equilibrium outcomes (Fearon 2007; Greif 1998, 2006). Because there is no external authority to enforce inter-temporal bargains, politically powerful groups may block changes that would be beneficial overall, or impose inefficient changes that benefit themselves at the expense of others. Fundamentally, therefore, a satisfactory understanding of these aspects of institutional change requires a recognition that the problem is not just choosing new rules, but the more restrictive problem of engineering a mutually beneficial shift to a new, self-enforcing equilibrium. We will return to this issue later.

A second, complementary strand of thought within the institutions-as-rules approach views the development of rules as an outcome of evolutionary competition among alternative institutional forms. Alchian (1950) argued that competitive pressure weeds out inefficient forms of organization among firms in competitive markets, because firms that develop more efficient organizational forms will be more profitable, and the use of these rules and forms of organization will therefore tend to spread through growth or imitation. Demsetz (1967) extended the evolutionary argument to the development of property-rights rules, hypothesizing that these rules develop and adjust as a result of "legal and moral experiments" which

"may be hit-and-miss procedures to some extent", but which only prove viable in the long run if they generate efficient outcomes. Hayek (1973) argues that groups or organizations that, by accident or design, develop less efficient rules will not survive competition with groups that develop more efficient rules. Therefore, through group selection, rules will evolve towards optimality.

The evolutionary approach finds its most prominent modern expression in Oliver Williamson's "Transactions cost economics" (TCE). According to this view, 'transaction costs' arise in many transactions because of the bounded rationality and opportunism of the transacting parties (Williamson 2000). Depending on the attributes of a particular transaction, some sets of rules ('governance structures') will lead to more efficient outcomes than others. The transactions-cost economics approach assumes that the most efficient institutional forms (those which 'minimize transactions costs') will emerge.[6] So, for example, if a change in production technology renders existing institutions inefficient, then over time, new, more efficient institutional forms will emerge to replace them.

Although the political-design and evolutionary approaches envisage quite different processes for the selection of rules, the two strands of research are best viewed as complementary. Both treat institutions as sets of rules (or "governance structures"); and both focus on how new rules are selected rather than how they are enforced. Different institutions are associated with different "transaction costs", including "monitoring costs" and "enforcement costs", but the nature of these costs is not part of the analysis.

The concept of "transactions costs" is widely used in New Institutional Economics. The term is generally used very broadly to include the costs of finding trading partners, negotiating and drawing up contracts, monitoring contractual partners' behavior and enforcing agreements, and other costs incurred in an effort to define, measure and enforce property rights or agreements to exchange property rights. Transaction costs may also include the costs of political activity, bargaining, legal action, and so on involved in deliberate efforts to create new rules, the costs of inefficiency resulting from commitment problems and other forms of political transaction costs, as well as all the costs involved in setting up, maintaining and changing the structure of rules and organizations, and monitoring the actions of the agents governed by those rules. In short, any difference between the value of output generated in the real world, where a real transaction is governed by real institutions, and an imagined world without any agency problems or information asymmetries (and therefore a world in which no governance is required), including any deviation from first-best production and exchange, can be called a "transaction cost".

Despite this breadth, the concept of "transaction costs" has achieved wide acceptance as an analytical tool in the theoretical literature on institutions, particularly within the institutions-as-rules framework. The usefulness of the concept is

---

[6]Williamson refers to this as the "discriminating alignment" hypothesis. Thus, for Williamson, "The overall object of the exercise essentially comes down to this: for each abstract description of a transaction, identify the most economical governance structure" (Williamson 1979, p. 234).

that it provides a measure of institutional efficiency. However, the use of trans-actions costs terminology risks clouding the issue of enforcement. To illustrate, consider an agency relationship between a manager and the workers within a firm. The sale of the agent's labor services involves a fundamental problem of exchange: the decision of whether to work hard is made by the agent, but it affects the welfare of the principal. Given this fundamental agency problem, different institutions will give rise to different patterns of behavior. The explicit and measurable transactions costs in such a setting might include the costs of hiring a manager to monitor the workers and measure their performance, as well as the costs of designing an organi-zation so as to enable this monitoring to occur, choosing a production process which facilitates such monitoring, installing surveillance equipment, and the legal costs of negotiating employment contracts, and suing or firing a shirker; and so on. In addition, if in the end it proves too costly to motivate the worker to act as she would in a first-best (zero transactions cost) world, then the resulting inefficiency would be another (implicit) transaction cost.

But while the concept of "transactions costs" can serve as a handy shorthand to describe how well these problems are solved, all of these "costs" ultimately derive from the agency problems and information asymmetries which give rise to the fun-damental problem of exchange in the (potential) transaction of interest. By separat-ing the "costs" of running the economic system – monitoring, enforcement, and so on – from the system itself, the institutions-as-rules approach clouds the issue of *why* people act as they do, and becomes a poor analytical substitute for an account of how behavior is actually motivated within alternative institutional regimes, none of which will approximate the zero-transactions-cost ideal. That is, the problem of designing efficient institutions is not fundamentally a problem of choosing rules so as to minimize "costs", but a problem of aligning incentives in a way which gene-rates the maximum possible benefit, given a fundamental problem of exchange. Higher efficiency (or a lower transaction cost) is a desired outcome of a successful solution to this problem, but it is not the problem itself, and focusing on transactions costs as a catch-all minimand risks masking the essence of the problem, which is one of aligning incentives.

# 3    Institutions as Rules: Applications

## 3.1    *Communities and Networks*

Community enforcement refers to a situation in which behavior within a group is governed by "rules" which are enforced by the members of the group themselves rather than a specialist third-party enforcer. One view holds that these kinds of informal rules are best taken as part of a fixed, exogenously-given cultural heritage (Williamson 2000). Other authors, however, consider that informal rules continu-ally adapt and evolve. For example, based on his studies of cattle farmers in Shasta county and New England whalers, Robert Ellickson (1991) hypothesizes that groups

within which information (gossip) circulates easily and informal power is broadly distributed will tend to develop efficient informal rules. Ostrom (1990) found that many communities manage to develop rules to successfully avert the tragedy of the commons in the management of common-pool resources, such as fisheries, forests, and common pasture. Other communities, however, do not, and Ostrom found that successful rules were more likely to emerge in groups with small numbers of decision makers, long time horizons, and members with similar interests.

As communities become larger, therefore, both Ellickson's and Ostrom's studies suggest that informal community enforcement is less likely to be able to support efficient outcomes. For example, as the online community of traders on eBay grew in the late 1990s, the "trust" sustained by a multilateral reputation mechanism based on user feedback had to be gradually supplemented by formal rules developed by eBay to discourage cheating, resolve disputes, and prevent illegal trades (Baron 2001).

## 3.2   Organizations

Organizations are akin to artificial communities of individuals brought together for a specific purpose – such as production, political activity, religious worship, recreation, and so on. While some organizations may begin as informal groups whose members later decide to develop a formal governance structure, others are created de novo by "entrepreneurs" with a goal in mind. As such, organizations are both cohesive entities which impact and interact with the broader world around them, and governance structures which develop formal rules to govern the interactions among their members and between members and outsiders. Within the institutions-as-rules framework, different authors have focused on each of these two aspects of organizations.

Some authors, notably Douglass North, have treated organizations primarily as unified entities that interact with the broader economic and political system within which they are embedded, and in particular, may act as "players" of the political game, attempting to alter broader institutional rules for the benefit of their members. This aspect of organizations will be discussed in Sect. 3.3 ("Politics"). The other aspect of organizations – their internal governance – is studied in economics primarily in the guise of the theory of the firm.

As is well known, the modern theory of the firm originates with Coase's (1937) insight that organizations and markets are alternative modes of organizing transactions, and the claim that the scope of activity carried out within organizations will therefore be determined so as to minimize "transactions costs". To explain the structure of an organization, therefore, we need to explain its function: what contractual problem it efficiently solves. But why would efficient organizations emerge? One possibility is that the structure of organizations is a product of rational design. If the organization's creators have a correct understanding of the effects of different organizational forms, then it may be reasonable to assume that they will design efficient organizations.

However, an alternative explanation for the emergence of efficient organizations is that evolutionary pressure forces firms to select efficient organizational forms by driving less-efficient organizations out of business. Alchian (1950) was an early proponent of this view, and it also implicitly underlies Williamson's "Transactions Cost Economics", which assumes that organizations (governance structures) will develop so as to achieve an optimal (efficient) match with the transactions they govern. The evolutionary approach has the advantage, noted by both Alchian and Williamson, that it enables us to assume that efficient institutions will develop even if the people designing them are boundedly rational. If a parameter change, such as a change in technology, renders existing institutions inefficient, then over time, by accident or design, some firms will develop more efficient sets of rules ("governance structures"), and through competitive pressure, these new institutions will gradually spread, so that the institutions governing the relevant transaction will evolve toward optimality.[7] Thus, the usefulness of the rational-choice framework does not rest on an assumption of rationality.

The validity of this approach, however, rests on the implicit assumption that there are deeper underlying institutions that lead to the selection of optimal (efficient) institutions. The issue of what exactly these underlying institutions are is frequently left unexplored, and thus the analysis can offer only a partial explanation for the observed configuration of rules. Nevertheless, for the purpose for which it was developed, namely examining the governance structures of firms operating in competitive markets within a modern economy, this approach works well and is an "empirical success story" (Williamson 2000, p. 607).

The assumption that organizations are organized efficiently (whether through evolution or design) also underpins much of the modern theory of the firm, including the literature on principal-agent problems within the firm, which studies how management can design optimal incentive systems to motivate workers; the property-rights approach following Hart (1995), which postulates that the boundary of the firm (ownership of assets) is determined in such a way as to minimize the inefficiencies which result from the inability to write complete contracts; and the theory of mechanism design.

Informal rules and norms, such as a "corporate culture", may also develop within organizations, including firms. The internal governance of organizations typically involves a combination of both formal and informal "rules". For example, one approach to overcoming the principal-agent problem between management and workers within a firm is through optimal wage and bonus structures based on contractible output measures. However, an alternative way to motivate worker effort, given the repeated nature of the relationship, is via the threat of firing a worker caught shirking (Bowles and Gintis 1993). While the formal contract, according to which

---

[7]Nelson and Winter (1982) built an evolutionary theory of the firm based on the evolution of routines – sequences of action which coordinate the activities of many individuals – rather than rules. Routines evolve as successful firms expand and their routines are imitated – perhaps imperfectly – by others, creating a tendency towards the adoption of efficient routines (although possibly with considerable inertia).

the worker is paid a wage for showing up to work, regardless of her effort – is enforceable in the courts, the worker's effort level is not contractible, and so the employment relationship is governed by both formal and informal rules: high effort is enforced informally through threat of non-renewal of the formal contract.

## 3.3 Politics, Informal Rules, and Institutional Change

The state, of course, is the most important source of formal rules, including laws, constitutional rules and decrees passed by representative bodies, voted on by citizens, or proclaimed by kings. Standard neoclassical economics assumes the existence of a well-functioning "state", and state activities such as taxation, regulation, and the provision of public goods are treated as well-functioning policy instruments in the hands of a benevolent policymaker. While this treatment of the state is useful for some purposes, it is woefully inadequate for others. It makes improbable assumptions about the state's ability to obtain and process the information needed to arrive at an optimal conclusion (Hayek 1945), and it obscures the fact that policy decisions are generally the result of bargaining and negotiation among organizations and individuals with divergent interests, and that implementing these decisions involves motivating and coordinating the organs of the state, such as regulatory agencies, courts, and the police.

A key function of the state, taken as a given in neoclassical economics, is to provide security of property rights and contract enforcement. In the absence of a state (anarchy), individuals must invest resources in the private production of security by acquiring a capacity for violence (Skaperdas 2006). The well-known Hobbesian justification for the creation of the state is that the presence of a higher authority enables people to replace the costly and inefficient spontaneous order of anarchy with a set of rules designed to improve overall welfare.

In Yoram Barzel's Hobbesian theory of the origins of the state (Barzel 2002), individuals begin in a state of nature without institutions, and they find it in their interests to create a state, as a monopolist of violence, to provide order. However, they wish to efficiently limit the state's scope of activity. This raises the question of why the state (which Barzel treats as a single actor) would obey the "rules" that its subjects create for it, rather than using its capacity for violence to expropriate those under its rule or expand the scope of its activity beyond that which is optimal. Barzel notes this danger, and postulates before the people create a state, they will also create collective-action mechanisms that constrain the state's actions by enabling them to overthrow the state if it becomes predatory. However, in keeping with the institutions-as-rules approach, Barzel treats the enforcement of these collective action mechanisms as exogenous.[8] As a result, the enforcement problem (keeping the state

---

[8]"Although the "social" arrangements used to enforce decisions by collective-action mechanisms seem to be of utmost importance, there is little that I, as an economist, can say about most of them. I simply assume that such arrangements exist and are put into use" (Barzel 2002, p.119).

honest) is merely pushed back one level; ultimately, the enforcement of the formal rules is taken as exogenous.

The problem of empowering the state to create order while constraining it from predation is of fundamental importance. Djankov et al. (2003) postulate that the "institutional design" of the state involves a fundamental tradeoff between "disorder" and "dictatorship": creating a more powerful state helps to reduce disorder and the risks of private expropriation, but at the cost of increasing the costs of dictatorship, corruption and expropriation by the state. Each society has a set of feasible combinations of dictatorship and disorder (an "institutional possibilities frontier"), which depends on a variety of societal characteristics including technology, culture, education, social capital, ethnic heterogeneity, history, factor endowments and the physical environment. In Djankov et al.'s basic model, societies choose an optimal political system (that is, one which minimizes the sum of the costs due to private and public expropriation) subject to the constraint of its institutional possibility frontier.

However, there are a variety of potential impediments to the selection of efficient political rules. Djankov et al. argue that countries which are former colonies might have inefficient rules if the rules were transplanted or imposed by their formal colonial masters rather than arising indigenously. La Porta et al. (2008) find that countries' legal origins affect economic outcomes. The civil law system, they argue favors a greater degree of state control and regulation, whereas the common law system relies more on market-supporting regulation and precedent-setting private litigation.[9]

Many authors emphasize that distributional conflict can lead to the selection of inefficient rules. For example, Libecap (1989) explores the development of the "property rights" rules that govern the use of a variety of resources such as fisheries, mineral rights (mining), and the use of public land. Different rules entail different distributional consequences, and individuals and groups therefore engage in bargaining, lobbying, and political action to try to alter the rules for their own benefit. As in Ostrom's schema, this rule-changing activity ("contracting") is itself a game governed by a higher level of political rules, and these higher-level rules, together with the activities and perceptions of the actors therefore shape the direction of institutional change of the lower-level (property rights) rules.

Acemoglu (2003) and Acemoglu and Robinson (2006) emphasize the importance of commitment problems as an impediment to the selection of efficient rules. Political incumbents might be willing to make concessions to disenfranchised groups in order to avert a costly or violent revolution, but if they cannot credibly commit themselves to honor their commitments to reform after the moment of crisis is passed, then whenever groups have the opportunity, they will seize power and craft rules to benefit themselves without regard for the other groups.

---

[9]See, however, Hadfield (2008), who casts doubt on the importance of the civil-law/common-law distinction, and provides a richer and more refined alternative set of key parameters for the classification of legal regimes.

Even if a society does initially select rules which are "efficient" in a static sense, these rules may ultimately turn out to be suboptimal in a dynamic sense. For example, Engerman and Sokoloff (2002) argue that the soil and climate in Europe's South American and Caribbean colonies were suitable for the production of cash crops, such as sugar, that could be efficiently produced on large slave plantations, resulting in highly unequal distributions of wealth, income, and human capital, which in turn enabled the elites to establish legal and political institutions that promoted their interests. In the North American colonies, in contrast, the initial factor endowments were more favorable to the production of crops and livestock that could be efficiently produced in small family farms. This led to the development of more egalitarian and democratic political institutions, higher levels of public goods provision (such as primary schooling), and greater levels of social mobility. Acemoglu et al. (2001, 2002) tell a related story, but with the focus on the disease environment and indigenous population density rather than soil and climate as the key exogenous variables explaining the initial development of state institutions. In places where Europeans found settlement difficult, they created "extractive states" aimed at transferring resources to the mother country. In areas more conducive to European settlement, they found it more profitable to build institutions aimed at protecting private property and encouraging investment. These institutions persisted even after independence, and led to a "reversal of fortune" in the nineteenth century, because regions that had previously been poor inherited institutions that later enabled the societies to industrialize.

These arguments give history a role in explaining the scope and functioning of the state. Institutions developed as an efficient response to circumstances in one time period may persist even if they later become inefficient. But why do institutions persist? Again, the basic answer within the institutions-as-rules approach is due to North (1990), who developed a theory of institutional change that combines deliberate changes in formal rules with evolutionary change in informal rules. In North's theory, given the current structure of formal and informal rules, entrepreneurs form organizations to take advantage of perceived opportunities. Over time, as they acquire skills and knowledge, they may find it worthwhile to attempt to change the structure of formal rules. When changes in formal rules occur, then the informal rules which "had gradually evolved as extensions of previous formal rules" (p. 91) adjust in response, and the end result "tends to be a restructuring of the overall constraints – in both directions – to produce a new equilibrium that is far less revolutionary" (North 1990, p. 91).

Thus, North argues that because of the persistence of organizations and informal rules, overall institutional change is "overwhelmingly incremental" (North 1990, p. 89), and that institutional change is a path-dependent process: "the consequence of small events and chance circumstances can determine solutions that, once they prevail, lead one to a particular path" (North 1990, p. 94). Current institutions provide incentives to create particular kinds of organizations and to invest in particular kinds of skills and knowledge. They also affect the distribution of wealth and political power, the preferences of the actors, and the stock of physical and human capital. All of these endogenous parameter changes in turn affect the costs and benefits of

alternative institutions, people's perceptions of new possibilities, and their ability to bring about or stifle institutional change. In all these ways, past institutions can influence the direction of institutional change (Libecap 1989; Pierson 2000; North 1990, 2005).

Building on North's work, a growing recent literature considers processes of institutional change that explore the interaction between formal and informal rules. In Roland (2004)'s theory, informal rules ("slow-moving institutions") are constantly evolving, and if these changes become incompatible with existing formal rules, then pressure for change builds up, leading to periodic abrupt and substantial changes in formal rules ("fast-moving institutions"). Brousseau and Raynaud (2008) build a theory in which new rules begin as informal, local and flexible orders, which compete for voluntary adherents. Successful rules spread, and as they spread, they become increasingly global and mandatory and "harden" into rigid formal rules. Aldashev et al. (2007) show that changes in formal rules can alter outside options and therefore bargaining power within informal relationships, and thereby shift customary informal rules in the direction of the formal law, even if it is never explicitly used.

One difficulty which arises in thinking about institutional change in this way – as an interaction between "formal" and "informal" rules – is that the nature of the "informal rules" is often left rather vague, and how they interact with formal rules – for example, which rule is followed when the two kinds of rules conflict – remains unclear. As noted above, the institutions-as-rules approach treats the question of how rules are enforced, and therefore why they are followed (or not followed), as a separate issue from their content. Thus by definition, if behavior does not conform to formal rules, by default it is attributed to – and assumed to be governed by – unobserved informal rules. Yet, since informal rules are generally implicit, it is hard to observe what these informal rules are, whether in fact they are indeed being followed (and if so, why), and what kinds of behavior they are affecting, and in what way. Attributing unexplained behavior to informal rules therefore amounts to a leap of faith that invokes a mysterious and scientifically untestable explanation for the observed behavior.

The problem is compounded by the fact that the term "informal rules" has been used to describe several quite distinct phenomena. Some authors treat informal rules as internalized "ethical" codes of conduct which are directly reflected in players' preferences (e.g., Ostrom 2005). For others, informal rules are rules which are not written down, or which are not enforced by the state. Still others identify informal rules as self-enforcing codes of conduct, shared cultural "focal points", or as "social norms" enforced within a community using a multilateral reputation mechanism – or as all of these things, as the occasion demands. For some (e.g. Williamson 2000), informal institutions change only over a period of centuries or millennia, so they may safely be taken as exogenous and fixed, while others, such as Roland, hold that gradual changes in informal rules are often an important part of the story of institutional change.

Ultimately, therefore, the institutions-as-rules approach is limited in its ability to explain institutional change because a key element inhibiting and shaping the

direction of institutional change, informal rules, originates outside the analytical framework. For example, Ostrom (2005, p. 138) notes that "many written statements have the form of a rule ... but ... do not affect behavior. Such statements are considered rules-in-form rather than rules-in-use." Yet, because she treats the enforcement of rules separately from their content, any explanation of what makes some rules "rules-in-use" while others rules are ignored is outside her framework. She notes that "in settings where a heavy investment is not made in monitoring the ongoing actions of participants... considerable difference between predicated and actual behavior can occur," (p. 21), but achieving this monitoring and enforcement is treated simply as a cost; the incentives of the monitors are not examined.

## 4   Self-Enforcing Institutions, or "Institutions-as-Equilibria": Conceptual Issues

The core idea in the institutions-as-equilibria approach is that it is ultimately the behavior and the expected behavior of others rather than prescriptive rules of behavior that induce people to behave (or not to behave) in a particular way. The aggregated expected behavior of all the individuals in society, which is beyond any one individual's control, constitutes and creates a structure that influences each individual's behavior. A social situation is 'institutionalized' when this structure motivates each individual to follow a regularity of behavior in that social situation and to act in a manner contributing to the perpetuation of that structure.[10]

The focus on regularities of behavior and the motivation to follow them responds to the observation that these factors, rather than rules, are the direct cause of distinct welfare-related outcomes. The corruption plaguing many political systems in the world is not caused by an absence of rules prescribing preventive measures. It is due to particular regularities of behavior.[11]

Focusing on motivation has the key advantage of avoiding the conceptual difficulties that come with treating institutions as rules. For example, the legal speed limit on highways in Massachusetts is 65 mph, but this limit is widely ignored. This is not to say that there are no "rules", however. Police officers do sometimes pull over cars traveling at 85 mph, but they never pull over cars traveling at 68 mph. What accounts

---

[10]This idea builds on the 'conventions' literature (e.g., Sugden 1989). See also Schotter (1981), Calvert (1995), Aoki (2001), Dixit (2004), Kingston and Caballero (2009) and Greif (1994, 1998).

[11]A regularity of social behavior does not imply uniformity of behavior as it is a characteristic of aggregates of individuals and not of each individual. Furthermore, social behavior is usually conditional on social roles and does not necessarily imply the same behavior by individuals with the same role. The behavioral regularity of 'males propose to a female and only when they can support a family,' for example, captures gender roles and implies that some males will never marry and the ages of those who do, will vary. Similarly, regularity of behavior is not necessarily frequent behavior. The process of impeaching a US president is regularized although rarely employed.

for the difference between the behavior specified by the "formal rule" and the behavior actually observed? From the institutions-as-rules perspective, the standard answer would be that the police and motorists must be following an "informal rule" – for example, that the true speed limit is 75 mph. But this invokes an exogenous and ad-hoc explanation for precisely what we would most like to explain.

Focusing on motivation complicates the analysis, however. One reason is that regularities of behavior are often caused by the net effect of multiple, and possibly conflicting, motivating factors. The fear of legal sanctions might motivate a teenage driver to slow down, but social pressure from his peers might have the opposite effect.

The evolving institutions-as-equilibria approach has not yet converged on an agreed definition of institutions. On the one hand, Calvert (1995), for example, literally equates institutions with game theoretic equilibria. "There is, strictly speaking, no separate animal that we can identify as an institution. There is only rational behavior, conditioned on expectations about the behavior and reactions of others. . . "Institution" is just a name we give to certain parts of certain kinds of equilibria" (pp. 22–23). The premise of this definition, however, is too restrictive. Game theory provides little guidance for identifying institutions or studying their dynamics. Greif (2006, Chaps. 2 and 5) defines an institution as a system of 'institutional elements,' particularly beliefs, norms, and expectations that generate a regularity of behavior in a social situation. These institutional elements are exogenous to each decision-maker whose behavior they influence, but endogenous to the system as a whole. The social 'rules' which emerge correspond to behavior which is endogenously motivated – constrained, enabled, and guided – by self-enforcing beliefs, norms and expectations. In addition, for an institution to be perpetuated, its constituent elements must be (1) confirmed (not refuted or eroded) by observed outcomes (2) reinforced by those outcomes (in the sense that its ability to be self-enforcing does not decline over time) and (3) inter-temporally regenerated by being transmitted to newcomers.

## 4.1 Self-Enforcing Expectations and Motivation

An empirically-oriented analysis relying on the institutions-as-equilibria approach focuses primarily on motivation provided by self-enforcing expectations (behavioral beliefs). Such an analysis usually begins by identifying the 'essential' physical, technological and social attributes necessary for the situation to be of interest. In the case of regularities of behavior among drivers, for example, essential attributes include that there many drivers who have property rights (or user rights) in cars, can benefit from driving compared to alternative modes of transportation, can observe other cars, and prefer to avoid accidents. Without any of these features, considering driving behavior is meaningless. Similarly, the analysis would be too general without more narrowly delineating the regularities of behavior we are interested in: is it the direction of traffic, priority-rules at intersections, speeding or passing?

By focusing on a situation's essential attributes, we initially set aside those potentially relevant social constructs that we initially wish to treat as exogenous to the analysis. In the case of driving, these might include such constructs as drivers' licenses, socialization to drive carefully, or a Highway Patrol Agency with the capacity to impose legal sanctions. Initially ignoring such potentially important constructs is not a statement about their irrelevance but a means to analytically examine whether they are relevant, why they are relevant, and to what effect.

The next step in the analysis is to focus on the set of self-enforcing expectations and the implied behavior that can prevail in this situation, by modeling the situation as a game (specifying the set of players, their possible actions, the order of moves, information, and payoffs) and finding equilibria. By "self-enforcing expectations" we mean that if the decision-makers share the expectation that others will generally follow the equilibrium behavior, then each of them will be motivated to follow it as well (the Nash criterion). From each decision-maker's perspective, the others' expected behavior constitutes the structure motivating her to conform to the behavior expected of her. But by conforming, she contributes to motivating others to conform too. Thus, the structure is self-perpetuating, and although it is beyond the control of each decision maker, it is endogenous to all of them taken together. Note that the "self-enforcing" requirement includes expectations about how others will behave in situations that would not transpire in equilibrium. For example, if a player does not steal because of a fear of punishment, the 'off-the-path' expectation of punishment must be credible (this the sub-game perfection requirement).

Having found equilibria in the minimal game, we can next examine how various social constructs can change the set of self-enforcing expectations by changing the expected responses by other players to particular actions. When these expectations are credible, the costs and benefits associated with actions in the minimal game are changed, and the set of potentially self-enforcing behaviors may be enlarged. For example, the creation of a "group" can create restrictions on entry to the situation (who the participants are) and change the pattern of relations (e.g., repeated interactions among the same individuals). Other kinds of social constructs might alter the information structure, or introduce a new actor with the ability to punish or reward players (e.g., a judge).

The introduction of new social constructs can change people's expectations (and therefore incentives and behavior) in many ways. Sanctions can be coercive (such as violence or imprisonment), social (such as ostracism), or economic. Guilt and the fear of expected punishment in the after-life are other means to link past actions to future rewards. The institutions-as-equilibria approach focuses on how such expectations are formed, why, and to what effect. Note that this involves much more than just the introduction of new "rules". In order to shift people's expectations, cognitive categories (e.g., "honesty", "cheating") need to be coordinated upon so that all players share coordinated expectations about punishment. If the desired behavior is to be self-enforcing in the modified, extended game that includes the new interactions, then the punishment should be sufficiently costly to make deterrence effective. Those who are to retaliate must have the information about who and when to punish, which potentially includes motivating those who know about the transgression to

inform others. People must also be motivated to punish, as the expectation that punishment will be inflicted has to be credible. Furthermore, they have to have the physical capacity to punish and those who are to be punished should not be able to evade punishment.

## 4.2 Rules and Organizations in the Institutions-as-Equilibria View

In the institutions-as-rules approach, rules are institutions and institutions are rules. Rules prescribe behavior. In the institutions-as-equilibria approach, the role of "rules", like that of other social constructs, is to coordinate behavior. Because there are multiple potentially self-enforcing expectations in a given situation, coordination mechanisms, including rules, play an essential role in generating regularities of behavior and social order. Rules fulfill this coordinating role by specifying patterns of expected behavior, and also by defining the cognitive categories – signs, symbols, and concepts – on which people condition their behavior. Actions have to be given meanings because, for example, 'cheating' is not naturally defined, but it must be defined before it can be discouraged. A road sign instructing a driver to yield at a pedestrian crossing has meaning, and motivates behavior, only because it is a component of a system ("rules of the road") that motivates behavior based on road signs.

The behavior that people can be motivated to follow depends on these cognitive categories and on the rules' ability to coordinate expectations based on these categories. Focusing on motivation exposes the limits on the realities that humans can use rules to construct. In order for a "rule" to matter, the behavior must be self-enforcing and it must be conditioned on observable aspects of the situation. If drivers cannot observe a pedestrian's age, they cannot condition their behavior on it. And it must be sufficiently costly to circumvent the categories. For example, a rule which conditions behavior on gender may not be self-enforcing if males can easily pretend to be females and vice versa.

Of course, the behavioral expectations and cognitive categories which people actually use to coordinate their behavior may be quite different from those specified by 'formal rules'. Nevertheless, we observe that explicit "rules" are often formalized and disseminated in a centralized manner. From the institutions-as-equilibria perspective, the creation of such formal "rules" can be interpreted as an *attempt* to achieve a coordinated shift of many people's expectations, while convincing the agents that these expectations are indeed widely shared. This mechanism can also be used, of course, to serve the interests of the politically powerful – those with the power to change formal rules. But if the new rules do not specify a self-enforcing pattern of behavior, they may not have their desired effect.

Organizations, too, are social constructs that change the set of self-enforcing expectations among the agents in the original interaction. Formal organizations, such as parliaments and firms, and informal organizations such as communities and

business networks, have a dual role both as institutions that govern their members' behavior, and as institutional elements within the broader institutions of society. Within the group, an organization can change the relevant rules of the game, such as information, actions, and payoffs sets, and can therefore increase the credibility and severity of sanctions, specify rules, and create shared knowledge. Organizations may also play a role in attempting to shape the preferences of community members, particularly children, through a process of socialization.

An organization and its members also interact, individually and as a group, with the outside world, and the beliefs, norms and expectations that govern the internal interactions between the members of the organization will often differ from those governing their interactions with outsiders. A police force, for example, has internal structures and rules to govern its members' behavior, but it also acts as an organization for enforcing other rules set by the government of the society of which it is a part. The reliance on organizations is fundamentally due to the fact that organizations have capabilities that are more than the sum of the individual capabilities of their members, due to their ability to coordinate their members' activities, economies of scale and scope in their efforts to change the rules of the game, and due to the organizations' longer time horizon and memory.

## 5 Institutions as Equilibria: Applications

### 5.1 Markets and Networks

The ability to engage in voluntary exchange encourages production, specialization, and innovation, and is a key prerequisite for economic efficiency. However, in all but the simplest market exchanges, enforcement problems arise due to the "fundamental problem of exchange" (Greif 2000). For example, in labor market and credit market transactions, there is an unavoidable separation between the quid and the quo, and at least one party therefore may have an opportunity to "cheat" the other: a borrower may choose not to repay a loan, or an employee may choose not to work hard. Markets can function only when this fundamental problem of exchange is overcome.

Neoclassical microeconomics tends to either assume away enforcement problems, or to take the presence of well-functioning market-supporting institutions as given. The institutions-as-rules approach is an improvement, in that it considers how market exchange may be supported by institutions (rules) that punish defection, such as a legal system, or informal codes of conduct. For some purposes this is adequate. However, these rules themselves ultimately require enforcement. For a convincing account of the institutional foundations of markets, therefore, we need to consider the enforcement not just of the market transactions themselves, but of the rules which govern those transactions. Saying that there is a norm against cheating, for example, is insufficient. It is critical to study how the norm is sustained as part of

a self-enforcing equilibrium outcome of a game in which the enforcement of the norm results from the behavior, and expected behavior, of the players. The same is true if enforcement is carried out by formal or informal organizations, such as a court, a credit bureau, or a community.

In situations where formal institutions do not function well, economic agents may rely heavily on informal enforcement. McMillan and Woodruff (1999), for example, found that in Vietnam, firms scrutinize prospective trading partners carefully, and rely on informal "private order" institutions, including both bilateral enforcement (ceasing to do business with a firm that cheats) and multilateral community enforcement (sharing information about cheaters). However, these mechanisms are a poor substitute for well-functioning formal enforcement mechanisms. McMillan and Woodruff do not investigate how the behavior which sustains these private order institutions is made self-enforcing: why do people share information and punish cheaters, given that doing so is costly?

Kandori (1992) uses a game-theoretic model to show that among a community of players who are randomly matched into pairs each period to play a prisoner's dilemma, a multilateral community enforcement mechanism can support cooperation if the players can observe a label which indicates (roughly speaking) whether their current trading partner is "a cheat", and which is honestly updated through some exogenous process. Several papers study how this kind of reputational information might be shared within a community. Gazzale (2005) shows that players may have an incentive to gossip because a reputation for gossiping can deter their future trading partners from cheating. Greif (1989, 1993, 1994) shows how information transmitted in correspondence among a commercial and social network of medieval traders (the Maghribi traders) supported a reputation mechanism that successfully dissuaded cheating. Merchants who cheated could expect that their actions would be widely reported within the network. Since merchants who were ostracized from the network for cheating had no further reputation to lose, they would be expected to (rationally) cheat in any future transactions; and therefore, each merchant in the network was motivated to punish cheaters by the expectation that others would also do so, so the punishment was self-enforcing.

Note that this approach directs attention away from the content of the "rules" about cheating to the networks and information flows that enable the expectation that other players will punish cheats to be sustained as part of a self-enforcing equilibrium. While the presence of these information-sharing structures (networks and communities) themselves can be taken as exogenous in the short run, if the informal punishment mechanisms sustained by the community are to survive as institutions, the maintenance of such networks must also ultimately be made endogenous to the analysis. In many cases, such as the Maghribi traders studied by Greif, the structure of the network is in large part an outcome of an historical process. However, even networks that are bequeathed by history need to be maintained. One way to study the origins and stability of such networks is to consider group members' incentives to retain their affiliation ex post, by submitting to punishment rather than attempting to evade it (Greif 1993). Another approach is to consider the networks themselves as an outcome of a prior game in which players (individuals, firms, or countries, for

example) deliberately form links with other players (friendships, supply links, or military alliances, for example) (Fafchamps 2004; Jackson 2006). In general, it is not necessarily the case that the networks that emerge through such a process will be "efficient" from the point of view of society overall, because individuals choosing to build links do not take into account the external effects of those links on other players. Greif (1993, 1994, 2006) emphasizes that when the group is an outcome of a historical process, there may not be a mechanism to coordinate inclusion of new members. Inefficient size is the likely outcome.

Even if formal enforcement mechanisms are available, they may not be employed in equilibrium. Kranton (1996) studies a model in which agents choose between trading within an informal network or in an anonymous market in which agency problems are absent – for example, because there is a well-functioning formal system to govern market exchange. The value of market interaction depends on the fraction of players who choose to buy and sell within the market. Therefore, two equilibria arise: if everyone makes use of their informal relationships to obtain goods, then the market is thin, the search costs of finding a trading partner in the market are high, and each individual has an incentive to use her network rather than the market to obtain goods. However, if instead many people choose the market, then finding a trading partner in the market becomes relatively easier, and the informal relationships break down as players' outside option improves. Therefore, either pattern of behavior – a market in which people carry out ephemeral, anonymous transactions, or a pattern of trading within long-lived, "trusting" informal relationships – can emerge as an institution corresponding to an equilibrium of the game. If the society begins in one equilibrium, in the absence of exogenous shocks, or some coordinating mechanism to engineer a coordinated shift in behavior, it may remain stuck in that equilibrium even if the alternative equilibrium would be more efficient. Thus, starting points matter, and history plays a role in equilibrium selection.

Kranton's analysis takes the market as given, but the presence of legal contract enforcement may itself be treated as endogenous. Greif (1994, 2006) has argued that in medieval Europe, the Genoese society characterized by individualistic cultural beliefs and interest-based communities experienced a higher demand for legal contract enforcement than the collectivist cultural beliefs and kin-based community of the Maghribi traders. As a result, the Genoese developed formal institutions including codified contract laws, double-entry book-keeping, family firms, bills of lading, and other antecedents of modern business practices.

## 5.2 Organizations

As the size of a community grows, relying solely on informal governance may tend to become problematic as players' ability to observe each other's actions, and to share information about transgressions, diminishes. As a result, there is a tendency to evolve from informal to more formal modes of governance.

For example, at the medieval Champagne Fairs, large numbers of merchants from all over Europe congregated to trade. Merchants from different localities entered into contracts, including contracts for future delivery, that required enforcement over time (Verlinden 1979; Milgrom et al. 1990). There was no state to enforce these contracts, and the large number of merchants as well as their geographic dispersion made an informal reputation mechanism infeasible. Greif (2006) argues that impersonal exchange was supported by a "community responsibility system". Traders were not atomized individuals, but belonged to pre-existing communities with distinct identities and strong internal governance mechanisms. Although particular traders from each community may have dealt with merchants from another community only infrequently, each community contained many merchants, so there was an ongoing trading relationship between the communities, taken as a whole. Merchants from different communities were able to trust each other, even in one-shot transactions, by leveraging the inter-community "trust" which sustained these interactions. If a member of one community cheated someone from another community, the community as a whole was punished for the transgression, and the community could then use its own internal enforcement institutions to punish the individual who had cheated.

This system was self-enforcing. Traders had an incentive to learn about the community identities of their trading partners, and to establish their own identities so that they could be trusted. The communities had an incentive to protect the rights of foreign traders, and to punish their members for cheating outsiders, so as to safeguard the valuable inter-community trade. Communities also developed formal institutions to supplement the informal reputation mechanism and coordinate expectations. For example, each community established organizations that enabled members of other communities to verify the identity of its members. Ultimately, the growth of trade that this institution enabled created the impetus for its eventual replacement by more formal public-order (state-based) institutions which could directly punish traders by, for example, jailing them or seizing their property.

However, public order never entirely replaces private order: markets in modern economies contain a mix of private-order and public-order institutions, and transactions may rely on both (Greif 2006). For example, a lender may obtain a credit report on a prospective borrower from a private credit-rating firm that lacks any enforcement power beyond the ability to share information, but it may also rely on state enforcement to seize the borrower's collateral if she fails to repay.

As noted in the previous section, within the institutions-as-rules tradition, it is widely held that the boundary between firms and markets as alternative modes of organizing transactions will be determined so as to achieve efficiency (minimize transaction costs). From the perspective of the institutions-as-equilibria approach, organizations appear as components of broader equilibria, interacting with other institutional elements, and constrained by the past. In particular, if there are many possible equilibria, then there may be different configurations of organizations associated with each of these equilibria, and the structure of organizations, including the boundaries between firms and markets, cannot therefore be deduced from a knowledge of the characteristics of the transaction alone.

For example, during the eighteenth century the institutions governing marine insurance transactions developed in different ways in different countries. In Britain, a coffee house (Lloyd's of London) gradually developed into a sophisticated marketplace for underwriting by private individuals who "under-wrote" the amounts they were willing to insure on policies presented to them by merchants or brokers. In contrast, in France, Holland and the US, private underwriting disappeared and was replaced by joint-stock corporations. The underlying marine insurance transactions were plagued by serious information asymmetries and agency problems, including the potential for various kinds of fraud. Each institutional form – private or corporate underwriting – had advantages and disadvantages in dealing with these agency problems.

Kingston (2007, 2008) argues that the industry was characterized by multiple equilibria. In Britain, an equilibrium based on private underwriting became institutionalized over time through the development of specialized institutions, in particular Lloyd's coffee-house, which became a hub for information about ships and their crews, political and economic developments, and the many other factors affecting the risk of a voyage, and also for information about the reputations of market participants. This ultimately meant that in Britain, the corporations suffered a "lemons" problem because of their inferior access to information about vessels and other developments affecting the risk of a voyage, enabling the private underwriters to dominate the market. In contrast, in the American colonies, although private underwriting had been developing rapidly, it never reached the level of complexity of Lloyd's. Instead, in the late 1790s there was a shift between equilibria as private underwriting was extinguished by competition from joint-stock corporations (Kingston 2011). Kingston (2007, 2008) shows how the timing of a series of historical events, involving both exogenous shocks (such as war) and endogenous parameter changes and learning processes, drove the process of institutional change (equilibrium selection), leading ultimately to a path-dependent bifurcation of institutional structure between Britain and the rest of the world. Each equilibrium, once established, proved stable. Thus, although the fundamental purpose of the transaction – sharing risk – was the same, by the end of the Napoleonic wars, the manner in which the transaction was accomplished was very different in different countries. The "governance structures" that emerged were the outcome of a historical process with multiple stable end-points, rather than being designed to "minimize transaction costs".

## 5.3   Politics

As noted in the previous section, a fundamental rationale for the existence of the state is that it can use its capacity for coercion to provide order and security. Many contemporary societies face the challenge of building states that effectively promote political stability, curtail political violence, and foster economic prosperity. This has proven no easy task, despite the fact that copying formal rules, including

constitutions and laws, is relatively straightforward. Why don't countries with inefficient or ineffective political structures simply copy the institutional structure of more successful ones? And why do countries that "transplant" formal rules frequently find that this fails to reproduce the desired patterns of behavior?

Within the institutions-as-rules framework, the explanation offered is that behavior is constrained not just by formal rules, but also by the "informal rules" present in a society. But as we saw earlier, this leaves unanswered the question of where these informal rules come from, and how the "rules", including the rules governing the behavior of the state itself, are ultimately enforced. The key puzzle is how to construct a state that is strong enough to provide order and protect the rights of its citizens, but in which political power-holders are nevertheless motivated not to use this power to abuse those rights: in other words, all actors, including "rulers", must obey "the rules". Thus, a stable, well-functioning political system should be viewed as a desirable equilibrium outcome rather than as a set of rules.

Bates et al. (2002) study a model in which players can choose to allocate their effort among three goals: production, leisure, and arming themselves to engage in violence. A capacity for violence enables players both to defend their own productive output and to "raid" the output of others. In anarchy (a situation with no "state"), there are two kinds of equilibria: one in which there is little violence, but also little production, so that most effort goes into leisure; and another equilibrium in which there is production, but also a lot of violence, as people who produce must also be willing to defend their output. As in the Hobbesian vision, the creation of a state as a specialist in violence can improve efficiency by enabling players to produce without fear of being raided by others. In return for taxing the output of the players, the state undertakes to punish raiding. Bates et al. (2002), however, go beyond Hobbes in probing the incentives of the state itself; they show that the state can be constrained from predation by the shadow of the future, since a failure to protect the property rights of the citizens can lead to reversion to a "warlord equilibrium" in which no taxes are paid, and the players arm themselves not just against raiding by other players, but against state predation as well. Thus, the balance of coercive powers can sustain the state as an equilibrium.[12]

Weingast (1997) shows how the rule of law emerged as an equilibrium outcome of a game between a ruler and his subjects in seventeenth-century England. The king (James) initially supported the interests of the landed Tories at the expense of the mercantile Whigs, who lacked the power to overthrow this ruling coalition.

---

[12]Of course, real-world processes of state-building do not start from the "clean slate" envisioned by Hobbes. Bates (2001) argues that historically, monarchical states emerged out of competition among feudal lineages as rural, agrarian societies based on kinship networks became increasingly urbanized and industrial. Olson (1993) provides an alternative parable for the origins of the state, arguing that the state emerged as those with the greatest capacity for violence found it privately more profitable to use this capacity to provide order in exchange for tax revenue, rather than simply to live by plunder.

However, after the king began to infringe on the rights of the Tories, the Whigs and Tories combined to overthrow him, and installed a successor (William). At the same time, they created a new constitution with the aim of preventing the king from future predation. The new constitution, Weingast argues, was fundamentally a coordination device that laid out the conditions that would trigger a coordinated reaction by the citizens against the king in future. Thereby, it enabled a shift from an equilibrium in which the king was able to transgress the rights of the Whigs with impunity to one in which the Whigs and Tories undertook to jointly resist any transgression by the king against either of their rights.

To achieve this, the new agreement needed to be self-enforcing. Both groups of citizens had an incentive to abide by the agreement, as a failure to do so would enable the king to abuse the rights of both groups in the future. The king was motivated torespect property rights by the credible expectation that both groups would react in concert to an infringement of the rights of either group. Thus, as in Bates et al.'s model, while the players may articulate "rules" to govern their behavior, it is the ultimately the threat of a breakdown of cooperation in an infinitely-repeated game that enables a non-predatory state to be sustained as an equilibrium outcome.

Greif (2006, Chap. 8) studies the process of state-building in medieval Genoa. Genoa's commercial expansion had been hindered by the threat of conflict between two rival feudal clans, which led each clan to waste substantial resources defending itself from the other. To achieve gains from cooperation, the warring clans agreed to invite a non-Genoese ruler/administrator, the *podesta*, to rule the city. The *podesta* held the balance of power between the clans, but was not militarily strong enough to impose his will on them and become a dictator. To avoid the danger of the *podesta* aligning himself with either clan, he and his family were forbidden from involvement in Genoese society or politics. Indeed, the *podesta*'s position depended on ensuring that neither clan became dominant, or they would have had no further need of him. Thus, the *podesteria* system was a set of self-enforcing institutions that promoted inter-clan cooperation and reduced the threat of conflict. The process of institutional change was shaped by the initial conditions, including the set of organizations (in this case, clans), and the feudal rules, beliefs, and norms inherited from the past.

Scartascini and Tommasi (2009) study a model of policymaking in which individuals can either pursue their interests via the formal political process or through violence, protests, bribery, and so on. They show that there may be multiple equilibria: one equilibrium in which all players choose formal channels, and another in which some players use the formal process, but many players "go to the streets". Moreover, the stability of these equilibria is reinforced by actors' investments over time. Their model can account for the differences in observed political behavior between countries with similar "formal rules" (such as the U.S. and Argentina) but these behaviors are explained as equilibrium outcomes rather than by invoking differences in unobservable "informal rules".

# 6  Institutions-as-Equilibria: The Frontier

## 6.1  Beliefs, Culture, and Institutional Trajectories

In the institutions-as-rules perspective, beliefs influence behavior through their impact on the choice of rules. In North (2005)'s framework, economic agents have "mental models" which reflect their understanding of the world and which they use to evaluate the desirability of particular rule changes. Over time, as they learn about the world, they revise their mental models, and this may alter their perceptions about the net gains from alternative possible rules, or lead them to perceive new possible rules, leading them to try to change the rules. Thus, "the key to understanding the process of change is the intentionality of the players enacting institutional change and their comprehension of the issues" (North 2005, p. 3).

The institution-as-equilibria perspective emphasizes additional causal relations between beliefs and outcomes. Beliefs motivate people's behavior by influencing the perceived costs and benefits of various actions, including expectations about others' behavior. Beliefs matter because individuals have potentially limited and different information, knowledge, and cognitive understanding about the environment and the strategies of other players.

Rules provide one means for people to coordinate their beliefs. Consider, for example, the seemingly unnecessary law specifying the direction of traffic (drive on the left, or drive on the right). Such a traffic law provides new drivers (or those visiting from abroad) with the knowledge required to make an informed decision based on a minimal understanding of the system. Furthermore, because such rules specify self-enforcing behavior, agents are motivated to acquire knowledge of the rules and follow them.

The analysis of the processes through which rules aggregate knowledge and information is in its infancy. A notable contribution is Aoki (2007) who proposes that as an existing equilibrium breaks down, cognitively limited agents perceive that their former strategies are no long optimal, without necessarily understanding why, and begin to experiment with new strategies. As their behavior and expectations change, institutional change – a movement to a new equilibrium – occurs. Eventually, agents' strategies and belief systems are brought back into alignment with each other as mutually consistent components of a new institutional equilibrium.

Greif (2006) proposes a social, rather than individualistic process of learning and convergence. Agents respond to the expected behavior of others as articulated in a known rule of behavior (either formal or informal). The traffic-law specifying a speed limit, for example, constitutes a social rule, known to the drivers and to which each of them responds. Their responses lead them to choose a speed higher than the legal maximum. But this is not explained by asserting that there is a hidden "informal rule" specifying the observed behavior. Rather, as players observe behavior and outcomes over time, the institutionalized 'rules of the road' which develop reflect the dispersed beliefs and information of individuals responding to a 'structure' which is simultaneously created by their aggregated responses to the structure itself.

The institutions-as-rules approach makes a clear distinction between formal rules, which are created in particular by the state, and culture, which consists of informal rules formulated by society. In contrast, the distinction between 'rules' and culture in the institutions-as-equilibria approach is one of kind and not one of essence. Both rules and culture influence behavior by giving rise to shared beliefs, norms, and expectations that generate regularities of behavior. This facilitates studying the inter-relations between institutions and culture.

Clearly, neither culture nor institutions are immutable. There are many historical examples of rapid cultural change. Yet, to the extent that one associates culture with institutional elements that prevailed prior to state formation or emerge independently of it, distinct cultures can lead societies along distinct trajectories of institutional development. The perpetuation and implications of both institutions and culture depend on the context, unintended consequences, and historical contingencies such as the sequence of various exogenous events, leadership, and the outcomes of military conflicts. Cultural change and culture's impact are not deterministic, but a specific culture can render some institutional trajectories more likely than others.

Benabou and Tirole (2006) give an example of how multiple equilibria can result from the two-way interaction between shared beliefs and public policies. They argue that in societies where many people hold a belief in a "just world" – the belief that economic success is highly dependent on effort – these people will favor low levels of redistribution and low tax rates. These policies increase the reward to economic effort, giving people an incentive to adopt (and teach their children) the "just world" ideology. If, instead, people believe that luck plays more of a role in determining individual success, they may favor higher levels of redistribution, which dampens the incentives for high effort, confirming the bases for their beliefs. Note that these ideological beliefs are more than just a reflection of different institutional structures. They are a fundamental part of each equilibrium.

Greif (1994, 2006) showed how distinct cultural beliefs led to distinct developments of contract enforcement institutions among eleventh century Jewish merchants operating in the Muslim world and the Latin-Christian Genoese. Collectivism among the former fostered reliance on enforcement based on a multilateral reputation mechanism while individualism among the latter fostered enforcement based on bilateral reputation and the law. The latter's reliance on the law, in turn, was facilitated by the fact that it was a man-made law and not a divine law. This fundamental distinction in legal conceptions was not instituted by states or rulers but reflected the distinct historical processes through which Christianity and Islam emerged.

## 6.2   Moral Norms and Endogenous Preferences

A second frontier issue in the institutions-as-equilibria approach is the inter-relations among institutions and preferences. Although some aspects of individual preferences, such as those directly related to survival, are primordial and selfish,

other aspects of preferences are shaped by society because humans have other-regarding preferences and seek moral justification for their behavior. The crucial element here is internalized moral norms or values that individuals are psychologically motivated to follow. An internalized norm against stealing, for example, places a wedge between the net utility value of five dollars earned and five dollars stolen. Such moral norms based on intrinsic motivation are different from 'social norms' which rely on extrinsic motivation provided by the threat of non-legal punishments. Moral norms influence behavior directly through their impact on preferences, and indirectly by influencing the expected behavior of those who are perceived to have internalized such norms.

People are born with the capacity and the propensity to internalize norms, and absorb norms through socialization by role models, parents, peers and organizations (such as schools and churches). Institutions can influence norms through their impact on these socializing agents. Tabellini (2008) provides a wonderful analysis of how norms of generalized or limited morality can evolve in the same situation depending on the incentives institutions provide to parents. Specifically, a parent faces a trade-off between socializing her child to have her norms or socializing the child to have the norms which would be optimal for the child in the future. Institutions influence this trade-off.

Akerlof and Kranton (2005) discuss the value of intrinsic motivation within organizations. They use the military as a compelling example of an organization whose members are primarily motivated by non-monetary incentives (such as honor). Such a theory of motivation has important implications for organizational design. For example, employing a supervisor to monitor a worker's effort may enable the firm to motivate the worker using high-powered monetary incentives, but Kranton and Akerlof argue that there is also a hidden cost: hiring the supervisor may also reduce the employee's sense of identification with the firm and its goals, thereby eroding the firm's "motivational capital" (worker's loyalty).

Akerlof and Kranton's work is related to an emerging literature on "endogenous preferences", much of which uses evolutionary arguments to investigate the role institutions play in molding not just people's behavior, but also their goals (Bowles 1998). These theories emphasize that while institutions, being man-made, are created through human action (whether intentional or not), institutions also play a role in reconstituting the goals and perceptions of the individuals they govern. The integration of these considerations in the institutions-as-equilibria perspective is in its initial stages. If successful, it will improve our ability to study norms as one element in a larger system in which people are moral, yet materialistic, and motivation is provided by endogenous beliefs, norms, and expectations.

## 6.3   *Origin, Dynamics and Complementarities*

The institutions-as-rules approach, as we have seen, studies institutions as (exogenous) constraints (rules) leading to (endogenous) behavior, while enforcement of the rules is treated as a separate issue. Institutional dynamics is fundamentally about

changing rules, and the analytical focus is on changing formal rules. In contrast, the institutions-as-equilibria approach focuses on the behavioral manifestations of endogenous motivation; how (endogenous) behavior generates (endogenous) institutions that perpetuate this behavior. Institutional dynamics is therefore fundamentally about changes in motivation and regularities of behavior, and the analytical focus is on changes in beliefs, norms, and expectations (Greif and Laitin 2004; Greif 2006).

Two causes of institutional change are particularly important. The first is an intentional attempt to bring about change by those who realize (or hope) that they can benefit from it. This kind of institutional change can result from the perception of new institutional possibilities, perhaps brought about by learning or by new interactions with outsiders. Because the existing institutions are equilibria, however, they generally cannot be changed unilaterally by a single actor. Bringing about a change may therefore involve overcoming collective action problems as well as overcoming the opposition of those who stand to lose from the change. Such collective action may occur through persuasion or through the use of new or existing organizations, or, less commonly, through the rise of a charismatic leader.

The second main cause of institutional change is "institutional disequilibrium" which results when an institution ceases to be self-enforcing. This can occur either due to exogenous shocks or due to endogenous changes in "quasi-parameters" (Greif and Laitin 2004): variables which change gradually over time as a result of the operation of the institution itself, and may ultimately pass a critical threshold so that the institution ceases to be self-enforcing. Whether the subsequent institutional change is gradual or abrupt, evolutionary or intentional depends, in particular, on whether the actors are cognitively aware of the process leading to change, who is aware of it, and how they can institutionally respond.

The details of the resulting new institutions, if they lead to the intended outcomes, are partially dictated by the function they have to serve. Yet, there are many institutions that can achieve the same objectives. From the institutions-as-equilibria perspective, past institutional elements influence the details of subsequent institutions because institutionalized beliefs, norms, and expectations are embodied in people's beliefs systems, preferences and memories, while existing organizations have enduring physical capacities, routines, and other resources. Although it may be technologically possible to create new beliefs, norms, expectations and organizations, doing so is usually costly, time consuming, and requires venturing into the cognitively unknown (Greif 1994, 1998, 2006).

There is therefore a *fundamental asymmetry* between institutional elements inherited from the past and technologically possible alternatives. Past institutional elements are the raw material on which new institutions are based. Sugden (1989) argues that people wishing to coordinate their strategies will generally adopt rules which are analogous to rules with which they are already familiar, for example, the "first come, first served" rule which is used to assign property rights in many situations (and thereby avoid potentially costly conflict). Greif (1989, 1994) argued that organizations inherited from the past and cultural beliefs determine selection among alternative institutions. Patterns of organizational membership determine whose identity is known to whom, and where information flows, while cultural beliefs

coordinate expectations. Campbell (2004) argues that actors often create new institutions through a process of "bricolage": recombining elements in their institutional repertoire to deal with new situations. New institutions often resemble older, familiar institutions because they contain elements inherited from or inspired by past institutions. Greif (2006) delineates how exactly past institutional elements influence subsequent institutions through their environmental, coordination, and inclusion effects.

Thus, the institutions-as-equilibria approach conceptualizes institutional dynamics as an accumulative historical process of inter-related institutional elements. Past institutional elements are incorporated into new institutions that emerge within the context of – and hence are complementary to – existing institutions. The results are *institutional complexes*, which are a set of institutions that govern various interactions, have common institutional elements, and are complementary to each other. Society's institutions have to be studied from a holistic, systemic perspective (Aoki 2001).

## 7   Concluding Notes

Recent scholarship has demonstrated the power of the rational choice framework for advancing our understanding of institutions and institutional change. And as our understanding improves, the conceptual frameworks employed to study institutions continue to evolve, enabling us to develop richer and more complete answers even as we probe deeper and more complex questions about the nature of institutions and processes of institutional emergence and change.

From the institutions-as-rules perspective, institutions are rules that are either optimal responses to the institutional environment or are determined by the interests of the political actors who make the rules. This approach has been fruitfully applied to explore how the "rules of the game" are formed in diverse settings. However, because enforcement of the rules is treated as exogenous, the institutions-as-rules approach works best in situations where there are, in fact, well-functioning and transparent enforcement institutions which can be taken as given, and in which the rules (whether formal or informal) are easy to observe, so that they may be expected to translate more or less directly into effects on human behavior.

If these conditions are not met, then ultimately, to explain how the rules are enforced (or not), and why they are followed (or not), the institutions-as-rules framework must be supplemented or replaced by a theory in which enforcement is treated as endogenous, and the incentives of all players to follow the rules are explained rather than assumed. The institutions-as-equilibria approach focuses on motivation provided by beliefs, norms, and expectations that both shape individual behavior whilst simultaneously themselves being a product of the strategic interplay between agents (individuals or organizations). Thus, both the content of the rules (behavior) and their enforcement (people's motivation for following them) can be studied within a unified framework. The key to institutional change, from

this perspective, is not just changing rules, but changing players motivations and patterns of behavior in a self-enforcing way.

These two approaches can be seen as complementary parts of the analysts toolkit. The institutions-as-rules approach seems appropriate for studying the development of institutions within an established structure that can enforce the rules, for example, in a stable democracy within which basic market-supporting institutions are already well-established. The institutions-as-equilibria approach might be better suited to studying the institutional foundations of markets and democratic political structures, and other situations in which enforcement of the "rules" must be considered as an endogenous outcome rather than taken as given.

# References

Acemoglu D (2003) Why not a political Coase theorem? Social conflict, commitment and politics. J Comp Econ 31:620–652

Acemoglu D, Robinson J (2006) Economic origins of dictatorship and democracy. Cambridge University Press, Cambridge

Acemoglu D, Johnson S, Robinson J (2001) The colonial origins of comparative development: an empirical investigation. Am Econ Rev 91:1369–1401

Acemoglu D, Johnson S, Robinson J (2002) Reversal of fortune: geography and institutions in the making of the modern world income distribution. Quart J Econ 117:1231–1294

Akerlof G, Kranton R (2005) Identity and the economics of organizations. J Econ Perspect 19(1): 9–32

Alchian A (1950) Uncertainty, evolution and economic theory. J Polit Econ 58:211–221

Aldashev G, Chaara I, Platteau J-P, Wahhaj Z (2007). Custom in the shadow of the formal law: an economic analysis. Working paper, University of Namur

Aoki M (2001) Towards a comparative institutional analysis. MIT Press, Cambridge, MA

Aoki M (2007) Endogenizing institutions and institutional changes. J Inst Econ 3(1):1–31

Baron D (2001) Private ordering on the Internet: the eBay community of traders, Stanford GSB research paper no. 1709

Barzel Y (2002) A theory of the state: economic rights, legal rights, and the scope of the state. Cambridge University Press, Cambridge

Bates R (2001) Prosperity and violence: the political economy of development. Norton, New York

Bates R, Greif A, Singh S (2002) Organizing violence. J Confl Res 46(5):599–628

Benabou R, Tirole J (2006) Belief in a just world and redistributive politics. Quart J Econ 699–746

Bowles S (1998) Endogenous preferences: the cultural consequences of markets and other economic institutions. J Econ Lit 36:75–111

Bowles S, Gintis H (1993) The revenge of homo economicus: contested exchange and the revival of political economy. J Econ Perspect 7(1):83–102

Brousseau E, Raynaud E (2008) Climbing the hierarchical ladders of rules: the dynamics of institutional frameworks. Working paper, University of Paris X

Calvert R (1995) Rational actors, equilibrium, and social institutions. In: Knight J, Sened I (eds) Explaining social institutions. University of Michigan Press, Ann Arbor, MI

Campbell J (2004) Institutional change and globalization. Princeton University Press, Princeton, NJ

Coase R (1937) The nature of the firm. Economica 4:386–405

Demsetz H (1967) Toward a theory of property rights. Am Econ Rev 57(2):347–359

Dixit A (2004) Lawlessness and economics: alternative modes of governance. Princeton University Press, Princeton, NJ

Djankov S, Glaeser E, La Porta R, Lopez-de-Silanes F, Shleifer A (2003) The new comparative economics. J Comp Econ 31:595–619

Easterly W, Levine R (2003) Tropics, germs and crops: how endowments influence economic development. J Monetary Econ 50:3–39

Ellickson R (1991) Order without law. Harvard University Press, Cambridge, MA

Engerman S, Sokoloff K (2002) Factor endowments, inequality, and paths of development among new world economies. Economia 3:41–109

Fafchamps M (2004) Market institutions in Sub-Saharan Africa. MIT Press, Cambridge, MA

Fearon J (2007) Fighting rather than bargaining. Working paper, Stanford University

Gazzale R (2005) Giving gossips their due: information provision in games with private monitoring. Working paper, Williams College

Greif A (1989) Reputation and coalitions in medieval trade: evidence on the Maghribi traders. J Econ Hist 49(4):857–882

Greif A (1993) Contract enforceability and economic institutions in early trade: the Maghribi traders' coalition. Am Econ Rev 83(3):525–548

Greif A (1994) Cultural beliefs and the organization of society: a historical and theoretical reflection on collectivist and individualist societies. J Polit Econ 102(5):912–950

Greif A (1998) Historical and comparative institutional analysis. Am Econ Rev 88(2):80–84

Greif A (2000) The fundamental problem of exchange: a research agenda in historical institutional analysis. Eur Rev Econ Hist 4(3):251–284

Greif A (2006) Institutions and the path to the modern economy: lessons from medieval trade. Cambridge University Press, Cambridge

Greif A, Laitin D (2004) A theory of endogenous institutional change. Am Polit Sci Rev 98(4)

Hadfield G (2008) The levers of institutional design: institutional determinants of the quality of law. J Comp Econ 36(1):43–73

Hart O (1995) Firms, contracts and financial structure. Oxford University Press, Oxford

Hayek F (1945) The use of knowledge in society. Am Econ Rev 35(4):519–530

Hayek F (1973) Law, legislation and liberty, vol 1: Rules and order. University of Chicago Press, Chicago

Jackson M (2006) The economics of social networks. In: Blundell R, Newey W, Persson T (eds) Volume I of advances in economics and econometrics, theory and applications: ninth world congress of the econometric society. Cambridge University Press, Cambridge

Kandori M (1992) Social norms and community enforcement. Rev Econ Stud 59:63–80

Keefer P, Knack S (1997) Does social capital have an economic payoff? A cross-country investigation. Quart J Econ 112(4):1251–1288

Kingston C (2007) Marine insurance in Britain and America, 1720–1844: a comparative institutional analysis. J Econ Hist 67(2):379–409

Kingston C (2008) Adverse selection and institutional change in eighteenth century marine insurance. Working paper, Amherst College

Kingston C (2011) Marine insurance in Philadelphia during the quasi-war with France. J Econ Hist 71(1):162–184

Kingston C, Caballero G (2009) Comparing theories of institutional change. J Inst Econ 5(2):151–180

Kranton R (1996) Reciprocal exchange: a self-sustaining system. Am Econ Rev 86(4):830–851

La Porta R, Lopez-de-Silanes F, Shleifer A (2008) The economic consequences of legal origins. J Econ Lit 46(2):285–332

Libecap G (1989) Contracting for property rights. Cambridge University Press, Cambridge

McMillan J, Woodruff C (1999) Dispute prevention without courts in Vietnam. J Law Econ Org 15(3):637–658

Milgrom P, North D, Weingast B (1990) The role of institutions in the revival of trade: the law merchant, private judges, and the champagne fairs. Econ Polit 1:1–23

Nelson R, Winter S (1982) An evolutionary theory of economic change. Harvard University Press, Cambridge, MA

North D (1990) Institutions, institutional change and economic performance. Cambridge University Press, Cambridge

North D (2005) Understanding the process of economic change. Princeton University Press, Princeton, NJ

Olson M (1982) The rise and decline of nations. Yale University Press, New Haven, CT

Olson M (1993) Dictatorship, democracy, and development. Am Polit Sci Rev 87(3):567–576

Ostrom E (1990) Governing the commons: the evolution of institutions for collective action. Cambridge University Press, Cambridge

Ostrom E (2005) Understanding institutional diversity. Princeton University Press, Princeton, NJ

Pierson P (2000) Increasing returns, path dependence, and the study of politics. Am Polit Sci Rev 94(2):251–267

Qian Y (2003) How reform worked in China. In: Rodrik D (ed) In search of prosperity: analytic narratives on economic growth. Princeton University Press, Princeton, NJ

Roland G (2000) Transition and economics: politics, markets and firms. MIT Press, Cambridge, MA

Roland G (2004) Understanding institutional change: fast-moving and slow-moving institutions. Stud Comp Int Dev 38(4):109–131

Scartascini C, Tommasi M (2009). The making of economic policy: institutionalized or not? IDB working paper no. 108

Schotter A (1981) The economic theory of social institutions. Cambridge University Press, Cambridge

Skaperdas S (2006) Anarchy. In: Weingast B, Wittman D (eds) Oxford handbook of political economy. Oxford University Press, Oxford

Sugden R (1989) Spontaneous order. J Econ Perspect 3(4):85–97

Tabellini G (2008) The scope of cooperation: values and incentives. Quart J Econ 123(3):905–950

Verlinden C (1979) Markets and fairs. In: Postan M, Rick E, Miltey M (eds) The Cambridge economic history of Europe, vol 3. Cambridge University Press, Cambridge, pp 119–153

Weingast B (1997) The political foundations of democracy and the rule of law. Am Polit Sci Rev 91(2):245–263

Williamson O (1979) Transaction cost economics: the governance of contractual relations. J Law Econ 22:233–261

Williamson O (2000) The new institutional economics: taking stock, looking ahead. J Econ Lit 38:595–613

# War, Wealth and the Formation of States

**Carles Boix, Bruno Codenotti, and Giovanni Resta**

The number and size of sovereign states across the world has varied substantially over time. Up until the late eighteenth century, and at least since the Middle Ages, the number of independent nations experienced a massive reduction. In Europe alone, Tilly estimates that in 1450 there were around 200 states, principalities and city-states, averaging 9,500 square miles (the size of Vermont) and about 350,000 inhabitants.[1] Yet by 1830 that number had declined substantially – to about 25.[2] After the unification of Germany and Italy, the number of European states bottomed at its lowest level ever, 17 states, in the early 1870s. That process of consolidation stopped and then reversed in the nineteenth century with the spread of the idea of self-determination and the access of new nationalities to independence. The independence of Latin America from Spain in the first half of that century resulted in the creation of about 15 new countries. Several nations in the Balkans, three former British colonies (Canada, New Zealand and Australia) and Norway followed in the decades before World War I. Finally, the collapse of the German, Austrian and Russian empires in 1918 pushed the number of states up to 70 by the mid 1920s. After a decline during World War II resulting from the German occupation of almost all continental Europe, the process of decolonization across the globe doubled that figure. The breakdown of the Soviet Union ushered another wave of secessions and put the number close to 200 sovereign units.

---

[1]According to Carneiro (1978), there were about 600,000 independent political communities on the Earth in the year 1000 BC.

[2]This figure assumes all Swiss cantons to be part of one confederation, takes the Papal states as just one political unit and only includes mid-sized German territories. Otherwise, the number would be closer to 80.

C. Boix (✉)
Department of Politics and Public Affairs, Princeton University, Princeton, NJ, USA
e-mail: cboix@princeton.edu

B. Codenotti and G. Resta
Institute for Informatics and Telematics, Consiglio Nazionale delle Ricerche, Pisa, Italy

Besides fluctuating in numbers, countries have differed (and continue to differ) in terms of their size. Figure 1 reports the evolution of the median and average size of countries as well as the maximum and minimum countries from 1815 to 1998. Dispersion is considerable. Most countries, however, are small. A fourth of all countries had a territory below 40,000 square miles – and in fact the size of the lowest quartile has declined to below 20,000 square miles in the last two decades. Fifty percent of all states have less of 100,000 square miles and only 25% have a territory over 350,000 square miles.

The purpose of this chapter is to offer a relatively comprehensive theory to explain the variation in number and size of states, blending standard explanations of international politics, which emphasize war and systemic factors, and domestic political factors, such as military might or the role of democracy and national sentiments. After reviewing the current research in Section 1, Section 2 sketches a theoretical framework in which state actors optimize their wealth and power according to principles of bounded rationality and uncertain information in an anarchical international system. The strategies of each state ruler (either military-expansionary or defensive-and-wealth-creating) vary with their territory's endowments, their own political institutions, the level of national identity and, naturally, the strategies of other actors. To probe the validity of this model, the following sections implement an agent-based model in which a set of rulers play multiple rounds under different structural conditions. This 'open game form' is much better adapted at capturing an essential part of the problem at hand: the fact that decisions at each stage of the game change the number (and payoffs) of players). We relate some of the results we obtain to existing historical data.
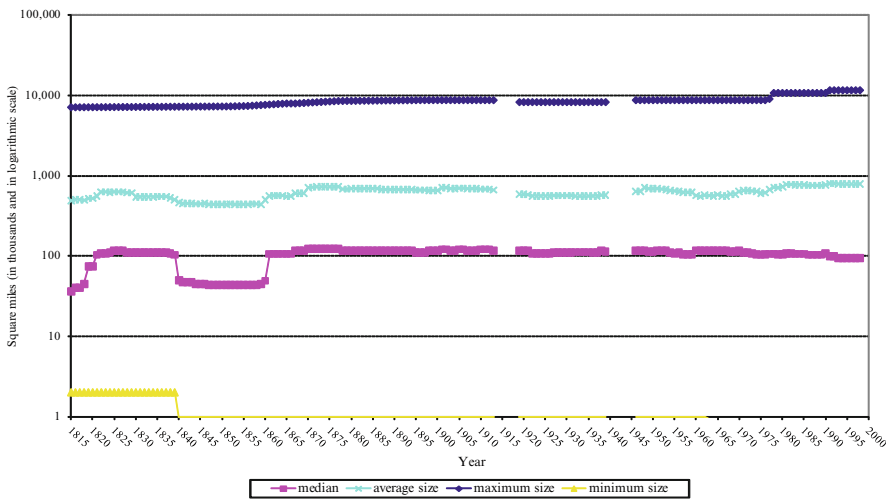


**Fig. 1** Evolution of size of country, 1815–1998

# 1    Current Research on Number and Size of States

The current scholarly literature developed to explain why and how the size and number of states varied over time falls into two categories: macrohistorical sociology and political economy.

   The thrust of the current research on the size of countries has been written within the tradition of macrohistorical sociology. At the turn of the twentieth century, Hintze (1975) was the first to point to war as the generator and creator of states. Tilly (1975, 1990) then traced the varying territorial size of states to the pressures of warfare conditional on the extent to which 'capital' or 'coercion' were employed to organize and fund warfare machines. Spruyt (1994) related the 'type' of state to the underlying social coalitions that sustained it.

   More recently, a few scholars, mostly employing the tools of microeconomics, have developed several analytical models to predict the size of countries as a consequence of both the degree of trade integration and the preferences of citizens over the level of redistribution and public good provision (Alesina 2002; Alesina and Spolaore 1997; Bolton and Roland 1997).[3]

   Although both the economic and sociological approaches provide invaluable insights on the formation of countries, they are ultimately limited by their respective methodological and substantive choices. The macrohistorical literature provides the theoretically most convincing building blocks (war, the technological shock of new and increasingly more expensive weapons, and the need of states to survive) to develop an empirically sound story of the evolution of the size and number of states. Still, its logical foundations are sketchy and the concepts it employs (such as 'capital' and 'coercion') imprecise (for example, 'coercion' is not a resource – akin to money – to be employed by the ruler to build a state, as Tilly suggests, but rather a mechanism to muster compliance). Moreover, its structuralist explanations are devoid of agency and hence mostly correlational in nature.

   In turn, the economics literature on state formation is attractive in the way it develops an analytical model and it is (relatively) careful in its treatment of how agents choose their political strategies. Its theoretical approach, which is based on the assumption that rulers (or citizens, in democratic polities) choose borders through some joint welfare maximization process given how the size of the state affects taxes and public goods provision, presents, however, two problems. On the one hand, it minimizes the role of security concerns and war in the formation of states. Trade, the creation of economies of scale and the choice of policies that maximize economic efficiency are indeed important concerns among policymakers. But, in the context of anarchy that characterizes the international system, security and military considerations play the most crucial role in explaining state behavior and state outcomes. Most, if not all, state rulers maximize their welfare through non-economic mechanisms, that is, through military expansion (exclusively or in conjunction with purely economic

---

[3]See a recent summary of the literature on state formation by Spruyt (2007).

means). And even when they do not contemplate military action to increase their wealth and power (for reasons we discuss and model later), they still need to prepare militarily and even take preemptive actions against potential aggressors. Given that other states may simply prefer to increase their wealth through military raids, even trading (and hence "peaceful" or peace-prone) states need to amass enough military power to secure the protection of their citizens' property rights against any external bandits. In short, security constitutes a central goal of states and war technology plays a key role in determining state borders. In fact, almost all the instances leading to the creation of unified nations or to the fragmentation of existing empires are related to military defeat (directly or after a protracted cold war): from the declaration of an independent United States to the collapse of Austria-Hungary and the split of the Soviet Union. There are very few cases in which nations have become independent through a peaceful process, completely unrelated to the military defeat or military exhaustion of the dominant state: Norway and Iceland seem to be the only two cases. It is also true that world trade may have facilitated the emergence of regionalist demands within nation-states but it seems only distantly related to the number of countries. The search for a unified trade space may have spurred the national unifications of Germany and Italy – but both events eventually required fighting several military campaigns. Even the creation of a common market in Europe may have been the result of security concerns (Rosato 2011). In short, a truly successful model of state formation must necessarily include violence as the central mechanism through which countries are created and reshaped.

On the other hand, and more pointedly, current economic models of state formation are static in nature. They posit a fixed number of agents that make decisions, generally in the context of a game-theoretic structure or a joint welfare maximization decision, on the "optimal" size of states. Yet the history of state and nation formation is an inherently dynamic one. The number of agents that had competed with each other has changed over time. The outcome of war in the first period alters the outcome of the following periods, since the winners of the first round of conflict can now count on newly acquired resources to launch a new wave of attacks (or defensive actions). In the mean while, war technologies and the economic techniques employed to generate income vary over time, affecting the number and payoffs of actors in unexpected manners. Finally, the structure of preferences of rulers and rules (e.g. in terms of their national identity) shift, often in direct relation to the outcomes in each round of the war game. All these transformations can be hardly modeled using closed game forms.

Given the state and shortcomings of the existing literature, we take a (partially) different theoretical and methodological approach to examine the process of state formation (i.e. the process of change in the territorial size of the state). From a theoretical point of view, we go back to the notion of war as the main driving force of state formation, while at the same time formalizing with more precise analytical foundations the general insights of the macrohistorical literature.[4]

---

[4]As it should be apparent shortly, we can also integrate economic considerations into the general model of war and territorial expansion we develop here.

Methodologically, we engage in agent-based simulations (and compare them to the evolution of the international state system), that is, once we create a virtual environment populated with a number of states that, behaving according to a set of theoretically-derived decision rules, interact over a given time period. The analysis of states' interactions in an agent-based world rather than in a closed game set-up has several advantages. First, it allows us to construct a model based on assumptions that approximate the empirical world quite closely – that is, we need to make fewer sacrifices to build the model than we would have to if we used the tools and solution concepts of standard game theory. Second, it enables us to track the dynamic structure of the game thoroughly – that is, over all the steps that lead from the starting moment of the world we create to a particular point in time (that we select or that seems to exhibit some robust equilibrium).[5]

## 2   Theoretical Framework

To explain the formation and evolution of a state system, we depart from the following stylized principles:

1. States, i.e. those organizations that have the monopoly of coercion over a given territory, live in an anarchical world in which there is no single authority capable of adjudicating among them.
2. All sovereign rulers are intent on the maximization of power, either as a goal in itself or as a means to secure other objectives such as, fundamentally, their survival or, sometimes, the accumulation of wealth. As stressed by standard international relations theory, in a context of anarchy states are necessarily compelled to accumulate power, even if they have no wish to conquer other state – simply to preempt any offensive actions of their neighbors.
3. Since power is a function, at least in the long run, of the country's population and productivity (Mearsheimer 2001), the maximization of power eventually implies the maximization of wealth.
4. The maximization of power (wealth) may be achieved through different strategies. The state may spend its resources to bolster its military capabilities, which can be then employed to expand its territory and the population under its control. If successful, this military strategy can then feed a self-sustaining campaigning of expansion. Alternatively, the state may spend its resources to maximize wealth in a particular territory (without much concern for its expansion) – this "non-military" strategy may imply, for example, establishing low-tax policies that give incentives to private agents to produce more.
5. Both strategies may be pursued – with rulers assigning different weights to each one. However, since states face some kind of budget constraint, there may

---

[5]For similar considerations, see Cederman (1997).

be some trade-offs in place in the choice of which strategy to be mainly pursued.

6. States choose strategies according to the degree of comparative advantage they have and according to the strategies other states pursue.

7. States' decisions are made under (varying) conditions of uncertainty about the behavior and capabilities of others and under limited knowledge about the evolution of areas that are beyond the closest neighbors.

8. Even though gaining an advantage over all other states may be desirable in principle, following a strategy of military expansion is never an unconstrained activity. The decision to wage wars to expand and conquer new territories is determined by, on the one hand, the costs of fighting external competitors and controlling the population of the territory to be occupied, and, on the other hand, the benefits or revenue that can be extracted from that new area. The costs of war and occupation vary with changes in the technologies used to produce violence and with the willingness of the occupied population to accept the new ruler. The benefits of expansion are a function of the distribution of the population, the latter's production technology, the administrative capacity of the state to tax its citizens and finally the types of economic activities (and assets) that states control (or could control). For example, any wealth that can be easily taxed (such as land and mineral wealth) will be much more sought after that any assets (such as trading or financial activities) whose control may be much harder to achieve. Accordingly, territorial expansion and the use of war will be differently distributed (across the world and over time) depending on the type of economy in place (at home and abroad).

9. Similarly, the decision to pursue "non-military" strategies will be a function of the population already under control, its productivity and the taxation capability of the state.

10. The decisions made by state rulers depend upon their political institutions. Authoritarian and democratic states differ in the distribution of the costs of war (for example, in a tyranny the ruler tends to face a much lower different probability of violent death than in a democracy) and in the distribution of the spoils of war.

11. States do not exist by fiat – but are generated endogenously. Individuals "choose" to establish some kinds of associations to defend themselves. Alternatively, some entrepreneurs (or some initial protective associations) succeed at expanding territorially. Conversely, existing states may split into several smaller states if certain subpopulations decide to challenge the center and establish their own independent political structure.

12. Finally, war and conflict may take place among two single states. But they may be equally fought by alliances of several states.

In short, we postulate a theoretical framework that starts from realist assumptions: an anarchical environment in which states as the only actors and where they interact in the context of a security dilemma. But we then proceed to endow our states with a set of rational calculations that derive from domestic and external

conditions affecting the benefits and costs of war amounts. This second step allows us to move beyond the inherent theoretical poverty of realism.[6]

## 3 Modeling the Behavior of Rulers

To understand the mechanisms that explain the variation in the size of countries over time (and in fact, across the world at any point in time), we consider an environment in which there are several rulers or states intent on maximizing their income (as a means in itself and as a means to have power). States are characterized by a set of parameters, such as population, technologies of production, military capabilities and so on. Given those factors they make decisions concerning their rate of expansion, etc. We examine those decisions in the context of a virtual environment and we gradually vary their parameters to understand how those changes affect the final outcome of the international state system.

Simulation modeling in political science already includes some notable contributions, spawning from the analysis of voting decisions to Axelrod's work on the emergence of cooperation in the context of a prisoner's dilemma game (Axelrod 1984).[7] In the field of international politics, war and the determination of political borders was first modeled through a simulation game by Benson (1961) and then Bremer and Mihalka (1977). The latter modeled the environment as an hexagonal grid in which countries are allotted some resources and information and then engage in decisions about war and conquest. This model was further extended by Cusack and Stoll (1990) and then explored and improved by Cederman (1997, 2003). Here we follow the same tradition of building an agent-based model to explore the process of state formation. Nonetheless, we depart from the existing work in several ways – mostly in characterizing different types of states in terms of their domestic or internal conditions (political regime, population densities, economic profiles, nature of assets). It is this richness in the economic and political parameters of our units that allows us to explain the causes of the different distributions of states that have emerged spatially and temporally in the world (and, in particular, in Europe).

### 3.1 Income Maximization

The final appendix describes in more detail the rule of the agent-based model we employ. Here we describe in general terms the main components of the model.

In the model, the ruler maximizes net income, that is, the difference between the revenues he will raise in the territory under his control and the costs he will have to

---

[6]Our closest predecessor is therefore Gilpin (1981).

[7]See a general review of the literature of simulation in political science in Johnson (1999) and a critical assessment of simulation as a method of inquiry in international relations in Pepinsky (2005).

incur to maintain the monopoly of violence over a certain territory (and hence domestic and external peace).

The goal of income maximization can be interpreted in strict "economic" terms – that is, we can think of the ruler as being simply interested in increasing his purse and therefore his consumption. However, as noted above, it also embeds the goal of securing the survival of the ruler. In the absence of a central authority and of general, enforceable rules at the international level, all sovereign rulers are uncertain about their neighbors' behavior and therefore about their own survival. In this context of anarchy, in which all states compete (or may compete) with each other, each ruler maximizes his resources and wealth to have enough power to overcome any challenge to his sovereignty.

There are several strategies to maximize income, ranging from generating a more productive economy to expanding and capturing new territories and subjects. Whether the state pursues wealth maximization through territorial expansion or chooses nonmilitary means to generate wealth will depend on the varying levels of technology, population growth rates, tax capacity and military costs of their territory and the territories of other rulers.

## 3.2   Costs of Violence

To establish a state, the ruler has to incur some costs to maintain a minimal domestic police, an army to fight other states and an administrative structure (to raise revenue, monitor the police and the army, and to adjudicate any disputes among the inhabitants of the territory under control).

These costs can be distinguished into two types. On the one hand, the ruler has to pay some 'fixed costs', that is, some start-up costs that are independent of the territory and population under control – for example, the costs of sustaining the central government and the central command of the army. On the other hand, the ruler must meet some 'variable costs' or, in other words, costs that are not fixed and that vary with the territory under control. Soldiers, who can be easily hired or laid off, as a function of the territory being defended, are variable costs. In between fixed and variable costs, there are quasi-fixed costs such as most weapons: although in the long run they can be adjusted to the size of the state, states normally have to spend heavily in guns to start with and they can only amortize that expenditure once they have expanded substantially.[8]

---

[8]In the long run, there are no fixed costs since all costs can be adjusted to controlled territory. However, we will assume throughout the discussion that we are discussing the short-run evolution of the state. Naturally, the threshold between short and long run is in a sense arbitrary.

Variable costs are shaped by three factors:

1. The easiness or difficulty with which the central government controls the population spread in the territory for purely 'technical' reasons. The army may have to march for longer periods of time to reach the bordering populations or communications become more sporadic, therefore weakening the effectiveness of the state.
2. Second, the cost of control may be shaped by 'political' domestic reasons. For example, if, as one moves away from the center, populations become increasingly different from the inhabitants at the core of the polity (they exhibit distinctive traits and preferences, make particular demands, etc.), the state will have to spend more to appease them.
3. Finally, control will be shaped by external political reasons. For example, as the country becomes larger, it clashes more directly with other states and political and military competition becomes more acute. If the enemies' armies have some advantage due to proximity to their center of decision or to the loyalty of a population that is more similar to them in preferences, the state will have to step up military expenditure at a faster pace than before.

## 3.3 State Revenues and "Comparative Advantage"

To decide what strategy to pursue, the ruler compares the income obtained from controlling new territories, net of the military costs of control, with the income derived from using the resources devoted to expansion to raise production at home. As the gains of conquering new territories decline relative to following a peaceful strategy and simply devoting all resources to home production, state rulers are more likely to adopt a non-expansionary policy. The choice of strategy will be therefore shaped by the technology and population of the home country compared to those of its neighbors.

Two additional parameters have a central effect on the decisions of rulers: first, the tax capacity of state; second, the mobility of assets. The higher the former, the more likely the state will be to invade and control other states. In turn, the higher the latter, the less interested the state will be to engage in wars of conquest since the value of the spoils of victory will be low.

## 3.4 National Identity and Secession

So far we have described an environment in which states can either wage war and expand (at the expense of other political units) or remain in peace. State systems can also witness the formation of new states – through a process of secession. To make things simple, we model secessions as random events subject to a set of political

constraints. Following some positive probability, a given space (smaller than an existing state) considers the possibility of secession. That probability is negatively related to the time that area has belonged to the current state – thus giving some weight to a past, particular 'national sentiment'. Provided that area can survive militarily as an independent entity, the process of secession starts – yet with different rules depending on the political regime in place in the original state. If that state is authoritarian, war is the fundamental means to determine whether secession occurs or not. If it is democratic, secessions tend to occur automatically (through the application of the right of self-determination).

## 4   Simulated Evolution of the Size of Countries

To understand how countries evolve and what may be the weight of each parameter of interest, we need to examine the strategic interaction of several rulers over time. As pointed in the introduction, we do so through the examination of how several rulers behave in a simulated environment given the rules we just laid out.

The environment in which agents act consists of a two-dimensional space of 50 x 50 squares populated by a maximum of 49 states (with contiguous borders). Each state is endowed with some initial population, determined by its area and a national density parameter. The density parameter, which is set up by the researcher (randomly or directly), varies according to the type of terrain in place – either mountainous, hilly or flat. Population grows at a given rate, also to be determined by the researcher. In addition to the demographic parameters, countries have some production technology, which, jointly with total population, determines total wealth. State rulers tax wealth according to two parameters: a tax rate, which represents the taxation capacity of the state, and a mobility parameter, which defines the extent to which national assets are more or less non-specific or mobile. As asset mobility or non-specificity increases, the share of wealth the ruler may tax declines.

In addition to these social and economic parameters, the model posits a military cost function. This cost function is identical for all countries – under the assumption that successful war technology is quickly copied by all states – and includes both a fixed cost component and a variable cost component that grows with distance from the state capital. The cost function may be specified to change over time tracking particular technological shocks in war waging. Given the demographic and economic structure of each country and the military cost function, rulers make decisions, under some informational uncertainty, about whether to challenge their neighbors or not. The war outcome is then determined by the resources of each side and the losses they bear. A detailed description of the decision rules of the states is given in the appendix to this paper.

Consider now what are the implications of the model to explain the evolution of the size (and number) of states. We first explore the effects of a shift in the cost function. We then examine what happens when the revenue schedule changes. Finally, we consider the joint interaction of changing costs and revenues.

## 4.1   Shifting Costs

The size of states varies, in the first place, as the schedule of costs shifts. The cost function may vary as a result of two changes – a shift in fixed costs or a change in variable costs – with different consequences. A shift in the level of fixed costs changes the minimal size any state has to meet to survive: we may call this level the 'survival threshold'. As fixed costs shift upwards, the size of the state has to be, by definition, bigger. Any modification of variable costs also alters the state size desired by the ruler and therefore the final number of states.

   Although we can think of shifts in fixed and variable costs as two separate events, both changes often take place simultaneously – broadly speaking, rising (declining) fixed costs are substituted for declining (rising) variable costs. Consider, for example, the case in which a technological shock leads to the invention of new, more sophisticated weapons – such as the invention of the cannon in Europe during the fifteenth century. Before the introduction of guns, that is, in the pre-cannon world, the provision of military power was labor-intensive. Soldiers carried at most spears and swords. Their capacity to control a given territory and its population was low or, to put it differently, the ratio of soldiers to population was high, especially in the periphery of the state. Thus, fixed and quasi-fixed costs were low. By contrast, variable costs were high and rose quickly with territory. The extension of fire weapons had two effects. In the first place, fixed and quasi-fixed effects went up. As a result, the survival threshold shifted to the left for each state – states need to be of a larger size just to be viable. In the second place, variable costs declined: with the help of guns, the ratio of soldiers to population declined and so the 'optimal size' of the state increased. State rulers had an additional incentive to enlarge their territories.

   Let us now turn to the analysis of state behavior and the general outcome at the level of the international level through the agent-based model. Figures 2 and 3 summarize the results of simulating the behavior of states as both fixed and variable costs change over time (4,000 units of time each). In Fig. 2 fixed costs are very low (in fact 0) while variable costs vary from high to very low (again, variable costs decline as the distance parameter declines in the model). In Fig. 3 fixed costs are set at high levels while variable costs also vary from high to very low. Both figures show that the number of states decline over time (as the resources in population and wealth increase and make conquest more attractive).[9] But the decline is strongly determined by variable costs. When the latter are high, that is, when the cost function has a steep function, the number of states hardly changes. However, when they fall, states have a much higher incentive to expand and over time the number of states converges to a minimum of 1. The pace of unification varies, however, considerably with variable costs. For high variable costs, the world remains extremely fragmented even at time 4,000. With very low variable costs, the process of unification is extremely fast – in our simulations there are less than five states at time

---

[9]Population is set to increase at some rate in the simulation.
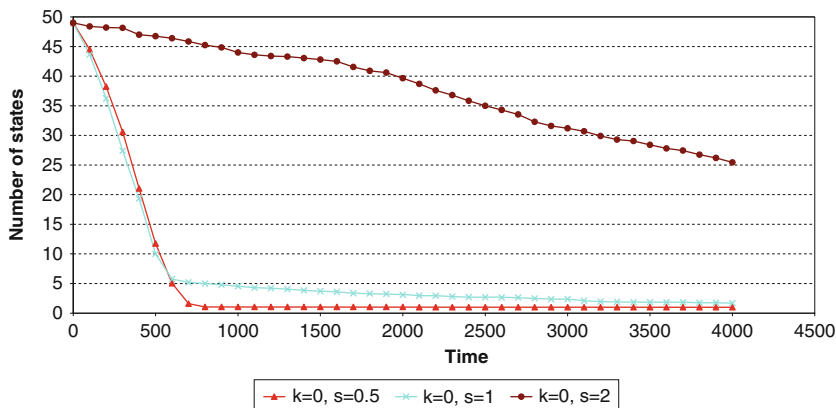
**Fig. 2** Number of states for low fixed costs (k) and different levels of variable costs (s). (Authoritarian States)
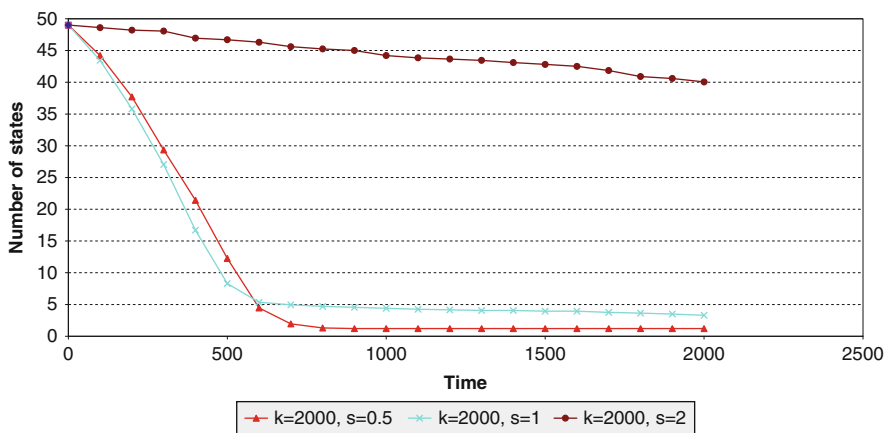


**Fig. 3** Number of states for high fixed costs (k) and different levels of variable costs (s)

600. By itself, the rate of change has interesting implications to explain the pace at which unification happens.

The impact of variable costs is also interesting in a different sense, that is, in interaction with other factors. Assume that there are some parameters (e.g. capital mobility) that, when taking some values, reduce the incentives to conquer new territories. Assume, further, that these values or thresholds occur at a given point in time – and not at the initial moment of the simulation. If variable costs become low very early in the simulation, before the other parameter reaches a threshold blocking unification, unification (through war and conquest) happens. By contrast, if variable costs decline very slowly so that hardly any state consolidation has occurred by the time the other parameter kicks in and reduces the incentives to engage in war, then

the pressures to unify decline and the world remains relatively fragmented even as variable military costs continue to decline.

The simulations summarized in Figs. 2 and 3 match the broad patterns of transformation affecting the world. At the dawn of mankind military control over the territory was very fragmented. As we shall shortly see this was in part related to population density (and wealth). But it had much to do with war technology. Fighting the enemy was such a labor intensive activity that any predator easily encountered very tough resistance from similarly equipped men as soon as he moved away from his home base. Once variable costs fell, however, conquest and control became possible, wars multiplied and the number of states (or political organizations with exclusive or near exclusive control a certain territory) declined.

For a period closer to us, these patterns also fit relatively well the evolution of the European state system since the Middle Ages to the nineteenth century. In the early Middle Ages (circa 1000 AD), political control was extremely fragmented and in the hands of several hundred feudal princes and magnates, cities and micro-states. This level of fragmentation matched the type of military structure in place. Armies were mostly composed of horsemen, organized in small bands, and had relatively primitive weapons. The costs of equipping a knight were probably high for those eminently agrarian territories and so only a few could serve on a permanent basis as combatants. Still, their output/input ratio (i.e. the number of killings and dominated populations per warrior) was low – the number of peasants a horseman could defend *and* control was rather limited. Across Europe, kings and upper-level nobles such as dukes and counts could only exercise nominal control over their territories. They relied on a hierarchical structure of barons and other low-level noblemen who, in turn, exercised direct de facto control over minute areas. In that configuration of power, monarchs acted as brokers, adjudicating disputes between their subordinates, thus in fact minimizing the possibility of a constant state of war between smallish bandits. The barons and feudal lords in turn pledged their loyalty to the crown – but their obligations were rather limited or ill-defined.

Over the following centuries, some European monarchs gradually asserted their preeminence over that feudal structure with varying rates of success. Still, it was not till fire weapons, and particularly the cannon, were introduced that the size of the state changed dramatically. Having fire power altered the nature of war permanently. At the beginning of the fifteenth century, it took Henry the Fifth of England 10 years to conquer the French region of Normandy. Thirty years later, the French monarch, now in possession of artillery pieces, conquered it back in 1 year, at the pace of one fortress per week. The feudal cavalry, which had dominated military action in the past, collapsed. States had to engage in a race to amass strong, well-disciplined armies to survive other foreign powers engaged in the same dynamics of international competition. In the late fifteenth century, France, Spain and England had been able to organize permanent armies in the order of 20,000–25,000 men – this was a significant departure from the medieval ages, in which battles were fought by a few thousand men at most. During the following centuries, however, the size of permanent troops spiraled to unprecedented levels. By 1530 Spain had a standing army of 130,000 men. In 1600 the number had risen to 200,000 and in

1630 to 230,000. Its more direct rival, France, had to engage in the same policy of general mobilization. French troops went up from 65,000 soldiers in the Italian campaign of 1498 to 155,000 around 1635 and then to about 279,000 in 1679.[10]

The upward shift in the costs of war led to the territorial expansion (and consolidation) of the existing states. The French monarch occupied Brittany, absorbed a good part of the former kingdom of Bourgogne and then invaded Italy at the turn of the century. In response, the previously independent kingdoms of Aragon-Catalonia and Castile joined, through a marriage contract, to form Spain in 1479. The Spanish kings replicated the same strategy by marrying their offspring with other European dynasties. By the second quarter of the sixteenth century Spain had succeeded in building a tight territorial 'cordon sanitaire' around France: its king held direct control over the Low Provinces (contemporary Belgium and Netherlands), several territories along the Rhine and Milan in Northern Italy, and had been elected emperor of Holy German Empire.

Still, notice that the evolution of downward trend in military costs only explains part of the phenomena we are interested in describing. It does not account for the considerable variation that we observe in the size of countries (e.g. France versus all the Italian city-states) even as the costs of war were forcing many countries to reconfigure their borders (and internal apparatus) to survive. It does not explain the growing number of countries since 1900. We now turn to changes in the revenue function to explain the first problem: the variance in territorial size across the world.

## 4.2    Shifting Revenues

The revenue schedule shifts as a result of, first, changes in the productivity of the territory, and, second, changes in the response of the economic agents to the effects of the tax rate (which we will refer to as tax elasticity of income).

The revenue function changes, in the first place, with the productivity rate of the territory. The productivity rate is the level of output produced per units of input employed in the production process. In each unit of land, production increases with an increase in population and/or with an increase in capital (both in the sense of having more tools per person or in a more efficient use of the same tools by the existing population). A higher productivity of the territory increases the incentive of states to expand.

As already pointed out before, rulers do not maximize output but rather (net) revenue. How much revenue can be collected will depend on the tax rate imposed by the state. In a world where there are only autocratic rulers intent on maximizing, they will tax constrained by two factors. The first factor is the tax capacity of the state – tax collection will increase with the organization of the state. The second factor is the distortionary impact that taxes may have on the economic decisions taken by individuals and therefore on the final size of the economy or Y.

---

[10]Data come from Finer (1975).

The relationship between total output and the tax rate can be expressed as $Y = Y(\tau)$, where $\delta Y/\delta \tau < 0$. Given this function and the tax elasticity of income, the state ruler will choose a certain $\tau^*$ to maximize public revenue. As total production becomes more sensitive to taxes, that is, $L$ is increasingly constrained to set a relatively low tax rate, the slope of $R$ flattens and the optimal size of the state declines.

Although the tax elasticity of $Y$ is a function of several factors, the type of assets (both in terms of their mobility and the easiness with which states can monitor them) are central among them, altering the size of the state as follows. In areas with more mobile types of assets (such as money or skilled people), which can escape high taxes easily, states would be smaller in size than in regions with fixed assets (land, mines, etc.), which can be heavily taxed by the ruler. Or, to put it differently, the level of aggressiveness of states will decline with asset mobility. Historically, asset mobility is correlated with the abundance of capital (and the relative scarcity of land) (Boix 2003). For that reason capital availability should lead to smaller states – just the opposite result of what we found by assuming that capital is more productive than land. To paraphrase Barrington Moore, no bourgeoisie, no peace: a world of industrialists and financial entrepreneurs may be (on average) a more peaceful world than an aristocratic one.

Consider again the results of our simulations. Figure 4 reproduces the evolution of the number of states over time under the same military cost function, unchanged technology and tax rates yet changing population levels. With no population growth, and given a very low initial density, no ruler has any incentive to expand. The number and size of states remains unaltered. With population growing at 0.05% every time period, wealth increases progressively and by time (or year) 2,000, when population density has multiplied by about 2.7 times, the number of states has declined to 35. Two thousand years later the number of states has declined to 26.
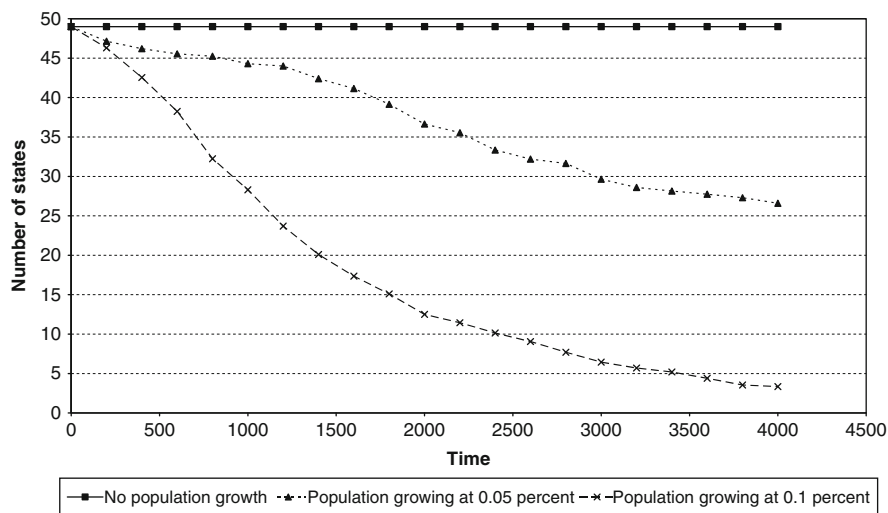


**Fig. 4** Number of states with varying population
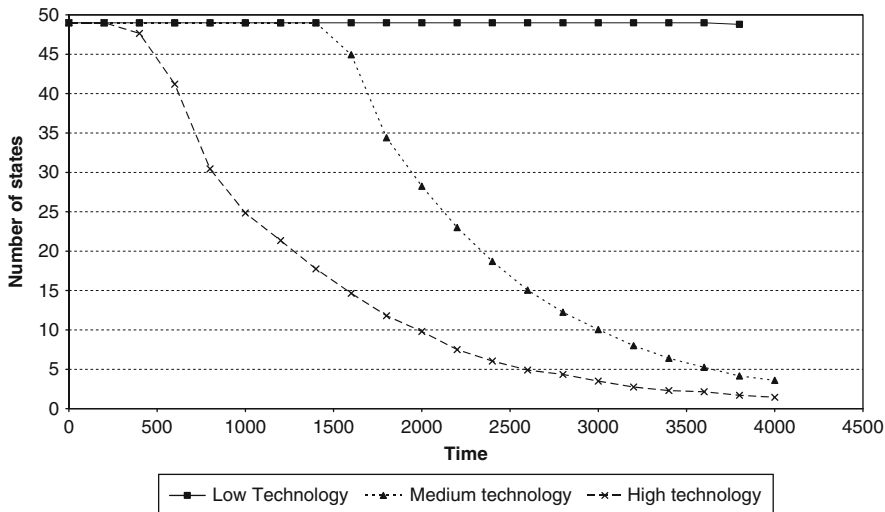
**Fig. 5** Number of states with different levels of economic technology

At a higher (and unrealistic pace of 0.1% annually), the world has become completely unified. Figure 5 shows similar results when we vary the technology parameter (that is, the parameter than transforms population into some income). With a very primitive technology, and hence little wealth creation, the number of states hardly changes (naturally, this outcome is a function as well of the level of military costs). As the technological capacity of countries changes, however, wealth multiplies and competition increases, leading again to a process of unification.

Figures 6 and 7 examine the impact of taxation on the number of states. Higher tax capacity makes conquest more profitable and accelerates the process of unification. In Fig. 6, the number of states at year 2,000 varies from over 45 for a low tax rate (20%) to 16 under a high tax rate (70%). Capital mobility has an even stronger effect on war. With complete asset immobility, an even for low taxation levels, the number of states falls rather quickly. Yet once capital becomes mobile and has an option to escape from the control of the conqueror, the incentive to invade declines rather quickly. Capital mobility can be understood in a broader sense that goes beyond mobility in a strict sense – for example, as the capacity to hide assets from the state or as having assets that cannot be exploited by the conqueror once part of the conquered population has been killed or wounded (something that may be unavoidable in a war).

Notice that population and technology affect both the incentive to attack and the capabilities the attacker has to win the war. By contrast, the type of asset mostly acts as a factor affecting the incentive to attack. This parameter captures and formalizes Tilly's insight that the evolution of the size of states may have been affected by the type of wealth they controlled. In land-abundant, capital-poor areas, all monarchs formed large, unified territories, such as Russia, Prussia and Spain,
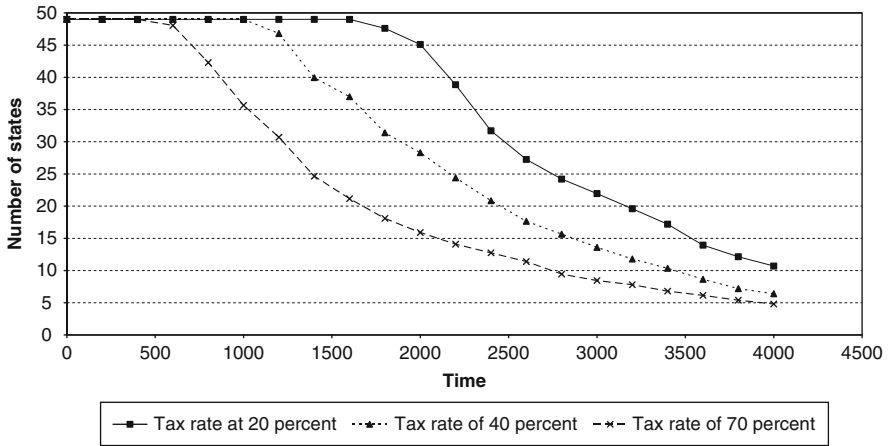
Fig. 6 Number of states with different tax rates
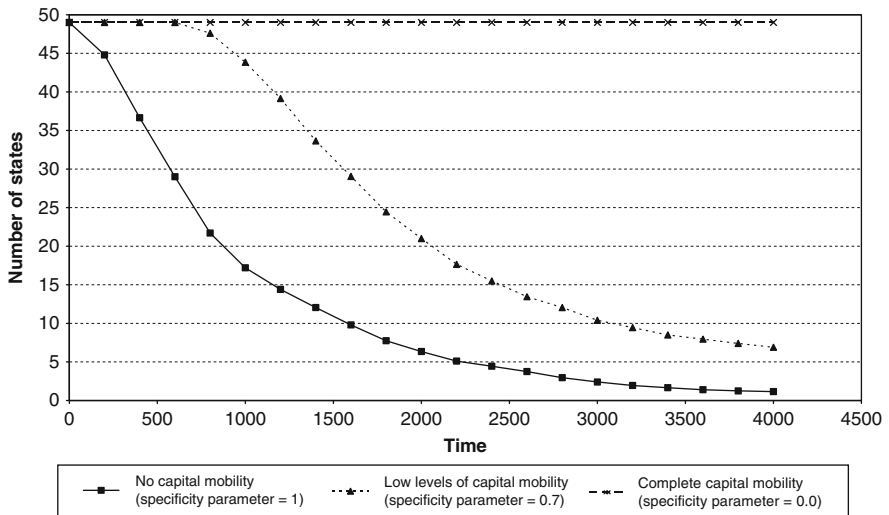


Fig. 7 Number of states with different levels of capital mobility (tax rate of 25%)

which were located at the periphery of the European continent. If they failed to do so, they disappeared: the Teutonic states or Poland stand as good examples of that fate. By contrast, unification was much slower in commercial regions – in part because they were rich enough to sustain strong armies but in part because their direct control must have been inherently hard to implement by invaders.

Table 1 shows the estimated number of independent political territories from 1000 to 1870 in the Iberian Peninsula, Italy, France and Britain. With the exception

**Table 1** Approximate number of independent political units in several European areas, 1000–1870

| Year | France | Iberian Peninsula | Italy | Britain |
|------|--------|-------------------|-------|---------|
| 1000 | >10 | About 13 | >20 | 3 |
| 1300 | >10 | 6 | | 3 |
| 1400 | >10 | 5 | | 2 |
| 1500 | 5–10 | 3 | 17 | 2 |
| 1600 | 5–10 | 1 | | 2 |
| 1700 | About 5 | 2 | About 12 | 1 |
| 1815 | 1 | 2 | 9 | 1 |
| 1870 | 1 | 2 | 1 | 1 |

of Italy, by 1600 those areas had experienced a process of consolidation. The French number is deceiving because it includes a set of small cities close to the German border – by the end of the seventeenth century France could be considered practically unified. Italy was a territory rich in commercial capital and its states were small. This was also the case of other medieval and modern city-states in central Europe such as Hamburg or the Dutch Provinces.

All these city-states were able to survive as small independent states until a new technological change made war too expensive (and shifted the 'survival threshold' upward): this shift happened with the French and Napoleonic wars, which mobilized massive infantry armies. This eventually led to the processes of German and Italian unification in the second third of the nineteenth century.

## 4.3 A Heterogeneous World

So far we have examined the impact that a change in several parameters may have on the evolution of states. However, we have treated each one of those parameters (population, technology, taxes and capital mobility) as identical across all countries. Let us now consider the effect of cross-country differences in the underlying parameters. Naturally, differences in wealth (due to population and technology) are somewhat trivial: richer countries exploit their resource advantage to defeat their enemies and to expand territorially.

Following the discussion in the previous subsection, it is more interesting to simulate the evolution of the world when countries vary in terms of their type of wealth – with some having mobile assets and others being rich on fixed wealth. Figure 8a depicts a world (of 49 states) with two types of states: the states in the core and the states in the periphery. In Fig. 8b we characterize the core as having perfectly mobile assets and the periphery as having fully immobile wealth. We then run the simulation for 10,000 time periods and reproduce the results at times 500, 1,500, 5,000 and 10,000.[11] By year 500 there are 19 countries left. The unification

---

[11]The remaining parameters are: initial density = 5, population growth = 0.001, technology = 3, inference error = 10, no fixed military costs and low variable costs (s = 0.1).
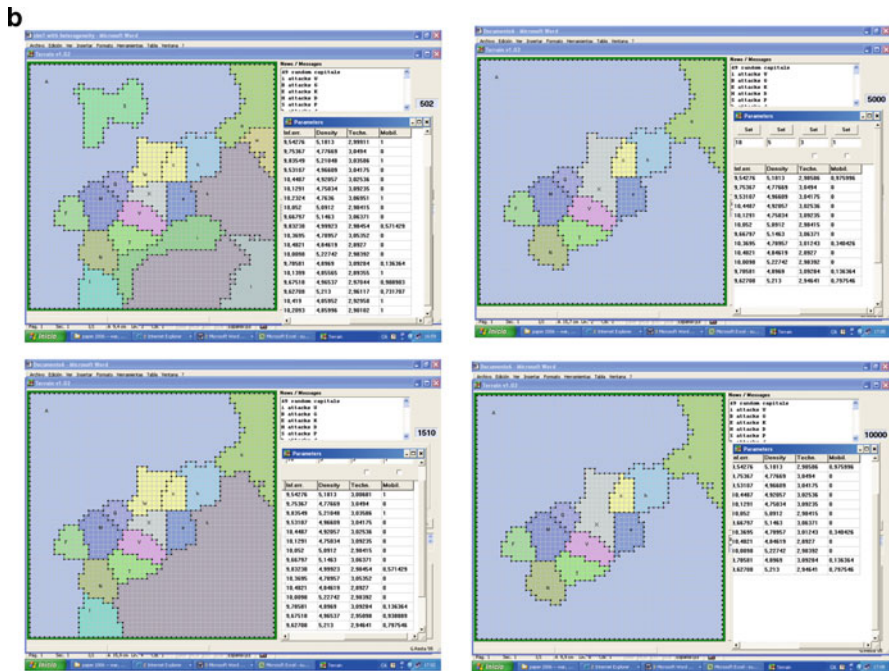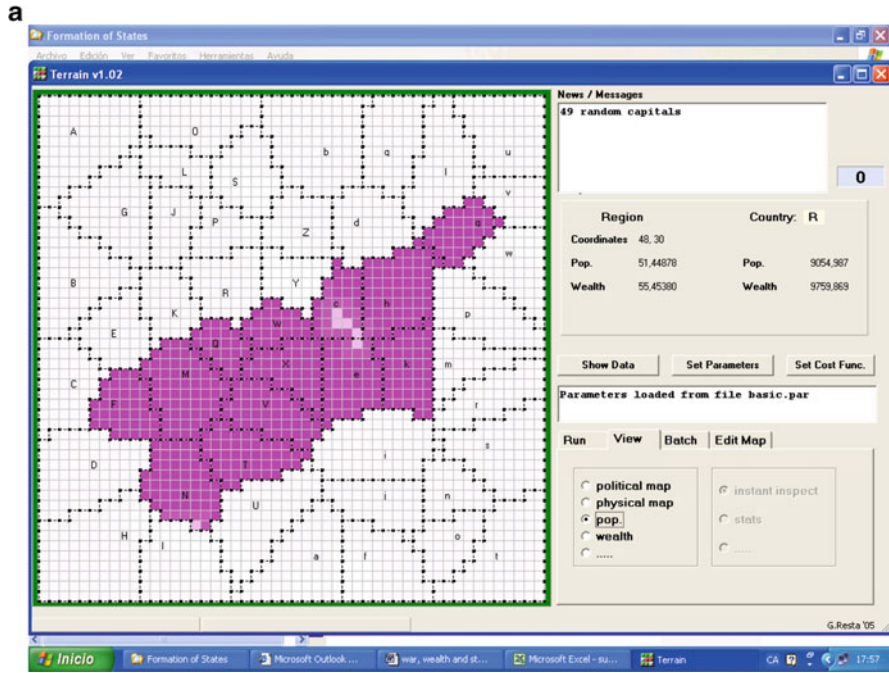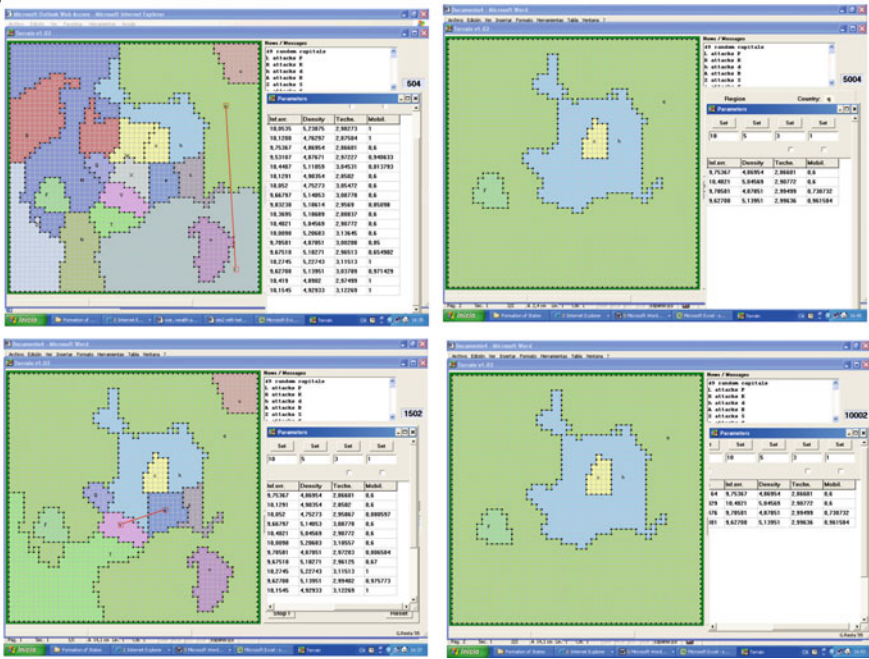
Fig. 8 (Continued)
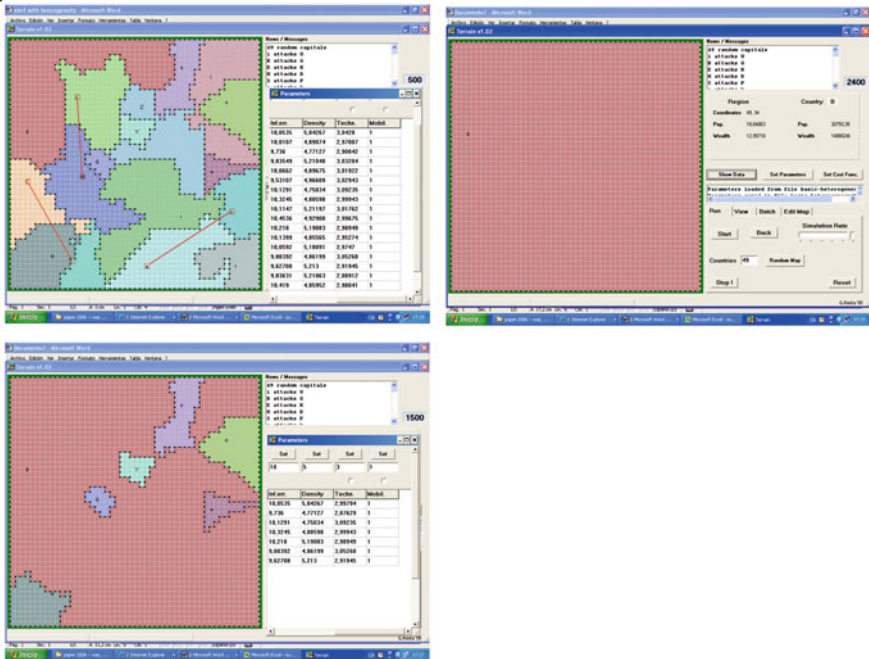
**Fig. 8** (**a**) Simulated world with core and periphery. (**b**) Complete capital mobility in the core. (**c**) Medium heterogeneity in the core. (**d**) No capital mobility

has taken place in the periphery although it is still incomplete. By year 1,500, there are two larger countries controlling the NW and SE of the world and the core is mostly intact. New expansions take place no much later and by 2,500 a single country (A) surrounds the core from all sides but one. One of the core countries (q) has expanded toward the NE but it is rich and its average wealth mobile enough that A has no incentive to attack. No military changes take place and the world looks identical throughout the end of the exercise.

In Fig. 8c the core has the more realistic trait of having medium levels of asset specificity (the parameter is set at 0.6). Unification also starts in the periphery and proceeds quickly. In the core states fight each other and consolidate in three. After 2,500 no change occurs either. The world is mainly unified but there is still some fragmentation.

Finally, in Fig. 8d we go back, for the sake of comparison, to a situation in which asset specificity is complete across the world. Unification takes place very fast, without respecting the pattern of fragmentation at the core that we observed before. By 2400 the simulated world looks like China after the period of warring states.

## 4.4  War Costs and World Government

The model on war and Leviathan's choice of territorial can encompass a wide variety of outcomes. Take a case in which there is a technology with extraordinary initial costs – such as a massive missile and space shield system – and zero or minimal additional costs (per unit of territory controlled). Assume as well as only one state can garner the revenue to pay its costs – in fact its technology base is so productive that a tax on the production on its core area is enough to pay for that military technology. That scenario should lead to a one-state government with the whole world (W*) under its control.

As a matter of fact, our model and simulation reveals how implausible the event of a world state is: the revenue schedule is not increasing on territory across the world. Roughly speaking, the globe contains three types of regions: a core of industrialized, highly productive countries (the North Atlantic and North Pacific basins), a few mineral-rich regions, and vast impoverished areas, riven from the core by mountains and deserts, and systematically lacerated by malaria, floods and famines. Any imperial power would establish its 'limes' just outside that latter, that is, what we want to call the sub-periphery. This is the outcome we get in Fig. 9. Here the periphery is just a very sparsely population, technologically underdeveloped area, located in the northern fringes of the space. The states in the core end up not expanding there, even after we run the simulation for 5,000 years or time units.

**Fig. 9** World government
and periphery. Countries in
blank have sparse population
and low technology
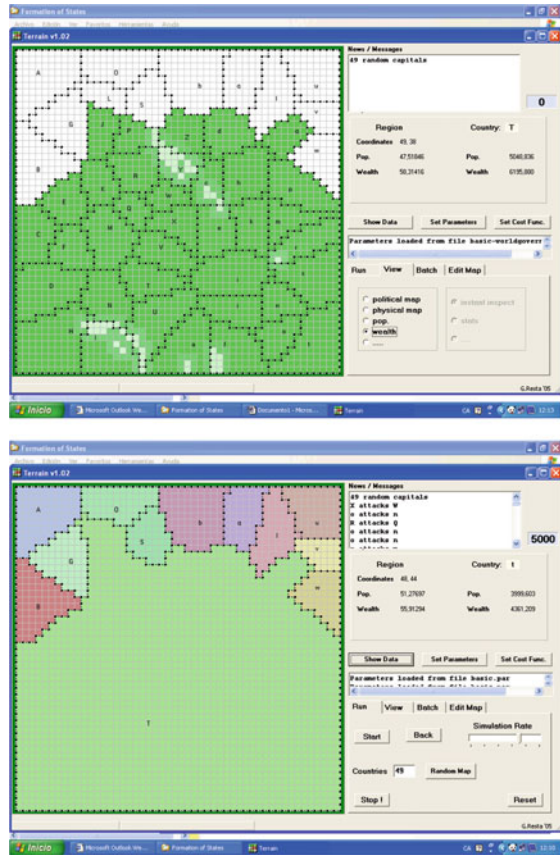


## 4.5 The Impact of Secessions

So far we have examined the evolution of the number (and size) of state in a world
in which secessions cannot happen. Figure 10 summarizes the evolution of the
number of states when secessions occur (given medium levels of variable costs or
s = 2), both when all countries are democratic and when all are authoritarian. The
underlying conditions are the same of the simulations displayed in Fig. 2.

The number of countries declines at a slower pace once secessions are possible –
nevertheless, military conditions still matter and there are enough wars and conquests
(given the parameters of the violence function) that overall there are fewer countries
at the end of the simulation than at the beginning. More important, in a world of
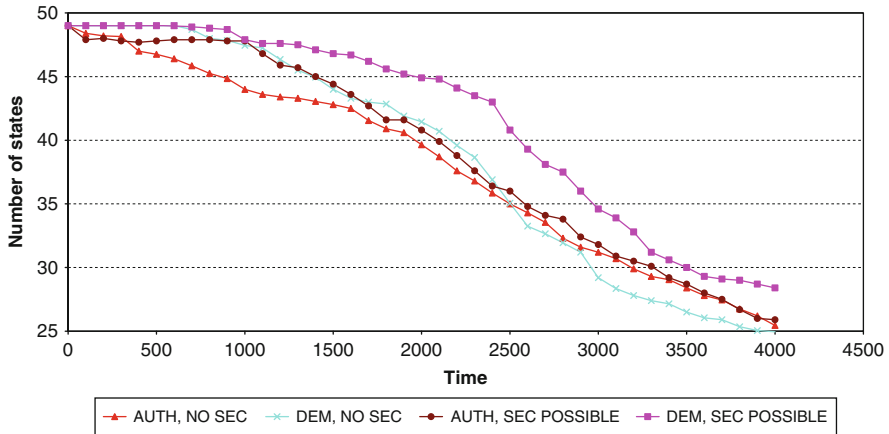
**Fig. 10**   Different regimes, secessions possible (k = 0, s = 2)

authoritarian states, the likelihood that several areas split does not affect the final number of states. It is only when countries are democratic that secessions are successful.

# 5   Conclusions

This chapter offers a simple model to account for the variation in territorial size and hence total number of states over time. States, that is, organizations that have the stable monopoly of violence over a given territory and population, decide over their optimal geographical area to maximize the welfare of their rulers. That decision is constrained by the costs involved in controlling that territory. After developing the model, the paper uses an agent-based simulation to probe its insights.

The main insights of the chapter are as follows:

1. The size of states increases (and their number decreases) as war technologies become capital-intensive. This explains the fundamental downward trend in numbers that we have witnessed from the emergence of states until our times.
2. In addition to war technology, wealth is also a central factor driving state formation and expansion. In scarcely populated, technologically underdeveloped areas the number of states remains high. Rulers have little incentive to expand and unification does not happen.
3. To explain variance in state size we need to consider the type of wealth of each country. Countries are smaller (larger) in capital-rich (capital-poor) economies. Rulers take into account the type of assets they may acquire. If wealth is difficult to tax, either because the state apparatus is inefficient or because its owners can hide it or move it abroad, then the incentive to conquer is much lower. As

a result, war is less frequent and fragmentation much more common. To put it more broadly, bourgeois economies are more peaceful than land-based, mineral-based economies.

4. Relatedly, world government may become possible in the future (given the evolution of military technology) yet only with a very low probability (given the distribution of economic activities throughout the globe).

5. The possibility of splits generates new countries. But this option only has permanent effects when the countries involved in the split are democratic. Otherwise, the result is determined by the military means and economic strength of the actors involved.

These stylized findings match the historical evolution of Europe and most of the territorial dynamics of state formation over time, at least until the nineteenth century. The last point accommodates the explosion of the number of countries we have witnessed in the twentieth century.

In general terms, this chapter represents a step toward developing a more sophisticated theory of international politics in two ways. First, it calls for endogenizing the number of states that operate in the international system. So far, most international relations theorists take the number of states or the so-called structure of the state system as given. They then make predictions about the behavior of states depending on the environment (unipolar, bipolar, etc.) they live in. We instead reflect on the ways in which the very anarchical environment in which they interact shapes the number and strength of the actors. This line of research follows previous work, by Gilpin, Cederman, Tilly and Alesina, interested in thinking about the causes of state size and formation. Still, we offer what we think is a broader framework in which war acts as the main (although not exclusive) cause of state formation and in which more realistic assumptions about state behavior (than what previous analytical models use) are combined with methodologically powerful tools to investigate shifts in the structure of state systems (and in the relative frequency of wars).

Second, we pay attention to the domestic conditions (demographic and economic) that motivate state behavior. These "second image" turn has important implications. It allows us to explain, in a systematic manner, certain systemic patterns that remain unaccounted for in traditional international theory, such as the presence of geographical clusters of peace-keeping nations or the fact that war and constant expansion does not exist in some areas (for example, along the US–Canada border). Using wealth maximization as the general goal of states allows us to still keep survival and security as central components directing state action. But the level of security threats (and the corresponding foreign policy of states) becomes conditional on the behavior and incentives of surrounding states. If all states in a given region maximize wealth through non-aggressive mechanisms, peace and border stability may become possible. (In the simulation we tentatively characterize wealth-maximizing states that do not choose expansion through the parameter of high capital mobility.) States are not less directed by self-interest in that case. Their goals are simply different than pure expansion through war and, given how others behave, they can pursue their most preferred pattern of action

in an unconstrained manner. Naturally, whenever their neighbors act aggressively, war becomes unavoidable and borders and states change as a consequence. All these different, conditional or strategic, types of behavior can be encompassed in the model we have written and are captured by the simulations we have run. The general product consists of a much richer set of outcomes, closer to how the real world looks like.

# Appendix

In the model, the parameters (that can be specified individually for each country) are: initial population density; rate of population increase; tax rate (from 0 to 1); a parameter indicating which percentage of error one country can incur when evaluating if it is convenient to attack another country; technology; and mobility (measuring how much of one country's belongings are transferred to another country in case the first one is attacked and defeated).

## *Initial Moment*

At the beginning of the simulation population is assigned to each region according to the density parameter. If the region is "hills" the density is decreased by 25% and if it is "mountains" it is decreased by 75%.

Given a certain initial density and the type of territory (plain, hill, mountain), the initial population is calculated as:

$Population_t$ = (density * number of plain cells) + (density * number of hilly cells * 0.75) + (density * number of mountainous cells * 0.25).

[In turn, savings of each state are calculated as:

$Savings_t$ = (population * technology * taxation).]

## *Evolution of Parameters*

Parameters evolve at each step according to following structure.

*Population*

The population of each unit region is incremented by a factor (1 + PopInc/1,000).

$Population_{t+1}$ = $Population_t$ * (1 + (Population increase/1,000) − in peace times. (Population increase is set by the simulator).

*Military costs*

Military costs are of two types: fixed costs K and variable costs $M * D^S$, where D is the distance of the cell from the capital of the country.

The parameters K, M, and S may depend on the current time step of the simulation. (See above for how to set them.)

Military costs can be manipulated to be of two types:

1. MC or Own Territory's Military Costs $= K + m * D^S$;
   where K is a fixed cost, m is a multiplier, D is the distance from the region to the capital. S approximates a function that modifies the cost of distance. The total military cost is computed by adding the military costs for each of its regions $m * D^S$.
2. MCplus or Military Costs of Controlling Other Territories: $m * D^S * z$;
   where z is a parameter (equal to or larger than 1) denoting how much harder it is to conquer and control the territory of other countries.

*Income*

$\text{Income}_{t+1} = \text{Population}_t * \text{Technology} * \text{Taxation}$          [Taxation $\leq$ Mobility Parameter).

[The parameters Tech and Taxes can remain as they were originally set in the country even after it has been conquered. Or they can change to take the values of the conquering country.]

*Net revenues*

$\text{Net Revenues}_{t+1} = \text{Income}_{t+1} - \text{MC}_{t+1}$

*Savings*

$\text{Savings}_{t+1} = \text{Savings}_t + \text{Net Revenues}_t$


## Simulation: War


*Decision to attack*

A country evaluates the possibility to attack a neighbor randomly, about every ten steps, unless there are already two ongoing attacks.

In this environment, any country A decides to attack D if the following conditions take place:

(1) D is A's neighbour;

(2) The 'savings' of A are larger than the estimated savings of D in the last stage of the simulation:

$S_A > S_D$ * informational_uncertainty,

where $S_A$ = Past Savings$_A$ + Net Revenue$_A$ – MC$_A$ – Mcplus $_A$

and $S_C$ = Past Savings $_C$ + Net Revenue $_C$ – MC $_C$

and "informational_uncertainty" equals 1 ± inference-error (as set by simulator, to capture uncertainty in information).

(3) Income of A under peace is smaller than Victory Spoils of war (Y$_A$_peace < VS)

where Y$_A$_peace = t * Pop(A) * tech(A) + x (1-t) * Pop(A) * tech(A)

and VS = prob(win)* [Y_ATT + Y_CONQ - Mc (A) - Mplus(A) ] + prob(tie) * [Y_ATT  - Mc (A) - Mplus(A)],

Y_ATT = [(Pop(A) - Death(A)) * tech(A) * tax(A)] + [ x (1-tax(A)) * (Pop (A) – Death (A)) * tech (A)]

Y_CONQ = (Pop(B) - Death(B)) * tech(B) * tax(B)

MC (A) = Fix(A) + Mult(A) * Dist(A)

Mplus(A) = Mult(A) * Dist(A,B) * z

Prob(tie)= 1- rel_power & prob(win)= rel_power if rel_power < 5 %

Prob(tie) = 0 and prob(win) = 1 if rel_power> 5 %.

Rel_power or relative power = ((Sav(A)-Sav(B))*perturbation/ Sav (A)

The attack lasts a random number of steps (between 10 and 30 steps).
*Duration of war*
An attack lasts a random number of steps (between 10 and 30).
*Formula to evaluate strength or relative power ratio of countries*
Relative power ratio (PR) = (S$_A$ − S$_D$) * random perturbation/S$_A$
*Consequences of war*
Both countries involved suffer a loss of 10% of population and accumulated wealth

If PR ≤ 1.05, war ends in a tie with no winner. (The attacker does not conquer anything. The defender does not gain any territory.)

If 1.20 ≤ PR < 1.05, partial conquest of D by A (unless the annexed territory of D includes its capital, in which case there is total annexation).

If PR > 1.20 or if the partial annexation involves the loser's capital, the winner absorbs completely the loser's country.

Depending on the mobility parameter, a fraction of the loser's accumulated wealth is transferred to the winner or it is lost.

If the attacking country has mistakenly evaluated the strength of its victim, the roles are reversed.

Depending on whether the parameters are "country-based" or not, the conquered regions either acquire the winner's default parameters or maintain their current values.

## *Rules of Secession*

*Seminal secession*

1. With some probability p, any cell decides to secede. Probability p is a function of (a) distance to capital and (b) of some parameter z. The distance to capital has a concave effect on probability to secede: it declines as cell is closer to capital and also closer to border of existing country. The parameter z is a time-declining parameter starting at 1 when the cell first belongs to the current country.

2. The seceding cell is the capital of the new country.

3. The political regime of the seceding cell is determined in the following way. If the cell belonged to a different country than the current one and has some 'memory' of that past, it will chose the regime in the past country. If it has no such memory (either it did not belong to any other country in the past or the 'memory' of belonging has decline to 0), then the seceding cell mirrors the political regime of the country it wants to secede from.

4. The seceding cell defines the potential space of secession as follows: all those (contiguous) cells whose net contribution (income – military costs) is positive.

5. After having identified the potential secession space, the seceding cell (which, again, acts as capital of the potential seceding space) determines whether that space can survive vis-à-vis its neighbors.

(The possibility of survival is calculated by looking both at the military strength and the probability of attack of neighbors. This means that there will be more secessions in areas that are contiguous to peace-prone countries.)

*Secession under democracy*

6. If both the seceding area and the existing country are democratic, the secession process works according to the following procedure:

a. If the 'seceding space' is of no interest to the existing country, the seceding space becomes a new country automatically. After a given number of rounds, the new country behaves as a 'normal' country and all rules (on war decisions etc.) apply to it with no exceptions. 'No interest' means that the seceding space does not generate net revenue for the old country (MC are larger than generated taxable income).

b. If the 'seceding cells' are equally valuable (i.e. generate net contributions) for the old country, they are allocated according to the following rule: a (weighted) formula determines under which capital the cell under dispute is better off; the formula takes into account the tax rate they pay and the distance from each capital (the higher the distance the worse the quantity/quality of goods received from the corresponding capital and therefore the higher the incentive to join the closer capital; but this is conditional on the tax effort the cell would make to each side).

Example:

```
Old capital              1/3                      2/3                  Seceding capital
|----------------------------|----------------------------|----------------------------|
Tax=20%                                                                          Tax=10%
```
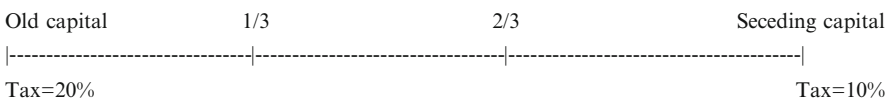
To the left of 1/3 to old capital.          To the right of 1/3 to the seceding space.

[*Note:* To set the tax of the new territory, tax t * random error. The tax t is the tax of the existing country if the seceding cell has always belonged to that country. Or the tax of the old country to which it belonged in the past.]

*Secession under authoritarian regime*

7. If any of the two countries is authoritarian, the secession process works according to the following procedure:

a. If there are no common valuable cells, the old country lets the country secede (since the seceding space is just generating net losses).
b. If there valuable cells to both parties, settlement comes through war.

At the end of the war is determined by looking at the military power of the old and the new countries:

MPOC (Military Power Old Country) = Income of all cells with positive contribution − MC of defending them.

MPNC (Military Power New Country) = Income of all cells with positive contribution − MC of defending them.

If MPOC > MPNC, the secession fails. Otherwise, the new countries becomes independent and obtains up to ½ of all the cells that are valuable to both countries. (Upper bound of ½ is arbitrary).

## Alliances

1. Alliances can only be defensive.
2. Alliances are only organized against potential aggressors.
3. At some random intervals, some countries estimate the strength of their neighbors. (This can occur at the same time those countries scan potential victims.)
4. If the scanning country W finds a stronger neighbor PA – with a strength defined over a given threshold (that the neighbor is two times stronger than the scanning unit or $S_{PA} > 2 * S_W$), the scanning country:
   a. Lists the neighbors of PA (for 'potential aggressor'), ranking them by strength.
   b. Offers an alliance to them – following the order of the list – up to the point where their accumulated strength surpasses the strength of PA.
5. The countries R that receive an alliance offer join it provided they meet the following conditions:
   i. If W and R are neighbors:
      a. Their relationship to PA in terms of strength is identical, that is, $S_{PA} > 2 * S_W$ and $S_{PA} > 2 * S_R$.
      b. W and R strengths are such that $S_c > 1.2 * S_W$. Otherwise W may prefer to avoid the alliance and have the opportunity to fight and conquer W.
   ii. If W and R are not neighbors:
      a. Their relationship to PA in terms of strength is identical, that is, $S_{PA} > 2 * S_W$ and $S_{PA} > 2 * S_R$. (former condition b does not apply here.)

6. Countries that are allies do not fight with each other.
7. Alliances remain in place for a fixed number of steps (25 as the default) unless the conditions in #5 above change.
8. Alliances are public knowledge.
9. Alliance only work to defend allied countries against the PA (potential aggressor) that has triggered the alliance to start with.
10. When looking at any country in an alliance against PA, PA assesses the strength of the alliance:
    i. As the sum of the strengths of the countries in the alliance.
    ii. Discounted by a factor that depends on the number of allies: 0.95 for a 2-state alliance, 0.9 for a three-state alliance, 0.85 for a four-state alliance, and 0.8 if there are 5 or more countries.
11. The losses or gains of allies are distributed according to a proportionality rule – that is, according to their contribution to total strength.

# References

Alesina A, Spolaore E (1997) On the number and size of nations. Quart J Econ 112:1027–1056

Alesina A (2002) The size od countries: does it matter? Harvard University. Unpublished manuscript.

Axelrod R (1984) The evolution of cooperation. Basic, New York

Benson O (1961) A simple diplomatic game. In: Rosenau JN (ed) International politics and foreign policy: a reader in research and theory. Free Press, Glencoe, IL

Boix C (2003) Democracy and redistribution. Cambridge University Press, Cambridge

Bolton P, Roland G (1997) The breakup of nations: a political economy analysis. Quart J Econ 112:1057–1090

Bremer SA, Mihalka M (1977) Machiavelli in machina: or politics among hexagons. In: Deutsch KW, Fritsch B, Jaquaribe H, Markovits AS (eds) Problems in world modeling: political and social implications. Ballinger, Cambridge

Cederman L-E (1997) Emergent actors in world politics: how states and nations develop and dissolve. Princeton University Press, Princeton, NJ

Cederman L-E (2003) Modeling the size of wars: from billiard balls to sandpiles. Am Polit Sci Rev 97:135–150

Cusack TR, Stoll RJ (1990) Exploring realpolitik: probing international relations theory with computer simulation. Lynne Reiner, Boulder, CO

Finer SE (1975) State- and nation-building in Europe: the role of the military. In: Tilly C (ed) The formation of national states in Europe. Princeton, Princeton, NJ

Gilpin R (1981) War and change in world politics. Cambridge University Press, Cambridge

Hintze O (1975) The historical essays of Otto Hintze. Edited by Felix Gilbert, with the assistance of Robert M. Berdahl. University Press, Oxford

Johnson PE (1999) Simulation modeling in political science. Am Behav Sci 42:1509–1530

Mearsheimer JJ (2001) The tragedy of great power politics. Norton, New York

Pepinsky TB (2005) From agents to outcomes: simulation in international relations. Eur J Int Relat 11:367–394

Rosato S (2011) Europe united: power politics and the making of the European Community. Cornell University Press, Ithaca, NY

Spruyt H (1994) The sovereign state and its competitors. Princeton University Press, Princeton, NJ

Spruyt H (2007) War, trade, and state formation. In: Boix C, Stokes S (eds) Oxford handbook of comparative politics. Oxford University Press, Oxford

Tilly C (ed) (1975) The formation of national states in Western Europe. Princeton University Press, Princeton, NJ

Tilly C (1990) Coercion, capital, and European states, AD 990–1990. Blackwell, Cambridge, MA

# Why Do Weak States Prefer Prohibition to Taxation?

**Desiree A. Desierto and John V.C. Nye**

## 1  Introduction

Proponents of eliminating the ban on illegal narcotics have long noted that prohibition is often ineffective and counterproductive. Recent work in the economics literature indicates that prohibiting "undesirable" goods such as drugs is inefficient [cf. Miron 2004, 2008; Becker, Murphy, and Grossman (BMG) 2006]. BMG is especially notable for discussing the problem of controlling the spread of an illegal good when enforcement is imperfect and avoidance costs are factored in. BMG show that prohibition can only be justifiable if the good is really undesirable, that is, if the marginal social value of consuming the good is very low or, in the case of inelastic demand, sufficiently negative. This is essentially because illegal good producers waste resources in order to avoid being detected, captured and/or penalized. They can, for instance, bribe prohibition enforcers – from police officers to courts – in order to continue supplying the market. Such avoidance costs, as BMG imply, are a dead-weight loss to society. On the other hand, legalization and taxation of the good are a more efficient way of curbing consumption, since taxes paid for the good are eventually plowed back to society. In other words, avoidance costs are a leakage, while taxes are not.

Desierto and Nye (2010, Two wrongs make a right: the market for illegal goods in the presence of corruption, unpublished manuscript) show, however, that avoidance costs are not necessarily a loss, if bribes are treated as additional income or revenue to enforcers. And, in contrast, taxes can be a loss since the government might use tax revenues inefficiently and might even illegally appropriate some of it. In this latter case, losses might be limited if the appropriated amount is also treated

D.A. Desierto (✉)
School of Economics, University of the Philippines, Diliman, Quezon City 1101, Philippines
e-mail: dadesierto@econ.upd.edu.ph

J.V.C. Nye
Economics Department, George Mason University, Carow Hall MSN 1D3, 4400, University Drive, Fairfax, VA 22030, USA
e-mail: jnye@gmu.edu

as additional revenue to tax enforcers, but corruption by tax enforcers is less likely than the corruption committed by prohibition enforcers. Bribery between illegal good producers and prohibition enforcers is likely to be more sustainable because both are complicit in the corruption activity, while legal good producers/taxpayers have little incentive to allow tax enforcers to illegally appropriate the tax.

Desierto and Nye thus show that in a second-best world in which corruption is present, prohibition is likely to be more efficient than taxation. This paper provides further support for this result.

None of the existing literature has dealt with the positive question of why undesirable goods are so often prohibited if enforcement is so lax. In contrast, our approach suggests why prohibiting undesirable goods like illegal narcotics or gambling is especially common in places with weak institutions and a reputation for corruption. Section 2 graphically analyzes equilibria in the BMG and Desierto and Nye models, and then compares the likely losses from prohibition versus taxation. In Sect. 3, we present some casual evidence indicating that not only is prohibition a more commonly employed method of curbing consumption of drugs, prostitution, and gambling than legalization and taxation, but that prohibition is more prevalent in countries with greater corruption. Section 4 concludes and thus provides a positive rationale for prohibition and an efficiency motivation for prevalence of prohibition in weak states.

## 2   Graphical Analysis

In BMG and Desierto and Nye, illegal producers can offer bribes to prohibition enforcers as part of total avoidance costs $AC$ that they incur.[1] The illegal producer chooses the level of $AC$ that minimizes its expected cost, given the level of prohibition/enforcement effort $E$ that the government undertakes. Meanwhile, given the amount $AC$ that the illegal producer spends, the government chooses its level of $E$ that maximizes social welfare $W$. The main difference between BMG and Desierto and Nye is that in the latter, the bribes to enforcers are internalized. That is, the government treats bribes as additional revenues, such that total social welfare is given by: $W = V + R + B$, where $V$ is the social value of consumption (net of all other consumption externalities), $R$ are the producers' net revenues, and $B$ are (net) bribe revenues of enforcers. (In BMG, $W = V + R$).

Solving simultaneously for $AC$ and $E$, equilibrium is achieved at the point where: $V_q - B_q = MR$, where $V_q$ is the marginal social value of consumption, $B_q$ is marginal bribe revenue to enforcers, and $MR$ is marginal revenue of producers. In contrast, equilibrium in BMG is captured by: $V_q = MR$. Thus, internalizing bribes

---

[1]Desierto and Nye, however, make the bribe amount explicit, by assuming that a fixed fraction $\beta$ of total $AC$ are in the form of bribes. This fraction captures the overall extent of corruption in the environment, as it is the permissible level at which enforcers can extract a bribe without being detected by the government.

tends to achieve higher $AC$ and $E$, precisely because the bribe benefits provide additional incentive to prohibition enforcers, which then require illegal producers to spend more to counteract the increased enforcement efforts. The optimal level of consumption is thus lower, making prohibition more effective in curtailing consumption, and more efficient since it need not incur additional losses in decreasing consumption further – the additional avoidance costs in the form of bribes are not 'wasted' but go to enforcers as bribe revenues.

The following graphs illustrate this result. We first consider the corner-solution case, in which either of two extremes is socially optimal – completely freeing the market (i.e. legalization), or full, all-out enforcement against the good which drives consumption to zero. We show that internalizing bribes makes the latter more likely to be optimal than the former. The next case depicts an interior solution, where some imperfect level of enforcement is optimal, allowing some positive level of consumption. Here it is shown that internalizing bribes actually decreases optimal consumption efficiently. We then compare this outcome to taxation as an alternative method of restricting consumption, and show that taxation is likely to be less efficient than prohibition with (internalized) bribery. The result holds even when tax collectors/enforcers are also corrupt (like prohibition enforcers), and the corruption is internalized. This is essentially because the internalization of corruption is limited when the good is legal than when it is prohibited – it is more difficult for tax enforcers to (illegally) appropriate taxes paid by *legal* producers than for prohibition enforcers to extract bribes from producers who have to keep quiet to stay underground.

## 2.1 Full Enforcement Versus Legalization

The graph depicted in Fig. 1 reproduces BMG's Fig. 2. Quantity $Qu$, although it satisfies $V_q = MR$, is not socially optimal as it violates the second-order condition. (To the right of $Qu$, $V_q$ falls slower than $MR$, so it makes sense to keep increasing output. To the left of $Qu$, $MR$ rises faster than $V_q$, so it makes sense to keep
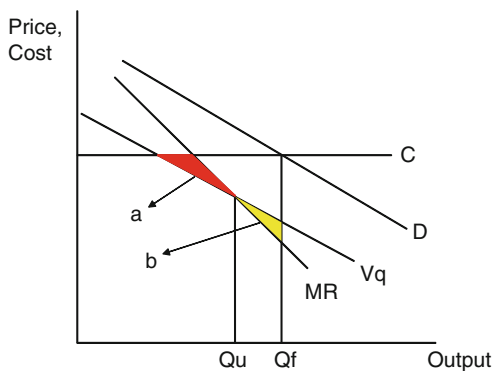


**Fig. 1** BMG equilibrium (corner solution)

decreasing output.) Thus, the only possible optimal consumption levels are either zero (i.e. full enforcement) or the free-market level $Qf$ (i.e. zero enforcement) at which demand $D$ is equal to the marginal cost $C$ of producing the good. This depends on the relative gains depicted by the triangles to the left and to the right of $Qu$ – if triangle $b$ is larger than triangle $a$, then it is more socially optimal to legalize the good, at which case $Qf$ is achieved. Otherwise, it is better to fully prohibit the good to curtail consumption to zero.

Figure 2 depicts equilibrium in the Desierto–Nye model, that is, when bribes to enforcers are internalized. At quantity $Qb$, $V_q - B_q = MR$, but this also violates the second-order condition. Thus, as in BMG, either the zero or free-market consumption level is optimal, depending on the relative social gains as captured by the triangles between $(V_q - B_q)$ and $MR$ to the right and to the left of $Qb$. Figure 3 labels these areas $c$ and $d$. In this example, $d$ is smaller than $c$, which makes full enforcement (zero consumption) socially optimal. More generally, however, the internalization of bribes always makes area $c > a$ and $d < b$. That is, the benefits of legalization shrinks vis-à-vis the benefits of (full) enforcement, thus making free
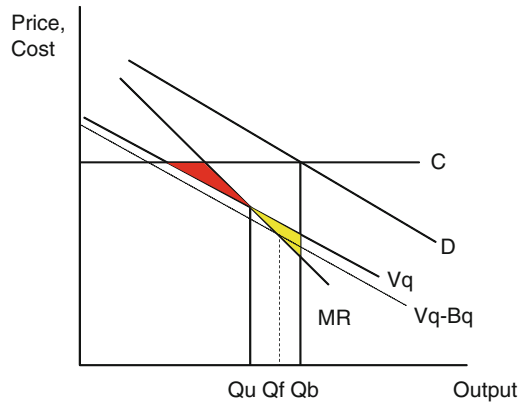


**Fig. 2** Desierto–Nye equilibrium (corner solution)
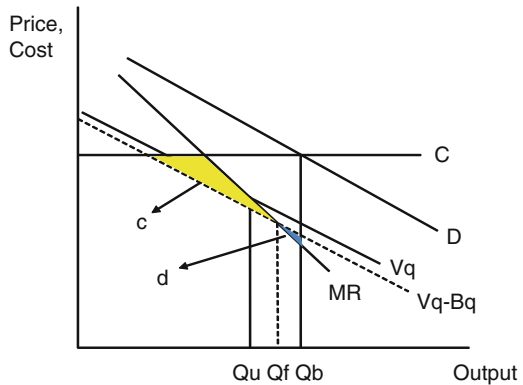


**Fig. 3** With bribery, legalization is less likely to be optimal

**Fig. 4** BMG equilibrium (interior solution)
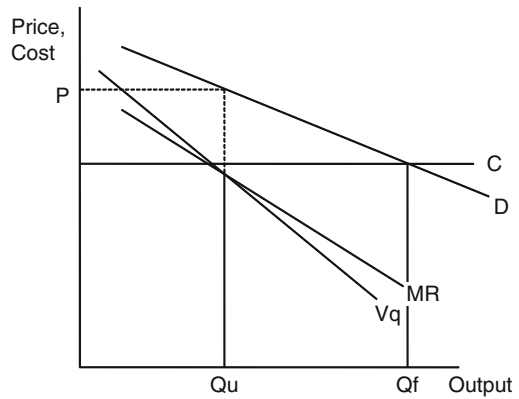


market consumption less likely to be optimal. This is because legalization foregoes the opportunity of enforcers to earn additional revenue by extracting bribes when the good is prohibited.[2]

## 2.2 Imperfect Enforcement

When the second-order condition is satisfied at $V_q = MR$, a non-zero but restricted level of consumption is socially optimal, and some imperfect level of enforcement is justifiable. In Fig. 4, the $V_q$ line is now steeper than $MR$. At low levels of quantity, each extra good is valued more by society (as consumption good) than by producers (as source of revenue). After $Qu$, each additional good becomes more costly to society than to producers.[3] Hence, the socially optimal consumption level is achieved at $Qu$, (which represents some imperfect enforcement level), with corresponding price $P$.

Note that producers are perfectly competitive and enjoy zero profits. Their total revenue, given by area $f + e$ in Fig. 5, is used to cover the total cost of manufacturing $Qu$, which is captured by area $e$, and 'avoidance costs', as depicted by area $f$. There is

---

[2]With $B_q > 0$, it is always the case that $(V_q - B_q)$ is less than, or below, $V_q$. In Figs. 2 and 3, the $(V_q - B_q)$ line has roughly the same slope as the $V_q$ line, but this is not necessarily the case. While the $(V_q - B_q)$ line cannot be flatter than $V_q$ (i.e. they cannot intersect), it can be steeper. This can happen if $Bq$ is more sensitive to output changes than $V_q$ – that is, if deriving bribe benefits is more costly than enjoying 'legitimate' consumption. In this case, the result is stronger, since area $d$ would be smaller, and area $c$ larger.

[3]Or, to state it analogously with the previous subsection: to the right of $Qu$, $V_q$ is falling faster than $MR$, so it is not socially optimal to increase quantity further. To the left of $Qu$, $V_q$ is rising faster than $MR$, so it is not optimal to decrease quantity.

Fig. 5 Avoidance costs
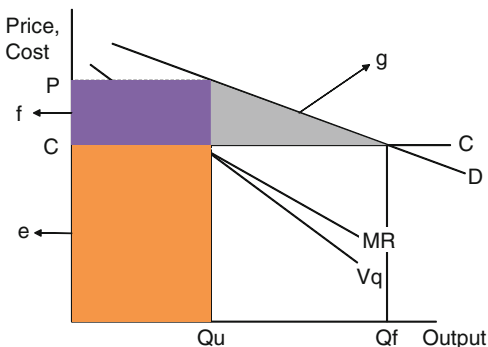


Fig. 6 Desierto–Nye
equilibrium (interior solution)
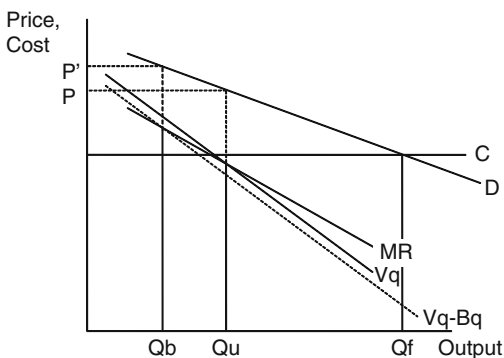


a loss of consumer surplus, captured by area $g$, since not all affordable demand is served.

Figure 6 depicts the interior solution in the Desierto–Nye model. It can be seen that when bribery is taken into account, the optimal consumption level given by $(V_q - B_q) = MR$ is lower at $Qb < Qu$. This is because enforcers have an incentive to increase effort. The higher effort is reflected in higher price $P'$ – that is, illegal producers spend more on avoidance per unit in order to counter the increased enforcement. Note that since some avoidance costs are spent on bribes, and the bribes are internalized, this fraction of avoidance costs is not really wasted, but just goes to enforcers as their revenues.

As Fig. 7 shows, restricting consumption to $Qb$ entails enforcers exerting a high enough effort (reflected in $P'$) and producers spending area $i + k$ to be able to sell $Qb$ amount to consumers without getting caught. (That is, of course, after manufacturing $Qb$ and incurring cost area $h$.) Out of $i + k$, area $k$ goes to enforcers as net bribes (while $i$ might be spent on litigation, relocation, and other ways to avoid being caught and punished). The gross bribe revenue is given by area $j + k$ – as though the producer spends an equivalent of this to pay for the bribe and thus incurs additional manufacturing cost $j$ plus premium $k$. In this manner, the extra

**Fig. 7** Prohibition is more efficient when bribery is internalized



production of $Qu - Qb$ is 'disguised' insofar as it is used only to pay for the bribe. Or to put it differently, the enforcer receives bribes of $(Qu - Qb)$ amount of drugs and can sell it to other producers at a premium and get profit equal to k.

Thus, while consumption is curtailed at $Qb$, all the production resources are equivalent to producing $Qu$. This, though, is not a waste – it just goes to enforcers. Enforcement is more effective and efficient when bribery is taken into account since curtailing consumption to $Qb$ does not incur any additional losses greater than the area $g$. In other words, if $Qb$ were the target consumption level (in both the BMG and Desierto–Nye model), not internalizing bribery (BMG result) would incur losses the total of $k + g$, while internalizing it (Desierto–Nye result) would only incur loss $g$.[4]

## 2.3   Imperfect Enforcement Versus Taxation

Suppose we again assume that $Qb$ is the target optimal consumption level. Then a high-enough tax rate $t$ can also achieve this, as illustrated in Fig. 8. The total revenue is given by area $i + h$, where $h$ covers manufacturing costs of $Qb$, while $i$ are paid as taxes to the government. The potential loss of consumer surplus is $k + g$, which is bigger than the loss $g$ incurred by prohibition with internalized bribery (in Fig. 7). However, if tax revenues are used efficiently by the government, this loss may be offset or minimized. Government investments and spending on public goods and services might be 'profitable' such that an equivalent of $i$ or more is transferred back to society. (In the example in Fig. 8, it turns out that a pure, 'one-for-one', transfer will still incur losses since area $i$ is smaller than area $k + g$. To eliminate all losses, the tax revenues $i$ have to be used profitably enough such that they have returns the equivalent of $k + g$).

---

[4]Recall that in Fig. 4, area $g$ is lost when the optimal consumption level is $Qu$. If optimal consumption were less than this, the loss would be greater than $g$.

**Fig. 8** Taxation



**Fig. 9** Is taxation as efficient?



The question, then, is how much of area *i* can be used efficiently to cover some or all of *k* + *g*, or, at least an area equivalent to *k*, to make taxation as efficient as prohibition with internalized bribery. The likely answer is that taxation may not be as efficient. In certain cases, area *k* might be bigger than area *i*, which means that even a pure transfer of tax revenues might not achieve the same efficiency as prohibition with bribery. It would then take more than a pure transfer, but this is unlikely if government projects/investments have low returns.

But suppose we let corrupt tax enforcers apportion tax revenues, and internalize these illegal appropriations as additional revenues to tax enforcers. Would the added incentive to corrupt tax enforcers induce them to exert more tax-collecting efforts so as to curb more consumption without generating further efficiency losses? That is, can we have the following situation in Fig. 9, which is exactly equivalent to the result obtained by prohibition with internalized bribery? If a total of *i* + *k* tax revenues can be collected, but area *k* is appropriated by tax enforcers as 'net illegal tax revenue', then the loss is just area *g* – the same as in the prohibition with bribery case – and taxation becomes as efficient as the latter. (It is also as effective, since

**Fig. 10** Taxation is less efficient than prohibition with internalized bribery
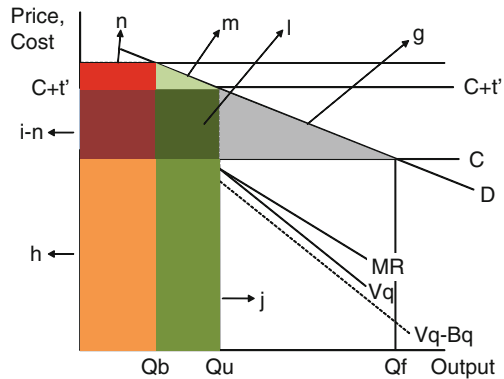


*Qb* is still optimal consumption – additional (*Qu* − *Qb*) is produced only to cover the equivalent of what tax enforcers would appropriate from tax revenues.)

This situation, however, is unlikely. The more likely result is shown by Fig. 10, where only area *l* can be illegally appropriated by tax enforcers, and only area *i* − *n* ends up in government coffers. That is, not only is the receipt of the government smaller, but also what tax enforcers can illegally appropriate. Thus, compared to prohibition with bribery (which has loss *g* in Fig. 7), taxation with illegal appropriation incurs a loss of *n* + *m* + *g*.

The reason for this is that the price of the good when some taxes are appropriated drops down uniformly to *C* + *t'*, while in the prohibition case, the effective price of the bribe is, as it were, discriminatory. (Notice in Fig. 7 that price goes up as the bribe quantity goes down.) In contrast, when tax enforcers' corrupt activities are internalized, legal producers have to produce up to *Qu* to 'fund' illegal appropriation, but price-discrimination is difficult, if not altogether impossible, because it is hard to 'disguise' *Qb* − *Qu*. 'Corruption' quantities, i.e. those beyond *Qb*, are priced the same as 'pure consumption' quantities, i.e. those until *Qb*, since all these units are homogeneous to legal producers and/or consumers. That is, the latter have very little incentive to allow corrupt enforcers to appropriate taxes, or equivalently, to allow *Qb* − *Qu* to go 'unnoticed'. Thus, this additional supply in circulation induces price to go down (uniformly) to *C* + *t'*. In the prohibition case, however, illegal producers are complicit with bribe-takers – they do not want to get caught, afterall – and, hence, each bribe unit in *Qb* − *Qu* can remain 'hidden' and can then be priced discriminately.

Furthermore, tax collectors can rarely appropriate tax revenues only from the good. If collectors are corrupt, they may appropriate all other kinds of tax revenue. Or they may end up appropriating revenues from some goods, but not from others. In this case, the benefits of tax enforcers from illegally appropriating tax revenues from the good in question are difficult to isolate and, hence, *Qb* − *Qu* are not distinguishable from *Qb*. In contrast, bribery revenues from prohibited goods are

more easily identified and isolated which makes $Qb - Qu$ distinct from $Qb$.[5] Again, this allows for price-discrimination of the corruption goods in the prohibition case, but not in the taxation case.

In other words, it is more difficult to internalize corruption when the good is legal and taxed, than when it is prohibited. This makes corruption more efficient in illegal, than in legal, environments.

## 3 Prohibition and Corruption

Prohibition is actually widespread. Among a sample of 101 countries that we have obtained, 100 prohibit drugs, 66 prohibit prostitution, and 33 prohibit gambling.[6] Of course, this fact alone need not be inconsistent with BMG's results – it might just be that the social marginal values of consumption for these goods are very low. Note, however, that some countries that are roughly comparable in terms of levels of development, culture, and/or geographical location, and thus have arguably similar social marginal values of consumption of undesirable goods like drugs, prostitution and gambling, can still have different approaches to curbing production/consumption of these goods. For instance, drugs are illegal in the US, Canada and most of Europe, but are legal in the Netherlands; prostitution is legal in France, but not in Switzerland; gambling is prohibited in Thailand, but not in the Philippines.

In addition, some illegal goods, e.g. gambling and pirated products, might pose little negative consumption externalities and might even have high consumption value and/or produce positive externalities, and yet they are still persistently prohibited. It is not clear, therefore, that the BMG result is the only explanation for why prohibition seems to be a sustainable equilibrium. As Desierto and Nye posit, bribes and rent-seeking by corrupt enforcers are easier to extract when markets are illegal.

Indeed, prohibition seems to be more common in countries that are more prone to corruption and that have weaker state capacities than those which merely tax. We obtain the 2009 Corruption Perceptions Index (CPI) from Transparency International for each country in the sample and compute the average CPI for countries that prohibit, and for those that legalize and tax, drugs, prostitution and gambling.[7] Tables 1–3 summarize the results and clearly suggest that corruption is higher (i.e. the average corruption score is lower) among countries that prohibit, than among those that legalize and tax.[8]

---

[5]Glaeser and Shleifer (2001) similarly note that its easier for an enforcer to detect possession of an illegal good than to verify whether taxes (on a legal good) have been paid.

[6]See Appendix for details.

[7]CPI scores reflect "political stability, long-established conflict-of-interest regulations and solid, functioning public institutions".

[8]Of course, there are many reasons why such association between the extent of corruption and the strength of prohibition might not be truly causal. For instance, developing countries that tend to be more corrupt might also have less liberal attitudes towards drugs, prostitution and gambling.

**Table 1** Drugs

|                                                    | Prohibition | Taxation |
|----------------------------------------------------|-------------|----------|
| Number ($n = 101$)                                 | 100         | 1        |
| Ave. corruption score (range: 9.4 to 1.1)          | 3.979       | 8.9      |

**Table 2** Prostitution

|                                                    | Prohibition | Taxation |
|----------------------------------------------------|-------------|----------|
| Number ($n = 101$)                                 | 66          | 35       |
| Ave. corruption score (range: 9.4 to 1.1)          | 3.53        | 4.96     |

**Table 3** Gambling

|                                                    | Prohibition | Taxation |
|----------------------------------------------------|-------------|----------|
| Number ($n = 101$)                                 | 33          | 68       |
| Ave. corruption score (range: 9.4 to 1.1)          | 3.31        | 4.26     |

# 4   Conclusions

Prohibition of undesirable goods can be an efficient response especially if govern-
ments are corrupt and enforcement is imperfect. Making a good illegal provides
prohibition enforcers opportunities to extract a bribe from illegal producers. Since
the latter are willing to pay the bribes, a sustainable equilibrium is achieved
whereby enforcers keep their efforts high and illegal producers spend more to
stay in business. While this is seen by BMG as a 'waste' of resources, Desierto
and Nye argue that this is not necessarily the case, if the bribes received by
enforcers are internalized as 'bribe revenues'. The result is that, for the same
amount of resources spent by illegal producers so as not to get caught and/or
punished, spending more of it as bribes to enforcers reduces optimal consumption
more than if such spending were apportioned to other 'external' avoidance activ-
ities. This is essentially because it is easier to internalize enforcers' bribe revenues
than, say, the additional income of lawyers hired by illegal producers. Such lawyers
cannot easily affect the amount of illegal goods that end up in the market, whereas
enforces directly determine it.

   The same result holds even when compared to taxation. Tax collectors can
indeed determine the amount of tax that goes to the government and, hence, if
they are incentivized by the possibility of illegally appropriating some of the tax
revenues, they might end up increasing tax enforcement efforts, which would then
decrease production/consumption of the good. However, this is a less efficient way
of curbing consumption because tax collectors cannot appropriate tax revenues as
easily as prohibition enforcers can extract bribes from illegal producers. In illegal
markets, bribes and rent-seeking are more easily masked, whereas legal markets are
more transparent. Thus, corrupt prohibition enforcers have greater incentive to

enforce against the 'undesirable' good than corrupt tax collectors. The latter care less about enforcing the tax on the good since their corrupt action can be more easily detected or reported and they need not focus on collecting taxes only on that good as there are other sources of tax revenues and tax collectors rarely specialize on any one undesirable good.

Highlighting the role of corruption can provide additional insights on why it is easier in some cases to justify prohibition over taxation. Of course, we have not specifically dealt with other costs and problems associated with prohibition itself (e.g. violence and loss of lives resulting from the enforcement of prohibition laws), but nor have we identified all sources of negative externalities from the consumption of the good (e.g. health treatment costs). The former would tend to increase the total social deadweight loss from prohibition, while the latter would tend to decrease it. While a more thorough, general-equilibrium analysis is needed to resolve the issue, studying the effect of corruption in illegal markets is a step that has not yet been sufficiently explored in the literature.

# Appendix

+Some form of partial prohibition.
   See the following sources for details:
   Drugs
   http://www.waikato.ac.nz/international/students/general/law.shtml
   http://www.cecc.org.nz/content/library/TAX_PERSPECTIVES.pdf
   http://www.frommers.com/destinations/denmark/0220020157.html
   http://en.wikipedia.org/wiki/Misuse_of_Drugs_Act_(Singapore)
   http://www.amsterdam.info/drugs/
   http://en.wikipedia.org/wiki/Drug_policy_of_Canada
   http://norml.org/index.cfm?Group_ID=4415
   http://www.citizensinformation.ie/categories/justice/criminal-law/criminal-offen-ces/drug_offences
   http://en.wikipedia.org/wiki/Law_of_Japan
   http://www.drugscope.org.uk/resources/faqs/faqpages/what-are-the-uk-drug-laws
   http://en.wikipedia.org/wiki/Drug_Enforcement_Administration
   http://travel.state.gov
   http://www.druglibrary.org/schaffer/LIBRARY/southam1.htm
   http://www.fco.gov.uk/en/travel-and-living-abroad
   http://www.smartraveller.gov.au
   http://www.pacificprime.com/countries/israel/
   http://www.gomadrid.com/practic/local-laws.html

| Country | CPI 2009 score (Highest – least corrupt) | | Drugs Prohibition | Drugs Taxation | Gambling Prohibition | Gambling Taxation | Prostitution Prohibition | Prostitution Taxation |
|---|---|---|---|---|---|---|---|---|
| New Zealand | Developed | 9.4 | • | | | • | | • |
| Denmark | Developed | 9.3 | • | | | • | | • |
| Singapore | Developed | 9.2 | • | | | •+ | | •+ |
| Switzerland | Developed | 9 | • | | | • | | |
| Netherlands | Developed | 8.9 | | • | | • | | • |
| Canada | Developed | 8.7 | • | | | •+ | | • |
| Luxembourg | Developed | 8.2 | • | | | • | | •+ |
| Germany | Developed | 8 | • | | | • | | • |
| Ireland | Developed | 8 | • | | | • | • | |
| Japan | Developed | 7.7 | • | | | • | • | •+ |
| United Kingdom | Developed | 7.7 | • | | | • | | |
| United States | Developed | 7.5 | • | | • | | •+ | • |
| Qatar | Developed | 7 | • | | | • | • | • |
| France | Developed | 6.9 | • | | • | | | •+ |
| Chile | Developing | 6.7 | • | | | • | | |
| Uruguay | Developing | 6.7 | • | | • | • | | |
| Estonia | Developing | 6.6 | • | | | • | •+ | |
| UAE | Developed | 6.5 | • | | | • | • | |
| Israel | Developed | 6.1 | • | | | •+ | | •+ |
| Spain | Developed | 6.1 | • | | • | • | • | •+ |
| Portugal | Developed | 5.9 | • | | | • | | |
| Botswana | Developing | 5.6 | • | | • | • | • | |
| Brunei | Developing | 5.5 | • | | | | • | |
| South Korea | Developed | 5.5 | • | | • | • | • | |
| Mauritius | Developing | 5.4 | • | | | | • | |
| Bahrain | Developing | 5.1 | • | | • | • | • | |
| Hungary | Developing | 5.1 | • | | | • | • | • |
| Jordan | Developing | 5 | • | | • | | • | • |
| Czech Republic | Developed | 4.9 | • | | | • | | |
| South Africa | Developing | 4.8 | • | | | • | • | |
| Malaysia | Developing | 4.5 | • | | | • | •+ | |
| Namibia | Developing | 4.5 | • | | | • | • | |
| Cuba | Developing | 4.4 | • | | • | • | • | |
| Turkey | Developing | 4.4 | • | | | • | • | • |

(*continued*)

| Country | CPI 2009 score (Highest – least corrupt) | Drugs Prohibition | Drugs Taxation | Gambling Prohibition | Gambling Taxation | Prostitution Prohibition | Prostitution Taxation |
|---|---|---|---|---|---|---|---|
| Italy | Developed — 4.3 | • | | | • | | • |
| Saudi Arabia | Developing — 4.3 | • | | • | | • | |
| Tunisia | Developing — 4.2 | • | | | • | • | |
| Georgia | Developing — 4.1 | • | | | • | • | |
| Kuwait | Developed — 4.1 | • | | • | | • | |
| Ghana | Developing — 3.9 | • | | | • | • | |
| Bulgaria | Developing — 3.8 | • | | | • | | •+ |
| Greece | Developed — 3.8 | • | | | • | | • |
| Brazil | Developing — 3.7 | • | | • | | | • |
| Peru | Developing — 3.7 | • | | | • | | • |
| China | Developing — 3.6 | • | | | • | • | |
| Serbia | Developing — 3.5 | • | | | • | • | |
| El Salvador | Developing — 3.4 | • | | | • | | •+ |
| India | Developing — 3.4 | • | | • | | • | |
| Thailand | Developing — 3.4 | • | | • | | • | |
| Malawi | Least developed — 3.3 | • | | | • | | |
| Mexico | Developing — 3.3 | • | | • | | | • |
| Rwanda | Least developed — 3.3 | • | | | • | | |
| Albania | Developing — 3.2 | • | | | • | • | |
| Liberia | Least developed — 3.1 | • | | | • | | |
| Sri Lanka | Developing — 3.1 | • | | | • | | • |
| Jamaica | Developing — 3 | • | | • | | | • |
| Senegal | Least developed — 3 | • | | | • | | |
| Zambia | Least developed — 3 | • | | | • | | |
| Argentina | Developing — 2.9 | •+ | | | • | • | |
| Niger | Least developed — 2.9 | • | | • | | | |
| Algeria | Developing — 2.8 | • | | • | | | |
| Egypt | Developing — 2.8 | • | | | • | | |
| Indonesia | Developing — 2.8 | • | | • | | | |
| Mali | Least developed — 2.8 | • | | • | | | • |
| Armenia | Developing — 2.7 | • | | | • | | • |
| Ethiopia | Least developed — 2.7 | • | | | • | | • |
| Kazakhstan | Developing — 2.7 | • | | | • | | |
| Vietnam | Developing — 2.7 | • | | | • | • | |
| Guyana | Developing — 2.6 | • | | | • | • | |
| Syria | Developing — 2.6 | • | | • | | • | |

| Country | CPI 2009 score (Highest – least corrupt) | Drugs Prohibition | Drugs Taxation | Gambling Prohibition | Gambling Taxation | Prostitution Prohibition | Prostitution Taxation |
|---|---|---|---|---|---|---|---|
| Honduras | Developing | 2.5 | • | | | • | | • |
| Lebanon | Developing | 2.5 | • | | | • | • | |
| Maldives | Developing | 2.5 | • | | • | • | • | |
| Mozambique | Least developed | 2.5 | • | | | • | • | • |
| Nicaragua | Developing | 2.5 | • | | • | • | • | • |
| Bangladesh | Developing | 2.4 | • | | • | | | • |
| Pakistan | Developing | 2.4 | • | | | | | |
| Philippines | Developing | 2.4 | • | | • | | • | |
| Azerbaijan | Developing | 2.3 | • | | • | | • | |
| Nepal | Developing | 2.3 | • | | | • | • | • |
| Ecuador | Developing | 2.2 | • | | | | • | |
| Kenya | Developing | 2.2 | • | | | • | • | |
| Russia | Developing | 2.2 | • | | | • | • | |
| Ukraine | Developing | 2.2 | • | | | • | | |
| Zimbabwe | Least developed | 2.2 | • | | • | | | |
| Papua New Guinea | Developing | 2.1 | • | | | • | | |
| Paraguay | Developing | 2.1 | • | | • | • | • | |
| Yemen | Developing | 2.1 | • | | | | • | |
| Cambodia | Developing | 2 | • | | • | | • | |
| Laos | Developing | 2 | • | | | | • | |
| Angola | Developing | 1.9 | • | | | • | • | • |
| Kyrgyzstan | Developing | 1.9 | • | | | • | + | |
| Venezuela | Developing | 1.9 | • | | | | | |
| Haiti | Least developed | 1.8 | • | | • | • | • | |
| Iran | Developing | 1.8 | • | | • | • | • | |
| Uzbekistan | Developing | 1.7 | • | | • | | • | |
| Chad | Least developed | 1.6 | • | | | | • | • |
| Iraq | Developing | 1.5 | • | | | | • | |
| Myanmar | Developing | 1.4 | • | | • | • | • | |
| Afghanistan | Least developed | 1.3 | • | | • | • | • | |
| Somalia | Least developed | 1.1 | • | | | | • | |

http://www.cato.org/pub_display.php?pub_id=10080

http://www.gov.bw/en/Visitors/Topics/Entry-Requirements/Entry-Requirements/

http://www.narcotics.gov.bn/DrugLaws.htm

http://www.cnn.com/2009/WORLD/asiapcf/05/05/asia.drug.offence.penalties/
index.html

http://mauritius.usembassy.gov/uploads/images/L-Sv5VAGEik02jWNzaqqow/
InfoSheet-Drug-Offenses-in-Mauritius.pdf

http://www.drugpolicy.org/docUploads/hungary.pdf

http://www.unodc.org/egypt/en/country_profile_jordan.html

http://www.emcdda.europa.eu/attachements.cfm/att_93232_EN_TDSI09001ENC.
pdf

http://www.doh.gov.za/docs/policy/drugsjan1996.pdf, http://travel.state.gov/travel/
cis_pa_tw/cis/cis_1008.html

http://havanajournal.com/politics/entry/the-war-on-drugs-in-cuba/

http://turkey.usembassy.gov/penalities_for_drug_offenses.html

http://stopthedrugwar.org/chronicle-old/430/italy.shtml

http://www.photius.com/countries/saudi_arabia/national_security/saudi_arabia_
national_security_crime_and_punishment.html

http://www.thesite.org/travelandfreetime/travel/beingthere/druglawsabroad

http://www.georgiadefenders.com/drugs.schedulepenalties.htm

http://www.fco.gov.uk/en/travel-and-living-abroad/travel-advice-by-country/
middle-east-north-africa/kuwait?ta=lawsCustoms&pg=3

http://bulgaria.angloinfo.com/countries/bulgaria/teenagers.asp#drugs

http://profiles.emcdda.europa.eu/html.cfm/index19688EN.html#nlaws

http://www.dfa.ie/home/index.aspx?id=8770

http://www.newssafety.org/index.php?option=com_content&view=section&id
=90&layout=blog&Itemid=100302index.html, http://www.ihra.net/Assets/489/
1/DeathPenaltyforDrugOffences.pdf

http://www.dfa.ie/home/index.aspx?id=451

http://www.nlgsolutions.com/packages/show_country.asp?countryid=MW

http://www.shipdetective.com/ports/ports_of_call/western_ports/mexico/criminal_
penalties.htm

http://www.voyage.gc.ca/countries_pays/report_rapport-eng.asp?id=250000

http://stopthedrugwar.org/chronicle/604/liberia_harsh_new_drug_law

http://news.bbc.co.uk/2/hi/south_asia/296124.stm

http://www.photius.com/countries/jamaica/national_security/jamaica_nationa_
security_justice.html

http://www.irinnews.org/Report.aspx?ReportId=75953

http://www.corpun.com/zmjur1.htm, http://travel.state.gov/travel/cis_pa_tw/
cis/cis_1062.html#Criminal_Penalties

http://www.nasarawastate.org/articles/538/1/NDLEA-destroys-6-000kg-of-illicit-
drugs-in-Niger/Page1.html

http://www.nidirect.gov.uk/index/travel-and-transport/travelling-abroad/if-things-
go-wrong/drugs-1.htm

http://www.gomideast.com/egypt/fastfax.html

http://www.iwpr.net/report-news/armenia-goes-soft-drugs
http://www.silkroadstudies.org/new/inside/research/narcotics_crime/FactSheet/2004/Kazakhstan.pdf
http://www.hawaii.edu/hivandaids/Honduras_Country_Brief__Drug_Situation_Report.pdf
http://www.travelinsurancedirect.com.au/travel_advice.php?country=Lebanon
http://www.iol.co.za/index.php?set_id=1&click_id=84&art_id=nw20070626143321534C948897
http://nicaragua.usembassy.gov/rpt_incsr_2010.html
http://www.drugpolicy.org/global/drugpolicyby/asia/casia/
http://www.voyage.gc.ca/countries_pays/report_rapport-eng.asp?id=19000
http://www.nepalvista.com/nepalfaq/drug.html
http://www.fpif.org/articles/beyond_the_drug_war
http://aglobalworld.com/international-countries/Africa/Kenya.php
http://www.hrw.org/en/news/2009/11/30/drugs-punitive-laws-policies-and-policing-practices-and-hivaids
http://www.mcaz.co.zw/legislation/DANGEROUS%20DRUGS%20ACT%20(Chapter%201502).pdf
http://www.webehigh.com/city/detail.php?CITYID=2504
http://aglobalworld.com/international-countries/Australia-and-the-Oceania/Papua-New-Guinea.php
http://www.servat.unibe.ch/icl/pa00000_.html
http://www.1uptravel.com/travelwarnings/kyrgyz-republic.html
http://www.latinamericanstudies.org/haiti/boom.htm
http://deathpenaltynews.blogspot.com/2010/05/iran-hangs-6-drug-traffickers.html
http://www.country-data.com/cgi-bin/query/r-14500.html
http://www.otal.com/ctbl/chadtravel.htm
http://www.niqash.org/content.php?contentTypeID=75&id=2093⟨=0
http://www.somalilandtimes.net/s1/2005/239/18.shtml
Gambling
http://www.gamblingcommission.govt.nz/gawebsite.nsf
http://www.eastonbh.ac.nz/?p=83
http://www.wwlegal.com/module-subjects-viewpage-pageid-53.html
http://www.asgam.com/article.php?id_article=1167&page=2
http://www.ejpd.admin.ch/etc/medialib/data/
http://www.worldgamblingreview.com
http://www.responsiblegambling.org/articles/legalization_of_gambling_in_canada_july_2005.pdf
http://www.homepokergames.com/lotterytax.php
http://gamingzion.com/gamblingnews/profitable-yet-untouched-gambling-market-lies-in-luxembourg-1147
http://www.gaminglaw.eu/actuality-313/focus-europe-germany-on-line-flutters.html
http://www.bettingmarket.com/eriesinnen090101.htm
http://www.lkshields.ie/htmdocs/publications/pub302_betting_duty.htm

http://en.wikipedia.org/wiki/Gambling_in_Japan
http://www.gamblingcommission.gov.uk/
http://www.worldcasinodirectory.com/american-casinos.asp
http://www.worldcasinodirectory.com/france
http://www.latestcasinobonuses.com/is-gambling-online-illegal-in-france.html
http://www.liebertonline.com/doi/abs/10.1089/1092188041427319?journalCode=glr
http://ec.europa.eu/taxation_customs/resources/documents/tax_sys_est_pol_hu_en.pdf
http://www.gamblingcompliance.com/node/37298taxation_customs/resources/documents/tax_sys_est_pol_hu_en.pdf
http://www.alloexpat.com/abu_dhabi_expat_forum/gambling-in-uae-t2460.html
http://gulfnews.com/about-gulf-news/al-nisr-portfolio/xpress/strict-laws-against-gambling-in-the-uae-1.572389
http://www.uv.es/ibanezs/SpanishTLRG.pdf, http://gamingzion.com/spain/
http://www.taxrates.cc/html/portugal-tax-rates.html
http://www.osac.gov/Reports/report.cfm?contentID=114778
http://gamingzion.com/gamblingnews/south-korea-police-corruption-leads-to-mandatory-surveillance-1182
http://www.articlesbase.com/law-articles/is-gambling-illegal-in-south-korea-1097050.html
http://www.netnewspublisher.com/mauritius-introduces-new-guidelines-for-gambling/
http://www.dcdm.biz/Budget2008/eBrief%202008-09.pdf
http://fedex.com/us/international/irc/profiles/irc_jo_profile.html?gtmcc=us
http://www.sars.gov.za/home.asp?pid=289
http://unpan1.un.org/intradoc/groups/public/documents/apcity/unpan005060.pdf
http://www.met.gov.na/programmes/legislation/gamblingact.pdf
http://havanajournal.com/culture/entry/cuban_lottery_as_ubiquitous_as_it_is_hush_hush
http://www.guardian.co.uk/uk/2010/jan/07/online-betting-regulation-overhaul
http://www.sakartvelo.com/Files/Economy/invest.html
http://www.ghanadistricts.com/home/?_=48&sa=4716&ssa=648
http://www.casinocitytimes.com/news/article/gaming-in-bulgaria-market-report-139880
http://www.natal-brazil.com/entertainment/casinos.html
http://www.radio86.co.uk/explore-learn/business-china/4991/all-bets-are-on-lottery-games-and-gambling-in-china
http://en.wikipedia.org/wiki/Gambling_in_India
http://www.antidrugs.gov.in/download/files/indian_drug_laws.pdf
http://www.thaipro.com/thailand_00/187_prostitutes_gamblers.htm
http://www.rra.gov.rw/rra_article289.html
http://jamaicacasinos.net/
http://fedex.com/us/international/irc/profiles/irc_ne_profile.html?gtmcc=us
http://www.formacompany.com/en/corporate-administration/gambling

http://www.hotel-online.com/News/PR2010_1st/Mar10_GuyanaCasino.html
http://www.lawcommissionbangladesh.org/reports/46.pdf
http://www.worldgamblingreview.com/gambling/pakistan/
http://www.bir.gov.ph/taxcode/1593.htm
http://gamingzion.com/gamblingnews/yemen-should-take-advantage-of-the-middle-east-lack-in-gambling-opportunities-1346
http://gamingzion.com/venezuela/
Prostitution
http://www.nzpc.org.nz/page.php?page_name=Law)
http://www.ecpat.org.nz/ProstitutioninAustriaandNewZealandNinaApril09Final.pdf.pdf
http://www.cafebabel.co.uk/article/17890/dutch-prostitution-from-sex-trade-to-trade-unionism.html
http://www.sosfemmes.com/english_sexwork/english_sexwork_comparing_systems.htm
http://en.wikipedia.org/wiki/Prostitution_in_Singapore#Commericial_sex_with_underaged-persons
http://findarticles.com/p/articles/mi_m0WDP/is_2004_June_21/ai_n6267378/
http://prostitution.procon.org
http://www.uri.edu/artsci/wms/hughes/netherl.htm
http://www.europarl.europa.eu/hearings/20040119/femm/document1_en.pdf
http://www.sosfemmes.com/english_sexwork/english_sexwork_comparing_systems.htm
http://en.wikipedia.org/wiki/Prostitution_in_Japan
www.public.iastate.edu/˜rhetoric/105H17/nnguyen/cof.html
http://www.fff.org/freedom/0998d.asp
http://www.hotline.org.il/english/pdf/The_Legalization_Of_Prostitution_English.pdf
http://www.state.gov/g/drl/rls/hrrpt/2007/100579.htm
http://wapedia.mobi/en/Prostitution_in_Europe#20
http://www.svu2000.org/women/prostitution.htm
http://www.news.com.au/business/prostitutes-at-57bn-world-cup/story-e6frfm1i-1225867016098
http://www.afrol.com/articles/18546
http://www.thebody.com/content/art27058.html
http://www.uri.edu/artsci/wms/hughes/italy.htm
http://www.atimes.com/atimes/Middle_East/EL09Ak01.html
http://www.state.gov
http://news.bbc.co.uk/2/hi/middle_east/458103.stm
http://www.nytimes.com/2007/10/05/world/europe/05iht-bulgaria.4.7773739.html
http://www.pbs.org/wgbh/pages/frontline/aids/countries/br.html
http://www.newint.org/issue252/facts.htm
http://gochina.about.com/od/tripplanning/a/prostitution.htm
http://www.allvoices.com/contributed-news/5683875-prostitute-arrested-in-serbian-assembly

http://www.illegaleconomy.com/prostitution_in_san_salvador.php
http://www.nyasatimes.com/national/have-your-say-mra-to-collect-tax-from-sex-industry.html
http://www.aegis.com/news/ips/2002/IP020419.html
http://www.newint.org/
http://www.docstoc.com/docs/17753819/Country-Report-for-Algeria
http://www-rohan.sdsu.edu/faculty/rwinslow/asia_pacific/armenia.html
http://swannet.org/node/556
http://www.sexwork.com/Thailand/traditiions.html
http://web.me.com/dwb217/iWeb/WorldViewBender/Blog%3A%20The%20View%20from%20Syria/9F4D594B-FFBF-4458-BDFA-C0BF504E14A3.html
http://www.dailystar.com.lb/article.asp?edition_id=1&categ_id=1&article_id=113834#axzzOoPsYZ7Ga
http://www.jendajournal.com/issue8/perschler-desai.html
http://www.sexwork.com/philippines/legalprost.html
http://www.nepalvista.com/nepalfaq/sex.html
http://www.europeansexunion.com/brothels-and-escort-services-in-the-ukraine.html
http://www.afesiplaos.org/upload/assets/OS_Garment%20factory%20girls_06.pdf
http://hanopolis.com/?articleNo=12821&story/Korean-sex-tourism-in-Mongolia-and-Uzbekistan

# References

Becker GS, Murphy KM, Grossman M (2006) The market for illegal goods: the case of drugs. J \Polit Econ 114(1):38–60

Glaeser EL, Shleifer A (2001) A reason for quantity regulation. AER Papers and Proceedings 91 (May):431–35

Miron JA (2004) Drug war crimes: the consequences of prohibition. Independent Inst, Oakland, CA

Miron JA (2008) The budgetary implications of drug prohibition. Report for the criminal justice policy foundation and law enforcement against prohibition, Maryland

# Self-Enforcing, Public-Order Institutions for Contract Enforcement: Litigation, Regulation, and Limited Government in Venice, 1050–1350

**Yadira González de Lara**

The inability of societies to develop effective, low cost enforcement of contracts is the most important source of both historical stagnation and contemporary underdevelopment in the Third World

North (1990, p. 54)

## 1 Introduction

Ever since the days of Adam Smith it has been recognized that an enhanced ability to exchange promotes economic growth. Some degree of market expansion and economic development can be and historically has been supported by private-order institutions based on reputation (e.g. Dixit 2004; Greif 2006). With a low fix cost but a high and rising marginal cost of expansion, private-order, reputation-based contract enforcement is, however, unviable at a large scale (Greif 1994; Li 2003; Dixit 2004, Chap. 3). Ultimately, impersonal exchange that would realize the gains from trade inherent in the modern market economy requires institutions that can enforce agreements based on the coercive power of the state (North 1990, p. 58; Dixit 2009, p. 13).

Such public-order, coercion-based institutions have been traditionally associated with the legal system. But, how can the state design a functional legal system and what exactly does it take it to effectively enforce impersonal exchange? Under what conditions, if any, does regulation improve enforcement efficiency by courts? Despite its enormous economies of scale, setting up a legal and regulatory system requires substantial fixed costs. Why would states incur into such costs if the volume of impersonal exchange prior to its establishment is necessarily low? We

Y. González de Lara (✉)
Universidad CEU-Cardenal Herrera, Elche, Spain
and
Universidad de Valencia, Valencia, Spain
e-mail: yadira.gonzalez@uch.ceu.es

know that efficiency considerations alone would not lead to its emergence (Greif 1994, 2004; Kranton 1996). Furthermore, if creating effective, low cost public-order institutions requires developing the administrative capacity of the state to confiscate wealth and inflict others punishment, those who run the state will be able to use this capacity to arbitrarily seize property and overtax income. In the absence of countervailing forces, creating effective public-order institutions for contract enforcement would therefore reduce the security of property rights and potentially undermine the market. How can one get the state to behave like an impartial third party that uses coercion to protect rather than abuse contract and property rights?

The historical experience of Venice during the late medieval period (1050–1350) offers an outstanding opportunity to elucidate on these questions. It was precisely during this age when Europe experienced its longest period of market expansion and sustained growth, as well as a process of state formation. Venice, in particular, developed impersonal securities markets for overseas trade, expanded its commerce along the Mediterranean and beyond, became one of the richest cities of Europe for centuries, and developed many distinctive characteristics of today's western institutions, including an effective, low-cost legal and regulatory system for contract enforcement and a limited government that constrained the state from abusing its coercive power.

In examining the institutional foundations of markets and democratic political structures in Venice, this historical institutional analysis combines the two main approaches within the study of institutions (for a discussion, see Greif and Kingston 2011). From the *institutions-as-rules* approach (e.g. North 1990) it borrows the focus on legal and political institutions but it integrates the study of the rules of the game created by the state with an explanation of how these rules are enforced as part of an institutional equilibrium, thus using the conceptual and analytical framework advanced by the *institutions-as-equilibrium* approach (e.g. Greif 2006).

Like the Law and Finance literature this chapter claims that legal and administrative institutions are essential parts of a broad system of corporate finance but it finds that welfare-enhancing regulations might extend beyond those mandating disclosure and specifying liability standards for gatekeepers (Kraakman 1986; La Porta et al. 2006; Hart 2009), simplifying judicial proceedings (Djankov et al. 2003; Bianco et al. 2005) or providing an (efficient) alternative to judicial enforcement when courts are expensive, unpredictable and/or corrupt (for an overview, see Kessler and Shleifer 2010). It also explores the interactions between the Venetian legal and regulatory system for contract enforcement and the emergence of a limited government, a coercion-constraining institution that motivated judges and regulators to use their coercive power for protecting rather than abusing investor rights.

The chapter thus relates to the Political Economy literature studying how property right institutions protect individuals from theft or expropriation by the government or elites (e.g. North 1981, 1990; Weingast 1997; North and Weingast 1989; Olson 2000; Haber et al. 2003; Acemoglu and Robinson 2004; González de Lara et al. 2008; Jha 2008). Following Greif (2006, 2008), however, it emphasizes the need to examine property right institutions and contracting institutions as a system whose effectiveness relies on their complementarities.

By examining the Venetian institutional system, this paper sheds light on various well-known but poorly explained historical phenomena, such as the exclusion of foreigners from Venetian trade, the tightness of Venetian citizenship regulations, its fluctuating commercial policy with sporadic restrictions on imports, the role of trading licenses in disciplining merchants, the transition from the sea loan to the commenda, and the co-emergence of well-functioning markets and polities. It also shed light on the process of market and state formation in the Third World, where both basic market supporting institutions and stable democracies need to be created.

To conduct this work, I draw on the almost 1,000 notary acts preserved in the State Archive of Venice and transcribed in full by Raimondo Morozzo della Rocca and Antonio Lombardo, henceforth MRL, for the period 1021–1261.9. To better interpret this fragmentary and notary-biased evidence, I compare it with secondary studies based on Genoese notary records of the twelfth century. These historical records enable me to, first, build up a context-specific game in which the state emerges as an institutional equilibrium and, second, evaluate various theoretical predictions generated by the game under the assumption that the state governed financial relations. Empirical confirmation of these predictions lends support to the hypothesis that the state functioned as an impartial third-party that used coercion to enforce contracts.

## 2 Impersonal Exchange

The spectacular economic growth of Venice during the late medieval period (1050–1350) was based on the expansion of its trade along the Mediterranean and beyond (De Roover 1965; Lane 1973; Lopez 1976; Cipolla 1993). This trade required large amounts of capital and involved high risks. A commercial round-trip voyage from Venice to Constantinople, the Crusader States or Alexandria took 6–9 months and overlapping sailing seasons precluded financing it with retained earnings from a previous voyage (Lane 1973, pp. 69–70, 120). Fitting costs were further increased due to the need of carrying a large armed crew, sailing in convoy with naval protection, and securing merchants' property rights abroad (Lane 1973, pp. 23–49, 68–85, 124–131). These protective measures notwithstanding, the *risk of the sea and people*, as the Venetians referred to the possibility of loss through shipwreck, piracy or confiscation of merchandise by foreign rulers, remained high (Lane 1973, pp. 77; De Roover 1965, pp. 44–46). The commercial risk was also high: profits varied widely depending on the tariffs and bribes paid in customs, the transportation and storage fees, the rates of conversions applied to various weights, measures and currencies, fluctuations in prices, the conditions of the goods upon arrival, and so forth (Lane 1967, pp. 95–111; Greif 1989, pp. 860–861). Trade in ordinary goods within Europe or the East did not set such high capital requirements as trade in luxuries between Europe and the East but was less profitable and still involved high risks (Lopez 1976, p. 95).

Crucial to this trade expansion was an impersonal financial market through which the Venetians invested their savings and shared the risks inherent into sea ventures

through bond-like sea loans and equity-like commenda contracts (Luzzatto 1952, pp. 59–80, 89–115; Lane 1966, pp. 56–57). The sea loan was a fixed payment loan with the particular feature that the investor assumed the risk of loss at sea or at the hands of hostile people and was therefore allowed a higher rate of return. Unlike in other localities where the commenda was considered a partnership, the Venetian commenda or *collegantia* was a credit instrument whereby an investor assumed liability for any loss in proportion to his capital investment and shared the commercial profits and risks with the merchant (Besta and Predelli 1901, St. Enr. Dand., 30–33, St. Ran. Dand., 16, St. Tiep., 1229, 16; Cessi 1938, St. Nov., III, 1–3).

The operation of impersonal markets among Venetian city dwellers is well reflected in the 435 sea loans and commenda contracts published by MRL for the period 1021–1261. Even though these contracts constitute a tiny sample, they show very flexible financial relations that transcended the boundaries of the family, business groups, social class, and other personal relations. First, although family ties were important, particularly during the eleventh century when 37.50% of the observed contracts involved kin, only 7.23 and 9.20% of the contracts established during the twelfth and thirteenth centuries were among family members.[1] A possible concern with this data is that financial relations among trustworthy relatives may have not required the services of a notary and Venetian brothers legally constituted a fraternal partnership (*fraterna*) without the need of formal contracts. Yet, notaries proved useful to trustworthy relatives for establishing legal title of their credits in potential disputes with third parties over the estate of a deceased merchant and the *fraterna* was of little importance until the mid thirteenth century (González de Lara 2008, pp. 258–259). We thus observe notarial acts showing a loan by a widow to her granddaughter 'for the needs of her home,' an investment by a widowed nun in her son's voyage, and commenda contracts between brothers (MRL 1940, # 146, 356, 785).

Second, although the available data do not indicate an investor's occupation (except in the case of priest and nuns, notaries and various high office holders) it provides various proxies for establishing whether he was a merchant himself. An investor is considered as belonging to a merchants' community or group if he is known to have functioned as a merchant, raising trading capital from other investors; resided overseas while functioning as investor; supplied funds to a merchant abroad; or received payment from a one-way-trip voyage at a different location from where he had supplied the funds. On the basis of this classification, we can conclude that both merchant and non-merchant capital was mobilized into trade at all times. Yet, the trend was towards a separation between investment and management. Whereas 82.22% of the investors known for the period 1021–1178 belonged to a merchants' community operating mainly within the East, only 41.42% and 11.11% of the investors known for the periods 1179–1219 and 1220–1261 were merchants themselves. Since the late twelfth century Venetian trade was increasingly opened to a

---

[1]Two individuals are considered to belong to the same family if they had the same surname or the contract mentions that they were relatives.

wide range of investors who remained in the city and had no control at all over the ventures they financed (Luzzatto 1952, pp. 70–72).

Third, unlike in other localities Venetian trade was never characterized by a *class* of rich investors and a separate *class* of poor merchants who rarely, if ever, functioned as investors (Luzzatto 1962, pp. 59–80, 89–116; Greif 1994, p. 928). Most investors were nobles and about half of them belonged to the ruling aristocracy, but over a third of the investors in our sample were non-nobles. Similarly, some merchants were relatively poor individuals, like Dobramiro Stagnario, a manumitted slave, and Romano Mairano, whose wife received a humble dowry (Luzzatto 1952, pp. 98–99, 108–116; Lane 1973, p. 52; Robbert 1999, p. 34) but most ventures were managed by noble merchants. Furthermore, financial relations seem not to have been driven by traders' social class: 25.98% of the contracts were entered between noble investors and non-noble merchants, but 37% were among nobles, 20% among non-nobles and 13.33% between non-noble investors and noble merchants; the remaining 3.69% were between Venetians and non-Venetians.

A merchant typically raised capital from various investors of different ranks and sometimes invested in other merchant's ventures for the sake of diversification. To undertake a trading voyage from Venice to its colonies in 1234 Rodolfo Suligo received funds ranging from 25 to 152 Venetian pounds from at least 15 noble and non-noble investors (MRL 1940, # 675–690, 804). In total he raised 1071 Venetian pounds and 5 pence, about 200 times the annual rent of a profitable shop in the market place of the Rialto in the year 1238 (MRL 1940, # 710; Robbert 1999, p. 37). During his trading carrier Romano Mairano – who appears 49 times as a merchant and three times as an investor during the period 1150–1199 – raised capital from 43 individuals belonging to 35 families, both noble and non-noble (Luzzatto 1952, pp. 108–116; González de Lara 2008, p. 159). Domenico Gradenigo entered into 28 commenda contracts during the period 1205–1226 with various members of his aristocratic, rich family and with 14 other investors, of whom only two financed him repeatedly (González de Lara 2008, pp. 159–160).

Diversification was pervasive. For example, Giovanni Serzi contracted with a different merchant in each of the eight sea loan contracts that have survived. In 1169 he funded four merchants who were sailing on three different ships from Armiro (Peloponnesus) to Constantinople and in 1170 he financed another four merchants under similar conditions (MRL, # 214–217, 219–223). Lazzaro Mercadante, another prosperous merchant, is known to have supplied capital through eight commenda contracts to seven different merchants during the period 1242–1258 and held as many as 25 credits of commenda when he died in 1281 (MRL, # 746, 759, 764, 771, 793, 839–840, 843; Luzzatto 1952, pp. 61–65). At the time of his death in 1268 the doge Raniero Zeno had about half of his fortune diversified in 132 commenda contracts (Luzzatto 1952, 81–87).

The broad participation of the citizenry, the flexibility within which merchants raised capital from many investors, and the large extent to which prosperous investors diversified their trade portfolios thus facilitated the mobilization of capital into risky trade investments that was crucial to the city's commercial success. According to Lane (1966, p. 62), there was an "ever-expanding volume of funds

seeking investment" in Venice. Yet, foreigners were excluded from the Venetian financial market. As we have seen, over 96% of the contracts were between Venetian citizens, who for the most part resided in the city of Venice itself, as opposed to any other settlement within the Venetian lagoons or abroad. Although this data might be biased, as foreigners were less likely to use Venetian notaries than citizens, a comparison with other equally biased data suggests that the Venetians indeed contracted mainly with each other. Genoese cartularies, for example, show many relations between citizens and noncitizens (Greif 1994, pp. 930–935).

## 3    The Fundamental Problem of Exchange

The operation of an impersonal financial market in Venice was key for growth but required the development of institutions that mitigated what Greif (2000) has labeled the fundamental problem of exchange. For individuals to enter into mutually beneficial exchange they need to be able to commit to fulfill their contractual obligations. Prospective creditors and shareholders would not invest unless assured that they would indeed receive a sufficiently high return on their risky investments. A Venetian merchant, however, could expropriate from investors in various ways. First, once overseas he could flee with all the capital entrusted to him. Second, even if he returned to Venice, he could render a false account and divert part of the profit. Finally, he could take excessive risks and/or shirk during the operation of the voyage. In the absence of effective institutions for contract enforcement, investors, anticipating expropriation, would not have entered into mutually beneficial debt-like sea loans and equity-like commenda contracts. Yet, we observe a diversified financial market, large in scale and scope among Venetian citizens. How could investors trust merchants? What were the institutional foundations of this market?

As discussed in González de Lara (2008), quite complex exchange can be realized by creating private-order institutions for contract enforcement. Loyalty among family members and other social control systems probably fostered market expansion, but financial relations in Venice were not confined to the family, a merchants' community or a particular social class. A reputation mechanism can mitigate the fundamental problem of exchange by ensuring that a merchant's future gains (or, more precisely, his discounted lifetime expected utility) from keeping his honest association with an investor or group of investors is larger than the gains he can obtain by cheating. A private-order institution based on bilateral reputation could have thus supported long-term relations of a sufficiently high per-period value between a merchant and a particular investor. Yet, investments by non-nobles who sporadically supplied small sums and the large extent to which investors diversified in Venice suggest that a merchant could commit to fulfill his contractual obligations even when his bilateral relations were expected to last only for a short period of time and be of little value. Furthermore, competition among Venetian investors to place funds rendered each investor's threat of terminating his bilateral association with a merchant who cheated him void. The operation of a private-order

institution based on multilateral reputation among the Venetians as a group is consistent with the observed scope and flexibility of the Venetian financial market. Yet, the relevance of such informal means for contract enforcement is challenged by the Venetians' documented reliance on notaries, courts, and legal codes and the absence of any evidence regarding private communication networks and collective punishment.

Ultimately, viable impersonal exchange of the sort observed in Venice requires institutions that can enforce agreements by the threat of coercion. According to some scholars, the legal system supported the emergence of a market economy in Europe during this period of time (De Roover 1965; Lopez 1976). Yet, asymmetric information and the geographical boundaries of the courts' jurisdictional power limited the ability of such public-order, coercion-based institutions to enforce financial contracts for overseas trade (Cipolla 1993, p. 164; North 1990, p. 57).

During the late-medieval period there was not a centralized legal system that was effective over a large geographical area, less over the whole Mediterranean (Greif 2004, p. 118). A Venetian court could and did force merchants within the (limited) territorial area over which it had legal jurisdiction to comply with their verifiable contractual obligations (González de Lara 2008, pp. 269–270). Yet, it could not exercise coercion over a merchant who fled. If a merchant embezzled an investor's capital and took refugee in another jurisdiction, a Venetian court could not force him to repay. Tracking down a merchant was prohibitively costly and even if he was located, inter-community litigation was not an option against a state in that Venice could not retaliate, for example, by interrupting trade (Greif 2004). Obviously, when considering the option to flee, a fraudulent merchant would choose a safe haven where Venice could not exert that pressure. The notorious case of Benetto Soranzo illustrates the failure of the Venetian legal system to gain extradition of *fugitives* as late as the fifteenth century. In 1455 Beneto fled Venice with all the surviving assets from his failed bank and took refugee in the lands of the duke of Modena. The Venetian authorities attempted in vain to persuade the duke to deny him asylum and when he finally intervened, Benetto merely moved to lands under the jurisdiction of the duke of Mantua (Mueller 1997, pp. 200–211).

Leaving collateral in Venice could have mitigated the problem of outright embezzlement but most merchants left few goods behind. As we have seen, many Venetian merchants were relatively poor individuals and those who were rich typically held most of their wealth in movable goods, which they could take with them, and had real estate holdings outside Venice and its domains (Pozza 1995). For example, at the time of his death in 1281 Lazzaro Mercadante had almost all his wealth invested in trade and owned some plots of cultivated land near Padua; in Venice he simply kept three houses of little value (Luzzatto 1952, pp. 61–65).

Relying on guarantors or kin to pay in place of insolvent merchants could also have mitigated the embezzlement problem. Yet, Venetian overseas trading contracts did not request naming guarantors, nor was the family held liable for a merchant's illegal actions. According to Venetian law, only sons under parental authority were liable for their father's debts (Besta and Predelli 1901, St. Enr. Dand., 68, St. Ran. Dand., 11–12; Cessi 1938, St. Nov. I.40). This is in sharp contrast with other

legislations that were also devised at the time. Genoese courts, for example, held all members of a merchant's family legally responsible for the merchant's verifiable transgressions, such as outright embezzlement of an investor's capital (Greif 1994, pp. 937–938).

A second problem with the legal system is that to enforce complex contracts, like the sea loan and even more the commenda, a court requires verifiable information. But, neither investors nor judges could provide it, among other things, because they remained in Venice and, so, could not directly monitor merchants abroad. In the absence of an institution generating and transmitting verifiable information, a merchant would have misreported a loss at sea or by the action of men in the expectation of being released from repayment, diverted part of a commenda's profits once they materialized, assumed high risks from which he was protected through limited liability and shirked.

By admitting the testimony of witnesses as evidence in judicial proceedings and conditioning repayment exceptions on the verification of losses from shipwreck, piracy or confiscation of merchandise abroad (Besta and Predelli 1901, St. Enr. Dand., 32; St. Tiep., 1229, 16; Cessi 1938, St. Nov., III.2), the legal system facilitated the enforcement of sea loans and, to some extent, commenda contracts. The court's reliance on witnesses to verify such losses is well reflected in the available data. In 1219, for example, the captain of and various merchants traveling on the ship *Lo Carello* testified that it had wrecked close to Negroponte and that the merchant Domenico Gradenigo had lost merchandise worth 110 hyperpers (MRL 1940, # 582).

Uncovering an accounting fraud, though, was far more challenging. It required verification of the tariffs and bribes a particular merchant had paid to pass customs, the transportation and storage fees he had arranged for, the rates of conversion he got applied among a plethora of weights, measures and currencies, the price at which he had bought and then sold his wares, whether these had been damaged on the voyage or pilfered by the crew, and so forth. Without verifiable information on such costs, prices and events, a Venetian court could have not supported the development of equity markets.

Asymmetric information regarding a merchant's choice of action and diligence while on voyage also impaired a court's effectiveness. To punish cheaters, the court needed to know when and to what extent a contract had been violated, for example, by unduly changing route to politically unsafe but highly profitable ports, storing merchandise in a precarious but cheap warehouse, betting on exchange rates, selling on credit with lax terms but superior prices, or simply shirking.

If kinship, private reputation, and the legal system did not provide the foundations of Venice's financial market, how could investors trust merchants not to expropriate all or part of their capital? What institutional arrangements reduced the measurement and agency costs of using bond-like sea loans and equity-like commenda contracts? Theoretical considerations and historical evidence suggest that the observed trust and implied information reflect the operation of a public-order yet reputation-based institution.

# 4   A Public-Order, Reputation-Based Institution for Contract Enforcement

In late-medieval Venice a public-order, reputation-based institution for contract enforcement prevailed. It was public-order in the sense that the state created the rents required to motivate a merchant to keep his affiliation with Venice, generated the information needed to detect a contractual breach, and punished a merchant if he cheated. It was reputation-based in the sense that the merchant was motivated to submit to the Venetian authorities and comply with his contractual obligations not only due to the threat of legal sanctions but also of losing his reputation with the state and thus access to Venice's lucrative trade. The Venetian institutional system thus combined coercion and reputation, but relied on public contract enforcement: threats backed by the coercive power of the state of legal sanctions and of exclusion from state-generated rents supported the Venetian financial market for overseas trade.

## 4.1   Providing Incentives to Keep One's Affiliation with Venice: Economic Rents and Barriers to Entry

The Venetian state took the military and diplomatic initiatives required to gain exclusive trading privileges, organized protective convoys, and restricted foreign entry. These regulations generated economic rents to which only Venetian citizens had access and so motivated merchants to return to Venice, where the legal system could force solvent borrowers to repay if they failed to do so spontaneously.

As González de Lara (2008) covers this issue in detail, a few considerations will suffice here. First, by obtaining exceptional trade privileges in *Romania*, the territory that belonged or once belonged to the Byzantine Empire, the Crusader States and Alexandria, Venice increased the expected profitability of its merchants beyond what they could have gained as residents of other cities. Furthermore, Venice's lordship over the Northern Adriatic ensured that all wares brought there were exchanged in the *Rialto*, the city's wholesale market, where Venetians would be the middlemen and have a chance to make a profit. Second, Venice outfitted a fleet to secure the seas for its merchants and rented them space on the state-owned galleys escorted by it, thereby providing them with protection at less cost than was available to their competitors. Last but not least, Venice's political stability assured its merchants that they and their offspring would continue enjoying rents from trade, while civil strife and foreign rule in rival cities hampered the long-term sustainability of their particular commerce.

To preserve per-citizen rents, Venice denied foreigners access to its lucrative trade and applied tight citizenship rules. Specifically, the state passed laws and regulations prohibiting foreigners from shipping merchandise from Venice and trading in its colonies (Lane 1973, pp. 7, 61). It also developed the administrative

structures required to enforce these prohibitions. For example, during the thirteenth century merchants planning to join a trading convoy from Venice to its colonies were required to register in advance and to obtain a license (Lane 1973, p. 49). More generally, the Consuls of the Merchants in Venice and colonial governors abroad had the specific duty of ensuring that foreigners did not participate in business reserved for Venetian citizens (Sacerdoti 1899, pp. 17, 44). The state also constrained citizenship and so access to trade by requiring immigrants to pay taxes in Venice for a long period of time, 10 years during the twelfth and thirteenth centuries and up to 25 years during the fourteenth century. At various points during the fourteenth century when an overabundance of Eastern wares in Venice was eroding profits, the state also established import quotas, thereby coordinating monopolistic practices that restored high prices (González de Lara 2008, pp. 266–267; Mueller 1997, pp. 151, 265, 503; 616). These regulations are striking, since they were not generally applied. For example, Genoa welcomed foreigners to its colonies, granted citizenship after 1 year of residence in the city, without taxation, and never restricted imports (Lopez 1982, pp. 33, 348).

Identifying these barriers to entry as essential elements of the Venetian institutional system for contract enforcement thus helps explain Venice's legislation. Governed by a public-order, reputation-based institution, the Venetians needed both to reserve the rents of their privileged trade for themselves and to keep these rents high by restricting access to citizenship and limiting the supply of imported wares in Venice when sale prices were low. The alternative view common in the political economy literature (e.g. Acemoglu and Robinson 2008) that elite merchants captured regulation is inconsistent with the distinctiveness, timing, and nature of the Venetian regulatory process. Had this been the case, the more powerful Genoese elite would have introduced similar barriers to entry and the Venetians would have implemented import quotas at all times and attempted to eliminate competition from less wealthy and politically influential citizens. The Venetian experience thus challenges the conventional wisdom that entry regulations are an unmitigated bad. Without restrictions on entry, Venetian rents from trade would have dissipated but without these rents, a merchant would have embezzled an investor's capital outright and fled to avoid legal sanctions.

## 4.2  Mitigating Asymmetric Information Problems: Formal Monitoring and Public Gatekeepers

The Venetian state also regulated the operation of trading ventures in a manner that reduced merchants' opportunities and incentives to breach their contracts and enhanced investors' ability to verify a breach. Specifically, by 1220s colonial governors, convoy admirals, ship scribes, tax collectors and many other public gate-keepers monitored merchants at all times, thereby preventing misconduct and generating the verifiable information required to adjudicate commercial disputes.

Venice obtained the right to hold permanent magistrates with administrative and judicial functions in Constantinople in 1186 and took possession of a chain of colonies with full extraterritorial rights in the former Byzantine Empire after the Fourth Crusade: a *podestà* was installed in Constantinople in 1205, a *castellano* in Coron and Modon in 1208, a *bailo* in Negroponte in 1216, and a duke in Crete in 1219. In the Crusaders States Venice had obtained large compounds with full extraterritoriality in the early twelfth century, but the Venetian population remained predominantly self-governing until about 1192, when a *bailo* was first sent to Acre. In 1208 Venice also gained consular representation in Alexandria. These colonial governors oversaw custom duties, administered warehouses and lodging facilities, enforced the use of Venetian measures, weights and coins, kept public records of the prices the Venetians paid for cotton and pepper while Venice maintained monopsonies in Acre or Alexandria, implemented all the regulations and trading controls established in Venice, and adjudicated commercial disputes abroad (Luzzatto 1952, pp. 62–60; Prawer 1973; Lane 1973, pp. 17–19, 49–51, 59–62; 99–100; Lopez 1982, p. 374; Ferluga 1992; Jacoby 1994; Ravegnani 1995).

By 1180s the state also organized trading convoys from Venice to its colonies. Unlike those previously planned by private individuals, these convoys were protected by state-owned galleys and "treated as community enterprises subject to governmental approval" (Lane 1973, p. 49). All the vessels and merchants planning to join a particular convoy were increasingly required to register in advance and to obtain a license. The first trading contract mentioning such a license dates from 1200; after 1220s the requirement prevailed and since 1266 it was compulsory by law. Licensed convoys were operated under admirals appointed and paid by the state and according to naval and commercial plans formulated by its governing Councils. All decisions regarding a convoy's sailing times, route, ports of call, and freight rates were thus delegated to the government in Venice or to the admiral in charge of the whole convoy overseas (Sacerdoti 1899, pp. 43–44; Lane 1973, pp. 68–70, 129–131, 145–146).

In addition, the Maritime Statutes of 1229 specified a ship's carrying capacity, arms and crew, the allocation of space for freight, equipment and officials, and the methods for loading and unloading cargoes. They also required that the crew swore under oath that they would not pilfer any shipment and that the ship owners choose among themselves a shipmaster to go with the ship. The shipmaster was held liable for any merchandise registered with the ship scribe, excluding losses at sea, from fire, or from the actions of hostile people. The scribe also had the duty to register the number, weight, and owner of any merchandise loaded and unloaded, record the contracts of all merchants on the voyage, and report any observed fraud. He was a semi-public official who was to be appointed by the shipmaster but with the consent of the Consuls of the Merchant, the top magistracy regarding trade oversight (Predelli and Sacerdoti 1902; Lane 1966).

On their arrival to Venice, ships were inspected by custom officials to make sure they had paid the appropriate dues. By 1180 naval patrols coerced traffic in the northern Adriatic and during the thirteenth century there were thirteen control points around the lagoon. At each, half dozen men with two or three vessels inspected all

passers to make sure that their cargoes were covered by permits to go where they were headed and required proof that they had been cleared in Venice (Lane 1973, 17–18, 49–51, and 59–62). In about 1200 Venice minted the silver *grosso*, which soon became an accepted medium of exchange and unit of account for overseas trade (Stahl 2000, pp. 204–213). Furthermore, by 1225 all commodity sales in the Rialto needed to be registered with the S*ensali dellla Messetteria*, a group of officials responsible for the collection of taxes and the provision of compulsory brokerage services (Rösch 1995, p. 453).

Consistent with the conjecture that Venetian regulations actually enhanced an investor's ability to verify a merchant's accounts while reducing the severity of moral hazard, we observe a progressive transition from debt-like sea loans to equity-like commenda contracts (for a full discussion of contractual design in Venice, see González de Lara 2009, 2011 and the references therein). As shown in Fig. 1, the sea loan prevailed until about 1180, when the state did not yet monitor trade. In the absence of (verifiable) information regarding a venture's true outcome, market participants were constrained to rely on debt-like sea loans, despite the implied inefficient risk allocation and perverse incentives on merchants to assume excessive risk from which they were protected through limited liability. As colonial governors, convoy admirals, ship scribes, custom officials and public brokers engaged on formal monitoring, bond-like sea loans progressively gave way to equity-like commenda contracts. By 1220s, when Venetian officials effectively monitored trade, generating the information required to detect a contractual breach and limiting a merchant's opportunities and incentives to breach his contractual obligations, the commenda prevailed. Since the Venetian regulations also rendered
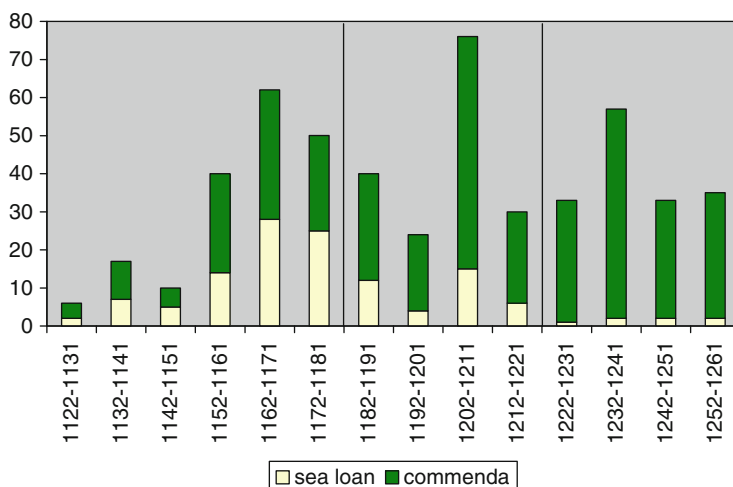


**Fig. 1** Documented sea loans and commenda contracts and their distribution over time
*Source*: The author, based on Morozzo della Rocca and Lombardo (1940, 1953). There is also evidence on one sea loan and eleven commenda contracts for the period 1021–1120. In addition, there are 38 contracts that cannot be classified with certainty and hence do not appear in the figure

a venture's outcome less sensitive to a merchant's effort, the low-powered commenda did not impair trade expected profitability. On the contrary, it "contributed greatly to the fast growth of maritime trade" (Lopez 1976, p. 76).[2]

## 4.3 Rendering Punishment and Reward Credible: An Impartial Legal and Administrative System

Last but not least, Venice developed an impartial regulatory and legal system, thus sustaining the shared belief among the Venetians that compliance with contracts would be rewarded with economic rents from trade, while a breach would be detected and punished through administrative and legal sanctions.

Venetian regulations were professedly designed to provide all city merchants equal access to the main routes of trade. These regulations were evaded many times but their mere existence at a time where elsewhere in Italy not such regulations existed indicates Venice's distinctive concern to distribute the economic rents from its trade widely among the citizenry (Lane 1967, pp. 582–583; 1973, pp. 145–146). Indeed, the distribution of wealth in Venice was not as concentrated as in other Italian city-states (Lane 1967, pp. 82–83; 1973, pp. 151–152, 332–334; Mueller 1997, pp. 491, 496; Rösch 1989; Greif 2006, pp. 286–287). Furthermore, trading contracts do not reflect political discrimination: both noble and non-noble merchants had access to the Venetian lucrative trade.

As discussed, the Venetian state used its administrative control over trade to exclude foreigners, thus motivating city merchants to keep their affiliation with Venice and preserving per-citizen rents. Apparently, it also excluded merchants known to have cheated other Venetians. That city merchants feared exclusion regardless of their political power or that of their financiers is well reflected in a document from 1242 showing one Omobone Barbo making commenda contracts for a convoy voyage with two noble investors. Both Omobone and one of the investors were actively involved in government and very rich. Yet, Omobone, expecting not to be allowed to join the licensed convoy, protected himself from his liability to sail with the capital received in commenda to the case in which he 'will be among those chosen men who are chosen according to the decree given by the lord Doge and his council' and specifying that 'if [he] will not be among those chosen, [he] will have the power to commit that merchandise or a part of it, with the witness of good men or with a charter, to some or to someone among those chosen' (MRL 1940, # 752, 753;

---

[2]Most economic historians have simply averted the transition from the sea loan to the commenda without providing any explanation for it. According to De Roover (1965, 55) and Lopez (1976, 73, 104), it can be attributed to the church's rising doctrine against usury (for a well-accepted critic, see Lane (1966). Williamson (2002) has documented a later revival of the sea loan and has associated it to disruptions in the flows of information caused by the Black Death. Since this transition was a widespread phenomena occurring under plausibly different institutional regimes, it is likely that private-order institutions played an important role in other localities (Greif 1994).

Pryor 1983, p. 141). The document does not specify why Omobone feared being disqualified from joining the licensed convoy but it might have been that a sentence was pending against him. Later in the fourteenth century administrative sanctions prohibiting 'fraudulent' merchant from entering the marketplace of the Rialto and the administrative and judicial site of San Marco de facto excluded them from Venetian trade (Cassandro 1937, pp. 99; see also Cassandro 1936, pp. 77–78; Lane 1973, p. 143; Castagnetti 1995, p. 101; Mueller 1997, pp. 124–125).

Similarly, the Venetian legal system was celebrated for providing 'a responsible justice equal to all' (Besta and Predelli 1901, p. 60; see also Lane 1973, p. 251). Venetian courts ruled according to the principle of speedy, informal, and equitable procedure and enforced their judgments as court orders. Offending merchants were subject to imprisonment and confiscation of their properties within Venice and its colonies. The procedure had already been established in the earliest Statutes ca. 1195 (Besta and Predelli 1901, St. Enr. Dand., cc. 7–14, 36, 66, 73; see also Cessi 1938, St. Nov., I, 51, 63.). If a merchant in arrears failed to pay or otherwise arrive at an agreement with his financiers within eight days after a court had sentenced him, he ought to remain within the territorial boundaries of the court for a month and thereafter ought to be incarcerated for another month. After these two months, his goods would be sequestered. To facilitate the enforcement of court sentences, the law prohibited offenders from both leaving Venice and disposing of their goods before their disputes were settled (Besta and Predelli 1901, St. Ran. Dand., cc. 8, 28; for the application of the law, see MRL 1940, # 466, 630–631, 853; 1953, # 20, 41). Litigation over various commenda contracts in 1195 and 1226 actually resulted in the forced sale of a merchant's estate and in the transfer of property to a plaintiff (MRL 1940, # 424, 626). In 1241 the personal property of a deceased merchant was auctioned and, because the amount retrieved was insufficient, the merchant's house was sold by the court (MRL 1940, # 743). Confiscation of real property due to failed credit was also extended to colonial estates. In 1178 Leone Falier gave power of attorney to act against one of his debtor's properties in the Venetian colony at Tyre (MRL 1940, # 295).

Since the records of the Merchant Consuls, who adjudicated commercial disputes, have been lost, we do not have evidence on legal sanctions due to accounting frauds. Yet, the court's enhanced ability to verify a merchant's trading accounts is manifest from the various compilations of the Civil Statutes. The earliest Statutes, which were written in about 1195 but codified previous customs, specifically called for the verification of a merchant's claims concerning losses at sea or from the action of hostile people but recognized the merchant's oath as legal proof of his commercial accounts, which he simply had to render by the date of due (Besta and Predelli 1901, p. 24; St. Enr. Dand., cc. 30–32). The Statutes of 1229, however, mandated a detailed account of each and every commercial operation, one by one in sequence, and the revision of 1233 made the court responsible for verifying them in case of litigation (Besta and Predelli 1901, St. Tiep., 1229, c.16; St. Tiep, IIIA, c.2.). The final Statutes of 1242 further entailed the investor to present reliable witnesses and established the merchant's obligation to compensate her for whatever she could prove to be owed (Cessi 1938, St. Nov., III.2, gloss 3).

# 5 Self-Enforcing Economic and Political Institutions

The discussion so far assumed that regulators, public gatekeepers and judges took the actions that were most beneficial to Venetian citizens instead of abusing their power or shrinking. But why was this case? Why did investors trust that the Venetian officials would do what was required for the above public-order institution to effectively enforce their contracts? What institutional elements within the system curtailed corruption, and to what extent? Was the system worthy of its cost?

To understand the complex mechanisms that were used to ensure officials' impartiality and diligence, one must turn to the political system. Venice's governing structures reduced each office holder's space of strategies in a way that curbed his ability to use the state power in his own interest at the expense of the rest of the society. Furthermore, a vigilant oversight supported a system of punishments and rewards that made each office holder's best strategy not to abuse his (limited) power or shirk from his duty.

Venice was established as an autonomous political unit in the eighth century after a period of Byzantine dominance. Initially it was ruled by a dictator-like doge but beginning in 1032, the autocratic prerogatives of the doge were progressively limited until he became only a top magistrate elected for life. The doge was at first responsible for enacting laws, implementing policy and administering justice. However, with the transformation of Venice from dukedom to commune during the thirteenth century, political, administrative and judicial power was distributed among a large number of interlocking councils and magistracies whose members were appointed for short terms and could not be reelected for a consecutive term. To further avoid the concentration of power in one individual or family, campaigning for office was outlawed, officers (including the doge) were nominated by randomly selected official committees, and only one family member was allowed on any such committee or office (Cessi 1963/1965).

Unlike modern limited governments, the division of power in Venice ignored completely the separation of the legislative, administrative, and judicial functions. Instead, various governing bodies were given overlapping jurisdictions so that each council or magistracy was checked by some other council or magistracy as to assure the rule of law. For example, the (three) Consuls of the Merchants were responsible for the safety of Venetian trading voyages and for the adjudication of commercial disputes. Yet the *Signoria* – composed by the Doge, the six Ducal Councilors and the three heads of the Forty – initiated maritime legislation, named the commanders of galleys and fleets, and had the authority to assign particular cases to a higher court of law. If any such board was proceeding contrary to the statues laid down for it, the State Attorneys could suspend proceedings and call for a meeting of the Great Council, in which all important families had representatives, to hear their charges. They in turn could be sued for dereliction of duty by the Heads of the Forty and in last resort the case would be judged by the Great Council (Lane 1973, pp. 88–117).

Within each magistracy or *officia*, the clerks exercised mutual monitoring and rendered various sets of accounts on the basis of which performance was evaluated. For example, to assess if the maritime codes had been violated or a fraud had been committed while on voyage, the Consuls of the Merchants relied on two ship scribes who they inducted since 1252. Besides, every official, including the Doge, was subject to strict independent reviews after his office came to term and was liable to prosecution for abuse of his office or dereliction of duty. This prosecution was carried out by a distinctively Venetian group of officials, the State Attorneys, who had investigating powers and to whom all office holders were to notify any observed wrongdoing (Cessi 1963/1965; Lane 1973, pp. 95, 251).

Severe punishment for failing duty was applied to all Venetian officials. Not even the head of the state escaped such punishment, for "doges were leaders, no lords, nay not even leaders, but honored servants of the State" (Petrarch, in a letter of May, 1355, cited by Lane 1973, p. 181). All office holders had to give an oath of impartiality and good behavior and were subjected to both hefty monetary sanctions and retirement from office if caught in whatever kind of fraud. For example, any official found guilty of having "put his hands in the state's goods" had to pay back the amount taken plus a fine of half the amount within three days from the conviction (Stahl 2000, p. 271). In addition, he was to be banned for ever from the specific office in which he embezzled the money and, if the amount taken was relatively high, from holding any public office.[3] Such penalties ensured that tax receipts were properly allocated to the provision of economic rents, verifiable information, impartial contract enforcement, and other public and collective goods. Lesser offenses were punished with smaller but significant penalties. For example, any crewman who failed to help in the recovery of a damaged Venetian ship, its equipment and cargo for the fifteen days prescribed by the Maritime Statutes of 1255 was to be deprived from his entire salary (Predelli and Sacerdoti 1902, p. 134 and St. Zeno c. XCIII).

Good performance, on the contrary, ensured the continuation of a very profitable public life. Although the same man could not serve in the same council or magistracy for two consecutive terms, nothing prevented an individual to rotate through the most important offices, so far as he proved himself honest and competent. Also, the re-election of public clerks was tied to the auditing of accounts by the councils or magistrates who chose them. Furthermore, all officers charged with enforcing regulations were induced to diligence by receipt of a portion of the fines levied in addition to their salaries (Lane 1973, pp. 98, 100).

Asserting that Venice distinctive limited government constrained political agents to protect rather than abuse the contract and property rights required for the

---

[3]This 1359 law basically standardized the punishment for embezzlement of state goods. To make sure that it was enforced, the State Attorneys were given the right to sell the property of the convicted official. For a 1385 case in which this procedures was applied to a noble, see Stahl (2000, p. 261). Evidence on the payment of fines for breaking the law dates back to at least the mid-twelfth century (MRL 1940, # 143, 163, 226 and 402). More generally, see Lane (1973, pp. 98–100).

operation of the financial market hides a key question. Why did Venetian prominent families and the doge, in particular, support the Venetian institutional system instead of attempting to establish an autocracy and misappropriate all Venetian rents?

Arguably, the Venetians were motivated to support rather than challenge the existing system by the belief that cooperating to render the polity of Venice conducive to trade would ensure each of them a sufficiently high share in the gains from their collective action, while any attempt to subvert Venice limited government would be successfully resisted and penalized with capital punishment. This belief was rendered self-enforcing by the operation and internal organization of the Venetian state, which, on the one hand, generated economic and political rents and allocated them among Venetian citizens widely and, on the other hand, guaranteed a balance of power among all the important Venetian families, thereby making it impossible for any one family or group to attain unchallenged supremacy. Given the gains the Venetians expected to obtain from supporting the prevailing institutional system and the fear of becoming victims of autocratic extraction, the best they could do was to join together to confront anyone's attempt to make himself a dictator and to impose on him the heaviest punishment, which in turn deterred each of them from trying.

On the one hand, trading profits and political authority remained highly diffused. As we have seen, trade regulations and state's control over the city and her colonies assured all Venetian merchants almost an equal chance to make a profit in oversea trade. Furthermore, any Venetian with little money to invest could obtain up to 40% interest on sea loans or three fourths of the net venture's return on commenda contracts without venturing overseas. Finally, the large number of governing bodies, high turnover in office and the regulation of the electoral procedures resulted in a wide distribution of political power and honors.

To preserve the economic and political rents that rendered the institutional foundations of markets self-enforcing and to make the resulting institutional system more likely to be a equilibrium, the Venetians established barriers to entry and created intergenerational links. Between the end of the thirteenth century and the early years of the fourteenth century, those families who had taken active part in political life during the former century became a hereditary aristocracy, whose male members of age would henceforth exercise a monopoly over political activity. This, of course, deprived the vast majority of Venice residents from any political power, but assured all the members of the existing ruling class that they and their offspring would continue enjoying economic and political rents (Lane 1973, p. 111; Chojnacki 1973, p. 56; for an alternative view, see Rösch 2000). Furthermore, Venice dwellers that were at the time comfortably well-off but who they themselves or their forefathers had not sat in the Great Council during the previous century, and were hence excluded from the aristocracy, soon came to form an hereditary privileged class of citizens-by-birth, who were admitted to foreign trade with almost the same rights as the patricians and from whose ranks all the public clerks were to be chosen. Nobles and citizens had all much the same interests and, although they did not obviously treat common people below the rank of citizens as well as they treated themselves,

they nonetheless granted them economic rights according to the guilds to which they belonged (Lane 1973, pp. 151–152).

On the other hand, a broadly equal dispersion of economic and political power made it imprudent for any leader or group to attempt to overpower the others. After the ignominious endeavor by doge Domenico Orseolo to seize power in 1032 and the subsequent constitutional reform, there were only two attempts in over seven hundred years to subvert Venice limited government. Both occurred at times of exceptional distress. Both miserably failed. In 1310 Bajamonte Tiepolo, taking advantage of the disruptions caused by the papal excommunication of the City and the general discontent with the ruling doge, led a plot to overthrow the doge and establish himself and his followers as despots in the Venetian domains (Lane 1973, pp. 115–116). Bajamonte was generously sentenced to exile in Dalmatia for fear of provoking a civil war if otherwise, but his house was razed to the ground and on its site it was raised a Column of Infamy bearing the inscription: "This land was the property of Bajamonte//And now, through his infamous betrayal,//Is held by the Commune as a lesson to others//So let these words proclaim to all, for ever" (Norwich 1989, p. 195). In 1355, after the "most disastrous decade Venetians had ever known," the doge Marin Falier was discovered conspiring to slaughter most of the nobility and make himself an autocrat (Lane 1973, p. 179). He was beheaded according to due process of law and his portrait in the Hall of the Great Council was substituted with a black curtain reading "Here is the place of Marin Falier, beheaded for his crimes" (Lane 1973, p. 183). Subsequent doges were followed in official procession by a sword-bearing symbolic executioner as a reminder of the punishment intended for any leader who attempted to assume dictatorial powers (Norwich 1989, p. 299).

## 6  Conclusions

In late-medieval Venice self-enforcing public-order institutions prevailed. Litigation and regulation conjointly provided investor protection thus enabling the Venetians to exchange through impersonal markets and ensuring a wide distribution of trading profits. This motivated them to cooperate in rendering the polity of Venice supportive to trade and to resist anyone's attempt to gain political control over the City and its economic resources, which in turn reduced anyone's incentives to challenge the prevailing institutional system. Venice's institutions for contract enforcement thus generated the political support they needed to perpetuate.

The nature and role of the state in Venice differed from those posited by economic historians and economists to pre-modern and modern states. The state has been modeled as a negligent ruler, who does nothing or little to maintain the rule of law, a predatory ruler, who abuses property rights, or a benevolent ruler, who provides an impartial legal system based on coercive power (for a discussion, see Shleifer and Vishny 1998). The Venetian state was neither negligent nor predatory on its citizens. Like a benevolent ruler, it played an active and salutary role in

promoting trade by providing third-party contract enforcement to the Venetians. Yet, it went beyond the coercive role traditionally attributed to the legal system, namely 'forcing solvent borrowers to repay if they failed to do so spontaneously' (Bianco et al. 2005, p. 225).

Venetian courts could and did force merchants within its jurisdiction to comply with their verifiable contractual obligations but could not exercise coercion over a merchant who embezzled capital and fled, nor could it enforce contracts based on asymmetric information. It was therefore necessary to motivate merchants to submit to the authorities and to limit their ability and interests to act opportunistically, but this required more extensive state intervention than is usually associated with good corporate governance and a benevolent ruler.

Venice, in particular, took the military and diplomatic initiative required to gain exclusive trading privileges, organized protective convoys and restricted foreign entry, thereby making Venetian commerce more profitable and secure than that of other rival cities and ensuring Venetian merchants a rent as long as they kept their city affiliation. Venetian regulations thus complemented rather than substituted for judicial enforcement. As the Venetian public-order institutions for contract enforcement relied both on the threat of coercion and the promise of future rents from trade, it was necessary to maintain per-citizen rents. Licensing and registration requirements that restricted entry were not the result of regulatory capture by the elites, as the political economy literatures commonly asserts (e.g. Acemoglu and Robinson 2008). Despite their static inefficiency, eliminating these barriers to entry would have undermined growth: anticipating outright embezzlement of funds, prospective investors would not have mobilized their capital to otherwise productive investments in overseas trade. The Venetian evidence thus lends support to the view that institutional reform based on first-best practices can easily do more harm than good (Rodrik 2008). Entry regulations are not only necessary to sustain relational contracting, as the institutional literature has noted (e.g. Dixit 2004, Chap. 2); they are also necessary to support dispute resolution in courts when courts have limited jurisdictional power.

Tight regulations and administrative controls over trade were also necessary to cope with asymmetric information. To enforce contracts, a court needs to know when and to what extent a contract has been violated, but neither judges nor investors could generated the required information, among other things because they remained in Venice and so could not monitor merchants on board and abroad. To enlarge the set of contracts that could be legally enforced, Venice instituted various public gatekeepers and placed merchants under their permanent oversight. Specifically, colonial governors, convoy admirals, ship scribes, custom officials, and public brokers generated the (verifiable) information required to prosecute merchants for profit diversion, prevented excessive risk taking by merchants, and rendered profits less responsive to their lack of application. Venetian regulations coping with asymmetric information thus went much beyond mandatory disclosure.

In Venice the combined used of regulation and litigation thus effectively provided public contract enforcement. It also paid for the high fixed costs of setting up the required administrative and legal system. Unlike in other historical and

contemporary episodes in which the threat of revoking a license and other regulations disciplined market participants and private gatekeepers on the basis of monopoly rights granted by the state (e.g. Greif et al. 1994; Glaeser et al. 2001), in Venice the required rents were created mainly by making Venetian commerce more profitable than it would have been otherwise. Venetian regulations thus increased the total gains to be realized from trade and generated the surplus (and political support) required to pay for their implementation. As a byproduct, merchants were motivated not to embezzle an investor's capital outright and never return to Venice. Likewise, as part of their rent-generating activity, various government officials monitored trade and acted as public gatekeepers, thereby enabling the legal system to enforce contracts otherwise characterized by asymmetric information at a low extra cost.

The Venetian institution for contract enforcement was exogenous to each individual, but endogenous to the society. Venice's governing structures both supported the operation of financial markets and restrained the tyrannical exercise of power. This ensured a wide distribution of trading profits and political authority among the Venetians and motivated them to cooperate and contribute to the maintenance of Venice' distinctive limited government.

While this paper accounts for the Venetian institution as being politically self-enforcing, it does not inquire into the process of equilibria selection. Why did Venice evolve along such a distinctive institutional trajectory? Did a unique historical experience, geographical position, and cultural heritage ultimately bring about what seems to be a particularly successful equilibrium? And if so, what prevented other Italian city-sates from adopting similar public-order reputation-based institutions? These questions lead the way to a future comparative and historical institutional analysis that may facilitate our understanding of both past economic developments and the political impediments to economic growth in contemporary developing countries.

# References

Acemoglu D, Robinson J (2004) Unbundling institutions. J Polit Econ 113:949–995

Acemoglu D, Robinson J (2008) Persistence of power, elites, and institutions. Am Econ Rev 98:267–293

Besta E, Predelli R (eds) (1901) Gli statuti civili di Venezia prima del 1242. Nuovo archivio Veneto, n.s. 1:5–117, 205–300

Bianco M, Japelli T, Pagano M (2005) Courts and banks: effects of judicial enforcement and credit markets. J Money Credit Bank 37:223–245

Cassandro GI (1936/1937) La Curia di Petizion. Archivio Veneto, ser. 5.19:72–144; ser. 5.20: 1–201

Castagnetti A (1995) Il primo commune. In: Ruggini LC, Pavan M, Cracco G, Ortalli G (eds) Storia di Venezia, vol 2. Ist. dell'Encicl. Ital. fondata da G. Treccani, Roma, pp 81–130

Cessi R (ed) (1938) Gli statuti veneziani di Jacopo Tiepolo del 1242 e le loro glosse. Memorie dell'Istituto veneto di scienze, lettere ed arti 30.2

Cessi R (1963/1965) Venezia ducale. Deputazione di Storia Patria per le Venezie 1, Venezia

Chojnacki S (1973) In search of the Venetian patriciate: families and factions in the fourteenth century. In: Hale JR (ed) Renaissance Venice. Faber and Faber, London, pp 47–90

Cipolla CM (1993) Before the industrial revolution: European society and economy, 1000–1700. Routledge, London (Third Edition; First Edition from 1976)

De Roover R (1965) The organization of trade. In: Postan MM, Rich EE, Miller E (eds) Cambridge economic history of Europe, vol 3. Cambridge University Press, Cambridge, pp 42–118

Dixit AK (2004) Lawlessness and economics: alternative modes of governance. Princeton University Press, Princeton

Dixit AK (2009) Governance institutions and economic activity. Am Econ Rev P&P 99:5–24

Djankov S, La Porta R, Lopez-de-Silanes F, Shleifer A (2003) Courts. Q J Econ 118:457–522

Ferluga J (1992) Veneziani fuori Venezia. In: Ruggini LC, Pavan M, Cracco G., Ortalli G (eds) Storia di Venezia, vol 1. Ist. dell'Encicl. Ital. fondata da G. Treccani, Roma, pp 693–722

Glaeser EL, Johnson S, Shleifer A (2001) Coase versus the Coasians. Q J Econ 116:853–899

González de Lara Y (2008) The secret of venetian success: a public-order, reputation-based institution. Eur Rev Econ Hist 12:247–285

González de Lara Y (2009) The birth of insurance contracts. WP CEU-UCH EE 2009-10, Universidad CEU-Cardenal Herrera

González de Lara Y (2011) Litigation plus regulation: the protection of investor rights in late medieval Venice. In: Koppell JGS (ed) Origins of shareholder advocacy. Palgrave Macmillan, New York

González de Lara Y, Greif A, Jha S (2008) The administrative foundations of self-enforcing constitutions. Am Econ Rev P&P 98:105–109

Greif A (1989) Reputation and coalitions in medieval trade: evidence on the Maghribi traders. J Econ Hist 49:857–882

Greif A (1994) Cultural beliefs and the organization of society: a historical and theoretical reflection on collectivist and individualist societies. J Polit Econ 102:912–950

Greif A (2000) The Fundamental problem of exchange: a research agenda in historical institutional analysis. Eur Rev Econ Hist 4:251–284

Greif A (2004) Impersonal exchange without impartial law: the community responsibility system. Chic J Int Law 5:107–136

Greif A (2006) Institutions and the path to the modern economy: lessons from medieval trade. Cambridge University Press, Cambridge

Greif A (2008) Coercion and exchange: how did markets evolve? (November 19, 2008). Available at SSRN. http://ssrn.com/abstract=1304204

Greif A, Kingston C (2011) Institutions: rules or equilibria. In: Schofield N and Caballero G (eds) Political economy of institutions, democracy and voting. Springer, Berlin/Heidelberg, Chapter 1

Greif A, Milgrom PR, Weingast BR (1994) Coordination, commitment and enforcement: the case of the merchant guild. J Polit Econ 102:745–776

Haber S, Razo A, Maurer N (2003) The politics of property rights. Political instability, credible commitments, and economic growth in Mexico, 1876–1929. Cambridge University Press, Cambridge

Hart O (2009) Regulation and Sarbanes-Oxley. J Acc Res 47:437–445

Jacoby D (1994) Italian privileges and trade in Byzantium before the fourth crusade: a reconsideration. Annuario de estudios medievales 24:349–368

Jha S (2008) Shares, coalition formation and political development: evidence from seventeenth century England, Stanford Graduate School of Business, Research Paper 2005

Kessler D, Shleifer A (eds) (2010) Regulation and litigation. University of Chicago Press, Chicago

Kraakman RH (1986) Gatekeepers: the anatomy of a third-party enforcement strategy. J Law Econ Organ 2:53–104

Kranton R (1996) Reciprocal exchange: a self-sustaining system. Am Econ Rev 86:830–851

La Porta R, Lopez-de-Silanes F, Shleifer A (2006) What works in securities laws. J Finance 61:1–32

Lane FC (1966) Venice and history: the collected papers of Frederic C Lane. The Johns Hopkins University Press, Baltimore

Lane FC (1967) Andrea Barbarigo, Merchant of Venice, 1418–1449. Octagon Books, New York

Lane FC (1973) Venice. A maritime republic. The Johns Hopkins University Press, Baltimore

Li JS (2003) Relation-based versus rule-based governance: an explanation of the East Asian miracle and Asian crisis. Rev Int Econ 58:651–673

Lopez RS (1976) The commercial revolution of the middle ages, 959-1350. Cambridge University Press, Cambridge

Lopez RS (1982) Il commercio dell'Europa medievale: il Sud. In: Postan MM, Rich EE (eds) Storia economica Cambridge II: commercio e inustria nel Medioevo. Giulio Einaudi, Torino

Luzzatto G (1952) Studi di storia economica veneziana. Cedam, Padova

Luzzatto G (1962) Storia economica di Venezia dall'XI all'XVI secolo. Centro Internazionale delle Arti e del Costume Venezia

Morozzo della Rocca R, Lombardo A (1940) Documenti del commercio veneziano nei secoli XI-XIII (2 vols.). Editrice Libraria Italiana, Torino

Morozzo della Rocca R, Lombardo A (1953) Nuovi documenti del commercio veneziano nei secoli XI-XIII. Deputazione di Storia Patria per le Venezie, n.s. 7, Treviso

Mueller RC (1997) The Venetian money market: banks, panics, and the public debt, 1200-1500. John Hopkins University Press, Baltimore

North DC (1981) Structure and change in economic history. Norton, New York

North DC (1990) Institutions, institutional change and economic performance. Cambridge University Press, Cambridge

North DC, Weingast BR (1989) Constitutions and commitment: the evolution of institutions governing public choice in seventeenth century England. J Econ Hist 49:803–832

Norwich JJ (1989) A history of Venice. Vintage Books, New York

Olson M (2000) Power and prosperity: outgrowing communist and capitalist dictatorship. Basic, New York

Pozza M (1995) I propietari fondiari in terraferma. In: Ruggini LC, Pavan M, Cracco G, Ortalli G (eds) Storia di Venezia, vol 2. Ist. dell'Encicl. Ital. fondata da G. Treccani, Roma, pp 81–130

Prawer J (1973) I veneziani e le colonie veneziane nel Regno di Gerusalemme. In: Pertusi A (ed) Venezia e il Levante fino al secolo XV. Olschki, Firenze

Predelli R, Sacerdoti A (eds) (1902/1903) Gli statuti maritimi veneziani fino al 1255. Nuovo archivio Veneto, n.s. 2: vol 4 and vol 5

Pryor JH (1983) Mediterranean commerce in the middle ages: a voyage under commenda. Viator Med Ren Stud 14:132–194

Ravegnani G (1995) La Romania veneziana. In: Ruggini LC, Pavan M, Cracco G, Ortalli G (eds) Storia di Venezia, vol. 2. Ist. dell'Encicl. Ital. fondata da G. Treccani, Roma, pp 183–232

Robbert LB (1999) Domenico Gradenigo: a thirteenth-century Venetian merchant. In: Kittel EE, Madden TF (eds) Medieval and renaissance Venice. University of Illinois Press, Urbana

Rodrik D (2008) Second-best institutions. Am Econ Rev P&P 98.2 (2008):100–104

Rösch G (1989) Der Venezianische Adel bis zur Schließung des Großen Rats. Zur Genese einer Führungsschicht (Kieler Historische Studien, 33). Thorbecke, Sigmaringen

Rösch G (1995) Le strutture commerciali. In: Ruggini LC, Pavan M, Cracco G, Ortalli G (eds) Storia di Venezia, vol 2. Ist. dell'Encicl. Ital. fondata da G. Treccani, Roma, pp 437–460

Rösch G (2000) The Serrata of the great council and venetian society, 1286–1323. In: Martin J, Romano D (eds) Venice reconsidered: the history and civilization of an Italian city state. Johns Hopkins University Press, Baltimore, pp 1297–1797

Sacerdoti A (1899/1900) Le colleganze nella pratica degli affari e nella legislazione Veneta. Atti del Reale Istituto Veneto di scienze, lettere ed arti 59:1–45

Shleifer A, Vishny RW (1998) The Grabbing hand: government pathologies and their cures. Harvard University Press, Cambridge

Stahl AM (2000) Zecca: the mint of Venice in the middle ages. The John Hopkins University Press, Baltimore

Weingast BR (1997) The political foundations of democracy and the rule of law. Am Polit Sci Rev 91(2):245–263

Williamson DV (2002) Transparency and contract selection: evidence from the financing of trade in venetian crete, 1303–1351. (July 2002). Available at SSRN: http://ssrn.com/abstract=320800 or doi:10.2139/ssrn.320800

# Judicial Stability During Regime Change: Apex Courts in India 1937–1960

**Alfred W. Darnell and Sunita Parikh**

## 1 Introduction

When India gained independence in 1948, the transition from colonial subject to a self-governing sovereign state was far from easy. Major changes in governance of the region occurred with the necessary incorporation of the Princely States and the partition of the subcontinent into India and Pakistan being especially tumultuous and costly. But by any standard, the institutional components of this transition were much less problematic. The gradual introduction of indigenous political participation during British rule, the adoption of a constitutional, parliamentary system built on the bureaucratic infrastructure of the colonial era, and the commanding dominance of the Congress Party as the ruling force, combined to provide a level of institutional continuity unparalleled in the history of decolonization.

Nowhere is this continuity more apparent than in the judicial arena. Although the Supreme Court of India was established in the new constitution and inaugurated in 1950, the Court's appointment procedures, legal powers, and institutional location in the governance structure closely followed the provisions of the Government of India act of 1935. Even members of the Court made the move from one government to the next, with the first Chief Justice having been the last Chief Justice of the Federal Court. In spite of this continuity, conventional accounts of how each court has functioned tend to portray them as qualitatively different: the Federal Court as ineffectual and the Supreme Court as elitist and obstructionist. What, then, is the basis for characterizing the two Courts as fundamentally different? Did the interests of the justices change, thus producing different behavior? It is conceivable, for example, that the change from a colonial to a parliamentary government significantly altered dynamics between the judiciary and executive branches to produce

A.W. Darnell and S. Parikh (✉)
Department of Political Science, Washington University in St. Louis, 1 Brookings Drive, St. Louis, MO 63130, USA
e-mail: adarnell@wustl.edu; saparikh@wustl.edu

differences in the actors' interests, thus leading to a difference in how the two courts functioned. Did changes in black letter law between colonial rule and a Constitutional government drive the transformation? Clearly changes occurred at this level of law and could underlie why the two courts are viewed differently. Or, are existing understandings of the differences simply wrong? Indeed, there could be some combination of accepted factors at work or dynamics not previously identified. In this paper, we offer some answers to these questions.

In addition to shedding a different light on India's judiciary, addressing these issues also can provide new insight into the reigning debates on the character of judicial behavior. Scholars have advanced many theories to explain judicial behavior. Some propose that judges are politically astute, strategic in their actions, and interested in achieving specific policy goals. Others suggest that judges are politically disinterested, sincere in their actions, and motivated by issues of law rather than policy outcomes. These characterizations focus on judges' attitudes toward the particulars of the cases they decide and the political ramifications of specific outcomes (or lack thereof), but they tend to neglect any consideration of judges' attitudes toward the establishment and maintenance of the institution in which they are situated. In the U.S. case, which has driven the development of theory, this is not surprising because the Court has been well established for two centuries and its social and political legitimacy is unassailable. The neglect is harder to justify for studies of judicial politics in new democracies and developing countries, however. In these contexts, both the Court's ability to issue decisions, the extent to which these decisions are honored in the implementation process, and the Court's insulation from political pressure are still being determined.

The Federal and Supreme Courts of India typically have been studied independently of one another. A comparison of the two, though, provides an ideal case in which to assess the impact of political change and transformation on institutional behavior. The Federal Court was established by the British government to address issues between the new federation envisioned in the 1935 Act, but since the Princely States never joined, much of its charter was unfulfilled. Nevertheless, the Federal Court issued several important decisions in its 13 year tenure and was regularly upheld in appeals to the Privy Council in London. It would be an exaggeration to call the Federal Court subservient to the colonial government, but it is rarely singled out for its independent behavior. By contrast, the Supreme Court was a true apex court with a wide jurisdiction and was intended to play a key role in the implementation of the Constitution. But the Supreme Court quickly came into conflict with Parliament and the governing Congress Party, and it found itself being criticized by politicians and elites for putting obstacles in the way of India's rapid social and economic development. Consequently, the Supreme Court's first four decades were punctuated by several periods of intense confrontation with the elected branch. If the Federal Court was often ignored or dismissed as ineffectual, the Supreme Court was criticized for exceeding its authority.

In order to evaluate our argument we analyze the decisions of the Federal Court and the early years of the Supreme Court. If judges are most concerned with protecting the institution, we should find that their decisions tend to support parties

that can affect the strength and stability of the Court; the party with the greatest influence is the British government in the colonial period and the central government in the period after independence. Although the colonial and democratic governments varied on many dimensions, they shared one critical characteristic: both had the ability to alter their courts if they became sufficiently dissatisfied with it. The colonial government had unilateral power in this area despite its limited extension of indigenous political participation, and the post-independence government was a parliamentary system in which the Congress Party dominated and faced no unified opposition. We therefore limit our analysis to those cases in which the Government of India (in the colonial period) and the Union of India (after independence) was a party.

The paper is organized as follows. In the next section we discuss research in the current literature on judicial politics that speaks directly to our questions. We then summarize briefly the history of the Federal Court and the first decade of the Supreme Court, emphasizing their composition, powers, and relations with other branches of government. We then compare judicial decision making by the two courts and analyze the similarities and differences in their treatment of the preferences of the executive or parliamentary branch. In conclusion, we return to our original arguments and consider the applicability of our findings for research in other contexts and other periods.

## 2    Judicial Behavior and Inter-Branch Relations

The literature on judicial politics includes two discussions that are especially pertinent to our analysis: studies of judicial decision making and studies of inter-branch bargaining. We begin by asking what factors predict how judges will decide cases, and the literature on decision making offers a wide range of motivations. At one end of the spectrum lie judges who are motivated almost entirely by issues in the law itself and are largely uninterested in either the policy outcomes under consideration or the political ramifications of their decisions. In the India-specific literature this approach has been termed "black-letter law" (Galanter 1989; Dhavan 1977). There is evidence to suggest that Indian judges had relatively apolitical backgrounds, even during the independence movement (Gadbois 1969), and that they were deeply concerned with issues like precedent and prior case law (Dhavan 1977). But it is a stretch to argue that their decisions were made in a completely apolitical or asocial context. The early decisions of the Federal Court as well as the concurrent writings of its Chief Justice demonstrate that the Court was intensely aware that its early decisions would shape its future status (Pylee 1966).

The literature on judicial behavior includes considerable research arguing that judges have preferences over policies as well as over points of law. Both the attitudinal and separation of powers (SOP) approaches assert that judges have policy preferences, although they take different paths to achieve them (Segal and

Spaeth 1993; Epstein and Knight 1998; Gely and Spiller 1990). The Indian case does not provide direct evidence that its judges had policy preferences independent of issues of law, but there is clearly evidence that they preferred certain outcomes over others in light of specific provisions of the statutory and constitutional questions those polices raised (Dhavan 1977; Galanter 1989; Pylee 1966; Austin 1999).

If judges have preferences over policy outcomes, interpretations of law, or some combination of the two, how do they ensure that their preferences are achieved? On this point the attitudinal and SOP models diverge sharply. The attitudinal model assumes judges are generally apolitical actors who decide cases based on their sincere policy preferences, while the SOP model posits that in order for judges to achieve their sincerely preferred outcomes they must take into account the behavior of other institutional actors who respond to judicial decisions. In anticipating the range of possible responses, at times judges may choose to act strategically and issue decisions that reflect second choices that are likely to be supported rather than first choices that might be challenged or repealed (Epstein and Knight 1998).

The importance of strategic interaction in the SOP model has led to an expanding literature that explores how strategic considerations and constraints on behavior affect judicial decision making, and it assumes explicit or implicit bargaining between the affected branches of government. In a strategic approach, judges will consider how legislatures and executives will implement their decisions and shape their rulings accordingly. Legislators and executive actors in turn will draft statutes and constitutional provisions that give judges more or less flexibility in interpretation. In a variety of time periods and political settings, scholars have shown how inter-branch bargaining narrows the range of possible judicial and political policies chosen (Vanberg 1998, 2001; Eskridge and Ferejohn 1992; Epstein and Knight 1998).

Researchers have offered a range of explanations for why and how the judicial and legislative-executive branches are constrained by this relationship. Rogers hypothesizes that because judges receive cases after laws have been passed and implemented, they have more knowledge than the original legislators and therefore the court's decision may be better informed (Rogers 2001). Several scholars point to the institutional and public legitimacy court decisions can give legislation and constitutional provisions, a consideration which may override legislative or executive preferences over a specific bill (Epstein and Knight 2000; Vanberg 2001; Epstein et al. 2001). And, of course, since legislative and executive branches can punish the court for decisions it dislikes, judges may constrain themselves in order to avoid being sanctioned (Rogers 2001; Gely and Spiller 1992; Epstein and Walker 1995).

A basic characteristic of the SOP model and the rational-actor approach is the assumption that because the actors behave strategically, they will strive to anticipate unwelcome responses. In equilibrium, we should not observe courts challenging legislatures and executives unless they are certain to prevail. But the Indian case contradicts this assumption. Both the Federal and Supreme Courts issued rulings that resulted in sanctions by the executive, and the Supreme Court engaged in long-term, obviously losing battles. We cannot explain this behavior within the

rational-actor paradigm; we would have to assume judges were either badly informed or acting irrationally. But neither explanation is fully satisfying.

We offer a different explanation, one that is probably most applicable to new courts but may also occur at times in mature, established courts. We assert that the current theories of judicial behavior and inter-branch bargaining pay insufficient attention to judges' preferences over the institution itself. Judges care about policies and issues of law, but they also have preferences over the strength and stability of the court. Indeed, it seems obvious that unless judges can assume a stable and powerful court, their policy or black-letter law preferences are much less likely to be achieved. In new courts, issues of stability, strength, and scope of decision making are often still contested (Vanberg 2001; Epstein et al. 2001; Orkeny and Scheppele 1996). Therefore, we hypothesize that while judges will not often seek to challenge the legislature and executive, they will be more likely to do so in cases that they perceive to be especially important in terms of institutional legitimacy or sphere of influence.

In order to assess this explanation, we examine rulings by the Federal and Supreme Courts of India, focusing in particular on cases that involve the government as a party. The Federal Court sat for 13 years and issued less than 150 decisions, so the subset of relevant cases is fairly small. We examine cases from its first session in 1938 through the creation of independent India in August 1947. For the Supreme Court we examine its first 13 years, when Jawaharlal Nehru was Prime Minister, but we limit ourselves in this initial study only to cases involving the central government. We then analyze how often and in which cases the Courts ruled against the government. Before turning to the cases, however, it is useful to discuss briefly the contexts in which these decisions were made.

## 3   The Federal Court of India

The Federal Court came into being under provisions of the British Parliament's 1935 Government of India Act, and as such was not the creation of a democratic regime. Nonetheless, the establishment of such a court was discussed extensively at the All-Parties' Round Table Conferences, which brought together all the British political parties as well as representatives of the major Indian political organizations, ethnic and religious groups, and princely states. The Government of India Act was designed to provide a framework in which Indians and colonial rulers governed together (although not equally). The Round Table Conferences had been convened to develop a new institutional framework to govern India. The British government, faced with increasing opposition to its rule in the 1920 s, invited representatives of indigenous political, ethnic, and religious groups to work with them to develop a system of partial power sharing.[1] Among the most important issues to be decided

---

[1]See Parikh (1997) for a more detailed discussion.

were the system of electoral representation and the formation of a federation which could incorporate both the provinces directly ruled by the British and the princely states. While the outcome of the conferences were inconclusive, the debates were sophisticated and wide-ranging, and the 1935 Act incorporated many of the suggestions raised there (Linlithgow papers, IOLR).

While the introduction of a Federal Court at the top of the judicial system was a response to a new political configuration, the history of the judiciary of the British Raj dated from the early days of the British presence in India and preceded formal colonial rule by nearly a century. The first court was established in Calcutta to introduce the rule of law even though the government was represented only indirectly by its agent, the East India Company. From that beginning, courts were expanded throughout the subcontinent and dispensed decisions based on English common law, Hindu law, and Islamic law. After India formally became part of the British Empire in 1858, the somewhat disjointed system of courts were reconciled into an integrated institution, much of which has survived intact to this day. By the 1920s there were high courts throughout the Indian provinces, with the High Courts of Bombay, Calcutta, and Madras the most prestigious. Appointments to the High Court benches were sought not only by Indian barristers, but also by respected British lawyers and jurists.[2]

Indeed, the prestige of the High Court system produced opposition to the idea of the Federal Court, with many of the sitting High Court judges fearing that the new court would undermine their prestige. They argued that the existing system, in which cases appealed at the High Court level were sent on to the Privy Council in England, could be adapted to suit the new circumstances. However, the forces that supported the new court prevailed, and the provisions of the 1935 Act empowered a Federal Court at the apex of an integrated judiciary. The Court was intended to have a very limited jurisdiction. Under the terms of the Act, British India was established as a federation for the first time; until then, provinces had had bilateral relations with the central Government of India, but had not been statutorily connected to each other. The princely states were not required to be members of the federation, but provisions for their voluntary inclusion were set forth. The very name, Federal Court, reflects the decision of the British government to avoid establishing a Supreme Court (Linlithgow papers, IOLR).

The Federal Court had carefully delineated areas of original, appellate, and advisory jurisdiction. The scope of its original jurisdiction seems fairly expansive at first glance, The 1935 Act specified that

> the Federal Court shall . . . have an original jurisdiction in any dispute between any two or more of the following parties, that is to say, the Federation, any of the Provinces or any of the Federated States.

However, the Act goes on to stipulate restrictions on this jurisdiction which take up more room than the previous section, and goes on to feature the following

---

[2]See Pylee (1966) for a précis of the development of the judicial system under British rule.

clause: "The Federal Court in the exercise of its original jurisdiction shall not pronounce any judgement other than a declaratory judgement." Even the stipulated jurisdiction of the Court bound only the provinces; the princely states were not required to subject themselves to the jurisdiction of the Court.

> The Court's original jurisdiction "was to cover the minimum grounds of dispute in a federation ... Nevertheless, ... [it] was certainly in conformity with the position of such a court as the sole defender of the federal compact" (Pylee 1966, p. 106).

The enumerated aspects of the Court's appellate jurisdiction appeared, like its original jurisdiction, to be quite extensive at first glance. It was given the power consider on appeal any civil or criminal case that involved "substantial question of law as to the interpretation of the Constitution Act [of 1935]" (Pylee 1966, p. 110). And the lower and High Courts were to consider its decisions as binding on them. But this was a less generous reading of appellate power than it first appeared. Once again, the princely states were not required to adhere to its rulings, and the provincial courts were only bound insofar as the particulars of the case fell under the provisions of the Act of 1935. Indeed, the British government expected few cases to arise this way (Linlithgow papers, Box 11, IOLR). As we shall see, however, both litigants and the Court expanded their reading of relevance under the Constitution Act to develop a jurisdiction that was broad and powerful in scope.

Finally, the third arena of jurisdiction was the obligation of the Court to provide advisory decisions to the Government of India about the constitutionality of new provisions formulated either by the Governor-General in Council or by one of the Federal Legislatures. This was expected to be infrequently used, but it was included in the expectation that a new constitution would occasionally give rise to ambiguities in pending legislation.

The provisions for judicial appointment, compensation, and retention were drawn from existing rules for High Court judges, and they were similarly crafted to provide a highly qualified judiciary. Candidates were required to have 5 years' experience as judge of a High Court; be a barrister of at least 10 years' standing in Britain; and have practiced as a pleader for at least 10 years in India. The appointments were formally made by the King-Emperor, in and by the joint decision of the Secretary of State for India in London and the Governor-General in India. Compensation was very generous by the standards of the time, given the cost of living in India: Rs. 5,500 per month for justices and Rs. 7,000 per month for the Chief Justice. Judges could only be removed by the King-Emperor on two grounds: misbehavior, or infirmity of mind or body.

One provision, that of retirement, was carried over from the rules for all judges within the Indian judiciary. Retirement was compulsory at the age of 65. This was 10 years later than the compulsory age of retirement in the Indian Civil Service, but it still removed judges who were intellectually and physically sound. Pylee notes that "the age-limit prescribed for the Federal Court judges was quite high and reasonable *vis-à-vis* the expectation of longevity in India" (Pylee 1966, p. 92). But given the qualifications necessary to be a candidate for the Court, justices were often nearing retirement age when they were appointed. This caused frequent

turnover on the Federal Court. At its inception the Court comprised a Chief Justice and two associate justices. The Government of India, following its usual procedure, named a British Chief Justice, a Hindu Associate Justice, and a Muslim Associate Justice. This pattern persisted until 1947, when the British Chief Justice resigned to return to Britain and the Muslim justices resigned to join Pakistan.

The first Chief Justice, Sir Maurice Gwyer, was a straightforward, consensus choice. He had been a key participant in the Round Table Conferences and was instrumental in drafting the Act of 1935. In addition, he had served as King's Proctor, Treasury Solicitor, and first Parliamentary Counsel to the Treasury. The first Muslim associate justice, Sir Shah Sulaiman, was Chief Justice of the Allahabad High Court for 6 years, and the Hindu associate justice, M. R. Jayakar, was a leading Indian lawyer who had been elected to the Bombay Legislative Council and the Bombay Legislative Assembly as a member of the Swaraj Party and the Nationalist Party respectively. These justices served for only a year; Sulaiman died in 1941 and Jayakar was appointed to the Privy Council. Later appointments were also distinguished, although by the middle and late 1940s the Governor-General and the Secretary of State both complained about finding qualified British candidates in the waning years of the Raj (Linlithgow papers, Box 10, IOLR). The Indian judges continued to be selected from the High Courts of the provinces, and the last Indian justices of the Federal Court became the first justices of the new Supreme Court in 1950.

Many members of the judiciary in India, particularly those in the High Courts, prided themselves on remaining above politics. The British judges, not surprisingly, saw themselves as agents of the Raj who were responsible for administering the rule of law in British India. But even Indian judges tended to avoid politics, or to embrace a moderate or conciliatory brand of political participation that frequently drew scorn from the left wing of the Indian National Congress and from strong nationalist and radical political actors (Dhavan 1977). As scholars of the early years of the Indian Supreme Court have noted, Indian judges of this period were British trained, from high-caste, economically elite backgrounds, and moderate or conservative in their political leanings (Gadbois 1968/1969). High Court judicial positions carried considerable prestige, and the years of British training and practice necessary to qualify for these positions created jurists who often resembled their British counterparts more than they did other Indians. In addition, the British legal tradition, which emphasized black-letter-law and legal formalism rather than judicial activism, reinforced an already strong tendency toward abstract decision-making (Galanter 1989).

These attributes meant that the Federal Court judges were perceived by both the British and Indian political leaders as conservative, removed from the heated politics of the independence period, and unlikely to be legal innovators. This perception gave relief to the British, who relied on the courts to support their efforts to quash the independence movement, and created antagonism and suspicion among the more confrontational members of the Indian National Congress and regional political movements, who assumed the judges would be allied with the Raj.

To summarize, the Federal Court came into existence as a necessary provision of the Government of India Act of 1935, which created a federal system for British India and the indirectly ruled princely states. It was not envisioned by its founders as a particularly powerful or expansive court, and the antagonism of the Indian National Congress toward the Act made it an object of suspicion to many Indian political leaders. However, the possibility for an effective Court was embodied in the statutes, and the quality of the justices appointed granted it additional stature.

## 4    The Supreme Court of India

The Constitution of India is drawn from three sets of sources. The bulk of the institutional provisions are taken from the Government of India Act of 1935 and adapted for an independent sovereign state. The civil rights and liberties sections were modeled on other constitutions, notably the U.S. and Irish examples. And finally, some provisions are drawn from Congress Party documents of the preceding decades. While there are several innovative sections in the constitution, the provisions for the judicial branch of government closely follow the Government of India Act. It is unclear why the Congress party, which was frequently dismissive of colonial institutions, was willing to adopt so thoroughly the colonial judicial system. Perhaps because the courts and their judges played little role in the struggle for independence, Congress leaders viewed judicial institutions as necessary to an independent state but unimportant compared to the elected branch.

When members of the Constituent Assembly took up the judicial provisions of the constitution, they were concerned with creating an autonomous judiciary that would command respect. But they also wanted to forestall a Supreme Court that would impede the Parliament in its efforts to direct India's economic, social, and political development. The result, in Austin's elegant phrase, was that they an idol and then fettered at least one of its arms" (1966, p. 174). The requirements for passing a constitutional amendment were relatively easy to meet, especially for a government with a comfortable majority in Parliament. For cases that did not involve state-level issues, an amendment could be passed by two-thirds of those attending the Parliament as long as more than half the members were present. If state-level issues were at stake, then half the state governments also had to pass the amendment. That this process has been fairly easy to achieve can be seen from the number of constitutional amendments passed in less than 50 years (currently more than 80).

In addition to the political constraints, legal norms and rules specific to the court as a political institution further limited its ability to develop strong and sustained positions that ran counter to parliamentary and bureaucratic preferences. According to Indian common-law tradition, Parliament has the right to delegate the power to pass and administer legislation to central and state administrative agencies and to state and local governments, as well as to quasi-governmental bodies. The courts can invalidate this delegation in only a few circumstances: if the delegated power is too sweeping, contradictory to the statute authorizing it, or "repugnant to the

general law" (Sarathi 1981, p. 399). Even more constraining is the related doctrine of "subjective satisfaction." Under this norm, the agent is given enormous discretion to choose how legislation is to be implemented. Almost any bureaucratic interpretation must be accepted by the courts, even if it can be construed as inappropriate or mistaken. "The emphasis [using the formula of subjective satisfaction] has been laid on the amplitude of the discretionary power rather than on the need to relate it to the purposes of the Act" (Dhavan 1977, p. 239). The only recourse available under statutory review is if the agent has intentionally and on *mala fide* grounds misinterpreted the legislation. Supreme Court has used this last option with considerable ingenuity to develop a pattern of review based on the argument that any interpretation that fails the *ultra vires* test "is in fact a *mala fide* exercise of power" (Dhavan 1977, p. 237), but the strategy has inherent limits.

The Supreme Court's jurisdiction followed closely the provisions laid down for its predecessor. It was granted the explicit power of judicial review, and it had original, appellate, and advisory jurisdiction. Unlike the Federal Court, however, the Supreme Court's caseload was heavy from the outset. The Court has no tradition of *certiorari*, but must take every case present and act upon it. While it disposes of many cases without comment, hundreds of decisions are officially reported, and the majority of reported cases have written judgments. While the Court is allowed some discretion in when it decides to rule on a case, in theory it must rule upon them all at some point. This means that Indian judges have far less ability to pick and choose cases that suit their strategic or sincere interests than their American counterparts. In addition, Supreme Court justices function as trial court judges much of the time. The Court was initially established with eight justices, but justices were added regularly from the mid-1950s, until today's full bench of 26 was reached. The bench system allows routine cases to be disposed with two- and three-judge panels, but cases with constitutional import must be heard by a minimum of five justices, and cases that overturn previous Supreme Court decisions must be heard by a larger bench than the original case. The largest bench ever convened was made up of the 13 judges who heard the *Keshavananda Bharati* case.

Like Federal Court justices, Supreme Court judges must retire when they reach the age of 65. Since Supreme Court judges who have come up through the judicial system are required to have served on High Courts for at least 5 years or been High Court advocates for at least ten, they are well into their fifties and even sixties by the time they are appointed. The average term of a justice is about six and one-half years, with the longest term being 13 years, 5 months and the shortest less than 4 months. Like their judicial colleagues on the Federal Court and the High Courts, Supreme Court justices in the first decades after independence were determinedly non-political, although they were very conscious of the importance of their position and the necessity of establishing the strength and legitimacy of the Court (Galanter 1984).

To summarize, the Supreme Court came into being with the adoption of the new Constitution of India in 1950. It was given wide jurisdiction and explicit powers of judicial review, but Indian traditions of legal deference to the executive mitigated its scope. While the dominant Congress party explicitly intended to have

a legitimate and independent Court, its leaders did not expect the Court to stand in the way of its plans for social and political development.

# 5   Decision Making on the Courts: Evidence and Discussion

We examined decisions of the Federal Court from 1938, when it began, to India's independence in August 1947. Although the Court continued to issue judgments from 1947 to 1950, the relationship between the Court and the executive was clouded by the lack of a constitution and the operation of a provisional government while the Constituent Assembly was drafting the new Constitution; 1947 therefore marks the end of the Court's relationship with the regime that established it. For the Supreme Court, we gathered decisions issued from its inception in 1950 through 1963, the last full year of Nehru's Prime Ministership. The Nehru era was marked by several factors: Congress party dominance at the center and in the majority of state governments; Nehru's unchallenged position as party leader after Sardar Vallabhai Patel's death in December 1950; strong public support for the government; and the absence of a unified opposition.

Recall that we focus primarily on the relationship between the judiciary and the branch of government which has the ability to challenge its decisions and even its institutional integrity, the executive. Therefore we limit ourselves to examining those cases in which the central government authority was a party to the case, either as appellant, defendant, or intervener. In the colonial period this category includes the Governor-General of India, the Viceroy, the Secretary of State for India, and the King-Emperor; the first three are executive officials of the Government of India. In the independence period the category comprises only the Union of India. These definitions result in a set of cases that is slightly under-inclusive, because it does not exhaust all the cases in which the center might have a stake.

We took note of two factors in the decisions: whether the Federal and Supreme Courts ruled in favor of the government or against it, and whether justices were unanimous or split. Those cases in which the Court ruled partially in favor and partially against the government were coded as "against." The results are given by year in Table 1.

Several features are immediately apparent in these data. As we expected, the Federal Court had far fewer decisions involving the government (33 in 10 years) than the Supreme Court, which had 129 in 14 years. One likely explanation for the difference is that the Federal Court's intended jurisdiction was never entirely realized, since the full federation could not be achieved when the princely states refused to join. In the first few years the Federal Court heard very few cases that involved the government; the increase occurred primarily as a result of government provisions related to World War II, in particular the Defence of India Act.

**Table 1** Dispositions of federal and supreme court cases, 1938–1947, 1950–1963

| Year | For govt. | Against govt. | Split dec. | Total |
|---|---|---|---|---|
| 1938 | 1 | 0 | 0 | 1 |
| 1939 | 1 | 1 | 1 | 2 |
| 1940 | 0 | 0 | 0 | 0 |
| 1941 | 0 | 0 | 0 | 0 |
| 1942 | 1 | 1 | 0 | 2 |
| 1943 | 2 | 4 | 2 | 6 |
| 1944 | 3 | 6 | 1 | 9 |
| 1945 | 4 | 3 | 1 | 7 |
| 1946 | 2 | 1 | 0 | 3 |
| 1947 | 3 | 0 | 0 | 3 |
| Total FC | 17 (51.5%) | 16 (48.5%) | 5 (15.1%) | 33 |
| 1950 | 2 | 3 | 2 | 5 |
| 1951 | 6 | 0 | 1 | 6 |
| 1952 | 3 | 0 | 0 | 3 |
| 1953 | 5 | 1 | 0 | 6 |
| 1954 | 5 | 0 | 0 | 5 |
| 1955 | 2 | 1 | 1 | 3 |
| 1956 | 1 | 1 | 0 | 2 |
| 1957 | 9 | 3 | 3 | 12 |
| 1958 | 4 | 1 | 0 | 5 |
| 1959 | 7 | 5 | 2 | 12 |
| 1960 | 8 | 2 | 1 | 10 |
| 1961 | 19 | 4 | 1 | 23 |
| 1962 | 9 | 12 | 2 | 21 |
| 1963 | 7 | 9 | 2 | 16 |
| Total SC | 87 (67.4%) | 42 (32.6%) | 15 (11.6%) | 129 |

The Supreme Court, also as expected, was quite busy during the Nehru regime. With a new constitution, an active, interventionist government, and a population accustomed to litigation, it is not surprising that new legislation was frequently challenged on the grounds that it violated the new constitution. Since the Supreme Court had original jurisdiction over the fundamental rights clauses according to Article 32, cases did not have to work their way through the lower courts. As a result, the Court was hearing important challenges to Parliamentary legislation by the end of 1950.

The two courts are fairly similar in terms of unanimous versus split decisions. The Federal Court issued decisions with dissents in 15.1% of their cases, while the Supreme Court was split even less often, in 11.6% of its cases. These data suggest that whatever the relationship of the Courts to their executives, the judges were overwhelmingly in agreement with each other. In the Federal Court period, four of the five split decisions involved the Court's rulings on the Defence of India Act, and the split reflected the judge's national backgrounds: in these cases, the Indian justices ruled against the government while the British judges voted in favor. But in far more cases, the three judges voted together, either for or against the government.

There is one aspect of the data that runs contrary to our expectations and to the conventional wisdom. The Federal Court ruled against the colonial government far more often than the Supreme Court ruled against the Congress Party government.

Nearly half the Federal Court's decisions rejected the government's argument (16 out of 33), while less than a third of the Supreme Court's decisions went against the government (42 out of 129). Despite the conventional wisdom that the Federal Court was relatively weak and ineffectual while the Supreme Court was aggressively obstructionist, this argument is not supported by a simple numerical tally; we would have to argue that the cases in which the Supreme Court rejected the government positions were somehow more important than those it upheld.

Given this challenge to the conventional wisdom, how can we explain the outcome? Why is the Federal Court's decision making considered unimportant? And why is the early Supreme Court perceived primarily as elitist and out of touch with popular sentiment (Galanter 1989; Dhavan 1977; Austin 1999)? To answer these questions we must go beyond the summary statistical data and look more closely at the cases themselves. For the Federal Court, we observe that the bulk of the cases rejecting the government's position as well as the majority of the split decisions occurred during the war years and involved disputes between individuals and the government over civil liberties issues.

The Court honored its mandate to hear only cases having to do with the constitutionality of existing and new laws. And given that the Court turned down as many appeals to constitutional protection as it supported, it cannot be argued that the Court grasped at every case that had a constitutional aspect in order to increase the number of decisions it made. But the Court's support of civil liberties arguments brought it into direct conflict with the Government of India, most often in cases that involved the Defence of India Act. This Act had been promulgated by the Government of India in the early 1940s as a wartime necessity to maintain order and crush sedition, and in this case sedition included the civil disobedience of the Quit India movement. The Government had argued that national security in wartime overrode civil liberties protections. But in its first case challenging a provision of the Defence of India Act, the Court unanimously ruled against the Government (*Keshav Talpade v. King Emperor*, 30 AIR 1943). In another case, (*Niharendu Dutt Mazumdar v. King Emperor*, 29 AIR 1942), the Court held that the speech of the appellant did not meet the definition of sedition under the act. The Government was infuriated by these decisions, but it accepted them and amended the relevant parts of the Defence of India Act where possible (Linlithgow papers, Box 12, IOLR).

The government's response was to accept the decisions in order to avoid challenging the Court in the first civil liberties cases. In 1943, however, when its first Chief Justice reached retirement age, the government chose a new British Chief Justice who was more likely to vote in their favor: Sir Patrick Spens was a respected judge, but he had never spent time in India and unlike his predecessor, had no particular attachment to the place or its institutions. He dissented on his first civil liberties cases, and during his term as Chief Justice the government increasingly chose to take cases to the Privy Council rather than allow the Federal Court to be the final word.

It is important to note that the Court ruled for the government slightly more often than it ruled against it, 51.5–48.5%. In particular, it upheld much of the new legislation passed by the provincial and central legislatures and it regularly affirmed the legislatures' rights to a wide scope of authority. We cannot argue, therefore, that

the Court grasped at every case that might allow it to assert itself over the executive. But it clearly perceived individual rights cases to be particularly important to its jurisdiction and sometimes more important than the government's assertions of national security and self-defense, even in a time of war.

Turning to the Supreme Court, we examine the cases in which the Court ruled against the government to understand how it is understood in the conventional wisdom to be confrontational and obstructionist. It is hard to argue that a Court that rules with the government 67% of the time insists upon challenging the executive, but this view is widely held. The first critical case heard by the Court was *A. K. Gopalan v. State of Madras* (1950 (1) SCR). Gopalan was a Communist who had been repeatedly jailed under the Preventive Detection Acts for speech that was allegedly threatening. He challenged his detention, arguing that it violated his fundamental rights as provided by the new constitution.

The Preventive Detention Act has an interesting history. The colonial government had issued Preventive Detention (PD) provisions during World War II and the Indian National Congress had repeatedly denounced the government for them. But in the tumultuous and conflict-ridden period following independence, Congress was now the party in government and found that PD was a convenient tool. The PD Act before the Supreme Court in *Gopalan* allowed individuals to be detained for up to a year without being informed of the reasons for their detention. While there were provisions for an advisory board to review the legitimacy of the detention, the government was not required to give the board the reasons for detention, and the board's recommendations were not binding.

The Supreme Court issued a complicated decision in *Gopalan*. Each justice wrote a separate opinion, and the majority upheld certain provisions of the PD Act while invalidating others. It did not declare the Act itself unconstitutional, but it ruled that the withholding of the reasons for detention and the failure to give the advisory board authority over its legitimacy violated the fundamental rights clauses of the constitution. Government leaders were nonetheless displeased with the Court. While they agreed that PD should only be used in cases of security and threats to the nation, they were loath to give up their ability to decide what constituted security and when information about detenus should be offered (Austin 1999). The Act was amended to take account of the Supreme Court's ruling in 1950 and in successive re-enactments, but the essential contours of the policy remained unchanged.

The *Gopalan* case is noteworthy because the Supreme Court was seen by the government as challenging its authority even as other observers criticized the Court for being insufficiently protective of fundamental rights and failing to strike down the Act in its entirety (Austin 1999). In the land reform and compensation cases, the Court went farther, and found itself in a battle with the executive over which branch was the guardian of the constitution.

Among the most important policy initiatives put forth by the Congress government was a commitment to the redistribution of land. The government's advocacy of a "social revolution" and socialist policies led to the policy of "zamindari abolition," where large absentee landlords, or zamindars, would have their land

seized and (in principle) redistributed to the people who actually worked it. Congress also claimed the right to appropriate non-agricultural land to the government or for redistributive purposes. Not surprisingly, landlords and urban property owners challenged these efforts in the Supreme Court as violations of their fundamental rights, and the Supreme Court ruled against the government in several of these cases.

Like the PD case, land reform and compensation cases posed a problem for the government. On the one hand, the redistribution of land and the public ownership of property were critical to the Congress vision of socialist economic and political development. But appropriation without compensation explicitly violated provisions of the constitution. The ongoing confrontation that resulted between the executive and the judiciary led to a debate over who had the ultimate authority over constitutional interpretation. Nehru asserted that Parliament, not the Supreme Court, has the "duty to see whether the Constitution so interpreted was rightly framed and whether it is desirable to change it . . . to give effect to what really . . . was intended or should be intended" (quoted in Austin 1999, p. 87). As for the challenges to fundamental rights, he remarked, "inevitably in big social changes some people have to suffer" (Austin 1999, pp. 87–88).

The Court took a different view. In a series of cases it upheld the individual's right to compensation and struck down governmental attempts to sidestep such compensation. The government passed legislation overriding the decisions, and the Court in turn consistently responded by issuing narrower but equally negative decisions. The resulting inter-branch debate was not fully resolved until 1980 and put the Court in institutional jeopardy in the 1970s during the Indira Gandhi administration. During these decades, the Court lost considerable public prestige because Nehru and the Congress party were successfully able to portray it as elitist and out of touch with popular needs (Rudolph and Rudolph 1987). The Court, on the other hand, considered itself to be protecting the essential part of the constitution, even as some justices worried about challenging the government.

## 6   Conclusion

Despite the institutional continuity between the Federal and Supreme Courts, scholars and other observers have tended to separate study of the two. Conventional accounts depict the Supreme Court as elitist and obstructionist and the Federal Courts as ineffectual, but we find that they share significant similarities. Chief among the continuities is a commitment to protecting individual rights against state incursions, even in the face of hostile and powerful executive branches. The conventional line also does not capture the disparity in the actual distribution of rulings, where we clearly see that in terms of numbers of cases, the Federal Court is more confrontational with its respective central government than the Supreme Court. The threat perceived by each central government, though, may have been

different, with India's new Congress-led government seeing challenges by the Supreme Court as more problematic for their policy agenda.

While further research is necessary to be able to state the justices' motivations with greater certainty, our findings here suggest that in nascent courts, justices seek to strengthen the Court's institutional standing and legitimacy. In India, the struggle between the central government and these apex courts developed around two areas of law which each branch had a significant investment in controlling: fundamental rights pursuant to individual freedoms and the balance between a government's pursuit of economic reforms and reasonable compensation for hardships experienced by populations inconvenienced by such policies. As we see, the Court's limited assertions of judicial authority in these areas may be met with hostility by strong elected or imperial actors who expect to be able to carry out policy preferences without opposition.

# References

Austin G (1966) The Indian constitution: cornerstone of a nation. Oxford University Press, Bombay

Austin G (1999) Working a democratic constitution: the Indian experience. Oxford University Press, Delhi

Dhavan R (1977) The supreme court of India: a socio-legal critique of its juristic techniques. N.M. Tripathi, Bombay

Epstein L, Knight J (1998) The choices justices make. Congressional Quarterly Press, Washington, DC

Epstein L, Knight J (2000) Toward a strategic revolution in judicial politics: a look back, a look ahead. Polit Res Q 53:625–661

Epstein L, Walker TG (1995) The role of the Supreme Court in American society: playing the reconstruction game. In: Epstein L (ed) Contemplating courts. CQ Press, Washington, DC

Epstein L, Knight J, Shvetsova O (2001) The role of constitutional courts in the establishment and maintenance of democratic systems of government. Law Soc Rev 35:117–164

Eskridge W, Ferejohn J (1992) Making the deal stick: enforcing the original constitutional structure of lawmaking in the modern regulatory state. J Law Econ Organ 8:165–189

Gadbois G (1968/1969) Indian supreme court judges: a portrait. Law Soc Rev 3:317–336

Gadbois G (1969) Selection, background characteristics, and voting behavior of Indian Supreme Court judges, 1950–1959. In: Schubert G, Danelski D (eds) Comparative judicial behavior. Oxford University Press, New York

Galanter M (1984) Competing equalities: law and the backward class in India. University of California Press, Berkeley

Galanter M (1989) Law and society in modern India. Oxford University Press, Delhi

Gely R, Spiller P (1990) A rational choice theory of supreme court statutory decisions, with applications to the *state farm* and *grove city* cases. J Law Econ Organ 6:263–301

Gely R, Spiller P (1992) The political economy of supreme court constitutional decisions: the case of Roosevelt's court-packing plan. Int Rev Law Econ 12:45–67

Linlithgow Papers, Eur. Mss. F125/157. India Office Library and Records, British Library, London

Orkeney A, Scheppele KL (1996) Rules of law: the complexity of legality in Hungary. Int J Sociol 26:76–94

Parikh SA (1997) The politics of preference. University of Michigan Press, Ann Arbor

Pylee MV (1966) The federal court of India. Manaktala & Sons, Bombay

Rogers JR (2001) Information and judicial review: a signaling game of legislative-judicial interaction. Am J Polit Sci 45:84–99

Rudolph L, Rudolph S (1987) In pursuit of Lakshmi. University of Chicago Press, Chicago

Sarathi VP (1981) The interpretation of statutes. Eastern Book Company, Lucknow

Segal J, Spaeth H (1993) The supreme court and the attitudinal model. Cambridge University Press, New York

Vanberg G (1998) Abstract judicial review, legislative bargaining, and policy compromise. J Theor Polit 10:299–326

Vanberg G (2001) Legislative-judicial relations: a game-theoretic approach to constitutional review. Am J Polit Sci 45:346–361

# Institutional Arrangements Matter for Both Efficiency and Distribution: Contributions and Challenges of the New Institutional Economics

**Fernando Toboso**

Rules are the results of human beings' efforts to establish order and increase predictability of social outcomes. Rules can be used to increase the welfare of many individuals or, if collective-choice processes are controlled by a well-organized subgroup, to benefit that group more than others.

(Ostrom and Ahn 2008)

## 1 Introduction

Are new institutionalists now systematically disregarding distributive aspects when approaching organizational issues in the political arena as was the case during the 1970s and 1980s? Do economic and political agents usually care about both efficiency and distribution? To provide an answer to these questions is the basic purpose of this paper. For so doing the paper deals with several contributions made by outstanding new institutionalists, mainly belonging to the so-called *political economy* branch of New Institutional Economics (NIE).

The analysis carried out demonstrates that not all NIE oriented authors are now doing so. Several contributions containing distributive aspects are examined, including some articles and books by Douglass North. By means of a well-known graphical analysis, the paper also emphasizes that we all clearly care about distribution, not just about efficiency, when participating in market transactions as well as in collective political decisions. The analysis also reveals very persuasively how institutional

F. Toboso
University of Valencia, Valencia, Spain
e-mail: Fernando.Toboso@uv.es; www.uv.es/ftoboso

reforms affect not only the total amount of transaction costs experienced by participants but also their relative rights and capacities to act and bargain.

As Eggertsson (1996: 16) wrote some time ago, the present paper concludes that the NIE approach also provides an opportunity for analyzing the institutional arrangements that affect the relative bargaining strength of different participants and consequently the relative income and wealth they obtain. Distributive considerations may be relevant as explanatory factors for the events under analysis even in situations in which transaction costs are zero, as this chapter shows. If economic transactions in all arenas are always very much influenced by participants' expectations about future personal gains, then it seems also straightforward that by incorporating these distributive aspects in the analysis its explanatory relevance may be increased.

The chapter is organizes as follows. Section 2 focuses on the evolution registered in the NIE analyses since the 1970s in order to show the enlargement of the approach that has already taken place, particularly in the so-called political economy branch of NIE. Section 3 examines some examples of NIE contributions in which distributive aspects are also taken into account. Through a revised graphical tool also employed by Eggertsson (1990), Sect. 4 emphasizes that distributive aspects are often present in our economic transactions in markets as well as in political negotiations aimed at reaching collective decisions. This section also shows that institutional reforms always have an impact on participants' relative rights and capacities to act and bargain, not just on transaction costs, then affecting the direction and terms of exchanges. Section 5 concludes that there is nothing in the basic assumptions, criteria and methods mostly used by NIE scholars that preclude them from incorporating distributive aspects into the analysis if considered relevant for the research purpose at hand. As indicated, several outstanding NIE oriented authors are already doing so.

## 2 On the Evolution of the NIE Perspective Since the 1970s

As is well-known, Ronald Coase is one of the NIE founders, and his classic works Coase (1937) and Coase (1960) are inevitable NIE references, together with the contributions of the property-rights theory in the 1960s and 1970s.[1] However, it was through the articles and books published over the 1980s and 1990s that a more complete and coherent set of central core concepts, assumptions and criteria could in fact be gathered from the many self-labeled NIE contributions of the time. It was particularly so since the international seminar series on the New Institutional Economics began in 1983 and these debates were published in the *Jour Inst Theo Econ*.[2] In 1997, when the International Society for the New Institutional Economics (ISNIE) was launched, a long way had already been traveled by those hundreds of scholars who participated in

---

[1]For further information about the antecedents see Scott (1984), Eggertsson (1990, Chaps. 8 and 9), Williamson (1985a, b, 1990), and Williamson and Winter (1991).

[2]The list of participants and topics can be seen at <http://www.mpp-rdg.mpg.de/oekinst.html>

the inaugural ISNIE meeting. Ronald Coase and Douglass North had already received the Nobel Prize award.[3]

Of course, those events already belong to NIE–ISNIE history. Historic are also those initial contributions over the 1970s and 1980s in which transaction costs, property rights, bounded rationality and a few more concepts were combined in an attempt to just solve some limitations of standard neoclassical reasoning. At that time it was also common to find papers in which scholars dedicated some of their time to explicitly remark that the antecedents of the new approach could more easily be found in the Neoclassical or Public Choice traditions[4] than in the so-called old institutionalism and their modern successors.[5] During those years many authors repeatedly expressed that their intention was to extend neoclassical microeconomic theory by relaxing some of its core assumptions in order to incorporate institutional factors and organizational aspects into their theories.[6] In some cases, however, we can found self-labeled new institutional contributions of the time in which the emergence of informal institutions, such as some shared behavioral rules, were explained with an approach explicitly based on a psychologistic methodological individualism (plus the usual rational choice assumptions) mode of analysis characteristic of the standard neoclassical contributions and, because of that, not fitting well in the above mentioned attempts to relax Neoclassical core assumptions.[7] Many debates on methodological issues took place at that time.[8]

It was also during the mid-1990s that Furubotn (1993: 8) expressed what seemed to be a shared preoccupation of the time by many new institutionalists. He wrote:

---

[3]Outstanding contributions over those years were Coase (1974, 1982, 1984, 1992), North (1986, 1988, 1989, 1990), and Denzau and North (1994).

[4]While Furubotn (1984: 3) refers to J.M. Buchanan and the public choice literature as something totally compatible to what new institutional economics should be when applied to constitutional choices, in his editorial preface of 1989 (Furubotn 1989: 3) he explicitly rejects de Pareto efficiency criterion that Buchanan uses in dealing with constitutional reform. During those years, referring to his own version of the new institutional economics, North (1990: 140) explicitly spoke against the public choice approach and the tools of the game theory, while in North (1986: 235) he ambiguously stated: "The new institutional economics that I have briefly described in the foregoing section builds on the literature of transaction costs, property rights and public choice, and it requires integration of this three bodies of literature".

[5]See Coase (1984: 230).

[6]Statements in this sense over the 1980s may be found in Coase (1984), Langlois (1986a), Hutchison (1984), Williamson (1984a, 1985b). Williamson sometimes refers to Commons as in Williamson (1985a: 187, 1990: 63).

[7]Axelrod (1984, 1986) are examples of contributions containing this kind of explanatory analyses as regards to the emergence of the "to cooperate if others do" behavioural rule.

[8]In 1984 an outstanding participant in those debates identified the methodological foundations of the new institutional economics with the "traditional foundation stones of neoclassical theory-viz., methodological individualism and the self-interest principle" (Furubotn 1984: 3). By contrast, Williamson (1985b: 197) stated in the same JITE issue the following: "bounded rationality is the cognitive assumption on which transaction cost economics relies". Eggertsson (1990: 10) made things even more complicated at those years with his (much criticised) differentiation between "neo" and "new" institutional economics, as he himself has acknowledged to me recently.

"If it was thought previously that the work of the new institutional economics could be accomplished simply by extending neoclassical theory, there is reason today to believe that something more is required. (. . .) Some insights yielded by neoclassical analysis will continue to be valuable, but theoretical movement seems likely to be in the direction of a more flexible and comprehensive model of political economy". Since the mid-1990s, and particularly since the launching of the ISNIE in 1997, it is evident to me that a significant evolution has taken place in NIE ideas and approaches in general. It is not by chance that the 2009 Nobel prize in Economics went to Oliver Williamson and Elinor Ostrom. As North (2005a), Menard (2004), Menard and Shirley (2005), Eggertsson (2005), Ostrom (2005) and Toboso and Arias (2006a) and others make it easy to check, many analyses are now built on a much more comprehensive approach, particularly in the political economy/public sector branch of NIE. In current NIE contributions, authors systematically take into account the relevant sets of legal rules and social norms, as well as many other organizational details, influencing the human economic transactions under analysis. The strict reductionist rules of methodological individualism for building explanatory analyses are now followed only exceptionally as the rules of institutional individualism (a mode of explanation firstly depicted by Joseph Agassi[9]) are being widely used either explicitly or implicitly. This means that those relevant formal and informal institutional aspects affecting the transaction under investigation are usually taken into account as explanatory variables. This is also the case when the research task consists of explaining those changes in formal institutions that we see everywhere. Organizational reforms cannot appropriately be explained through non-institutional models as institutional arrangements, particularly the formal ones, are nested realities. There are always rules for reforming other rules. Some of them are written, others not. In many cases the informal ones are the most effective.[10] Of course, most self-labeled new institutionalists still concentrate on revealing how institutional arrangements in one arena or another affect transaction costs and the total output reached by participants. Suggested reforms for downsizing transaction costs usually follow these analyses.

These ideas are not new. They were already debated during the 1980s and 1990s as the papers from the seminar series on the NIE that were published in the JITE show. In fact, it was in the 1990s that the NIE research program acquired its current distinguishing characteristics. Some articles and books by North during those years can help us to further explain the NIE approach and also show how it already departed from the neoclassical perspective at that time.

In North (1988, 1990, 1991a, b, 1993, 1995), for example, the *efficiency view* that characterized some of North's previous works was abandoned. Efficient changes are sometimes promoted and sometimes blocked in the political arena. Sometimes they are blocked for centuries, as the case of the medieval manor institution shows. In North's own words at that time: Institutions are not necessarily

---

[9]Agassi (1960, 1975). See also Toboso (2001, 2008).

[10]On multilevel institutional frameworks see Tsebelis (1990) and Williamson (1996c, 2000, 2003).

or even usually created to be socially efficient, rather they, or at least the formal rules, are created to serve the interests of those with the bargaining power to devise new rules. (North 1990: 16) He also abandoned his previous purpose of explaining all institutional changes in terms of self-interested human actions responding to changes in prices, technologies and natural conditions alone. Bounded rationality, as a behavioral assumption, entered North's analyses as well as many others.[11] Furthermore, it is through these publications that preliminary insights were provided on the role mental models and ideologies can play. His concern with cultural inertia and path-dependence should not surprise anyone given that North is an economic historian.

Although capital accumulation and technological progress are relevant factors, the sources of contrasting economic performance between societies, he wrote, lie within the institutional structures that define incentives for saving, investment, production and trade and that also influence production and transaction costs. Because in standard neoclassical analyses most, if not all, institutional factors are removed and zero transaction costs are usually assumed, these analyses are of little help to North for explaining growth and development. Although North's analyses at that time were no doubt of an institutionalist kind, individual action remained essential for his explanations. Of course, individuals can act independently or they can coordinate their strategies and efforts through organizations. People in organizations always try to profit from the existing institutional environment, but they also attempt to modify the given institutional structure in order to achieve a more favorable one.

Although this is not the place to carry out a survey, what seems evident is that more and more relevant insights are being provided as the approach is being made wider and wider.[12] Recently, Williamson (2003) has stressed that, in contrast to the neoclassical resource allocation approach in economics, what NIE represents is a move from the lenses of choice under physical, monetary and technological constraints to the lenses of contract and organization for systematically analyzing people economic transactions and agreements of all kinds. North (2005b: 22) has also stated that "in contrast to standard (neoclassical) theory that draws its inspiration from physics, modeling the process of change must derive its inspiration from evolutionary biology. But in contrast to Darwinian theory in which the selection mechanisms are not informed by beliefs about the eventual consequences, human evolution is guided by the perceptions of the players in which choices -decisions- are made. . . in pursuit of their goals".

Transaction costs, credible commitments, modes of governance, persuasive abilities, social norms, ideological values, decisive perceptions, gained control, enforcement mechanism, assets specificity, human assets, social capital, asymmetric

---

[11]See for example Williamson (2003).

[12]In the ISNIE website <www.isnie.org>, it can be read: "The New Institutional Economics (NIE) is an interdisciplinary enterprise combining economics, law, organization theory, political science, sociology and anthropology to understand the institutions of social, political and commercial life. It borrows liberally from various social-science disciplines, but its primary language is economics. . .".

information, strategic behavior, bounded rationality, opportunism, adverse selection, moral hazard, contractual safeguards, surrounding uncertainty, monitoring costs, incentives to collude, hierarchical structures, and even bargaining strength, etc are now analytical conceptualizations much used when considered relevant for the research task at hand, even if some concepts abound more in some research areas than in others.[13]

If economic theories and reports must be relevant and instructive from the point of view of practitioners in firms, markets, governments, etc., they cannot be all built on an identical ex ante methodological jacket. Accounting for relevant particularities requires, of course, a methodologically consistent approach or research program formed by a set of basic *hard core* conceptualizations, principles, and criteria. However, it also requires a varied and wide set of *protective belt* concepts to choose from depending on the situation under investigation.[14]

These central core NIE conceptualizations and criteria are now being also used in research areas in which 15 years ago they were just marginal. New institutionally oriented journals have been launched, and an increasing number of books and articles are published each year as on line search engines show. In the realm of development economics, for example, Bardhan (2004) has recently written: Earlier preoccupations with the forces of capital accumulation or technological progress have been widely replaced by a belief that the institutional framework of an economy is crucial for an understanding of the process of development or lack of it. Several Annual Reports by the World Bank have also focused on the importance for economic development of general institutional arrangements.[15] And even Oates (2005), in a recent article entitled Toward a Second Generation Theory of Fiscal Federalism, refers to the lessons that can be obtained from NIE contributions.[16]

Therefore, it can be said that the NIE perspective has evolved substantially since the 1970s, particularly as regards to the analyses of public policy and institutional reform issues. Many contributions represent now much more than simple extensions of the neoclassical approach attempting to incorporate transaction costs. These costs as well as efficiency considerations are still important but they are not all. Many new institutionalists engage now in multi-disciplinary collaboration aimed at breaking traditional disciplinary divides in order to account for more and more situational factors, including political, cultural, and path dependence-historical ones. As a result, a variety of approximations can be found even among authors dealing with similar phenomena if they take place in different historical, cultural, or locally specific

---

[13]See Williamson (2000), See also Toboso (1995, 2006).

[14]Among many others, Coase (1974: 181, 1982: 7, 1992: 718) has recurrently emphasized these and other aspects. Menard (2001) deals also with key methodological issues of NIE theories.

[15]See for example, World Bank (2002), Institutions for Markets; World Bank (1997), The State in a Changing World; World Bank (1994), Institutional Change and Public Sector in Transition Economies.

[16]Several other examples could be mentioned for showing the increasing attention paid to these contributions. Schmid (2001) emphasizes that several scholars with an institutionally oriented economic approach have received the Nobel award.

situations, as Kingston and Caballero (2009) have recently shown, for example, concerning analyses of institutional change. It is not by chance that Ostrom (2007) explicitly mentions the necessity to go beyond panaceas in order to build a strong interdisciplinary science of complex, multilevel systems that will enable future diagnosticians to match governance arrangements to specific problems embedded in a social–ecological context, thus avoiding making simple, predictive models for deducting universal solutions, panaceas, to complex, situational collective problems.

## 3  Some Examples of NIE Contributions Containing Distributive Aspects

To achieve the stated aims of the present paper, now we must turn to and examine some of the NIE contributions in which distributive aspects are taken into account. As mentioned earlier, here I do not intend to conduct a complete survey on the topic but to provide some examples, starting with some contributions by North. Although North often focuses on efficiency issues trying to evaluate how alternative institutional arrangements affect transaction costs and the total output obtained by participants, he has sometimes paid explicit attention to distributive issues, to participants' bargaining strength and to non-voluntary transactions.

In North (2005a: 112), he points out the *violent struggle among competing groups for control of the polity and economy* that took place in all new Latin American republics created after the defeat of the Spanish and the rise of independence movements. Even if many countries adapted a version of the United States Constitution as a model for independence, the consequences were radically different because of the colonial heritage they suffered from. The entire pattern of settlement, trade, and development was geared to the extraction of precious metals for the Spanish Crown. Through an authoritarian system not based on self-government, the Crown granted exclusive monopoly privileges to selected merchants, and trade was confined to a small number of ports in the whole of South America. The evident and well documented purpose was to facilitate the extraction and sending of precious metals to Spain, not to promote the development of people living there, North writes.

Without a heritage of democratic self-government with well-defined and enforced political and market rules, and a legitimated distribution of property and economic assets, independence disintegrated in a violent battle between groups for capturing the polity for controlling economic transactions. Those groups that emerged victorious established authoritarian regimes to secure order and the phenomenon of *caudillismo* became pervasive. But new conflicts appeared with those who had inherited rights from the royal regime as they were in fundamental conflict with the interests of the new governing elites and their supporters. Huge land grants had been given to wealthy individuals as well as to the church elites. A series of

local monopolies in production and trade had also been authorized to be managed by some family groups. The result was, and still is, continuous political instability, extensive rent-seeking behavior at expenses of productive activities, adverse income distribution, a really poor provision of public goods, and a huge number of people suffering from extreme poverty. North concludes that the endowment arguments for explaining growth and development must be fundamentally supplemented by the powerful consequences the path dependent colonial heritage have on both informal and formal institutional arrangements.

All these arguments are, of course, debatable. However, they reveal that some analyses by North also contain references to distributive issues, conflict of interest, and groups of people with diverging capacities to influence the rules of the political and economic games being played, setting aside their differences as players under the existing rules. In North (2005a: 165) he also writes: As noted above, alteration of the economic rules entails winners and losers and it is essential to be aware of them. ... In 1990 North (1990: 16), had already explicitly stated: "Institutions are not necessarily or even usually created to be socially efficient, rather they, or at least the formal rules, are created to serve the interests of those with the bargaining power to devise new rules".

Some contributions by Barry Weingast may also be mentioned as examples of analyses in which non-voluntary influences and distributive issues are taken into account. In the classic and very cited paper by Weingast and Marshall (1988), these authors mention that the control over the agenda within their jurisdiction by the Committees in the USA Congress implies that each committee has veto power over the proposals of others. Agenda power allows committees to bias the outcome in favor of the alternatives they prefer most. The agreements made among members of a committee or between members of different committees are simply enforced by the property rights over seats system already mentioned. A legislator on committee $i$ gives up influence over the selection of proposals in the area of committee $j$ in exchange for members of committee $j$'s giving up their rights to influence proposals in the area $i$. Institutionalizing control over the design and selection of those proposals that will come to a vote substitutes for purchasing the votes of others in an explicit market. Since committees afford their members a disproportionate influence over the policy choice within their jurisdiction, representatives from farm districts, for example, are much more likely to bid for seats on agriculture committees than they are for seats on urban, housing, or merchant marine committees.

There can be no doubt that distributive considerations are present in the analysis made by Weingast. The diversity of interests among legislators, Weingast says, creates gains from exchange within the legislature. But under a logrolling system, votes are usually *sold* and *bought* for a *price*, with the equilibrium prices determining vote trades and hence the set of bills passed. Legislators are better off, Weingast writes, by giving away votes on issues that have lower marginal distributive impact on their districts (and therefore on their electoral fortunes) in exchange for votes on issues having larger marginal impact. But what if bills that are going to come are not known in advance? Or what if future events modify the payoffs of bills already exchanged? Or what if a legislator changes his mind and his perceptions of an issue that was previously subject to exchange? In a pure exchange system or simple

logrolling the time dimension cannot be fully taken into account, consequently the enforcement of agreements remains exogenous. It is unlikely that agreements will cover more than one legislative session. A variety of exchange problems arises because the value of today's legislation significantly depends on next year's legislative events, Weingast writes. Even if no change in seats has occurred, members of future sessions may face different incentives from those faced when the trade occurred and may seek, for example, to amend, abolish, or simply ignore previous agreements. Moreover, these settings inhibit the ability of reputation to serve as the sole enforcement tool. There must be little doubt that the analysis Weingast makes in this paper is a good example too.

As already mentioned, all these contributions are brought here just as examples. And of course, several other articles and books can be mentioned. Libecap (1989a, b), Winiecky (1994, 1998), Winiecki (1996), Bardhan (2000, 2001, 2004), Greif (2005a, b), Nye (1997), Mokyr and Nye (2007), Knight (1992) and Knight and North (1997) are some other examples. Even in the industrial-business organization branch of NIE we can find some examples of analyses in which distribution issues, bargaining strength and control aspects are mentioned. Williamson (1996a, 1997) and Menard (1997, 2004: 39–45) are outstanding examples as both refer even to the power influences that emerge from hierarchical relationships and that are necessary to assure control over the assets. Current NIE conceptualizations and methods have been enriched so much over the last decades that several authors refer also to situations in which, for example, some groups of people may oppose a specific market institutional reform even if there is a generalized perception that total transaction costs could be reduced as a result.[17] In a similar sense, I would finally like to mention here the views of Horn (1995: 16) where he states "if enacting legislators' commitments, as well as the benefits provided to their constituents, are uncertain when subsequent legislatures come, they may have an incentive to protect those benefits by even attempting to implement inefficient institutional arrangements that increase the transaction costs of reversing those policies". Again, it is relevant to mention here that the book by Horn was published in the Political Economy of Institutions and Decisions collection of books edited by North and Alt.

It must now be obvious that outstanding authors working in the NIE tradition do pay attention to distributive aspects in some of their contributions. For these authors it is evident that institutional arrangements have also a significant impact on distribution, that is on who gets what and how much of it in any arena. If team action is considered, now it is also evident for them that arrangements used for organizing team action also affect the share in those efforts and costs needed to accomplish the common purpose. As different alternatives for reducing production and transaction costs always exist and they often produce (expectedly) different distributional impacts on the affected participants, proposals for organizational reform are always a source of some dispute. In the polity this often causes conflict ending not in a general agreement but in a final vote with winners and losers, as Libecap (1993: 32) already emphasized time ago.

---

[17]See Libecap (1989a, b), Greif (2005a, b, 2008), Winiecki (1996) and Bardhan (2000, 2001).

Finally, let me cite Eggertsson ([1996](): 16), former president of the ISNIE too, who wrote almost 25 years ago that the NIE perspective "also provides an opportunity for explaining the institutional arrangements than affect the relative power of workers and employers, and exploring how these power relationships emerged and how they are maintained... The framework does not suggest that all institutional change is explicitly designed to increase aggregate wealth as many critics seem to believe. Purposive institutional change reflects both the power and interests of those who control institutional change and the process for making decisions in the political sphere".[18]

## 4 Institutional Arrangements Also Matter for Distribution: A Graphical Analysis

The above comments and citations reveal that several outstanding new institutionalists do in fact account for distributional aspects in some of their analyses. Now, it is time to go a step further in order to fully accomplish the purpose of the present article. Do economic and political agents systematically care about distributive aspects too, not just about efficiency? Or are these aspects usually irrelevant for explaining events and decisions we all take in the economy and the polity? I think the obvious answer is that we all care about distribution when participating in economic transactions (selling and buying, investing, making international alliances, outsourcing, etc.). The attempts to increase efficiency by reducing production and transaction costs in private firms, for example, often are just instrumental measures in order to protect or even increase the market share and obtain better distributive results than competitors.

What about collective actions aimed at institutional reform adopted through central, regional or local levels of government in western democracies? Are these reforms partially influenced by distributive considerations on the part of relevant actors? Do changes in the institutional–legal environment directly impact on participants' current and future distributive achievements? Again, I think that the obvious answer is yes in all these cases too.

The stylized graphical analysis that follows is particularly suitable for showing how pervasive these distributive aspects are in human economic decisions adopted in markets, governments and many other institutional settings such as clubs, non-profit organizations, unions, professional associations, etc.

Because the ideas I am emphasizing here have a *non-orthodox flavor* if compared with traditional NIE contributions on transaction costs and efficiency, I will use a traditional, orthodox analytical tool in economics to increase their degree of persuasiveness for economic audiences. With a minor revision, the Edgeworth Box will serve my purpose as all economists are familiar with these concepts. Though designed for emphasizing the benefits of voluntary market exchanges, the tool can be adapted for persuasively revealing the pervasiveness already mentioned.

---

[18]See also Eggertsson ([1995](): 48), See also Toboso and Compés (2003).

For so doing, instead of considering the potential exchange of two private goods (goods, services and/or money) under stable rules of the game, the box needs to be modified to incorporate an activity that generates external effects and also to leave open the possibility of a change in the institutional–legal environment confronted by participants. This means that the traditional Edgeworth Box has to be drawn without a lid. No other change is needed.

I firstly found such a revised Edgeworth Box in Eggertsson (1990), who refers to a conference paper by Haddok and Spiegel. It is true that Eggertsson uses the said *broken box* for a different purpose than the one I am pursuing here. He mainly emphasizes that the potential benefits of voluntary market exchange may not be fully realized if transaction costs are high for participants. Therefore, efficiency could be increased if the said institutional–legal framework could be altered in the direction of reducing transaction costs and encouraging market exchange. For example, Eggertsson states, *by enforcing exclusive rights to the commons or introducing individual marketable quotas in ocean fisheries.*[19] As often emphasized in the NIE tradition, institutional–legal arrangements matter for transaction costs, for efficiency and for economic growth. In Fig. 1, which is taken from Eggertsson (1990: 106), high transaction costs may block exchange and participants are therefore not able to move from S to S* or from F to F* depending on the initial assignment of decision
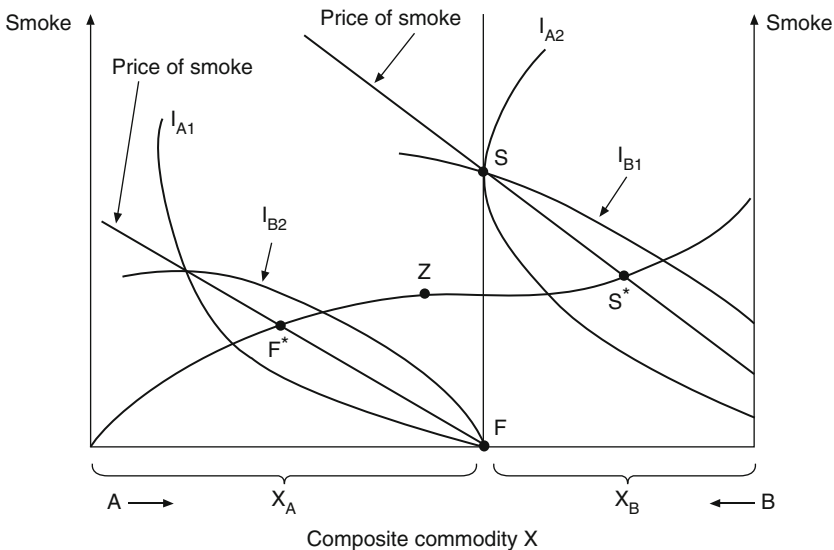


**Fig. 1** Property rights and the cost of exchange in a revised Edgeworth Box
*Source*: Scanned from Eggertsson (1990: 106) who takes it from a conference paper by Haddok and Spiegel

---

[19]Though the purpose of Eggertsson in the said pages is not focusing on distributive aspects, in page 109 he just mentions that "the assignment of property rights to an individual produces a kind of wealth effect that influences his or her valuations".

rights upon the activity that generates the external effect (in our case, producing a kind of public good – smoke – that is jointly consumed by the two participants).

But these are not the aspects I am emphasizing in this paper, as mentioned before. Instead, I am going to use the said broken Edgeworth Box to stress that institutions also matter for distribution. Institutional–legal arrangements are not neutral as regards to distribution as they also affect the initial assignment of rights and entitlements and, consequently, the subsequent capacities of involved participants to act and bargain, ceteris paribus. And this is always true regardless of the amount of transaction costs resulting from the said institutional–legal environment. Altering the said institutional structures and the rules of the game they define generates consequences not only on transaction costs and the total output experienced by participants (efficiency aspects) but also on the share of that output each one gets: his–her current and future relative income and wealth.[20]

In order to show this graphically, I have to draw a new top-broken Edgeworth Box that is shown as Fig. 2. This diagram also represents a very stylized situation as was the case in Fig. 1. In a nutshell, the main aspects of the situation are the following. Two persons who only care about their own interests have to work in the same room with nobody else around. There is only a formal institutional framework formed by a single legal rule framing their interaction (smoking allowed) Both participants own a quantity of a private product for exchanging purposes. The stylized situation contains also an external effect that affects negatively to the well-being of the non-smoker (the level of smoke generated in the shared room) The diagram allows for making a *comparative statics* analysis of alternative institutional–legal settings which is particularly suitable for showing, as already mentioned, that institutional arrangements also matter for distribution.

In our stylized example, if smoking is allowed, it is the smoker (let us assume that he is a man) who owns the right to initially decide how much smoke will be in the shared room. This means that he will smoke as much as he likes. And the other person (let's assume that she is a non-smoker girl) will suffer from a working environment full of smoke. Using the Edgeworth well-known concepts, we might represent the two participants as having maps of potential indifference curves such as the ones of Fig. 2, for example. As each curve represents alternative combinations of good X and cigarettes smoked (level of smoke) that make a participant equally satisfied (similar level of personal utility), it is straightforward that if the initial distribution of good X were the one indicated in the said diagram (Xa, Xb), the assumed legal–institutional framework would lead (if person A only cares about himself and he smokes a lot) to point $S_1$ of Fig. 2 according to Edgeworth assumptions and

---

[20]In order to concentrate on the aspects being investigated here, many other institutional and non-institutional aspects will be left aside. This is the case of aspects related with the possibility that some participants might have the right to behave as veto players. As the analysis built is of a comparative statics nature, those aspects related with causes and processes aiming at institutional reform will not be considered either. On veto players and political, nested games played at different institutional levels see Tsebelis (1990, 2002) As will be mentioned next, in this section institutional change is assumed exogenous for the analysis.

Broken Edgeworth Box

External effect

Level of smoke in
the shared office

External effect

Level of smoke in
the shared office

$P_1^*$

$CI_1^a$

CC

$CI_1^b$

S1

$f_1$

$z_1$

$E^*$

$h_1$

$f_1'$

$t_1$

A

Smoker

X'a

Xa

X'b

Xb

B

Non-smoker

Exchangeable private good
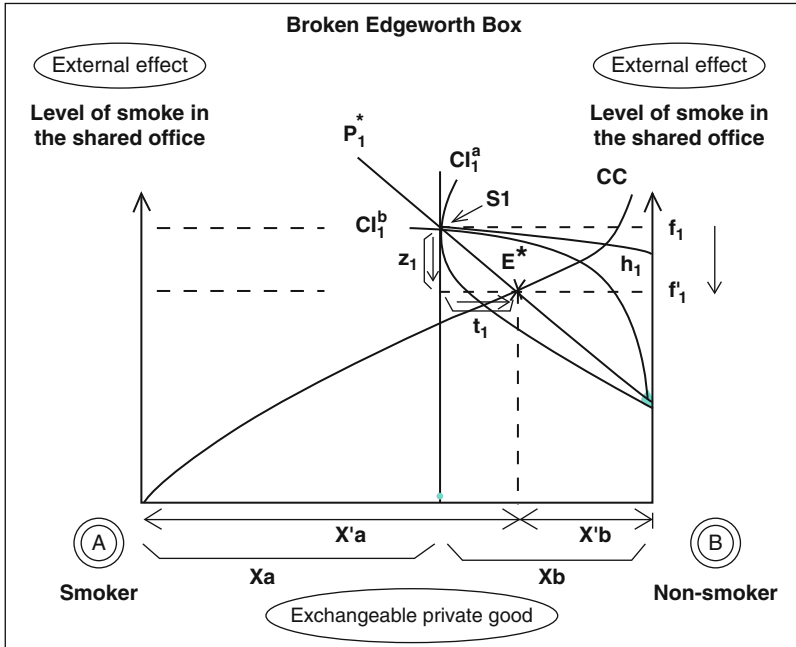
Institutional framework 1:  smoking allowed
Initial assignment of rights/entitlements ........ S1
Free versus bargained exchanges...... E*

**Fig. 2** Distributive consequences of institutional arrangements: a graphical analysis
*Source*: Adapted from Eggertsson (1990)

criteria. This means that the smoker would place himself at the highest indifference curve ($CI_1^a$) he can reach. The non-smoker girl (agent B) would confront a very different situation as regards her well-being or total utility obtained from the situation: a room full of smoke. She would be placed in a low indifference curve in the said stylized Edgeworth diagram: ($CI_1^b$) If any impediment existed for them to talk and reach a voluntary agreement, this would represent the final situation with person B experiencing very low utility from the situation (a very low indifference curve, $CI_1^b$) because of such a high level of smoke in the shared room ($f_1$).

If the two participants care about their own interests but (modifying the previous assumption) still are open to dialog and exchange, it is reasonable to think that an agreement might be achieved. The prevailing institutional framework very much influences (together with other factors that are not being considered in our stylized example[21]) the nature-direction and the terms of the exchange, with person B (the non-smoking girl) having to pay person A for him to smoke less. How much will person B

---

[21]I am not considering, for example, what Ostrom (2009), Ostrom and Walker (1991), Williamson (2003) and others in the NIE tradition refer to as "cheat talking", "opportunistic behavior", "costly external enforcement" and the like.

pay for each cigarette the other person smokes less will depend on many factors that are out of the stylized situation considered. Figure 2 just shows an intermediate situation in which the bargaining capacity of both parts is quite similar (the price of exchange is represented by $P_1^*$) Person A reduces the level of smoke from f to f'$_1$ and the non-smoking girl gives a part of her private good to him (an amount of $t_1$) But the terms of the exchange could even be more favorable for the smoker who is in a better bargaining position than person B because of the initial assignment of rights. I will consider this in Fig. 3.

The story can be extended, of course, as the *orthodox* analytical tool employed also allows for an account of institutional change to be made; an account through which the distributive impacts of institutional reform can be shown very persuasively too. Institutional reforms do not impact on transaction and production costs only. Institutional reforms often have differential distributive impacts upon participants too. That is why monetary payments are often needed to compensate



**Fig. 3** Distributive consequences of institutional reform: a graphical analysis
*Source*: Adapted from Eggertsson (1990)

those participants who fear to see their rights, capacities and/or opportunities deteriorated after the reform (or expect to profit less than others in the new institutional setting created).

Figure 3 is extremely useful for showing how these distributive consequences of institutional reform are generated. For my purpose here, let's assume that an exogenous change in the legal environment of the situation takes place and smoking at work facilities is no more allowed except if all persons involved in the same room agree to the contrary.

Under the new institutional framework, the legal right to decide about the level of smoke in our stylized situation would automatically be transferred from smoker A to non-smoker B if both are normal people that voluntarily obey the law. And this new assignment of rights drastically changes the situation confronting our two participants, also affecting the direction and terms of a potential exchange they might reach under the circumstances (institutional and non-institutional).

In the new institutional setting, it seems evident that the best choice for person B is to free the room of smoke, as her well-being (utility) diminishes when the level of smoke increases. The revised Edgeworth box helps to graphically illustrate the new situation. S2 would represent the new starting point as the non-smoker reaches there her most advanced indifference curve if she only owns Xb of the private exchangeable good. The smoker would remain with his initial amount of X (Xa) and cero cigarettes smoked. It is evident that the non-smoker is better in the new institutional setting than in the previous one. The Edgeworth box also serves to illustrate this as the non-smoker appears now placed in a much more advanced indifference curve ($CI_2^b$) if compared with $CI_1^b$ in the previous institutional setting. The smoker is having a bad time, particularly if he cannot leave the room to smoke. In the Edgeworth box this means that he gets initially placed in a very low indifference curve ($CI_2^a$) when compared with his initial situation in the previous institutional setting ($CI_1^a$).

If both participants decided to open negotiations to see whether an agreement is possible, it is evident that the new institutional framework has changed the direction and terms of the exchange, in case. The new institutional setting would, in case, require the smoker A to pay person B for allowing him to pollute less the work environment. The institutional setting also influences the bargaining capacities of traders, now being agent B in a better position to get a more favorable exchange, in case. Instead of the average $P_2^*$ that might result from the standard Edgeworth box assumptions, our stylized example makes a price of exchange more feasible such as $P_2^{**}$, which shows more favorable terms of exchange for the non-smoker. In the previous institutional setting, it was the smoker who was in a better bargaining position.

It must now be obvious that the stylized example here examined and the analytical tool employed are extremely suitable for persuasively stressing that institutional reforms also tend to produce differential distributional impacts on participants that may be compensated by complementary payments or not. Institutional–legal arrangements greatly affect the starting point from which people enter market negotiations and consequently the feasible alternatives they confront as well as

their relative bargaining capacities, ceteris paribus. Letting aside some other relevant institutional and non-institutional factors not incorporated in the stylized example examined, through this graphical analysis it can be shown very persuasively how the direction and the terms of the exchange are, in part, influenced by the institutional arrangements framing the exchange. Of course, institutional factors are important but they are not all. As Ostrom (2005: 29) also emphasizes, I like to end this paper by stressing that *the focus on the components of institutions in this volume should not be interpreted to mean that I feel that institutions are the only factors affecting outcomes in all action situations.*

## 5 Final Remarks

Coase, North, Williamson, Menard, Eggertsson, Libecap and many other scholars in the NIE tradition have been able, particularly since the mid 1990s, to make it clear for almost everybody that *institutions matter for economic performance.* Formal and informal institutional arrangements are seen now as key factors influencing the total amount of transaction costs confronted by participants in all arenas and, consequently, the total output collectively obtained. These have not been, however, the aspects that have attracted my attention in the present paper. Here I have emphasized that institutional arrangements also matter for distribution and that we all care about distributive aspects when transacting through markets, governments and other institutional settings such as clubs, for example. I have also shown that some well-known NIE oriented authors are paying attention to these distributive aspects in some of their published articles and books.

Institutional–legal arrangements are not neutral as regards to distribution. And the same can be stated concerning institutional change and reform. Because of the expected differential distributive impacts of institutional reform, some groups may oppose a new legal reform even when there is a generalized perception that total transaction costs might be diminished and, therefore, the total efficiency of interactions taking place in such a market or setting increased, ceteris paribus. These agents might even reject a compensation payment and be in favor of an alternative reform for reducing transaction costs if they expect to get a bigger share in current and future results through this alternative reform. In many occasions there are barriers that create transaction costs for some agents but, at the same time, represent mechanisms other agents use to ensure higher incomes for themselves or for those they love or support.

The analysis carried out in the paper also allows to conclude that distribution aspects are so pervasive that they cannot be disregarded if complete explanatory pictures of the economic transactions taking place all over have to be built. There is nothing in the basic assumptions, criteria and methods mostly used by scholars attending ISNIE conferences that preclude them from incorporating distribution aspects into the analysis. Though many authors in the NIE tradition like to focus on transaction costs and efficiency, here I have shown that some outstanding NIE

scholars do in fact pay attention to distributive aspects in some of their contributions. Of course, many other scholars in other traditions or schools of thought, both in economics and political science, like to focus on these aspects, but it has not been my purpose here to deal with these other analyses.

The orthodox but revised Edgeworth Box here drawn, has allowed me to easily decouple some of these aspects thanks to the stylized, but real, example examined. The institutional–legal framework greatly affects the starting point from which people enter market negotiations and, consequently, the feasible alternatives they confront as well as their relative capabilities to act and bargain, ceteris paribus. The terms of the exchange are, in part, influenced by the legal–institutional arrangements framing interactions in all arenas, putting aside some other relevant factors deliberately excluded from consideration in the stylized example used. As the graphical analysis here employed also shows, institutional reforms greatly alter the participants' rights, opportunities and capacities to act and bargain, thus affecting their current and future income and wealth. The analysis provided reveals that institutions matter for distribution too, not just for efficiency. Although all this is well-acknowledged and recurrently emphasized by authors doing research under alternative schools of thought in economics and political science, accounting for these other contributions was not among the purposes of the present chapter.

# References

Agassi J (1960) Methodological individualism. In: Oneil J (ed) Modes of individualism and collectivism. Heinemann, London

Axelrod R (1984) The Evolution of Cooperation. Basic Books, New York

Axelrod R (1986) An Evolutionary Approach to Social Norms. Am Pol Sci Rev 80:1095–1111

Agassi J (1975) Institutional individualism. Br J Sociol 26:144–155

Bardhan P (2000) Understanding underdevelopment. Challenges for institutional economics from the point of view of poor countries. Re-edited in: Menard C (ed) (2004) The international library of new institutional economics. Edwar Elgar, Cheltenham

Bardhan P (2001) Distributive conflicts, collective action, and institutional economics. In: Meier GM, Stiglitz JE (eds) Frontiers of development economics: the future in perspective. Oxford University Press, Oxford

Bardhan P (2004) Scarcity, conflicts, and cooperation. Essays in the political and institutional economics of development. MIT Press, Cambridge, MA

Coase RH (1937) The nature of the firm. Economica New Series 16:386–405

Coase RH (1960) The problem of social cost. J Law Econ 3:1–44

Coase RH (1974) Economists and public policy en Fred Weston J large corporations in changing society. New York University Press, New York

Coase RH (1982) How should economists choose. Warren Nutter lecture in political economy. The American Enterprise Institute for Public Policy Research, Washington, DC

Coase RH (1984) The New Institutional Economics. J Inst Theor Econ 140:229–231

Coase RH (1992) The institutional structure of production. Am Econ Rev 82:713–719

Denzau AT, North DC (1994) Shared mental models: ideologies and institutions. Kyklos 47:3–31

Eggertsson T (1990) Economic behaviour and institutions. Cambridge University Press, Cambridge

Eggertsson T (1995) Economic perspectives on property rights and the economics of institutions. In: Foss P (ed) Economic approaches to organizations and institutions. Dartmouth, Aldershot

Eggertsson T (1996) A note on the economics of institutions. In: Alston LJ et al (eds) Empirical studies in institutional change. Cambridge University Press, Cambridge

Eggertsson T (2005) Imperfect institutions. Opportunities and limits of reform. University of Michigan Press, Ann Arbor, MI

Furubotn EG (1984) The New Institutional Economics. Editorial Preface. J Inst Theo Econ 140(1):1–6

Furubotn EG (1989) The New Institutional Approach to Economic History. Editorial Preface. J Inst Theo Econ 145(1):1–5

Furubotn E (1993) The New Institutional Economics. Recent progress; expanding frontiers. J Inst Theor Econ 149:1–10

Greif A (2005a) Commitment, coercion, and markets: the nature and dynamics of institutions supporting exchange. In: Menard Cl, Shirley M (eds) Handbook of New Institutional Economics. Springer, Berlin

Greif A (2005b) Institutions and the path to the modern economy: lessons from medieval trade. Cambridge University Press, Cambridge

Greif A (2008) Toward political economy of implementation: the impact of administrative power on institutional and economic developments. In: Helpman E (ed) Institutions and growth. Harvard University Press, Cambridge, MA

Horn MJ (1995) The political economy of public administration. Institutional choice in the public sector. Cambridge University Press, Cambridge

Hutchison TW (1984) Institutional Economic Old and New. J Inst Theo Econ 140(1):20–33

Kingston C, Caballero G (2009) Comparing theories of institutional change. J Inst Econ 5(2): 151–180

Knight J (1992) Institutions and social conflict. Cambridge University Press, Cambridge

Knight J, North DC (1997) Explaining the complexity of institutional change. In: Weimer D (ed) Organizational, and structural changes in industries and firms. Kluwer, Norwell, MA

Libecap G (1989a) Distributional issues in contracting for property rights. J Inst Theor Econ 145: 6–24

Libecap GD (1989b) Contracting for property rights. Cambridge University Press, Cambridge

Libecap GD (1993) Politics, institutions, and institutional change. Comment. J Inst Theor Econ 149:29–35

Menard C (1997) Internal characteristics of formal organizations. In: Menard C (ed) Transaction cost economics. Recent developments. Edward Elgar, Cheltenham

Menard C (2001) Methodological issues in New Institutional Economics. J Econ Method 8:85–92

Menard C (ed) (2004) The international library of the New Institutional Economics (7 vols) Edward Elgar, Cheltenham

Menard C, Shirley M (eds) (2005) Handbook of New Institutional Economics. Edward Elgar, Cheltenham

Mokyr J, Nye JVC (2007) Distribution coalitions, the industrial revolution, and the origins of economics growth in Britain. South Econ J 74:50–70

North DC (1986) The New Institutional Economics. J Inst Theor Econ 142:230–237

North DC (1988) Ideology and political–economic institutions. Cato J 8:15–28

North DC (1989) A transaction cost approach to the historical development of polities and economies. J Inst Theor Econ 145(4):661–668

North DC (1990) Institutions, institutional change, and economic performance. Cambridge University Press, Cambridge

North DC (1991a) Institutions. J Econ Perspect 5:97–112

North DC (1991b) Towards a theory of institutional change. Quart Rev Econ Bus 31:3–11

North DC (1993) Institutions and credible commitment. J Inst Theor Econ 149:11–23

North DC (1994) Economic performance through time. Am Econ Rev 84:3–17

North DC (1995) The New Institutional Economics and third world development. In: Harris J et al (eds) The New Institutional Economics and third world development. Routledge, London

North DC (2005a) Understanding the process of institutional change. Princeton University Press, Princeton

North DC (2005b) Institutions and the performance of economies over time. In: Menard C, Shirley M (eds) Handbook of New Institutional Economics. Edward Elgar, Cheltenham

Nye JV (1997) Thinking about the state: property rights, trade, and changing contractual arrangements in a world with coercion. In: Drobak JN, Nye JV (eds) The frontiers of the New Institutional Economics. Academic, San Diego, CA

Oates W (2005) Towars a second generation theory of fiscal federalism. Int Tax Public Finance 12:349–373

Ostrom E (2005) Understanding institutional diversity. Princeton University Press, Princeton, NJ

Ostrom E (2007) A diagnostic approach for going beyond panaceas. Proc Natl Acad Sci USA 104:3–15

Ostrom E (2009) Building trust to solve commons dilemmas: taking small steps to test an evolving theory of collective action. In: Levin S (Ed) Games, Groups, and the Global Good. Springer, New York

Ostrom E, Ahn TK (2008) The meaning of social capital and its link to collective action. In: Svendsen GT, Svendsen GL (eds) Handbook on social capital. Edward Elgar, Cheltenham

Ostrom E, Walker J (1991) Communication in a commons: cooperation without external enforcement. In: Palfrey TR (ed) Laboratory research in political economy. University of Michigan Press, Ann Arbor, MI

Schmid AA (2001) The institutional economics of Nobel prize winners. In: Biddle J et al (eds) Economics broadly considered. Routledge, London

Scott KE (1984) Corporate governance and the New Institutional Economics. J Inst Theor Econ 140:136–152

Toboso F (1995) Explaining the process of change taking place in legal rules and social norms: the cases of institutional economics and New Institutional Economics. Eur J Law Econ 2:63–84

Toboso F (2001) Institutional individualism and institutional change: the search for a middle way mode of explanation. Cam J Econ 25:765–783

Toboso F (2006) Old organizational issues from a new institutional economics perspective: some introductory remarks. Rev Anal Econ 21:3–11

Toboso F (2008) On institutional individualism as a middle way mode of explanation for approaching organizational issues. In: Batie S, Mercuro N (eds) Alternative institutional structures. Evolution and impact. Routledge, London

Toboso F, Arias XC (eds) (2006a) Organización de gobiernos y mercados. Análisis de casos desde la nueva economía institucional. PUV, Valencia (Edited jointly with Universidad de Vigo)

Toboso F, Arias XC (eds) (2006b) En el ámbito público la clave también está en las instituciones. In: Toboso F, Arias XC (eds) Organización de gobiernos y mercados. Análisis de casos desde la nueva economía institucional. PUV, Valencia (Edited jointly con Universidad de Vigo)

Toboso F, Compés R (2003) Nuevas tendencias analíticas en el ámbito de la Nueva Economía Institucional. La incorporación de los aspectos distributivos. Trim Econ 280 (Oct–Dec): 637–671

Tsebelis G (1990) Nested games. University of California Press, Berkeley, CA

Tsebelis G (2002) Veto players: how political institutions work. Princeton University Press, Princeton, NJ

Weingast BR, Marshall WJ (1988) The industrial organization of congress. Or why legislatures like firms are not organized as markets. J Polit Econ 96:132–163

Williamson OE (1984a) Perspectives on the Modern Corporation. Quart Rev Econ Bus, 24, Winter 1984, pp. 64–71

Williamson OE (1984b) The Economics of Governance: Framework and Implications. J Inst Theo Econ 140(1):195–223

Williamson OE (1985a) Reflections on the New Institutional Economics. J Inst Theor Econ 141: 187–195

Williamson OE (1985b) The economic institutions of capitalism. Firms, markets, and relational contracting. Free Press, New York

Williamson OE (1990) A comparison of alternative approaches to economic organization. J Inst Theor Econ 146:61–71

Williamson OE (1996a) Efficiency, power, authority and economic organization. In: Groenewegen J (ed) Transaction cost economics and beyond. Kluwer, Norwell, MA

Williamson OE (1996b) The politics and economics of redistribution and efficiency. In: Williamson OE (ed) The mechanisms of governance. Oxford University Press, Oxford

Williamson OE (1996c) The institutions and governance of economic development and reform. In: Williamson OE (ed) The mechanisms of governance. Oxford University Press, Oxford

Williamson OE (1997) Hierarchies, markets and power in the economy: an economic perspective. In: Menard CL (ed) Transaction cost economics. Recent developments. Edward Elgar, Cheltenham

Williamson OE (2000) The New Institutional Economics: taking stock, looking ahead. J Econ Lit 38:595–613

Williamson OE (2003) Examining economic organization through the lens of contract. Ind Corp Change 917–942

Williamson OE, Winter SG (1991) The nature of the firm: origins, evolution, and development. Oxford University Press, Oxford

Winiecki J (1994) Shaping the Institutional Infrastructure. Econ Inquiry 32:66–78

Winiecki J (1996) Why economic reforms fail in the Soviet System: a property rights-based approach. In: Alston LJ et al (eds) Empirical studies in institutional change. Cambridge University Press, Cambridge

Winiecki J (1998) Formal and Informal Rules in Post-Communist Transition. J Pub Finan Pub Choice 16(1):3–26

World Bank (1994) Institutional Change and the Public Sector in Transitional Economies. Washington

World Bank (1997) The State in a Changing World. Washington

World Bank (2002) Institutions for Markets. Washington

# Institutional Foundations, Committee System and Amateur Legislators in the Governance of the Spanish Congress: An Institutional Comparative Perspective (USA, Argentina, Spain)

**Gonzalo Caballero**

*Studies of the US Congress predominate while articles on legislatures in other countries are rare*

*(Jones et al. 2002, p. 657).*

## 1    Introduction

There is ample literature on the political economy on the US Congress, but the American case is "a rare outliner in the population of national legislatures". In fact, party-centered systems dominate most of the world's democracies, and the legislatures of these countries differ considerably from the US Congress (Jones et al. 2002). Understanding legislative performance in different institutional frameworks requires a new effort that studies the different structures of legislative organizations around the world.

It is true that in the last two decades there have been significant advances in various research programs that study the political economy of democratic systems other than the United States system. The contributions of Laver and Schofield (1998) on the politics of coalition in Europe, Schofield and Sened (2006) on elections and legislative politics (for the cases of Israel, Italy, Netherlands and Britain), and Schofield (2009) on several cases such as parliaments and elections in Russia, Turkey and Canada, are good examples of the advances made in our knowledge on different democracies. Nevertheless, in other research programs, a comparative deficit persists because most of the efforts have been focused on the US experience, and this is the case of the study of legislative organization

G. Caballero
Faculty of Economics, University of Vigo, Campus as Lagoas, 36310 Vigo, Spain
e-mail: gcaballero@uvigo.es

from the New Institutional Economics. Specifically, the young Spanish democracy has been poorly studied from a comparative institutional perspective (Field and Hamann 2009), and the Spanish Congress that emerged in Spain resulting in the Democratic Constitution of 1978 constitutes a good research lab for our interest on the governance of Congress in different institutional frameworks. Therefore, this paper will focus on the first chamber of the Spanish bicameral Parliament, because it is the chamber with the primary legislative responsibility (Uhr 2006).

In a relevant research program on legislative institutions and organizations, Weingast and Marshall (1988), Jones et al. (2002), McCubbins (2005), Caballero (2006a, b, 2007) and Spiller and Tommasi (2003, 2007) have studied different elements of the industrial organization of Congress, analyzing the relationship between institutions, incentives and transactions in the legislative organizations.

This paper comparatively studies the institutional foundations of the industrial organization of the Congress of Deputies in Spain. The legislative organization, the committee system and the role of individual deputies in the Spanish Congress will be analyzed from a transactional and institutional approach, and we will proceed in a comparative way. The two reference models for our comparison are the industrial organization of the American Congress and the Argentine Congress. In this respect, we formulate an institutional comparative analysis that explains the structure of governance of the Spanish Congress in comparison with the cases presented by Weingast and Marshall (1988) and Jones et al. (2002).

While the American Congress represents a prototype model of Congress in which congressmen have strong property rights that facilitate the legislative transaction (candidate-based electoral politics, powerful committees with individual property rights), and the Argentine Congress applies another model of legislative organization (party-based electoral politics, weak committees, power of regional political leaders), this paper presents the industrial organization of the Spanish Congress, which is characterized by party-based electoral politics, weak committees and the power of national leaders of each political party.

Legislative behavior and the organization of legislative institutions are affected by political and electoral rules. We are interested in distinguishing between "party-centered electoral rules" and "candidate-centered electoral rules", since it is key for the incentives of congressmen. Moreover, the institutional structure of committees is relevant for the structure of property rights of individual congressmen. Our hypothesis is that electoral rules and committee systems are two of the main institutional determinants of political property rights in legislative organization, and they determine the structure of governance of legislative organization.

Our starting point will be the paper by Weingast and Marshall (1988) on the industrial organization of the American Congress. There has been ample literature on the US Congress since this paper was produced, and several approaches and arguments on legislative organization in US have emerged in recent decades. For example, Krehbiel (1991) pointed out the relevance of the informational aspects of legislative organization, while Weingast and Marshall (1988) focused on distributive aspects. On the other hand, Owens (1997) explains the return of party government in the US House of Representatives after an era of committee government, and

he considers that a return to more autonomous, less party-dominated committee system would be unlikely. In this respect, Uslaner and Zittel (2006) survey some changes in legislative behavior in the United States Congress. Nevertheless, from our perspective, Weingast and Marshall (1988) adopted an approach that was specifically focused on political transactions, committee system and legislative organization in Congress. Even if their model of powerful committee system does not correctly represent the current legislative performance of US Congress, it represents an interesting archetype of legislative organization for comparative analysis. We use Weingast and Marshall's model as a category of industrial organization of Congress, which is useful for the comparative analysis with the Spanish case. We will name it "the traditional model of industrial organization" of the American Congress. Moreover, we will introduce the case of the Argentine Congress for our institutional comparative analysis (Jones et al. 2002).

This paper is a first step in a more ambitious research program that attempts to present a map of hybrid structures of legislative organization that exist in different democratic countries around the world. This paper has been written as an institutional comparative narrative, although in future more theoretical and empirical findings should push forward this research program on legislative organization from the New Institutional Economics. Section 2 presents the foundations of the New Institutional Economics and Transaction Cost Politics for the analysis of legislative organization. Section 3 shows the structure of industrial organization of the American Congress according to the traditional analysis of Weingast and Marshall (1988). Section 4 shows the organizational structure of the Argentine Congress according to Jones et al. (2002). Section 5 presents the institutional and electoral rules of the Spanish democracy after *The Franco era* and the performance of the Spanish Congress is analyzed. Section 6 analyses the industrial organization of the Spanish Congress. Section 7 formulates an institutional comparative analysis of the different models of Congress that have been previously analyzed. Section 8 presents the conclusions.

## 2 The New Institutional Economics, Transaction Cost Politics and Legislative Organization

The theoretical framework of the New Institutional Economics (NIE) combines the *coasean* notion of transaction costs with the *northian* notion of institutions, such that institutions are a medium for reducing transaction costs and obtaining greater efficiency in economic performance. On the one hand, Coase (1937) generated a micro-analytical approach of organizations which gave rise to "transaction cost economics" (Williamson 1975, 1985); while on the other hand, Coase (1960) generated a macro-analytical approach that studied the relationships between institutions and economic performance, as well as institutional change processes (North 1990a). The NIE incorporated both approaches, which are mutually inter-related, that is to say, the NIE studies institutions and how institutions interact with organizational structure within the economy (Menard and Shirley 2005).

In a world with zero transaction costs, the parties concerned would carry out all such transactions that would result in efficiency gains[1] (Coase 1960). However, contrary to this hypothetical world where negotiation cost nothing, economic markets are characterized by the presence of positive transaction costs, and therefore no transaction is carried out whenever such costs surpass the expected gains from such transaction. The level of transaction costs will depend on the characteristics of each specific transaction as well as on the nature of the institutional environment in which the transaction is being carried out. In this respect, every society will have its own "rules of the game", which will determine the cost of carrying out transactions (North 1990a). According to the Northian approach, institutions are human devised constraints that shape political, economic and social interaction. Institutions consist of formal rules, informal rules and enforcement mechanisms, and they provide the incentive structure of an economy (North 1990a, 1991).

Transaction Cost Politics (TCP) is a research program that has emerged as an application of the theoretical approach of the New Institutional Economics to political analysis (North 1990b; Dixit 1996; Caballero and Arias 2009; Spiller and Tommasi 2007). TCP sustains that political institutions matter, that they can be analyzed and that their effect is to economize transaction costs. TCP uses political transaction as the unit of analysis and explains the evolution of political relationships as transactions and contracts. It highlights the relevance of institutions in political markets characterized by incomplete political rights, imperfect enforcement of agreements, bounded rationality, imperfect information, subjective mental models on the part of the actors and high transaction costs. The institutional structure of polity or current regime acts as a set of rules that structures incentives, determines the volume of transaction costs and results in a bias to political output.[2]

A first approach to the theoretical bases of TCP is characterized by the following proposals: (1) the application of the transactional approach to the political field leads us to consider political interaction as a set of (implicit or explicit) contractual relations. In this respect, public policies are the outcome of transactions among policy-makers. (2) Institutions are the rules of the political game, and they determine the incentive structure of the agents, and therefore determine a high level of public policy *output*. (3) Organizational structures of governance are quite relevant when explaining the relations between institutions and outcomes. (4) Transaction costs tend to be greater in the political field than in the economic field and therefore the design of an efficient institutional structure becomes more complex in the

---

[1]Coase (1960) presents the world of zero transaction costs as the non-existent world that is analysed by the neoclassical economists. The mainstream in economics had forgotten that transaction costs exist. Coase (1999) tells that the world of zero transaction costs is the world of the modern economic theory; it is not the "coasean world".

[2]Transaction Cost Politics (TCP), besides considering the contract as an analysis unit, also studies the enforcement mechanism of contracts, compares the different *governance* structures and adopts the bounded rationality supposition (Epstein and O'Halloran 1999). In TCP, North (1990b) and Dixit (1996) are the two fundamental contributors who provided the theoretical bases for the program, while Weingast and Marshall (1988) is one of the most relevant precedents.

political world. (5) In recent times, we are witnessing the progressive vision of public policies as a result of a series of inter-temporal political transactions. (6) TCP provides a central role to the notion of credible commitment, which justifies the importance of *reputational capital* and the organizational formulae of the State (Caballero and Arias 2009).

In the NIE program, the analysis of political rules implies a "first order economizing" (Williamson 2000), and studies how institutions shape the structure of incentives of the political actors in political markets, acting as a bias on the making of public policy. In this way, a study of the role of institutions such as "transaction cost economizers" in political interaction is begun, and the structural analysis efforts of the new institutionalists on State governance have come about, which include matters such as decentralization, congress, the bureaucracy or agencies (Menard and Shirley 2005). In this respect, the political economy is reborn with power over the theoretical bases of the New Institutional Economy and the links between economic theory and political theory are strengthened (North 1999).

The study of organizational, transactional and institutional matters of the State includes the analysis of the legislative market. Traditionally, legislatures have been considered as the principle policy-making institutions in modern societies (Carey 2006). The NIE and TCP are interested in the governance of Congress. By embarking on the political economy of the legislative organization, parties and committees can appear as substitutes for organizing the functions of any Congress. While a "committee parliament" opts for a system of property rights that favors the transactions between individual members of congress in order to pass bills (Weingast and Marshall 1988), when political parties are the key to legislative organization, a hierarchy is established with a centralized leadership. In as much as the parties are able to control their deputy members via a hierarchy, there exists a differing mechanism to the property rights mechanism established by the committees in order to enforce the agreements. In this case, hierarchy substitutes market in legislative transactions.

The political and electoral institutions of each country are reflected in the corresponding legislative organization of the Congress. That is to say, the legislative structure reflects the institutional matrix of the country conforming to the "*mirroring principle*" (McCubbins 2005). Therefore when in a parliamentary system power is concentrated around a government's president who controls the executive and legislative branch as the leader of the majority party, the base is established for legislative markets in which the individual representative or parliament member may lack property rights with respect to the political agenda. Particularly in the electoral systems with complete, closed and blocked lists does this constitute a system of incentives which favors the political hierarchy, since the congressmen that seek the possibility of reelection are encouraged to follow their party's guidelines that have been establish by the party leadership, because it is this party leadership who decides who will be included in the electoral lists of the party in the subsequent elections. Likewise, insofar as the parliament cannot liberate itself from the control of the majority party whose leader is the president of the executive branch, the Congress seems to be an actor whose counterweight function is diminished.

## 3   The Traditional Model of Industrial Organization of the US Congress: Candidate-Centered Electoral Rules, Strong Committees and Professional Legislators in a Presidential System

In an outstanding contribution on the political economy of the parliamentary process, Weingast and Marshall (1988) analyzed the industrial organization of the American Congress. This paper assumes that the model of governance showed by Weingast and Marshall (1988) constitutes an archetype that will be used in our comparative analysis as the "traditional model of industrial organization of the US Congress", such as was explained in the introductory section.

This traditional model is made on three basic assumptions that can characterize the US legislative experience according to Weingast and Marshall (1988). Firstly, Congressmen represent the (politically responsive) interests located within their district, because their constituents are the principal of the agency relationship. Secondly, party leaders place no constraints on the behavior of other congressmen. Thirdly, proposed bills must command the support of a majority of the entire legislature.[3] Therefore, the congressmen need to make agreements with other congressmen to pass the projects that are interesting for the district from which they are elected. An explicit or implicit vote market exists.

To understand how this special system for votes exchange works, it is necessary to refer to the Legislative Committee System that characterizes the Congress in the US.[4] The rules of this governance mechanism are a substitute for an explicit market for votes.

Because institutional details matter, let us go a little bit further to examine the main rules characterizing this Committee System. First of all, it must be said that committees are composed of a number of seats or positions, being each position held by an individual legislator. Associated with each committee, there is a specific subset of policy issues over which it has jurisdiction. It is within each committee's jurisdiction to possess the monopoly right to propose alternatives to the status quo before the legislature. Committee proposals must of course command a majority of votes to become public policy.

Secondly, it must be emphasized that a property rights system already exists over committee seats called the "seniority system". Under this system any committee

---

[3]The literature on the US Congress included several approaches with different conclusions on the relevance of congressional parties (Shepsle and Weingast 1995). In fact, the debate continues until nowadays in contributions such as Cox and McCubbins (2005) or Krehbiel (2004).

[4]The paper by Weingast and Marshall (1988) continued the research tradition of Shepsle (1978) on the American committees. The study of committees in legislative organization in the US Congress has generated a broad literature from different approaches. This literature includes Shepsle and Weingast (1987), Krehbiel (1991), Cox and McCubbins (1993), Shepsle and Weingast (1995), Maltzman (1997), Baron (2000), Polsby and Schickler (2001), Beniers and Swank (2004), Krehbiel (2004) and Kim and Rothenberg (2008).

member holds his position as long as he chooses to remain on the committee (there is however only one condition: his re-election). Leadership positions within the committee are allocated by seniority, and rights to committee positions cannot be sold or traded to others.

Third, when by transfer, death or defeat there is a vacant seat on the committee, a bidding mechanism exists whereby the vacant seat is assigned. Legislators seek assignment to those committees that have the greatest marginal impact over their electoral fortunes. There are committees that are valued by all, and the higher the competition in a bid for seeking a seat in those committees, the smaller the chance of success. The congressmen that do not succeed in their application will be assigned to less important committees. In this way, the process of assignment operates as an auto-selection mechanism and committees are not representative of the preferences of all the members of Congress (they show extreme preferences).[5]

It is evident that if committees have agenda control over their own jurisdiction to propose a bill to Congress, they have veto power on the proposals from others. The restrictive access to the agenda constitutes a mechanism by which each committee can avoid declining the agreements ex-post.[6]

Under these rules, a legislator of committee A can cede his intention to influence the selection of jurisdiction of committee B. In return the members of committee B may waive their right so as not to influence the proposals of the jurisdiction of A. The "institutionalization of rights on the agenda control" substitutes the explicit market exchange mechanism. Legislators seek a seat on those committees which are more highly valued for them, instead of trading votes. Having a position in a committee is a type of property right mechanism that reduces transaction costs and favours independent negotiations among congressmen regardless of their party affiliation.

The agenda control that the committee members have implies that successful coalitions should include the members of the relevant committee, because their votes are necessary to allow the bill to be debated in Congress. Committees are, then, decentralized units for adopting decisions that are composed by those legislators

---

[5]In their paper, Weingast and Marshall (1988) explain how the committee assignments constitute a bidding mechanism. In this sense, "there are certain committees (e.g., Post Office) that no one wants. Those who fail to get one of their requested slots are generally put on one of these committees. Requesting the most valuable slots, therefore, increases the probability of ending up with Post Office... Which freshman will opt instead to request the more powerful committees?. Since this option involves a lottery between the most valuable committee and one worth virtually nothing, only those freshmen who value it most highly in comparison with the sure thing of getting on their policy committee will bid for it. This lottery implies that revealed preferences reflect true preferences... The pattern of committee assignments looks remarkably like an optimization process that maps members into those committees they value the most".

[6]The committee system provides substantial protection against opportunistic behaviour, thereby providing durability to policy bargains. Only the committee with jurisdiction can bring it to the floor for a vote. This control over the agenda within its jurisdiction implies that a committee has veto power over the proposal of others. In other words, the restricted access to the agenda serves as a mechanism to prevent ex-post reneging (Weingast and Marshall 1988).
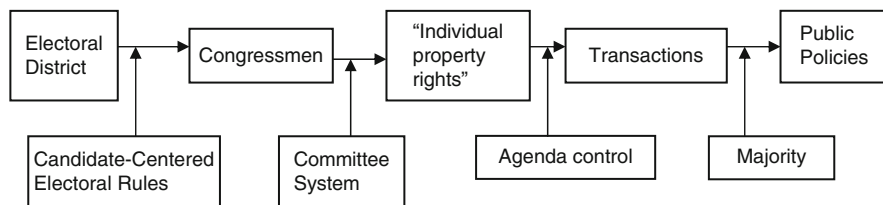
**Fig. 1** The industrial organization of Congress in USA

more interested in the jurisdiction of the committee. It is also evident that members of committees usually receive a non-proportional part of the benefits of the programs under their jurisdiction (Weingast and Marshall 1988). Committee members are in an agency relationship with the complete Congress. In Congress, of course, bills are passed by majority. Figure 1 summarizes the entire process.

In this respect, the US Congress plays an active role in policymaking and committees are key players in the legislative organization of Congress. Moreover, US congressmen exhibit remarkable longevity and tend to specialize in committees (Polsby 1968; Jones et al. 2002).

## 4 The Argentine Congress: Party-Centered Electoral Rules, Weak Committees and Amateur Legislators in a Presidential System

The industrial organization of the Argentine Congress constitutes an interesting case study for institutional comparative analysis of legislative organization. It represents a model of Congress whereby party-centered electoral rules and weak committee structure imply the existence of amateur legislators and the inability of the Argentine Congress to function as an effective check on the executive branch. Jones et al. (2002) and Spiller and Tommasi (2007) presents an institutional analysis of the Argentine Congress, and they conclude that it does not play an active role in Argentine policy-making. This section presents the main characteristics of the legislative organization of the Argentine Congress according to this literature.

Argentina is a federal republic with 23 provinces, and it has a presidential system of government and a bicameral legislature.[7] The Argentine Chamber of Deputies (Congress) has 257 members that are chosen for 4 years from closed party lists using proportional representation. Each district selects several congressmen via PR

---

[7]Gallo et al. (2006) presents an interesting review of the political-economic environment of Argentina since the late 1980s until 2005.

(multi-member districts) and half of the Chamber is renewed every 2 years[8] (Jones et al. 2002; Spiller and Tommasi 2007).

With proportional representation and closed lists, political parties play a relevant role. Argentina is a federal system and provincial governments are very powerful. In Argentina, provincial party bosses control the making of the provincial party list, and each individual legislator depends on the party list to be re-elected. The provincial political bosses determine which congressmen have the possibility to be re-elected, and the re-election of legislators are not in the hands of the voters, but rather in the hands of the provincial governor/party bosses (Jones et al. 2002). In this respect, the Argentine political game is centered on political parties (Fig. 2).

The assignment of each congressman to a specific committee is not determined by the constituency interests and electoral incentives; the committee Chairs are assigned by the bosses of each political party. The rules of Congress allow multiple assignment and Argentine congressmen belong to a multiplicity of committees. The composition of congressional committees reflects the proportion of seats held by the political parties in the plenary session, and the party's leadership distributes its committee assignments. Congressmen will serve on different committees in order to keep in good standing with the local party leadership (Jones et al. 2002).

Legislators have little incentive to develop legislative policy expertise since a professional legislative career is not a goal of legislators. Legislators only average one term in office. According to Jones et al. (2002), since 1983 the average reelection rate for the Argentine Congress has been 20%, but the members of the Chamber are politician with long political careers. As a result, being a legislator is only a stage in the political life of politicians in Argentina. They can be considered as "amateur legislators, professional politicians".

The consequence of party-centered electoral rules and closed lists imply that legislative behavior is affected by the power of party bosses rather than the preferences and interests of the constituency. According to the "mirroring principle", the internal organization of the Argentine Congress reflects the political and electoral institutions of the country. The result is an "amateur Congress", where legislators leave
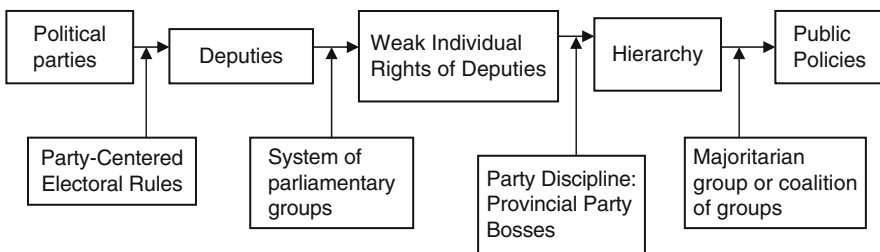


**Fig. 2** The industrial organization of Congress in Argentina

---

[8]On the other hand, the Argentine Senate has 72 members and each province (and the federal capital of Buenos Aires) elects three senators.

Congress early to continue in other political activities, and the Argentine Congress works more as a blunt veto player (Jones et al. 2002).

# 5  Political Institutions in the Spanish Democracy: Party-Centered Electoral Rules and Parliamentary Performance

After nearly 40 years of the Franco dictatorship in Spain, the 1978 Spanish Constitution established a democracy and a new political order in the country (Heywood 1998; Caballero 2008). The political reform of democratization implied a new set of political and electoral rules that affect the "first order economizing" of Williamson (2000). Spain today is a constitutional monarchy with a parliamentary structure (Field and Hamann 2009), and the principal formal institutional framework of the Spanish society was established in the 1978 Constitution.

The Spanish political system consists of a parliamentary model that has two Chambers: the Congress of Deputies (Congreso de los Diputados), which is the country's main Chamber, and the Senate (Senado).[9] Moreover, Spain experienced a process of political decentralization that created 17 new regional Parliaments since 1978, and an increasing number of legislative responsibilities have been attributed to them.[10]

This paper is focused on the Spanish Congress, which according to the "mirroring principle" reflects the political rules of the country. The parliamentary elections to Congress are governed by the following rules. Firstly, a proportional representation via the d'Hondt formula is applied in the electoral system, and candidates are presented in blocked and closed lists. Secondly, there are 350 deputy seats in Congress. Deputies are elected in 50 electoral provincial districts, while Ceuta and Melilla (the two Spanish Cities in North Africa) each have the right to elect one congressman. Thirdly, at least two deputies are assigned to each district (Ceuta and Melilla only have one each), and the distribution of the remaining seats is allocated via a population criteria. Fourth, there is a formal threshold of 3% of valid votes at the electoral district level for a party list to obtain representation (Montero 1998).

Table 1 shows the number of deputies of each party that were elected in Spain since the end of *the Franco era*. In these decades of democratic life in Spain, the national governments have alternated between single-party majority and minority, although there has never been a formal coalition government (Field and Hamann 2009).

---

[9]The Senate has 208 directly elected senators, and a variable number (about one-fifth of the chamber) indirectly selected by the assemblies of the 17 regional authorities (autonomous communities). Regarding to the Senate, in the direct elections a majoritarian electoral system is used.

[10]The process of political decentralization in Spain has constituted the model of the State of Autonomies, which implied an original institutional solution (Caballero 2005; Toboso and Scorsone 2010).

**Table 1** Political parties and number of elected deputies in Spain, 1977–2008

|          | 1977 | 1979 | 1982 | 1986 | 1989 | 1993 | 1996 | 2000 | 2004 | 2008 |
|----------|------|------|------|------|------|------|------|------|------|------|
| PCE/IU   | 20   | 23   | 3    | 7    | 17   | 18   | 21   | 8    | 5    | 2    |
| PSOE     | 118  | 121  | 202  | 184  | 175  | 159  | 141  | 125  | 164  | 169  |
| UCD      | 165  | 168  | 11   | –    | –    | –    | –    | –    | –    | –    |
| CDS      | –    | –    | 2    | 19   | 14   | –    | –    | –    | –    | –    |
| AP/CP/PP | 16   | 9    | 107  | 105  | 107  | 141  | 156  | 183  | 148  | 152  |
| CIU      | 11   | 8    | 12   | 18   | 18   | 17   | 16   | 15   | 10   | 10   |
| PNV      | 8    | 7    | 8    | 6    | 5    | 5    | 5    | 7    | 7    | 6    |
| ERC      | –    | –    | –    | –    | –    | 1    | 1    | 1    | 8    | 3    |
| Others   | 12   | 14   | 5    | 11   | 14   | 9    | 10   | 11   | 8    | 8    |
| Total    | 350  | 350  | 350  | 350  | 350  | 350  | 350  | 350  | 350  | 350  |

*Abbreviations: PCE/IU* Spanish Communist Party/United Left, *PSOE* Spanish Worker Socialist Party, *UCD* Democratic Center Union, *CDS* Social and Democratic Center, *AP/CP/PP* Popular Alliance/Popular Coalition/Popular Party, *CIU* Convergence and Union, *PNV* Basque Nacionalist Party, *ERC* Republican Left of Catalonia

The Spanish Constitution establishes that deputies within the Congress select the president of the government. The president afterwards appoints his government ministers, but the Congress does not approve the government that is appointed by the president. In this respect, the head of the majoritarian political party enjoys a considerable influence over both the executive and the legislative. This influence is, of course, higher when no coalition is needed to form a majority. In practice, the president of the executive branch is also the leader of the majoritarian party. Therefore, it is evident that the government can pass the desired bills without the presence of powerful "veto players" (Tsebelis 1995), particularly if an absolute majority has been obtained by a single political party.

In the parliamentary system, the majoritarian political party does not have the inherent checks in a system in which an effective separation of powers exists. Moreover, as closed and blocked lists exist in Spain, people vote for the name of the political party rather than for single candidates. In fact, the studies of the Centre for Sociological Research in Spain indicate that only 4% of the voters say that they vote primarily based on whom the candidates are that each party presents in their districts.[11]

Concerning the specific rules contained in the Congress internal Regulatory Statute, it must be said that the legislative task is organized through several internal organs and commissions. Firstly, the Management and Administrative Organs include the President of Congress, the Board of the Congress and the Spokesperson's

---

[11]In Spain, the small size of the Congress and the high number of electoral districts mean that the average size of a district is very reduced (6.73 seats by district), and 39 of the districts have seven or fewer seats. This is a very low number if proportional systems are considered. In fact, only Ireland has smaller districts in Western Europe than Spain. The rules here examined produce a majoritarian bias in the small districts, while in the broader districts the proportionality is adequately verified. In this way, the Spanish system of districts with few seats affects the system of parties in a way that reduces the number of parties that obtain parliamentary representation. It implies a low level of fragmentation.

**Table 2** Number of committees in Spanish Congress (I–VIII Legislature)

|                                    | I    | II  | III | IV  | V   | VI  | VII | VIII |
|------------------------------------|------|-----|-----|-----|-----|-----|-----|------|
| Legislative permanent committee    | 37   | 11  | 11  | 11  | 14  | 15  | 14  | 16   |
| Non-legislative permanent committee| 10   | 5   | 5   | 5   | 4   | 6   | 6   | 7    |
| Investigation committees           | 1[a] | 3   | 1   | 2   | 3   | 2   | 1   | 1    |

*Source*: own elaboration
[a]This was a mixed investigation committee

Meeting. Secondly, the Work Organs directly exercise the legislative function, and the Committees are the most relevant of these types of organs. They are institutionalized organs that handle the issues of its jurisdiction and are formed by small groups of deputies (approximately 40) who come from the different parties in a similar proportion to the party weight in the chamber.[12] Committees prepare the issues to be discussed later in the plenary session and, in some special and specific cases, they can act as a substitute for the plenary.

Committees work on the topics that later will be discussed at the plenary session of the Chamber. Committees can be permanent or non-permanent. Permanent Committees are cited in the internal Regulatory Statute of Congress, and they include legislative permanent committees and non-legislative permanent committees. Non-permanent Committees are ad-hoc organs, and they are created to address a particular issue, therefore they are abolished when their specific work is finished, and in any case when the legislature expires. Moreover, Congress can create investigation committees on a particular topic, mixed committees and sub-committees. Table 2 shows the number of Committees established in the Spanish Congress in the different Legislatures since 1979 (I–VIII Legislatures). In the current IX Legislature, which started in 2008, 19 legislative permanent committees and six non-legislative permanent committees were created.

On the other hand, parliamentary groups are sets of congressmen that are grouped to realize a collective action in the Congress according to their political affinity. No congressman can be a member of more than one parliamentary group. In practice, even when it is not a requirement established in the Regulatory Statute of Congress, each parliamentary group only incorporates the deputies that are affiliated to its particular political party. The exception is the mixed group. In the constituent legislature there were 9 parliamentary groups; 10 in the I Legislature; 6 in the II and in the III; 7 in the IV, V, VI and VII; 8 in the VIII and 6 in the IX Legislature. Parliamentary groups are in charge of implementing various initiatives such as the proposals for new bills, the totality amendment and the non-legislative propositions.

As it is obvious, the main function of the Congress, as a legislative chamber, is the passage of law. Title V of the Regulatory Statute presents the process of making and passing of law. The legislative initiative is presented in the Congress of deputies

---

[12]Committees can be permanent or non-permanent. In the first case, they necessarily have to be quoted in the Chamber Regulations, and they can have a legislative character or a non-legislative character. On the other hand, the non-permanent committees have an ad-hoc character and are created to carry out a particular task.

or in the Senate. The Spanish Constitution grants this initiative to the Government, the Congress, the Senate, the Regional Parliaments and to those citizens groups that are no less than half a million. In practice, the government is the player that is responsible for the passing of more bills (legislative bills), on the basis of its support from the majority of the chamber (Table 3).

On the other hand, the number of bill proposals submitted by the parliamentary groups or by other agents (such as the regional parliament, the Senate or the citizenship initiative) is quite high, but the number of them that are passed is very low. In Spain, the weight of the executive on the legislative process is verified by the recent experience shown in Table 3. The legislative initiative from the Government constitutes a clear difference with the US model in which committees can initiate the process by themselves.

Regarding the process by which bills are drafted and passed in the Congress, bill proposals go through a period in which they can be entirely or partially amended. When a total amendment is submitted, the amendment is debated and voted on in a plenary session: only when this amendment is rejected will the process continue. At this point, the proposal/project goes to the corresponding parliamentary committee, where a Reporting Sub-Committee is in charge of studying the amendments. Then, the plenary session of Congress debates and votes the legislative texts and various amendments.[13]

Of course, in addition to this legislative function, the Congress is in charge of overseeing the executive branch, for which the Regulatory Statute (Titles VI, VIII, IX, X and XI) includes various items such as the vote of no confidence, the trust motion, the "interpelaciones", the questions, the appearances, the non-legislative proposals, the motions and the resolutions.

## 6 The Industrial Organization of the Spanish Congress: Weak Committees and Amateur Legislators in a Parliamentary System

In determining the structure of the Spanish Congress, there are some political and electoral rules of the Spanish democracy that have a clear effect on the organization of Congress, and there are other just organizational norms of Congress that are relevant too. In this respect, Cox (2000) makes an important distinction between exogenous rules (those that cannot legally be changed by the legislature by itself) and endogenous rules (that can legally be changed by the legislature itself). In any case, in this section we are going to introduce the industrial organization of the Spanish

---

[13]Senate can pass "vetos" (totality amendment) or particular amendments, but later the text comes back to the Congress, which is definitively the decision-maker (it needs a qualified majority and some procedural conditions to pass a project that has been rejected in Senate).

**Table 3** Number of submitted and passed bills, Spain

| Bills | | I Legislature | II Legislature | III Legislature | IV Legislature | V Legislature | VI Legislature | VII Legislature | VIII Legislature |
|---|---|---|---|---|---|---|---|---|---|
| Legislative bill from the executive | Submitted bills | 347 | 209 | 125 | 137 | 130 | 192 | 175 | 152 |
| | Passed bills | 213 | 187 | 108 | 109 | 112 | 172 | 173 | 140 |
| Bill proposals from the parliamentary groups | Submitted bills | 193 | 108 | 139 | 165 | 140 | 300 | 322 | 235 |
| | Passed bills | 26 | 14 | 9 | 18 | 17 | 28 | 16 | 18 |
| All other bill proposals (Senate, CC.AA., citizenship initiative) | Submitted bills | 23 | 23 | 23 | 42 | 44 | 50 | 60 | 82 |
| | Passed bills | 10 | 6 | 4 | 8 | 18 | 20 | 3 | 9 |

*Source*: Own elaboration based on Caballero (2007)
CC.AA. are the autonomous communities (17 regional political units)

Congress following the structure of analysis of Weingast and Marshall (1988) for the US case.

The perspective on the Spanish Congress developed in this section rests on the three following assumptions:

A. Deputies represent the interests of their political parties. Each congressman is immersed in an agency relationship with multiple principals (Dixit 1996). The most important principal of each congressman in this relationship is the head of his political party at the national level. The head of the political party is who directly or indirectly determines the possibility of re-election of each deputy since the electoral system is based on closed and blocked lists. This system reduces the role of any deputy as an independent defender of the interests of his district. For this reason, the interest groups consider that the capture of an individual deputy has no great interest, since his freedom is very limited by party discipline and the necessity to cooperate with his fellow party members. The interest groups will try to capture or influence the leaders of the political party and the head of the parliamentary group.

B. Parties place constraints on the behaviour of individual representatives. Political party leaders have great power and try to restrict the behaviour of the remaining deputies in several ways. This implies that the individual ability of free decision-making is very limited for those individual deputies. Relevant decision-making corresponds more to the choices preferred by the collective heads of political parties than to individual preference, and in the case of conflict, by ordinary deputies.

C. Majority rule is a binding constraint. If passing a bill in the Congress requires the support of the majority of congressmen (simple, absolute or qualified majority in the various cases), agreements among the deputies of the same province or region will not be enough if all others do not support the proposal, particularly the head of the majoritarian parliamentary group. Therefore, negotiations aimed at passing a bill will have to be made within the majoritarian group via a set of relationships in which transactions between equals do not exist due to the fact that there are several agreed upon hierarchical rules. In fact, those congressmen with a relevant position in the structure of the party organization often enjoy a higher power to negotiate and establish the priorities of the majoritarian political party than those who do not occupy such a position.

In conclusion, a hierarchical relationship exists in which deputies usually delegate the decision-making process toward the head of the parliamentary group. This is why ordinary congressmen relinquish the exchange transactions in which the head of the group is not present. In fact, the hierarchical system and the internal discipline of each parliamentary group make independent dialogue and agreement between individual deputies from different political parties almost impossible.

Transaction costs are really high in legislative markets due to non-contemporaneous benefit flows and non-simultaneous exchanges. Political hierarchy in the parliamentary group is the way to solve the transacting problem in the Spanish legislative market. Agreements are channelled through the collective heads of

hierarchical parliamentary groups. In fact, the empowerment of parliamentary groups in the Regulatory Statute of Congress was against the position of individual deputies, and the number of decisive subjects in Congress was reduced to a handful of people.

The industrial organization of the Spanish Congress implies a system of Legislative Committees that are characterized by the three following aspects:

1. Committees are composed of a number of seats occupied by some deputies and each committee is associated with the jurisdiction on a subset of policy issues. Although the committees have no competence to initiate legislation, the legislative bills are discussed and amended within the committees. The committee proposals on bills and projects must be discussed and voted on later by the plenary session of Congress (except when the special procedure of legislative competence is applied; in this case, committee proposals go directly to the Senate).
2. The distribution system of the seats of committees among the parliamentary groups is by apportionment, that is to say, the proportion of seats of the plenary session is maintained in each committee. Moreover, each group can freely appoint deputies to the seats that correspond to the group and decides which deputy will be the group leader in each committee. Groups have property rights on the committee seats. This implies that each group freely assigns the seats to its deputies, and the collective head of the group can change the assignment of deputies. Each parliamentary group cannot trade committee positions with other parliamentary groups.
3. When there are vacant seats in a committee (by resignation, death or new election), the parliamentary group members choose who will be assigned. Each group tries to maximize its performance in the parliament, assigning its deputies in a way which is coherent with its maximization. The collective head of the parliamentary group coordinates the affiliation of each deputy to the different committees, and can adjust this allocation whenever it is considered necessary for improving the parliamentary group's performance. In fact, changes in the allocation are common and it is even possible that the parliamentary group substitutes a member of a committee particularly for one subject, debate or session.

Committee members vote in a way that is coherent with the decisions of their parliamentary groups. In this way, they have a narrow margin of discretion and follow the rigid voting discipline established by the collective head of the group (Sánchez de Dios 1999). The discretional choice of any individual deputy who does not form part of the collective head is directly proportional to his power in the parliamentary group, and inversely proportional to the interest in the subject of the head of the parliamentary group (Fig. 3).

The presence parties have in committees is proportional to their political representation in Congress. This implies that the majority formed in the chamber is repeated in all the committees. When there is an absolute majority in the chamber, the majoritarian political party controls all the committees too.

For this reason, committees do not have a "separation of purpose" from the plenary session in the sense of Cox and McCubbins (1999), that is to say, the committee's
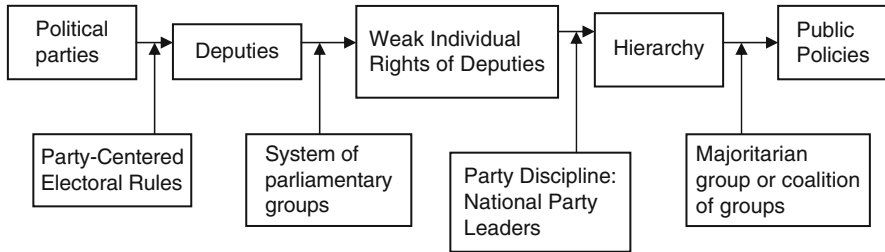
**Fig. 3** The industrial organization of Congress in Spain

**Table 4** Number and duration of the sessions of committees

|  | II Legislature | III Legislature | IV Legislature | V Legislature | VI Legislature | VII Legislature |
|---|---|---|---|---|---|---|
| Number of sessions | 564 | 645 | 866 | 874 | 1,082 | 1,123 |
| Duration (in hours) | 2,158 | 2,322 | 2,823 | 3,097 | 3,584 | 3,760 |

*Source*: Spanish Congress of deputies

preferences are the same as the parliamentary preferences. Because the same preferences control the plenary session and the committees, committees are not independent of the plenary guardianship and the parliamentary groups act as the power mechanisms that impose those preferences. As a result, committees are not independent as "non majoritarian institutions" as in Majone (2001) (a level of autonomy that other agencies have, such as the Central Bank or the European Commission). This makes the committees weak regarding professionalism, independence, specialization and the assignment of property rights. On the other hand, committees have higher quotas of democratic representation. In spite of the growing number of sessions of committees and the growing duration of these sessions (Table 4), the organizational structure of Congress has continued to weaken the possible relevance of committees.[14]

There is little doubt that through the hierarchical structures of the political parties, party leaders in the executive also have a quasi-monopolistic control of parliamentary life via the majoritarian parliamentary groups. The political party system in Spain dominates parliamentary life in Congress, and parliamentary groups turn into sub-units of the political party organization.

The political rules of the Spanish system and the industrial organization of the Congress establish a system of incentives that involves a series of biases on the profile

---

[14]Moreover, other causes of the weakness of the committees can be pointed out: the small number of workers that they have, their scarcity of resources and the scarce specialization of their members.

of the Spanish deputies. According to our analysis: (1) The character of the deputy corresponds more to that of a "party politician" than that of a "district representative". (2) A higher rate of parliamentary turnover is encouraged since permanence in office does not constitute a stimulus to guarantee a position or some sort of institutionalized leadership as in the case of the United States committees. (3) The rate of turnover of the Spanish deputies makes the formation of a specialized and professional legislative body difficult. (4) One can arrive at the hypothesis of "political professionals, legislative amateurs" for the Spanish case, according to Jones et al. (2000). We will analyze the Spanish deputies in order to test this hypothesis.

The Spanish deputies join the Congress of Deputies with high political fanfare, having previously dedicated themselves to politics as their primary activity. In fact, 80% of the deputies elected in 1996 were already dedicated to politics before entering and forming part of the parliament (Uriarte 2000). Among those that were already dedicated to politics as professionals, 25% were dedicated politicians before 1982 and another 25% started between 1982 and 1985, as Table 5 illustrates.

In any case, it should be noted that the deputies that joined Congress have already been affiliated previously with their political party for a long time period. The percentages in Table 6 illustrate the long political careers of the Spanish deputies.

In the Spanish case, individual deputies have no property rights on legislative Committees, and seniority does not attribute neither seats nor positions of leadership in Congress. The permanence in the Chamber does not concede more individual rights to congressmen, and it is not a valued asset for legislative-making. In fact, the rate of non-reelection in Congress has been very high in the recent democratic experience. Table 7 shows the rate of removal in the Spanish Congress; with the

**Table 5** Incorporation year to politics as main activity of the Spanish deputies elected in 1996

| | |
|---|---|
| Before 1977 | 6% |
| 1977–1978 | 9% |
| 1979–1981 | 11% |
| 1982–1985 | 24% |
| 1986–1988 | 10% |
| 1989–1992 | 20% |
| 1993–1995 | 21% |

*Source*: Uriarte (2000)

**Table 6** Duration of party affiliation of elected deputies in Spain

| | |
|---|---|
| More than 20 years | 28% |
| Between 10 and 20 years | 42% |
| Between 5 and 10 years | 16% |
| Between 2 and 5 years | 9% |
| Less than 2 years | 2% |
| No answer or non-affiliated | 3% |

*Source*: Uriarte (2000) for the VI legislature

**Table 7** Rate of removal of the Spanish congressmen

| 1979 | 1982 | 1986 | 1989 | 1993 | 1996 | 2000 | 2004 | 2008 |
|------|------|------|------|------|------|------|------|------|
| 47.2% | 63.6% | 19.1% | 48.9% | 35.5% | 46.7% | 56% | 54.8% | 38% |

*Sources*: Guerrero (2004), Caballero (2007) and author's calculations

**Table 8** Parliamentary life of deputies in Spain

| Number of periods of legislature | Percentage of deputies |
|------|------|
| 1 | 52.1 |
| 2 | 25.6 |
| 3 | 11 |
| 4 | 6.6 |
| 5 | 2.2 |
| 6 | 1.9 |

*Source*: Morán (1996)

exceptions of 1986, 1993 and 2008 more than 45% of the deputies were removed in each election.[15]

This high level of parliamentary turnover is reflected in the legislative terms that a deputy remains in Congress. Table 8 collects this information for the deputies elected between 1993 and 1997. More than half the deputies of this period were elected for only one legislature which confirms the brief parliamentary experience of most of the Spanish deputies (Morán 1996). In fact, nowadays there is only one deputy that is a member of Congress since the first democratic elections.

This trend of short legislative careers is understandable in an institutional framework where long-term agreements are not a concern for individual deputies, since it is the collective head of the parliamentary group who maintains the permanence criteria and lengthens the time horizon of parliamentary activity. The collective head of the parliamentary group specializes in legislative matters, and verifies that the high rate of deputy removal prevents the individual deputies from developing their legislative career until they attain a high enough level to compete with the group head.

In conclusion, the Spanish experience demonstrates two traits that characterize the Spanish deputies: on the one hand, a long career of political activity (reflected in their long political lives as party militants and even by their long dedication to politics as the main activity), and on the other hand, a high rate of turnover of deputies in Spain (in such a way that a strong majority of deputies pass through Congress on a "quasi-occasional" basis). This is why the Spanish deputies can be generally characterized as "professional politicians, amateur legislators" (Jones et al. 2000). In fact, in the Spanish case there exists only a minority of deputies that become professionalized in the parliament, heading their respective parliamentary groups and exercising legislative tasks.

---

[15]This is especially noticeable if we take into account that the immense majority of the deputies say that they want to continue as members of parliament. According to Uriarte (2000), 85% of deputies answered it in the VI Legislature.

# 7 An Institutional Comparative Analysis of Congressional Organization: US, Spain, Argentina

## 7.1 Models of Governance of Congress: Spain Versus USA

The structure of the industrial organization of Congress is varied in Spain and in the US. They have different models that reflect the institutional framework of each country, and specifically the type of electoral rules (Caballero 2006a,b, 2007). Figure 4 summarizes the main differences of the institutional comparative analysis between both cases.

A. The Party Deputy in Spain versus the District Congressman in the US. The Spanish electoral system with closed and blocked lists converts the head of the political party into the "principal" in a relationship of agency of the deputy, and assuming that this deputy seeks reelection, he is motivated to follow the instructions of the head of the political party. The voters choose between political parties and the party decides who will form part of its electoral lists. On the contrary, in the case of the United States, the electoral system converts the individual candidate into the key component for the voter, and the members of congress are encouraged to address the interests of their voters of their electoral district, because these voters determine the possibilities of reelection.

B. Party Discipline in Spain versus Congressmen that are not controlled by political parties in USA. The objective of permanency in the post by the deputies makes it possible for the head of the political parties to be able to impose behavioral discipline among their deputies in Spain, while in the United States, the parties do not have this control mechanism over its members.

C. The majoritarian parliamentary group dominates Committees in Spain versus the Seniority System in US Committees. In the Spanish model, the assignment of committee seats between groups is carried out according to a proportionality criteria with respect to the distribution of seats in the plenary in such a way that

| Spanish Congress of Deputies | American Congress |
|---|---|
| Deputies represent their political party | Congressmen represent districts |
| Internal discipline in parties | Parties do not control congressmen |
| Majoritarian group dominates committees | Seniority system in committees |
| Deputies have no individual rights | Congressmen have property rights |
| Parliamentary groups make decisions | Committees are key |
| Hierarchy with a leader | Legislative transactions via committees |
| Parliamentary renovation | Long duration of congressmen |
| Group and Party Parliament | Committee Parliament |

**Fig. 4** The industrial organization of Congress: Spain versus USA
*Source*: Own elaboration

the head of the majority group (or coalition of groups that are held together by agreements with the government) maintains control over the plenary where it has the majority, over each one of the committees (which represents the distribution of power in the chamber on a smaller scale) and over the deputies of their party which are subject to party discipline. On the contrary, the United States system establishes an assignment mechanism to committees that awards previous permanence to this committee as well as seniority, in such a way that the composition of the committees does not necessarily reflect the existing political majority in the chamber.

D. Deputies without political property rights in Spanish Committees versus Congressmen with property rights in US Committees. The industrial organizational model of the Spanish Congress does not grant property rights to the deputies for their committee seats: the head of each parliamentary group decides which deputies are assigned to each committee, being able to change them when it so desires. The seniority system in practice in the House of Representatives in the United States is turned into a mechanism that grants property rights or permanence for the seat as well as the leadership position in the committee. In the United States model, there are individual rights which the political party cannot interfere with.

E. Prominence of parliamentary groups in Spain versus the relevance of Committees in the USA. The organizational structure of the Spanish parliamentary process places the prominent role of the chamber on the head of the parliamentary groups, which grants a privileged position of control over the plenary as well as to each one of the committees to the majority group. The Spanish Congress is more of a "parliament of groups", while the industrial organization of the United States Congress places the prominent role on committees that have the exclusive on proposal ability to the plenary and the political position of the committees is the fruit of the will of their members (which does not necessarily coincide with the will of the house majority).

Membership in committees is of lesser motivating value for the deputies in Spain, while in the United States, it constitutes a motivating factor of great importance. This results because the committees in Spain are not able to establish a system of property rights as in the United States, and this is why they are not able to determine the legislative transactions. In the United States, each congressman specializes in one committee, while in Spain, the deputies are assigned to various committees. In Spain, membership in various committees has little electoral cost for a deputy, and the greater number of committees he belongs to, the lesser is his degree of specialization.

F. Hierarchy System in Spanish Congress versus Legislative transactions in American Committees. The Spanish model of Congressional organization reflects a political system with great power concentrated around the figure of the president of the government who, as the leader of the majority party, tries to control the parliamentary majority. In this way, legislative transactions and agreements are carried out via a hierarchical system. As long as the executive and the majority of the legislature represent the same political preferences, the role of the Congress

is clearly reduced. On the other hand, the system of property rights regarding the US committees reduces the high transaction costs of legislative exchange, being that the United States Congress establishes a system of committees that allow transactions between congressmen in order to establish majorities that permit changing the *status quo*. In this case, there is no relationship of subordination of committees with respect to the United States President.

G. Parliamentary Turnover in Spain versus Long-Duration Congressmen in the USA. The industrial organization model of the Spanish Congress does not convert seniority in the Chamber into an asset that grants permanency rights or leadership in committees, furthermore the individual members of parliament are subject to the group leadership, resulting in a system that does not encourage permanency in parliament for long periods of time. In fact, the Spanish deputies actually place more value on the access to other political posts within the party, in the government or other administrations. In addition, the bosses of the parliamentary groups have normally caused a high turnover of positions in the Committees and spokespersons of the groups in the Spanish committees. Thus, after each electoral process, the rate of turnover of some Committees (such as Economy, Foreign Affairs, Defense, Justice, Interior, or Budgets) surpassed two thirds, and sometimes the renewal reaches 100% (Guerrero 2004). This high turnover is also evident in the composition of the permanent legislative committees, where the continuity of members of parliament is reduced, as illustrated in Table 9. In the case of the United States, it is very different since the congress members obtain leadership positions via a seniority mechanism. If we add to this the fact that in this case the members of congress have some individual property rights, we can understand the incentives of the members of congress to remain there, and even for the voters to back the members of congress with long congressional careers. The House of Representatives in the United States reached high levels of institutionalization as permanence in the House became

**Table 9** Number and percentage of Spanish deputies that remained in the permanent legislative committees in the VI legislature

| Committee | Number of deputies repeating in committee | Percentage |
|---|---|---|
| Economic affairs | 18 | 44 |
| Public finance | 15 | 37 |
| Agriculture | 11 | 27 |
| Health | 8 | 20 |
| Industry | 18 | 44 |
| Infrastructures | 12 | 22 |
| Social welfare | 16 | 39 |
| Foreign affairs | 21 | 51 |
| Defence | 13 | 32 |
| Law and security | 17 | 41 |
| Constitutional affairs | 10 | 24 |
| Education and culture | 8 | 20 |
| Public administration | 8 | 20 |

*Source*: López Nieto (2001)

more and more attractive and turnover of House members became less and less frequent (Jones et al. 2000).

H. Parliamentary Group and Party Parliament in Spain versus Committee Congress in the US. While the organization of the Congress of deputies in Spain converts the parliamentary groups (integrated in a party hierarchy) into the key elements for the making of public policies in the legislature, the United States Congress concedes the leadership role to the committee system. The Spanish model does not grant "de facto" property rights to the individual deputies and greatly leaves it to the discretion of the parliamentary group leadership in order to determine its organizational structure. In this way the party hierarchy is able to impose its will, with the leader and the rest of the party leadership at the head. Specifically in a system such as the Spanish system, in which the majority party controls the executive branch and the legislature (particularly when there is an absolute majority), the institutional structure of the Congress concedes all the weight of the Chamber performance to the majority parliamentary group according to the limitations established by the Constitution and the Congressional Regulations. Therefore, the majority party group leadership controls the majority of the chamber and decides how to organize the internal workings of the group, such as the role of the committees and the plenary sessions. This way the majority group can determine when decisions are made, who is granted power and what de facto functions and operations each Congressional organ will have. The parliamentary group is converted into the key element of the Spanish Congress, and its true organizational structure depends upon its relationship with the executive branch and upon the corresponding political party. In the case of an absolute majority, the leadership of the majority parliamentary group decides in practice if the legislative committees have a role to fill or not. The fact that the majority is able to impose its dominance without hindrance by institutional mechanisms such as in the United States legislative committees, postulates that the role of the Spanish legislative committees is not very relevant since they do not enjoy property rights over the policy decision agenda. The political property rights fall upon the parliament majority group that is not confronted with possible vetoes neither in the plenary nor in the committees (excluding internal cohesion problems).

## 7.2   A Comparative Perspective with the Argentine Congress

Jones et al. (2002) clearly shows the main conclusions of a comparative institutional analysis between the American House of Representatives and the Argentina Congress:

– While members of the US Congress exhibit remarkable longevity, the Argentina deputies last on average only one 4-year term.
– While members of the US Congress tend to specialize in committees, Argentine legislators belong to a multiplicity of committees.

– While the US Congress has an active role in policy-making, the Argentina Congress plays more the role of a blunt veto player.
– While the US Congress is the center of policy-making, in Argentina crucial political bargaining is held far away from the national legislature.

On the other hand, while Spain has a parliamentary political system and Argentina has a presidential system, both coincide in having an electoral system with closed and blocked lists where the D'Hondt rule applies and a Congress in which the individual deputies do not have property rights in the parliamentary committees. That is to say, both cases share two important elements of the institutional framework: party-centered electoral rules and weak committee systems. The results are the existence of amateur legislators with a high rate of turnover and a party Congress with a limited role in policymaking.

In particular, the Spanish and Argentine Congress coincide, among other traits, in the following: (A) The lack of power and individual rights of the deputies; (B) the political parties control the reelection possibilities of the deputies; (C) the political parties control the Congress through the parliamentary groups; (D) There is a lack of specialization in the parliamentary committees; (E) The members of parliament belong to various committees; (F) The deputies acts more as "party politicians" than as "district deputies". (G) The deputies have on average a short career as members of parliament; (H) The Congress has limited power as a "veto player".

Nevertheless, in spite of the fact that Spain and Argentina have a similar Congressional structure and electoral system, one difference between both cases is that in the Argentine case, the provincial political leaders (especially the governors) are the ones who have the power to elaborate the electoral lists. While in the Spanish case, this power is wielded by the upper echelons of the party leadership at the national level (especially in the case of the governing party). Jones et al. (2002) points out that Argentina is a federal system and this may appear to be the explicative factor of power of the provincial leaders when elaborating the electoral lists. However, Spain has also experienced a process of political decentralization of a federal nature which has created 17 regional governments, each one with its own regional parliament and Autonomous Community president, although on the contrary, the power of elaborating the electoral lists for the Spanish Congress is not in the hands of the regional leaders, but in the hands of the national leaders. The key is that Spain has a parliamentary political regime and Argentina has a presidential model. This implies that in the Spanish case the chief executive is elected by the assembly and the chief executive remains in office subject to legislative confidence, but this too implies that elections to Congress will be focused on the candidate to the presidency of the government. On the other hand, Argentina is a presidential system that Carey (2005) characterizes as a hybrid regime where the president is popularly elected and is endowed with meaningful powers. In this case, legislative elections are independent of presidential elections.

In a parliamentary system such as Spain's in which the Congressional elections are focused on the figure of the candidate to the presidency of the government, control over the lists are more centralized at the national level than in the Argentine model, in which the presidential system implies that the legislative elections have a greater provincial district component and the reference to the national leadership is less.

# 8 Conclusion

Institutional Comparative Analysis on property rights, transactions and hierarchies in the governance structure of Congress is relevant for the understanding of the role of individual deputies and parliamentary committees in the legislative performance. This structure has different models of organization and performance, as we have shown by comparing the cases of the Spanish Congress, the House of Representatives in the United States and the Argentine Congress.

The Spanish Congress is composed chiefly of organized hierarchical parliamentary groups than that of deputies understood as individual political agents. It is in the "deputy leaders" where all the power is concentrated by their guidance of the parliamentary groups. Transactions are carried out via the delegations in the hierarchical structure of the parliamentary group. The Spanish political system can be characterized as a State of parties in which the political elites control the hierarchical political parties. The weight of the parties in public life ends up being stretched towards a parliamentary configuration and the political parties penetrate the organization of Congress through a hierarchical structure over the parliamentary group. As a result, the parliament ends up as a parliament of groups.

Contrarily, the traditional model of industrial organization of the United States Congress implies a parliament with property rights for congressmen over the political agenda, with a system of powerful committees, long-term members of congress and with an influential structure in the making of public policy.

The incorporation of the Argentine model to the comparative analysis allows us to study a structure of governance which is more similar to that of Spain than that of the United States, in spite of the fact that Argentina and the United States both have a presidential regime, while Spain has a parliamentary political system. It is this way because the Argentine electoral system is proportional to the D'Hondt rule; it establishes closed and blocked electoral lists and the parties control both parliamentary life and committees. Both traits are shared by both the Spanish and Argentine Congresses but not by the United States. Nevertheless, the fact that the Spanish system is parliamentary with party-centered electoral rules implies that the power over the electoral lists lies with the national leadership of the various parties, while with a presidential system with province-districts as in Argentina, the provincial party leaders and the governors are who have the most power to elaborate the electoral lists and decide which deputies will have the possibility to be reelected.

In summary, this chapter has applied an institutional political economy approach to the comparative study of parliamentary chamber organization. Stemming from the theoretical foundations of the New Institutional Economics and Transaction Cost Politics, this chapter has brought about advances in the knowledge of the legislative organization of the Spanish Congress from a comparative perspective. This paper advances the analysis of the Spanish case which had been poorly studied in comparative terms. In addition, this paper has advanced upon a line of research on comparative institutional analysis of parliamentary governance, although future efforts shall be necessary in order to obtain new developments. The agenda of future research

should incorporate theoretical developments, new empirical evidence and the incorporation of other existing Congressional models around the world. In this way we will advance in our understanding of the political economy of democratic institutions.

# References

Baron DP (2000) Legislative organization with informational committees. Am J Polit Sci 44(3):485–505

Beniers KJ, Swank OH (2004) On the composition of committees. J Law Econ Org 20(2):353–378

Caballero G (2005) Instituciones, federalismo defensor de mercados y Estado de las Autonomías. Un análisis de segunda generación. El Trimestre Económico 286:283–328

Caballero G (2006a) La economía política de la organización industrial del Congreso de los Diputados en España: Derechos de propiedad, transacciones y jerarquías. El Trimestre Económico LXXIII(291):637–666

Caballero G (2006b) The industrial organization of Congress in USA and Spain: a comparative institutional analysis. Revista de Análisis Económico 21(2):105–123

Caballero G (2007) Comisiones, grupos parlamentarios y diputados en la gobernanza del Congreso de los Diputados. Revista de Estudios Políticos 135:67–107

Caballero G (2008) El cambio institucional de la economía del franquismo a la democracia: un análisis histórico institucional. Política y Gobierno XV(2):353–401

Caballero G, Arias XC (2009) The program of transaction cost politics in the map of the new institutionalism. In: Annual meeting of the public choice society, USA

Carey JM (2005) Presidential versus parliamentary government. In: Ménard C, Shirley MM (eds) Handbook of New Institutional Economics. Springer, Dordrecht, pp 91–121

Carey JM (2006) Legislative organization. In: Rhodes RAW, Binder SA, Rockman BA (eds) Oxford handbook of political science. Oxford University Press, Oxford, pp 431–454

Coase RH (1937) The nature of the firm. Economica 4:386–405

Coase RH (1960) The problem of social cost. J Law Econ 3(1):1–44

Coase RH (1999) An interview with Ronald Coase. ISNIE Newslett 2(1):3–10

Cox GW (2000) On the effects of legislatives Rules. Legis Stud Quart 25(2):169–192

Cox GW, McCubbins MD (1993) Legislative leviathan. University of California Press, Berkeley, CA

Cox GW, McCubbins MD (1999) The institutional determinants of economic policy outcomes. In: Annual conference of the ISNIE, Washington, USA

Cox GW, McCubbins MD (2005) Setting the agenda. Cambridge University Press, Cambridge

Dixit AK (1996) The making of economic policy: a transaction-cost politics perspective. The MIT Press, Cambridge

Epstein D, O'Halloran S (1999) Delegating powers. A transaction cost politics approach to policy making under separate powers. Cambridge University Press, Cambridge

Field BN, Hamann K (2009) Democracy and institutional development. Palgrave Macmillan, UK

Gallo A, Stegmann JP, Steagall J (2006) The role of political institutions in the resolution of economic crises: the case of Argentina 2001–2005. Oxford Dev Stud 34(2):193–217

Guerrero E (2004) El Parlamento, Editorial Síntesis

Heywood P (1998) Power diffusion or concentration? In search of the Spanish policy process. West Eur Polit 21(4):103–123

Jones MP, Saiegh S, Spiller PT, Tommasi M (2002) Amateur legislators-professional politicians: the consequences of party-centered electoral rules in a federal system. Am J Polit Sci 46(3):656–669

Kim J, Rothenberg LS (2008) Foundations of legislative organization and committee Influence. J Theor Polit 20(3):339–374

Krehbiel K (1991) Information and legislative organization. University of Michigan Press, Ann Arbor, MI

Krehbiel K (2004) Legislative organization. J Econ Perspect 18(1):113–128

Laver M, Schofield N (1998) Multiparty government. The politics of coalition in Europe. The University of Michigan Press, An Arbor, MI

López Nieto L (2001) Las Cortes Generales. In: Alcántara M, Martínez A (eds) Política y Gobierno en España. Tirant lo Blanch, pp 215–242

Majone G (2001) Nonmajoritarian institutions and the limits of democratic governance: a political transaction-cost approach. J Inst Theor Econ 157:57–78

Maltzman F (1997) Competing principals: committees, Parties and the Organization of Congress. University of Michigan Press, Ann Arbor, MI

McCubbins MD (2005) Legislative process and the mirroring principle. In: Ménard C, Shirley MM (eds) Handbook of New Institutional Economics. Springer, Dordrecht, pp 123–147

Menard C, Shirley M (2005) Handbook of New Institutional Economics. Springer, Dordrecht

Montero JR (1998) Stablishing the democratic order: electoral behaviour in Spain. West Eur Polit 21(4):53–79

Morán LM (1996) Renewal and permanency of the Spanish members of parliament, 1977–1993, WP 1996/81. Instituto Juan March, Madrid

North DC (1990a) Institutions, institutional change and economic performance. Cambridge University Press, Cambridge

North DC (1990b) A transaction cost theory of politics. J Theor Polit 2(4):355–367

North DC (1991) Institutions. J Econ Perspect 5(1):97–112

North DC (1999) In anticipation of the marriage of political and economic theory. In: Alt J, Levi M, Ostrom E (eds) Competition and cooperation. Conversations with nobelists about economics and political science. Russell Sage, New York, pp 314–317

Owens JE (1997) The return of party government in the US house of representatives: central leadership – committee relations in the 104th Congress. Br J Polit Sci 27:247–272

Polsby NW (1968) The institutionalization of the U.S. house of representatives. Am Polit Sci Rev 62(1):144–168

Polsby NW, Schickler E (2001). Landmarks in the study of Congress since 1945: sketches for an informal history. In: Annual meeting of the APSA, San Francisco

Sánchez de Dios M (1999) Parliamentary party discipline in Spain. In: Bowler S et al (eds) Party discipline and parliamentary government. Ohio State University Press, Columbus, OH, pp 141–162

Schofield N (2009) The political economy of democracy and tyranny. Oldenbourg, Munich

Schofield N, Sened I (2006) Multiparty democracy: elections and legislative politics. Cambridge University Press, Cambridge

Shepsle KA (1978) The giant jigsaw puzzle: democratic committee assignments in the modern house. University of Chicago Press, Chicago, IL

Shepsle KA, Weingast BR (1987) The institutional foundations of committee power. Am Polit Sci Rev 81(1):85–104

Shepsle KA, Weingast BR (1995) Positive theories of congressional institutions. The University of Michigan Press, Ann Arbor, MI

Spiller PT, Tommasi M (2003) The institutional foundations of public policy: a transactions approach with applications to Argentina. J Law Econ Org 19(2):281–306

Spiller PT, Tommasi M (2007) The institutional foundations of public policy in Argentina. A transaction cost approach. Cambridge University Press, Cambridge

Toboso F, Scorsone E (2010) How much power to tax do regional governments enjoy in Spain since the 1996 and 2001 reforms? Reg Fed Stud 20(2):157–174

Tsebelis G (1995) Decision making in political system: veto players in presidentialism, parliamentarism, multicameralism and multipartism. Br J Polit Sci 25:289–325

Uhr J (2006) Bicameralism. In: Rhodes RAW, Binder SA, Rockman BA (eds) Oxford handbook of political science. Oxford University Press, Oxford, pp 475–494

Uriarte E (2000) La política como vocación y como profesión: análisis de las motivaciones y de la carrera política de los diputados españoles. Revista Española de Ciencia Política 3:97–124

Uslaner EM, Zittel T (2006) Comparative legislative behavior. In: Rhodes RAW, Binder SA, Rockman BA (eds) Oxford handbook of political science. Oxford University Press, Oxford, pp 455–473

Weingast BR, Marshall WJ (1988) The industrial organization of congress; or, why legislatures, like firms, are not organized as markets. J Polit Econ 96:132–163

Williamson OE (1975) Markets and hierarchies. The Free Press, New York

Williamson OE (1985) The economics institutions of capitalism: firms, markets, relational contracting. The Free Press, New York

Williamson OE (2000) The New Institutional Economics: taking stock, looking ahead. J Econ Lit 38:595–613

# Coalition Governments and Electoral Behavior: Who Is Accountable?

**Ignacio Urquizu-Sancho**

> A coalition is like a mule. It has no pride of ancestry and no hope of posterity

James Callaghan, British Prime Minister, Chancellor of the Exchanquer, Home Secretary and Foreign Secretary

## 1 Introduction

In 1998, the German Socialist Party (SPD) and the Green Party formed the first red-green German coalition. It was the first time that these parties agreed to form a coalition cabinet. Four years later, the SPD lost 2.43% points of total support whereas the Greens won 1.91% points. Moreover, in spite of these electoral results, after 2002 elections, both incumbent parties continued holding the cabinet. During that legislature, the GDP growth fell from 2% in 1998 to 0% in 2002,[1] inflation increased from 0.93 to 1.59% and unemployment dropped from 8.1 to 7.7%. The electoral results are intriguing because both parties faced the same problems, but their electoral payoffs differed: the small party was rewarded while the big one lost electoral support. Thus, we may wonder: Why have these electoral results occurred? Or, to put it another way, if both incumbent parties were responsible for economic performance, how do we explain the voters' behavior? These questions frame this research.

The second issue that is considered by this chapter, is that of the electoral results of incumbent parties as single actors. Perhaps, this is the principal oversight that is observable in studies of voting behavior and accountability. Most scholars analyze the cabinet as if it is a single actor and they do not take into account the intra-government electoral results. Thus, they contrast government and opposition, considering each as single actors (Lewis-Beck 1986, 1988; Norpoth et al. 1991; Powell and Whitten 1993;

---

I. Urquizu-Sancho
Department of Sociology VI, Complutense University of Madrid, Madrid, Spain
e-mail: iurquizu@march.es

Whitten and Palmer 1999; Powell 2000; Nadeu et al. 2002; Bengtsson 2004; Barreiro 2007). The unique exceptions are Wilkin et al (1997) and Tucker (2006).

However, if we want to more appropriately study accountability, we ought to consider political parties themselves as the object of study and even more so when cabinets are formed by two or more parties. Accountability is a question about parties, not about governments. As Elster affirms, "accountability is individual rather than collective" (Przeworski et al. 1999: 255). Or, in other words, "collective responsibility has leaked out of the system. Officeholders appear responsible only for their personal actions and activities, not for their part in the collective enterprise of governing" (Fiorina 1981: 210). Therefore, any study of accountability ought to consider parties as the main actors.

This chapter is divided in the following sections. First, I present the theoretical puzzle. The second section deals with the causal mechanism that explain accountability in coalition governments. Third, I describe the data. And finally, I present the empirical evidence.

## 2    The Theoretical Puzzle

For a long time, philosophers and social scientist have wondered how citizens can protect themselves from the power of governing class. In order to do that, democratic regimes have established control mechanisms that restrict the freedom of politicians. The mechanisms at play may be twofold: vertical and horizontal. Vertical mechanisms suggest that citizens control politicians through elections. Horizontal mechanisms refer to the control between institutions: the legislative controls the executive, the constitutional courts control legislation and so on. Nevertheless, if we combine both types of mechanisms, sometimes the outcomes may be unsatisfactory. The main aim of this chapter is to analyze that relationship.

Elections have been studied from two points of view: as mechanisms of selection and as mechanisms of sanction. They imply different assumptions. One the one hand, elections may be described as an instrument for selecting the 'good types'. This view is known as prospective models of voting. The main idea is that voters look to the future and they entrust a program. Politicians are supposed to implement this electoral manifesto. The main assumption is that politicians are distinctive.

On the other hand, elections may be described as a process of assigning responsibilities. From this point of view, the main objective of an election is to sanction: voters have the capacity of 'fining' the incumbent. The main idea is that citizens use the past, the incumbent's performance, to judge politicians. These models have been described as retrospective voting.

This chapter focuses on elections as mechanisms for assigning responsibilities. As I develop in the following lines, a theoretical problem arises when the distribution of powers is not clear. But, before dealing with this problem, we need to know how the process of assigning responsibilities works.

If we take into account the broad range of definitions of elections as mechanisms of sanctioning (Pitkin 1967; Fiorina 1981; Przeworski et al. 1999; Strom et al. 2003) and use principal-agent theory, an agent would be accountable when the principal could punish or reward her because of her performance. Thus, with complete information, a principal can observe the agent's performance and then, decide either to support of reject the agent. In accountability models, the main assumption is that politicians are alike or, in other words, that there is no distinction between candidates.

Accountability may be defined as the criterion of re-election.[2] Which criterion does the principal use for supporting the agent? The majority of the models assume that there is a 'threshold of well-being', $k_t$, which guarantees re-election (Ferejohn 1986; Persson and Tabellini 2000; Adsera et al. 2003; Przeworski 2003). Citizens will support the incumbent if she achieves the minimum threshold vale $k_t$, otherwise electors will punish her. As Maravall says, "their threshold for re-electing the incumbent will be arbitrary" (Maravall 2007: 18). Hence, "voters will punish the government at the polls when economic performance is poor" (Maravall 2007: 23). The question that arises is: what is 'poor'? Nevertheless, there is not an unique criterion of punishing or regarding incumbents. Barro's model assumes that an office-holder would be re-elected if the public spending was equal to voters' preferences of public outputs (Barro 1973). In Austen-Smith and Banks' model (1988, 1989), incumbent would be re-elected if policy outcomes and electoral promises matched up. Otherwise, if voters observed large disparities, they would vote for the challenger.

Hence, each model establishes its criterion. If we analyze the broad range of researches, we shall observe that the authors establish a relationship between economic variables and electoral support. Thus, most of scholars focus on economic criterion. Or, in other words, the main idea is that "economic conditions have an enduring place in the voter calculus" (Lewis-Beck 1986: 104). However, empirical evidence does not correlate well with this idea and the relationship between the state of the economy and electoral results is not simple (Stokes 1996; Przeworski et al. 1999; Maravall and Przeworski 1999; Maravall 2007).

Finally, to consider elections as a mechanism of sanctioning presents several problems. First, the vote is a rough tool for punishing or rewarding all the performance government. Second, politicians should have incentives to stay in government.[3] Third, voters have less information than politicians. In this case, citizens

---

[2]Re-election is an important issue. If politicians do no have incentives for re-election, accountability will not work. Or, in Barro's Word, "the electoral process is an instrument which, through the threat of non-reelection, can be used to induce officeholder to" (Barro 1973: 26) act as voters' desires. What are these incentives? The earnings that politicians extract (Barro 1973; Adsera et al. 2003; Przeworski 2003), the value of office (Ferejohn 1986) or the value of the state of the world (Ferejohn 1986) are incentives for politicians. If they are not favorable to their interests, politicians will not invest effort in reelection.

[3]For instance, some political systems do not permit re-election.

face a problem of 'moral hazard'. Voters may not be able to make a complete retrospective assessment of the past performance. When voters look at the past, they may not have full information. Thus, officeholders may shirk their responsibilities. And finally, voters should have capacity to assign responsibilities.

This research focuses on the last two problems: a trouble arises when we assume that elections are a question of assigning responsibilities and the distribution of powers is not clear (Powell and Whitten 1993; Powell 2000; Nadeu et al. 2002; Bengtsson 2004). Scholars have concluded that "if citizens in a democracy cannot identify responsibility for policy, they cannot use elections precisely to hold policy-makers retrospectively accountable for their actions" (Powell 2000: 51). Thus, divided power may be a challenge for elections as mechanisms for accountability.

This is the case of coalition governments. When several parties share power, it is not therefore absolutely clear who is responsible for the policies. At this point, the previously mentioned problem appears again. Thus, the consensus of academic literature on elections and coalition governments was that voters are not able to assign responsibilities to multiparty cabinets. Following those arguments, scholars have concluded that citizens will have difficulties in holding politicians accountable. We can read findings like: "the more political parties in the governing coalition, the less strong the economic voting" (Lewis-Beck 1986: 109) or "where clarity of responsibility is low, the economic factors will be blurred" (Powell and Whitten 1993: 405). This hypothesis has been termed as 'clarity of responsibility' and it is commonly discussed in the literature (Lewis-Beck 1986, 1988; Powell and Whitten 1993; Mershon 1996, 2002; Przeworski et al. 1999; Whitten and Palmer 1999; Anderson 2000; Powell 2000; Nadeu et al. 2002; Strom et al. 2003; Bengtsson 2004). However, as we shall see in this chapter, far from this issue being settled, several loose ends remain that require explanation.

Therefore, coalition cabinets are seen as a problem for democracy. However, if this argument is correct, we would then wonder: how would we explain the electoral results of multiparty governments? In other words, why do voters support one party if the responsibility is unclear? If scholars were right and people could not assign responsibilities to divided power, the electoral results would be random. Is chance the main explanatory variable? However, as Einstein said, "God does not play dice with the universe".

In spite of the large amount of literature about this issue, scholars have fully not resolved the main questions surrounding them. First, they have not developed a theoretical model and causal mechanisms that explain how accountability works in divided power. Moreover, previous studies have achieved similar results to this study, but they have not developed theoretical explanations (Fisher and Hobolt 2010).

Second, empirical evidence may be improved. Most scholars analyze the cabinet as if it is a single actor and they do not take into account the intra-government electoral results. Thus, they contrast government and opposition, considering each as single actors. However, as I said before, accountability is a question about parties, not about governments. Therefore, any study of accountability ought to consider parties as the main actors.

The main aims of this chapter are to resolve these issues and to shed more light on how voters assign responsibilities to coalition governments. Next subsection deals with the theoretical models and causal mechanisms.

## 2.1   The Causal Mechanisms

The first open question is that we do not know exactly how accountability works in coalition governments. Scholars have not developed theoretical arguments that explain how voters behave when pass judgment on this type of cabinet. Researchers have assumed, for example, that voters may not know who is in charge of formulating specific policies. However, they have not developed the causal mechanisms that explain the theoretical outcomes. For instance, they do not explain how citizens may weigh up information when they consider the fate of coalition governments.

We need to develop a simple model of accountability. Elections may be presented as 'political contract': citizens delegate to politicians. However, this contract is not a simple agreement. In microeconomy, contracts have been analyzed using principal-agent models: the main goal of a contract is that an agent carries out a task that benefits the principal (Laffont and Martimort 2002; Macho and Pérez 2005).

In these models, an important problem is information. If the relation between principals and agents emerges in a world with perfect information, principals – voters – will know the agents' – politicians – level of effort and they will only pay them if they strive. However, asymmetric information is quite frequent. Asymmetric information implies that the principal is unaware of the level of effort that the agent invested – that is, the 'moral hazard' problem – or the agents know something about their features that the principal is unaware – that is, 'adverse selection' – (Berganza 2000; Laffont and Martimort 2002; Macho and Pérez 2005).

Theoretical models have concluded that in spite of asymmetric information, that contract may work. First, agents have incentives to strive because their payoff depends on their performance. Second, the principal uses the unique verifiable variable that he observes: performance. Performance works as statistical inference (Macho and Pérez 2005: 58). That is, in the 'political contract', where politicians are the agents and voters are the principal, well-being gives information about the effort of agents, and this information forms part of the contract. The probability that agents have made a high effort when economic performance is observed, affects the optimal payoff that agents receive. In sum, economic performance will be the verifiable variable that voters use for guaranteeing that the contract has been carried out. Those findings agree with other models of accountability (Ferejohn 1986).

Multiparty cabinets add a new problem of information: citizens may not know who is in charge of incumbent performance. If principal-agent model has problems of information, the multiparty cabinets will multiply those problems. For that reason, the solution centers around information, and this is the key feature of the following theoretical arguments.

The argument that information matters is not new in the literature about accountability (Fiorina 1981; Ferejohn 1986; Adsera et al. 2003), Thus, for instance,

Adsera et al. conclude that "political control of public officials turns out to depend on (...) the degree of information of citizens" (Adsera et al. 2003: 478–479). More precisely, if degree of information increases, accountability improves. In a bipartisan system, we can assume that people have some intuition about the government did. However, in a multiparty system, obtaining information can be more difficult.

Academic literature has proposed some mechanisms that can improve asymmetric information in principal-agent models. If the principal faces a problem of 'moral hazard', the solution will be a set of constraints and incentives. However, if the principal faces a problem of 'adverse selection', the solution will be signals. That is, agents send signals that permit the principal to distinguish between the true features of agents. Multiparty governments produce both problems and this is the reason that accountability is more difficult in those types of cabinets. On the one hand, voters are not sure whether politicians strive during the legislature – 'moral hazard'. On the other hand, citizens have different parties in the government and they cannot distinguish between them –'adverse selection'. How can we solve this problem?

I consider that 'moral' hazard problem has been resolved in the previous theoretical arguments. This is because, in spite of multiparty cabinet, voters use incumbent performance to know whether politicians have made an effort during the legislature. Or, in other words, well-being will be the verifiable variable that voters use to guarantee that the political contract has been fulfilled by the coalition parties. But, how do we know how is responsible for that well-being? The answer is signals. However, signals involve several problems. First, who gives out these signals? And, second, why are some signals more credible than other signals?

I start with the speakers. Signals may be emitted by the agents – incumbent parties. Coalition partners have incentives to supply information and then, to increase accountability. A similar argument is presented by Ferejohn. He points out that sometimes politicians have incentives to increase accountability. Why should they do that? Because they want to get more resources for administration. If citizens can hold politicians accountable, people will be willing to increase politicians' resources (Przeworski et al. 1999: 140–141). We can observe similar behavior in multiparty cabinets, although the theoretical reason is different. Parties are vote-seeking. They are constantly thinking about next election. When politicians share government, that interest, votes, will not disappear. Therefore, parties may have incentives to supply information – and then, to increase accountability – because they want to be different from the cabinet's partner. That is, in coalition governments, a contradictory issue appears: parties are partner and future competitors simultaneously. Therefore, politicians may have incentives to supply information.

Let's assume the simplest scenario where we only have multiparty government and opposition parties. I do not assume anything about the features of those actors. Incumbent parties may send signals about what they did and opposition parties may supply information as well. In this situation, voters receive several messages from different agents. How do they distinguish credible signals from unbelievable messages? Austen-Smith deals with that problem and he concludes that "the harder it is to verify information, the less likely it is that such information can be communicated credibly in speech" (Austen-Smith 1992: 57). Then, after listening to all

messages, voters would be confused and would have two possibilities: don't assign responsibilities or simplify the messages. The literature has emphasized the importance of the first alternative and has not considered the second one. However, voters may process information and simplify it. How do they do that? They may focus on the most visible party, the Prime Minister's party, and blame it because of the performance. It is not unrealistic to assume that Prime Minister's party is in the spotlight: "in the majority of cases it is his party that dominates economic policy and makes the relevant economic decisions" (López-Nava 2007: 24).

Perhaps, we would be able to think that voters focus on the biggest parties. However, if we think about the definition of accountability, it is not a question about size, it is a question about tasks. Citizens try to blame or to reward incumbent parties because of their performance. The Primer Minister's party holds the most important task, to manage the government, and, in view of confused information, voters may focus on that party and assign responsibilities.

This is the hypothesis that I want to study in this research: Is accountability a question about task or about size? To sum up, if just think in a world of multiparty cabinets and opposition parties, voters will simplify information and focus on Primer Minister parties.

## 3    Data and the Statistical Model

In order to analyze the previous theoretical arguments, I decided to construct my own data. The database that I use, is formed by all the governments from 1945 to 2006 in 22 OECD parliamentary democracies. The countries are: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Iceland, Ireland, Israel, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland and United Kingdom.

The first problem that emerges is: when do consider that a government is over? I have found that any of the possible answers is arbitrary and entails problem. As a result, I have decided to use the same definition used in Woldendorp, Keman and Budge's database (1998). They consider that a new government exists when one of these events happens: elections, voluntary resignation of Primer Minister, resignation of Primer Minister due to health reasons, dissension within the government, lack of parliamentary support, intervention by the Head of the State and broadening of coalition. Moreover, I am not just using their criterion, I using their data as my main source of information as well. Woldendorp et al collected information on all the governments between 1945 and 1996 in 20 democracies.[4] I have increased the database with two countries – Spain and Portugal – and have updated it to 2006. In

---

[4]Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Iceland, Ireland, Israel, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Sweden, Switzerland and United Kingdom.

**Table 1** Type of governments

|           | Majority        | Minority        | All             |
|-----------|-----------------|-----------------|-----------------|
| Single    | 122 (16.97%)    | 112 (15.58%)    | 234 (32.55%)    |
| Coalition | 407 (56.61%)    | 78 (10.85%)     | 485 (67.45%)    |
| Total     | 529 (73.57%)    | 190 (26.43%)    | 719 (100%)      |

**Table 2** Electoral payoffs by governments (electors)

|           | Majority      | Minority      | All           |
|-----------|---------------|---------------|---------------|
| Single    | −2.87 (6.22)  | −1.69 (6.93)  | −2.31 (6.58)  |
|           | N = 118       | N = 109       | N = 227       |
| Coalition | −3.42 (6.95)  | −1.83 (5.46)  | −3.15 (6.74)  |
|           | N = 364       | N = 75        | N = 439       |
| Total     | −3.28 (6.78)  | −1.75 (6.36)  | −2.86 (6.69)  |
|           | N = 482       | N = 184       | N = 666       |

order to do that, I consulted Kessing's Contemporary Archive,[5] Zararate's Political Collection[6] and Montabe's work (1997).

Table 1 shows a picture of the database taking into account the type of governments. I have classified cabinets by taking into account the number of parties – i.e. single versus coalition governments – and parliamentary support – i.e. majority versus minority. If we compare that sample with other studies, it will be noted that I have increased the number of cases.[7] However, my data is not different from other databases, although majority governments represent a bigger portion than other samples.

A relevant finding is that politicians share the government more frequently than not. Coalition governments make up 67.45% of the cases in my sample. However, the literature on electoral behavior and accountability has paid little attention to that type of governments.

One we know the distribution of governments, I deal with the electoral results of those cabinets.[8] Table 2 summarizes the electoral payoffs by governments.[9] The electoral payoffs are calculated among electors or, in other words, among people who participated in the elections. The electoral outcomes match up with Strom's (1990: 128) findings. On the one hand, majority multiparty governments lose more

---

[5]http://www.kessings.com

[6]http://www.terra.es/personal2/monolith/home.htm

[7]For instance, Strom (1990) uses 15 democracies in his study. Powell (2000) considers 20 democracies, Powell and Whitten (1993) analyze 19 democracies. Moreover, I have expanded the period of analysis too. For instance, Fisher and Hobolt (2010) covers from 2001 to 2006 and one election per country.

[8]The main sources of information are Mackie and Rose (1982, 1997) and Caramani et al. (2000).

[9]I present the average, the standard deviation in brackets and the number of cases that have in my sample.

votes than other cabinets. And on the other hand, minority single-party governments lose the least votes of all.

Moreover, there is a strong significant difference between minority and majority cabinets. If I perform the t test on the difference of means, it reveals that the electoral results of minority governments are different from majority cabinets at a statistically significant level. Coalition and single-party governments are not significantly different, although we are close to rejecting that both averages are statistically different.[10]

The electoral results of governments may be calculated among citizens too. That is, we may assume that abstention is part of the rewards and penalties. Therefore, we may calculate the electoral results considering total population. This does not mean that accountability explains the abstention entirely. However, we cannot forget that possibility. Table 3 shows the electoral outcomes of cabinets among citizens. The differences between types of government that we observed in the previous table are similar to the results of Table 3, and we get the same findings.

Those results have been shown in other studies (Strom 1990; Powell 2000). However, we know much less about parties. As I said before, the literature of political science has studies government as if they were single actors, paying little attention to the electoral results of parties. For that reason, I have constructed a second database where parties are the unit of analyses.[11] Table 4 summarizes the dependent variable: the electoral results of incumbent parties.[12]

**Table 3** Electoral payoffs by governments (citizens)

|  | Majority | Minority | All |
|---|---|---|---|
| Single | −2.3 (5.04) | −1.33 (5.51) | −1.83 (5.28) |
|  | N = 115 | N = 108 | N = 223 |
| Coalition | −2.58 (7.76) | −1.44 (5.32) | −2.39 (7.42) |
|  | N = 363 | N = 73 | N = 436 |
| Total | −2.51 (7.2) | −1.38 (5.42) | −2.2 (6.77) |
|  | N = 478 | N = 181 | N = 659 |

---

[10]The t is 1.54 with 664 degrees of freedom.

[11]I have considered any political organization that has participated in a government. When a party participates in a coalition government and a single-party government in the same legislature, I have selected the coalition case. Coalition cabinet prevails over single party government. Since the main aim of this research is to study coalition governments, I have followed a strategy that widens the sample of coalition cabinets as much as possible. Moreover, when I find different coalition governments in the same legislature, I have considered the cabinet that survives for a longest period.

[12]Perhaps, the reader may wonder why the electoral results of single party governments are different from previous Tables 2 and 3. As I argue above, in the government data set I consider a new unit of analysis, for instance, when the Prime Minister changes. This means that the only change is the Chief of government, and that the members of the cabinet remain the same. Thus, one electoral payoff may count in two or three units of analysis.

**Table 4** Electoral payoffs by parties (electors)

|  |  | Majority | Minority | Total |
|---|---|---|---|---|
| Single |  | −2.84 (5.897) | −0.79 (7.78) | −2.01 (6.77) |
|  |  | N = 80 | N = 55 | N = 135 |
| Coalition | All parties | −1.01 (4.07) | −0.23 (3.84) | −0.89 (4.04) |
|  |  | N = 604 | N = 90 | N = 700 |
|  | Prime Ministers | −1.24 (4.91) | −0.84 (4.03) | −1.17 (4.77) |
|  |  | N = 186 | N = 32 | N = 219 |
|  | PMs' partners | −0.87 (3.62) | 0.11 (3.72) | −0.74 (3.65) |
|  |  | N = 411 | N = 58 | N = 474 |
| Total |  | −1.22 (4.36) | −0.44 (5.64) | −1.08 (4.61) |
|  |  | N = 684 | N = 145 | N = 835 |

**Table 5** Electoral payoffs by parties (citizens)

|  |  | Majority | Minority | Total |
|---|---|---|---|---|
| Single |  | −2.45 (4.95) | −0.84 (6.16) | −1.79 (5.51) |
|  |  | N = 77 | N = 54 | N = 131 |
| Coalition | All parties | −0.81 (3.68) | −0.08 (3.29) | −0.7 (3.65) |
|  |  | N = 591 | N = 86 | N = 683 |
|  | Prime Ministers | −0.91 (4.66) | −0.28 (4.09) | −0.8 (4.58) |
|  |  | N = 182 | N = 30 | N = 213 |
|  | PMs' partners | −0.71 (3.1) | 0.02 (2.82) | −0.6 (3.09) |
|  |  | N = 402 | N = 56 | N = 463 |
| Total |  | −0.998 (3.881) | −0.37 (4.61) | −0.89 (4.02) |
|  |  | N = 668 | N = 140 | N = 808 |

One important finding, which contradicts previous conclusions, is that from the point of view of parties, participating in a coalition government is not worse than participating in a single-party government. In Tables 2 and 3, we saw that coalition cabinets had higher electoral costs than single party governments. Nevertheless, if we use parties as the unit of analysis, we conclude quite the opposite: single party cabinets have more electoral costs than coalition governments. Thus, parties that take part in a multiparty cabinet lose, on average, 0.89% of votes. However, parties that participate in single party government lose, on average, 2.01% of their votes. If we run the mean comparison test, we observe that these differences are highly statistically significant.

Moreover, Table 4 presents the electoral results of parties taking into account their role in the coalition government: Prime Minister versus partner. We observe that to hold the Prime Minister portfolio is more 'dangerous' than to hold other portfolios. On average, Prime Minister parties lose more votes than their partners. These differences are statistically significant, as well.

Finally, as I did with governments, we may assume that abstention is relevant for the electoral payoffs. For that reason, I have calculated the electoral results of parties among total population. Table 5 shows the data. The results are similar to Table 4 in that big differences are not observable. Prime Minister parties lose more votes than their partners, single party governments lose more votes than coalition

parties and majority governments lose more votes than minority cabinets. Majority single-party governments lose the most of all.

The questions that arise are: Why do we observe those electoral results? How do we explain the outcomes? What factors do voters consider for punishing or rewarding parties?

## 3.1 The Independent Variables

In this section I am describing the independent variables: how I created them and the sources of information. The independent variables that I shall use in the empirical evidences are economic and political. Table 6 shows all of them.

The economic independent variables are inflation,[13] unemployment,[14] economic growth[15] and public expenditure.[16] They are measured in different ways. Thus, inflation and unemployment are collected as the difference between the rates in two successive elections; economic growth is measured as the relative difference of real

**Table 6** Independent variables

| Variables | N | Mean | Std. dev. | Minimum | Maximum |
| --- | --- | --- | --- | --- | --- |
| *Economy* | | | | | |
| Inflation | 232 | −0.2642 | 30.2916 | −357.5525 | 257.0205 |
| Unemployment | 139 | 0.1382 | 2.5503 | −8.2 | 8.8 |
| Economic growth | 298 | 5.4876 | 4.7101 | −9.1608 | 23.6874 |
| Total public expenditures | 161 | 4.7229 | 12.0255 | −44.0018 | 50.9042 |
| Public expenditures on health care | 133 | 3.9482 | 18.8906 | −59.4737 | 90.449 |
| Public expenditures on education | 59 | 1.1665 | 8.5081 | −20.8887 | 24.036 |
| *Politics* | | | | | |
| Electoral payoffs | 835 | −1.076 | 4.6089 | −27 | 21.42 |
| Left parties | 934 | 0.3062 | 0.4611 | 0 | 1 |
| Coalition governments | 719 | 0.6745 | 0.4688 | 0 | 1 |
| Proportional electoral system | 361 | 0.3712 | 0.4838 | 0 | 1 |
| Multilevel electoral system | 361 | 0.2468 | 0.4314 | 0 | 1 |
| Mixed electoral system | 361 | 0.0803 | 0.2721 | 0 | 1 |
| ENEP | 360 | 4.0786 | 1.4257 | 1.99 | 10.29 |

---

[13]The source is World Development Indicators (WDI) from Alvarez, Cheibub, Limongi and Przeworski (ACLP). It covers from 1960 to 2000.

[14]The source of information is World Development Indicators (WDI) from Alvarez, Cheibub, Limongi and Przeworski (ACLP). It covers from 1975 to 2000.

[15]The source is WDI and covers from 1949 to 1996.

[16]I have collected total public expenditures, public expenditures on Elath and public expenditures on education. The source of information is World Development Indicators and covers from 1965 to 1999. However, it depends on the country. In those variables there are several gaps in the information available. For instance, in the case of health, to collect data before 1984 is extremely difficult.

GDP per capita[17] between election year and the 2 years prior to elections, and public expenditures as a percentage of GDP between two successive elections.

The main difference between them is the timing. Inflation, unemployment and public expenditures are observed in the whole period the incumbent was in office, whereas economic growth takes into account the GDP per capita growth rates of the 2 years preceding the election. This chapter does not focus on accountability timing. The main aim is to know whether citizens have the capacity of assigning responsibilities in spite of coalition governments. This is the reason that I do not concern myself with whether or not voters are far-sighted or myopic. Moreover, the literature about accountability timing is not conclusive. On the one hand, some scholars have pointed out that voters take into account long periods of economic information (Peltzman 1990; López-Nava 2007). On the other, Anchen and Bartels (2004) have come to the opposite conclusions: voters are myopic and they only consider short-term results. In view of that debate, I consider that both long and short-term measures are useful. After running several statistical analyses, I am using the measurements that explain more about the electoral results of parties.

The political variables that I have collected are:

(a) Electoral payoffs. This refers to the electoral gains or loses of parties between two elections. Thus, I calculate the differences between two consecutive elections. The main sources of information are Mackie and Rose (1982, 1997) and Caramani et al. (2000).
(b) Ideological variable: leftist parties. This variable has been created from Swank's categorization.[18] Swank classifies party ideology into six categories: left libertarian, left,[19] secular center,[20] centrist Christian Democratic,[21] right[22] and right-wing populist. The variable "left parties" assumes values 1 if parties are left libertarian or left and value 0 for the remaining values.
(c) Type of governments: coalition governments. This is a dummy variable. It takes value 1 when the government is multiparty and vale 0 otherwise. The sources of information are the same as I used to create variable "electoral payoffs".
(d) Electoral systems. This variable has been collected from Matta Golder's database (2006). I have split it into three dummy variables: proportional, multilevel and mixed, and majoritarian electoral system is the category of reference. We may observe that the majority of electoral systems are proportional −37.12%, whereas similar proportion of majoritarian and multilevel electoral systems

---

[17]Per capita income in 1996 purchasing power parity (PPP) dollars.

[18]Duane Swank, *Comparative parties data set*. In the cases of Iceland, Israel and Luxembourg, I did it taking into account his categories.

[19]Communist, socialist, social democratic, labor and other various left-wing parties (e.g., left-libertarian parties).

[20]Non-catholic parties of the center.

[21]Non-conservative catholic parties.

[22]Far-right (e.g., neo-fascist, right-wing populist), classical liberal, conservative Christian Democratic and other various right-wing parties.

exists $-29.09$ and $25.76\%$ respectively. Those variables are used to explain the origin of different types of governments.

(e) ENEP: effective number of electoral parties. This is based on the following formula from Laasko and Taagepera:

$$\frac{1}{\sum v_i^2}$$

where $v$ is the percentage of the vote received by the $i$th party.[23] The source is Matt Golder's database (2004).

## 3.2    The Statistical Model

I consider that I need to control for the possibility of self-selection bias (Przeworski 2007). What does it mean? We may create a database randomly. It would be a good sample of reality. However, the world is not random. This means that the origin of any object has an explanation. Perhaps the researcher does not observe these factors, but they exist. Thus, coalition governments are not exogenous actors and there are several variables that may explain their existence. These variables may affect the independent variables of other statistical models too. Therefore, I ought to correct self-selection bias. How do I do that? By developing Heckman models. In maths,

$$Y_i = \beta_i X_i + \gamma \lambda_i + u_i \tag{1}$$

$$Z_i = f(\alpha_i W_i + e_i) \tag{2}$$

$$Z\left\{ \frac{1 \text{ if } z_i \text{ is single} - \text{party government}}{0 \text{ otherwise}} \right\} \tag{3}$$

$$\lambda_i = \frac{\phi(\alpha_i W_i)}{\Phi(\alpha_i W_i)} \tag{4}$$

$$\lambda_i \geq 0 \tag{5}$$

where (1) is the outcome equation and (2) is the selection equation. That statistical model is known as two-steps (Heckman 1974, 1979; Breen 1996). How does it work? First, we calculate (2). This is a binomial probit where $Z_i$ is the dependent variable and the $W_i$ is the matrix of independent variables that explain the existence

---

[23]Independents or others are treated as a single party.

of different types of governments. As we see in (3), $Z_i$ assumes value 1 if the cabinet is single-party and value 0 for the remaining values.

Second, we calculate the hazard rate, or inverse Mill's ration, $\lambda_i$. This is calculated in (4), using the information from function (2) and (3). In few words, the hazard rate is the probability of an event occurring given that it has not occurred prior to this time. In maths, it's the quotient between the probability distribution function – $\phi(\alpha_i W_i)$ – and the survival function – $\Phi(\alpha_i W_i)$. The only restriction on hazard rate, and implied by the properties of both functions, is that $\lambda_i$ may not be negative and it may be greater than one. The hazard rate will correct the possible self-selection bias in (1).

Third, function 1, or outcome equation, is the statistical model that analyzes my theoretical arguments. Thus, $Y_i$ and $X_i$ are the matrix of dependent and independent variables that I have presented before. In order to correct for self-selection bias, I shall introduce $\lambda_i$ as independent variable and $\gamma$ is the coefficient that describes its effect. Finally, $u_i$ and $e_i$ are the random disturbances.

In the Appendix, Table 15 presents the results of the selection equation (2) that I shall use. The outcomes are probit effects, although we may focus on the signs. The independent variables are the type of electoral systems and effective number of electoral systems. Thus, I explain the existence of single-party governments using institutional variables. The results fit what I expected: single-party governments are more likely in majority electoral systems[24] and as the number of electoral parties decreases, the probability of observing single-party government decreases. Using that equation, I get the hazard rate (4) that I shall introduce in the outcome equation.

## 4 Empirical Evidence

### 4.1 When Voters Evaluate Single-Party Governments

What do voters take into account when they evaluate a government? As I said before, this is an open question in the social sciences literature. Most researchers have concentrated their efforts on establishing a relationship between the economy and the electoral results. Economic voting is widely discussed phenomena in the literature (Fiorina 1981; Kramer 1983; Chapel and Keech 1985; Lewis-Beck 1986, 1988; Norpoth et al. 1991; Powell and Whitten 1993; Przeworski et al. 1999; Whitten and Palmer 1999; Anderson 2000; Sánchez-Cuenca and Barreiro 2000; Royed et al. 2000; Norpoth 2001; Nadeu et al. 2002; Bengtsson 2004; Duch and Stevenson 2005a, b; Barreiro 2007). The main idea is that voters use the economy to

---

[24]All dummy variables are negative and the category of reference is majority electoral systems. This means that single-party governments are less probable in proportional, mixed and multilevel electoral systems than in majority electoral systems.

evaluate government performance. This chapter is a straightforward extension of that literature.

To test the economic voting hypothesis, I have developed the following function:

$$V_{it} = \beta_i X_{it} + \delta_{it} G_{it} + \gamma \lambda_{it} + u_{it}$$

$V_{it}$ indicates the electoral payoff by party $i$ in each of $t$ elections. $\beta_i$ are the coefficients that describe the effects of economic variables. $X_{it}$ is the matrix of the following economic variables: inflation, GDP per capita, unemployment and government expenditures. $\delta_{it}$ is a dummy variable that assumes value 1 if the party ideology is on the left and value 0 if the party ideology is on the center and right. I have introduced the party ideology variable because I consider that government expenditures – $G_{it}$ – affect the electoral results of parties depending on their ideology (Barreiro 2007). The main idea is that left parties benefit from the increase of budgets. It does not mean that leftist governments spend more than rightist cabinets. In fact, using may database, we can see that left-wing governments spend less that left-centre cabinets. Moreover, left-wing governments spend the same amount of money as right-wing cabinets. Therefore, ideology may not explain the differences between governments. The idea of interaction – $\delta_{it} G_{it}$ – is how citizens evaluate public spending. I assume that leftist voters reward an increase in public spending with more probability than rightist electors. $\gamma$ is the coefficient that describes the effect of $\lambda_{it}$, or hazard rate. As I said in previous section, Mill's ratio corrects the possible self-selection bias.

I would expect that if economy improves, that the electoral results will improve too. This means that if inflation and unemployment rise, electoral performance will decline. However, if GDP per capita increases, the electoral performance will improve too. Moreover, I would argue that inflation, GDP per capita and unemployment are the key economic variables. Government expenditures are related to party ideology.

Table 7 shows the statistical analyses. I have developed four models. Models 1 and 2 use as dependent variables the electoral results of parties among electors. However, in models 3 and 4, I consider that abstention accounts for part of voters' reward and penalties. This means that the electoral payoffs are calculated among citizens. Model 1 and 3 take Ordinary Least Squares (OLS) as the statistical method and it leaves out spatial and time controls. However, models 2 and 4 introduce these controls by country.[25] The modified Wald test revealed a heteroskedasticity problem. For that reason, I have estimated the econometric analyses using Feasible Generalized Least Square (FGLS) (Castilla 1998; Hsiao 2003; Baltagi 2005). The statistical analyses fits as well as I had hoped. In all models, inflation, unemployment and

---

[25] I have used fixed effects models because it is the correct analysis when "we are focusing on a specific set of N" (Baltagi 2005: 12). Moreover, the F test suggests that this is the appropriate specification, and the correlation between the dependent variables and residuals is close to 0 in both analysis.

**Table 7** Single-party governments. Analysis on electoral payoffs

| Variables | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Inflation | −1.664** | −0.942** | −1.222** | −0.552* |
|  | (0.699) | (0.38) | (0.506) | (0.303) |
| GDP per capita | −0.151 | 0.126 | −0.175 | 0.147 |
|  | (0.539) | (0.313) | (0.378) | (0.241) |
| Unemployment | −2.025** | −0.921 | −1.407** | −0.512 |
|  | (0.931) | (0.631) | (0.681) | (0.526) |
| Left party | 3.109 | 3.19 | 1.572 | 2.59 |
|  | (3.773) | (2.215) | (2.976) | (1.637) |
| Government expenditures | −0.097 | −0.067 | −0.086 | −0.092 |
|  | (0.115) | (0.049) | (0.084) | (0.063) |
| Left* expenditures | 0.461* | 0.369* | 0.388* | 0.343* |
|  | (0.234) | (0.154) | (0.187) | (0.136) |
| $\lambda$ | −0.449 | 1.626 | −0.627 | 0.803 |
|  | (3.775) | (2.208) | (2.774) | (1.665) |
| Intercept | −6.889 | −11.438** | −4.425 | −8.855** |
|  | (10.089) | (5.601) | (7.571) | (4.119) |
| N | 36 | 36 | 36 | 36 |
| n |  | 13 |  | 13 |
| $R^2$ | 0.32 |  | 0.299 |  |
| F | 1.51 |  | 1.56 |  |
| Wald $\chi^2$ |  | 21.2*** |  | 15.54** |
| Method | OLS | FGLS | OLS | FGLS |

Significance levels *** 1% ** 5% *10%; robust standard error in brackets

the interaction between ideology party and government expenditures are highly significant and have the expected signs. Thus, when inflation and unemployment increase, the electoral payoffs decrease. Moreover, when government expenditures increase and the party ideology in on the left, the electoral payoffs increase. Once I use fixed effects – models 2 and 4, the empirical results just get worse: unemployment stops being significant.

One surprising outcome is that a GDP per capita increase is not statistically significant and has the opposite expected sign. The literature on the economic vote has stressed the relevance of economic growth (Lewis-Beck 1988; Norpoth et al. 1991; Powell and Whitten 1993; Whitten and Palmer 1999; Barreiro 2007). Nevertheless, my results question this widespread hypothesis. Perhaps, we may think that a problem of multicollinearity exists because the economic variables are highly correlated. For instance, the correlation between unemployment and the GDP per capita increase is −0.484. For that reason, I have run statistical analyses where GDP per capita is the only independent variable. In this simple analysis, the explanatory variable has the right sign, but I have not found any significant links. Therefore, this result seems to suggest that economic growth is not as important as the literature of economic vote presupposes.

In the light of this outcome, we may wonder why scholars have stressed the importance of GDP. The economy involves many variables and citizens may use them as a signal of incumbent performance too. Hence, it is unclear why GDP has to

be the most relevant economic variable. In the literature, we do not find theoretical arguments that seek to demonstrate why GDP ought to be the key factor in economic voting.

In short, the empirical evidence of Table 7 confirms the economic voting hypothesis for single party governments. That is, economic variables explain the electoral results of parties that govern alone. We may infer from these outcomes that voters use the economy for assessing single party governments. The questions that arise are: do we observe the same behavior in multiparty cabinets? Is economy relevant when voters assess coalition governments? The next section deals with these questions.

## 4.2   When Voters Evaluate Coalition Governments

As I have argued previously, I believe that scholars have not correctly dealt with the process of assigning responsibilities to multiparty cabinets. They have downplayed the role of parties in their empirical analysis, treating governments as single actors. I believe that more appropriate way to study this topic is to consider parties as the unit of analysis. This is the strategy that I follow in this section.

However, if we consider parties as individual actors, there is a statistical problem: in each election, the dependent variable – the electoral payoffs – changes whereas the independent variables remain constant. That is, after coalition cabinet, several incumbent parties compete for the votes and obtain different electoral results. Although the economic indicators are the same for all parties. To put it another way, I cannot explain variability within a government when the explanatory variables keep constant. For that reason, I have decided to split the sample into groups and to analyze them separately. I have followed two criteria for classifying into groups: their role in the government – portfolio – and their size.

I apply the voting function that presented in previous section. The only change is that in (2), $Z_{it}$ assumes value 1 when the type of government is multiparty and value 0 otherwise. The remaining functions are equal.

### 4.2.1   The Role of Parties and Accountability

When parties form a coalition cabinet, they divide the portfolios. Each portfolio deals with different subject: education, economy, judiciary and so on. The main aim of this subsection is to check whether there is a relationship between the electoral results of parties that hold those responsibilities and economic indicators.

I start with Prime Minister parties. Table 8 shows the regression results. As in the previous section, I have run different models with different assumptions. Finally, models 1 and 2 use as dependent variable the electoral results among voters, whereas models 3 and 4 use the electoral payoffs among citizens. Secondly, models

**Table 8** Prime Minister party. Analysis on electoral payoffs

| Variables | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Inflation | −0.015* | −0.015 | −0.013* | −0.014 |
| | (0.009) | (0.013) | (0.007) | (0.01) |
| GDP per capita | 0.242* | 0.184* | 0.143 | 0.069 |
| | (0.143) | (0.108) | (0.109) | (0.083) |
| Unemployment | −0.579** | −0.622*** | −0.445** | −0.5*** |
| | (0.259) | (0.191) | (0.185) | (0.139) |
| Left party | −1.044 | −0.421 | −0.972 | −0.258 |
| | (1.352) | (0.877) | (1.11) | (0.805) |
| Government expenditures | 0.124 | 0.132* | 0.084 | 0.099* |
| | (0.084) | (0.069) | (0.068) | (0.052) |
| Left* expenditures | 0.063 | 0.193** | 0.082 | 0.18** |
| | (0.146) | (0.095) | (0.113) | (0.078) |
| $\lambda$ | 1.531 | 1.227** | 0.752 | 0.781 |
| | (1.026) | (0.573) | (0.838) | (0.519) |
| Intercept | −6.898** | −6.095*** | −4.296 | −4.299*** |
| | (3.276) | (1.717) | (2.691) | (1.623) |
| N | 65 | 65 | 65 | 65 |
| n | | 19 | | 19 |
| $R^2$ | 0.132 | | 0.097 | |
| F | 1.95* | | 2.61** | |
| Wald $\chi^2$ | | 17.72** | | 22.81** |
| Method | OLS | FGLS | OLS | FGLS |

Significance levels *** 1% ** 5% *10%; robust standard error in brackets

1 and 3 are linear regressions that leave out spatial and time controls, while models 2 and 4 estimate using fixed effects.[26]

The results of Table 8 are contrary to scholars' expectations. It is widely assumed in the literature that if power is divided, citizens will not be able to assign responsibilities. However, Table 8 shows that if we split coalition governments into incumbent parties, Prime Minister parties seem to be accountable to voters: their electoral results are explained by the economic performance. The empirical analysis shows relevant statistical relations. First, in all models unemployment has the expected sign and is highly statistically significant. Second, models 1 and 3 show that inflation affects significantly the electoral payoffs of parties. Third, in models 1 and 2 GDP per capita increase is statistically significant. In this case, this variable, that is seen relevant in the literature, explains the electoral results of Prime Minister parties. Fourth, when I control by country, the interaction between ideology and government expenditures works. Hence, economic variables explain part of the electoral results of Prime Minister parties that participate in coalition governments.

---

[26]The correlation between residuals and dependent variable is close to 0. It supports the use of fixed effects estimate. The modified Wald test revealed a heteroskedasticity problem. For that reason, I have estimated the econometric analysis using Feasible Generalized Least Square (FGLS).

In sum, in spite of multiparty cabinet, we find strong relationship between electoral performance and economic indicators. Or, to put it another way, we cannot reject the idea that Prime Minister parties seem to be accountable to voters. Moreover, if we compare these results with outputs of single-party cabinets, we may argue that economic voting works better in Prime Ministers from coalition governments than in single-party governments. This finding contradicts the hypothesis of 'clarity of responsibility' that is widespread in the literature (Lewis-Beck 1986, 1988; Powell and Whitten 1993; Mershon 1996, 2002; Przeworski et al. 1999; Whitten and Palmer 1999; Anderson 2000; Powell 2000; Nadeu et al. 2002; Strom et al. 2003; Bengtsson 2004).

However, the outcome fits my theoretical arguments. One of my hypotheses was that in a world of asymmetric information, voters may simplify messages focusing on Prime Minister parties. The preliminary results confirm my arguments. But, we may wonder if these outcomes are produced in case of other incumbent parties too. If so, my theory would have to be rejected because the other incumbent parties are accountable to electors as well.

Table 9 shows the empirical evidence for the Deputy Chairman parties.[27] As in previous statistical analyses, I have run for different statistical models. These

**Table 9** Deputy chairman party. Analysis on electoral payoffs

| Variables | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Inflation | −0.174 | −0.151 | −0.257 | −0.18* |
|  | (0.204) | (0.113) | (0.185) | (0.093) |
| GDP per capita | −0.145 | −0.073 | −0.128 | −0.102 |
|  | (0.11) | (0.082) | (0.091) | (0.073) |
| Unemployment | −0.163 | −0.136 | −0.242 | −0.078 |
|  | (0.299) | (0.168) | (0.27) | (0.15) |
| Left party | −1.326 | −1.115 | −1.024 | −1.389* |
|  | (1.208) | (0.684) | (1.002) | (0.637) |
| Government expenditures | −0.027 | −0.009 | 0.026 | 0.002 |
|  | (0.096) | (0.043) | (0.079) | (0.045) |
| Left* Expenditures | 0.252** | 0.204** | 0.136 | 0.138** |
|  | (0.103) | (0.074) | (0.084) | (0.064) |
| λ | 1.156 | 1.604** | 0.772 | 0.983 |
|  | (1.254) | (0.806) | (1.172) | (0.681) |
| Intercept | −3.627 | −5.109** | −3.012 | −3.076 |
|  | (3.861) | (2.357) | (3.648) | (2.085) |
| N | 35 | 35 | 35 | 35 |
| n |  | 12 |  | 12 |
| $R^2$ | 0.208 |  | 0.197 |  |
| F | 1.78 |  | 1.41 |  |
| Wald $\chi^2$ |  | 23.78*** |  | 27.96*** |
| Method | OLS | FGLS | OLS | FGLS |

Significance levels *** 1% ** 5% *10%; robust standard error in brackets

---

[27]In a multiparty cabinet, the probability that the same party holds simultaneously the Prime Minister and Deputy Chairman portfolios is low. In my sample, it happens in 27.37% of the cases.

results show that there is a weak relationship between the electoral performance and economic indicators. First, the interaction between government expenditures and party ideology is significant in three out of four models. Second, inflation is only significant when I assume that abstention is part of rewards and penalties, and I control by country. Therefore, the results are not significant enough to conclude that Deputy Chairman parties are accountable to voters because of the state of the economy.

I have undertaken the same analysis for the Ministry of Finance and Economy parties. Table 10 shows the empirical evidence. As in previous analysis, the four models have the same characteristics. These results are quite similar to previous outputs: I do not find significant relationship between electoral payoffs and the state of economy. Only GDP per capita increase is statistically significant and has the correct sign in models 1 and 2. The remaining variables are irrelevant or show the opposite expected outcomes – inflation in models 1, 2 and 3. Therefore, I may conclude that these parties are not either accountable to people because of their economic performance.

Finally, I analyze the electoral results of Ministry of Education and Ministry of Health parties. In these cases, I introduce some changes in the voting equation. Now, I replace government expenditure with relative expenditure increase on education and on health care – depending on the statistical model – as a share of GDP in two successive elections. The main idea is that voters may be particularly concerned

**Table 10** Ministry of Finance and Economy party. Analysis on electoral payoffs

| Variables | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Inflation | 0.122*** | 0.099** | 0.07* | 0.06 |
|  | (0.039) | (0.045) | (0.038) | (0.037) |
| GDP per capita | 0.272* | 0.236* | 0.145 | 0.119 |
|  | (0.142) | (0.099) | (0.116) | (0.079) |
| Unemployment | −0.309 | −0.174 | −0.172 | −0.159 |
|  | (0.276) | (0.182) | (0.204) | (0.147) |
| Left party | 0.302 | 0.62 | 0.404 | 0.539 |
|  | (1.032) | (0.602) | (0.905) | (0.571) |
| Government expenditures | 0.079 | 0.053 | 0.026 | 0.046 |
|  | (0.103) | (0.064) | (0.076) | (0.052) |
| Left* expenditures | 0.156 | 0.13* | 0.133 | 0.082 |
|  | (0.117) | (0.074) | (0.096) | (0.06) |
| $\lambda$ | 2.073** | 2.097*** | 1.411** | 1.567*** |
|  | (0.832) | (0.353) | (0.69) | (0.351) |
| Intercept | −9.709*** | −9.596*** | −7.137 | −7.4*** |
|  | (2.67) | (1.102) | (2.243) | (1.098) |
| N | 54 | 54 | 54 | 54 |
| n |  | 15 |  | 15 |
| $R^2$ | 0.266 |  | 0.189 |  |
| F | 2.97** |  | 2.67** |  |
| Wald $\chi^2$ |  | 90.32*** |  | 41.26*** |
| Method | OLS | FGLS | OLS | FGLS |

Significance levels *** 1% ** 5% *10%; robust standard error in brackets

about this part of the budget and take into account when they assign responsibilities to parties that hold these subjects. The rest of the variables do not change. In the Appendix, Table 16 summarizes the statistical results. Models 1, 2, 3 and 4 take as dependent variables the electoral payoffs of Ministry of Education arties and Models 5 and 6 deal with Ministry of Health party.[28] Of the economic variables, only growth in GDP per capita is significant in model 5. As regards the remaining economic variables, I cannot dismiss that their influence is zero. That is, I do not find significant relation between those economic indicators and the electoral results. Therefore, it seems to me that these parties are unaccountable too.

Up until this point, I have dealt with coalition parties as if they were independent actors and their electoral results were independent of their coalition partners. However, this is a strong assumption. Generally, the fate of coalition partners is related or, in some cases, develop together. This means that explanatory variable simultaneously affect incumbent parties. To put it mathematically, estimating the equations separately will waste the information that the same set of parameters appears in all functions. Seemingly unrelated regression (SURE) allow an estimation of this idea (Greene 2003: 339–377). Or, in mathematical language,

$$v_1 = \beta_1 X_1 + \delta_1 G_1 + \gamma_1 \lambda_1 + u_1$$
$$v_2 = \beta_2 X_2 + \delta_2 G_2 + \gamma_2 \lambda_2 + u_2$$
$$\dots\dots$$
$$v_M = \beta_M X_M + \delta_M G_M + \gamma_M \lambda_M + u_M$$

where M is the number of equations. The variables are the same in previous analyses and each equation deals with one party. Tables 11 and 12 show the empirical evidence for four parties: Prime Minister – (1), Deputy Chairman – (2), Ministry of Finance and Economy – (3) – and Ministry of Education – (4).[29] The Breusch–Pagan test of independence reveals that there is a strong correlation between random disturbance. Therefore, these for equations are related.

The results are similar to previous analyses. First, the electoral results of Prime Minister parties are explained by the following economic and political variables: economic growth, unemployment and the increase in government expenditure when a party is on the left. Second, explanatory variables are statistically non-significant for Deputy Chairman, Ministry of Finance and Economy and Ministry of Education parties.

In sum, if I simply consider the role of parties, in coalition governments, I shall be able to conclude that Prime Minister parties are the only members of the cabinet that voters hold accountable. In the remaining political formations, I am not able to

---

[28]Only models 2 and 4 introduce fixed effects. Another relevant difference between models is the dependent variable. Models 1, 3, 5 and 6 use as dependent variable the electoral payoffs among electors, whereas models 2 and 4 use the electoral payoffs among citizens.

[29]Unlike the other equations, this function takes as an independent variable the relative expenditure increase on education.

**Table 11** Analysis on the whole government (by portfolio) I

| Variables | Equation (1) | Equation (2) | Equation (3) | Equation (4) |
|---|---|---|---|---|
| Inflation | −0.14 | −0.103 | 0.188 | 0.172 |
| | (0.243) | (0.219) | (0.23) | (0.227) |
| GDP per capita | 0.272* | −0.165 | 0.224 | −0.099 |
| | (0.166) | (0.151) | (0.172) | (0.149) |
| Unemployment | −0.861** | −0.132 | 0.007 | −0.17 |
| | (0.335) | (0.274) | (0.279) | (0.284) |
| Left party | −1.038 | −0.102 | 0.903 | −1.84** |
| | (0.871) | (1.131) | (1.09) | (0.819) |
| Government expenditure | 0.086 | | | 0.009 |
| | (0.078) | | | (0.02) |
| Left * expenditure | 0.329*** | 0.114 | 0.179 | −0.001 |
| | (0.095) | (0.122) | (0.112) | (0.029) |
| $\lambda$ | 2.452** | 1.189 | 1.928* | 2.58** |
| | (1.188) | (1.076) | (1.11) | (1.127) |
| Intercept | −9.704*** | −4.076 | −9.742*** | −6.752** |
| | (3.459) | (3.117) | (3.259) | (3.231) |
| N | 34 | 34 | 34 | 34 |
| $R^2$ | 0.233 | 0.155 | 0.189 | 0.209 |
| Wald $\chi^2$ | 30.91*** | 4.53 | 9 | 12* |
| Method | OLS | OLS | OLS | OLS |

Significance levels *** 1% ** 5% *10%; robust standard error in brackets

**Table 12** Analysis on the whole government (by portfolio) II

| Variables | Equation (1) | Equation (2) | Equation (3) | Equation (4) |
|---|---|---|---|---|
| Inflation | −0.299 | −0.183 | 0.007 | −0.07 |
| | (0.193) | (0.185) | (0.184) | (0.17) |
| GDP per capita | 0.134 | −0.172 | 0.101 | −0.131 |
| | (0.132) | (0.127) | (0.136) | (0.112) |
| Unemployment | −0.689*** | −0.122 | −0.076 | −0.232 |
| | (0.263) | (0.231) | (0.224) | (0.213) |
| Left party | −0.833 | 0.006 | 0.551 | −1.714*** |
| | (0.643) | (0.91) | (0.781) | (0.626) |
| Government expenditure | 0.057 | | | 0.009 |
| | (0.058) | | | (0.015) |
| Left * expenditure | 0.251*** | 0.054 | 0.101 | −0.001 |
| | (0.07) | (0.099) | (0.083) | (0.022) |
| $\lambda$ | 1.816* | 0.764 | 1.512* | 2.001** |
| | (0.944) | (0.909) | (0.892) | (0.846) |
| Intercept | −7.842*** | −3.122 | −7.827*** | −5.681** |
| | (2.745) | (2.631) | (2.608) | (2.425) |
| N | 34 | 34 | 34 | 34 |
| $R^2$ | 0.255 | 0.131 | 0.139 | 0.244 |
| Wald $\chi^2$ | 31.14*** | 4.3 | 5.66 | 14.61** |
| Method | OLS | OLS | OLS | |

Significance levels *** 1% ** 5% *10%; robust standard error in brackets

identify significant relationships between economic indicators and electoral results. Therefore, if we classify coalition parties taking into account their type of portfolio, accountability only takes place in the case of Prime Minister parties. In the case of

Deputy Chairman, Ministry of Finance and Economy, Ministry of Education and Ministry of Health parties, I have not found empirical evidence that supports the idea that voters or citizens assign responsibilities to these parties in relation to economic performance.

These outcomes fit my theoretical model. One of my theoretical conclusions was that voters may focus on the most visible party, the Prime Minister party, and blame it because its incumbent performance. The empirical evidence confirms that statement. It may be also suggested that in coalition cabinets, accountability has been channeled through to Prime Ministers.

Moreover, these findings suggest that the literature has not been completely right. It seems that the process of assigning responsibilities to coalition governments is not simple. However, we may wonder: what happens if we change the criterion of party classification? In this subsection I have split parties according to their portfolios, but parties may be also be classified according to their size. Will results be the same? I shall answer this question in the next subsection.

### 4.2.2  The Size of Parties and Accountability

Another way to classify parties is to consider their weight in the government. Now, the criterion of classification is their size. In the following subsamples, I have divided parties in relation to their number of seats. Thus, for instance, the bigger parties are those that have more seats in the parliament among incumbent parties. I apply the same economic voting functions.

I start with the biggest coalition parties.[30] As in previous subsection, I have run four statistical models. Table 13 shows the results. Out of the explanatory variables, I only observe a relevant relationship between unemployment and electoral results – in model 3, it is not statistically significant. GDP per capita increase is just significant in model 2. Thus, economic performance has a weak influence in the electoral payoffs of these parties. The outputs are poorer that the results of single-party cabinets and Prime Minister parties. It seems that accountability does not work effectively with the biggest coalition parties.

How does accountability work for the second biggest parties of coalition cabinets? Table 14 shows the empirical evidence. In all models, I can only identify a weak relationship between the economy and electoral payoffs. The interaction between left parties and government expenditures is the only single variable that explains the electoral results of the second biggest parties. The remaining variables do not have any influence on the electoral payoffs. In conclusion, accountability foes not work properly on the second biggest parties of multiparty cabinets. In other words, I cannot

---

[30]One of the possible statistical problems could be that the biggest parties are Prime Minister parties too. Thus, the following results would be redundant because they would have been presented above. However, in my data base, among Prime Minister parties, 81.01% of them were the biggest. Therefore, I am not measuring exactly the same.

**Table 13** The biggest party in the coalition. Analysis on electoral payoffs

| Variables | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Inflation | −0.001 | −0.007 | −0.003 | −0.009 |
| | (0.017) | (0.015) | (0.016) | (0.011) |
| GDP per capita | 0.232 | 0.213* | 0.117 | 0.095 |
| | (0.195) | (0.111) | (0.158) | (0.078) |
| Unemployment | −0.543* | −0.535** | −0.332 | −0.351** |
| | (0.286) | (0.21) | (0.236) | (0.146) |
| Left party | −0.259 | 0.812 | −0.161 | 0.75 |
| | (1.45) | (0.898) | (1.224) | (0.727) |
| Government expenditure | 0.162 | 0.146** | 0.097 | 0.106** |
| | (0.132) | (0.07) | (0.114) | (0.05) |
| Left * expenditure | −0.065 | 0.032 | −0.019 | 0.06 |
| | (0.117) | (0.095) | (0.093) | (0.075) |
| $\lambda$ | 1.712 | 2.021*** | 1.08 | 1.281*** |
| | (1.191) | (0.494) | (1.125) | (0.424) |
| Intercept | −7.961** | −8.841*** | −5.785* | −6.394*** |
| | (3.616) | (1.576) | (3.424) | (1.369) |
| N | 70 | 70 | 70 | 70 |
| n | | 19 | | 19 |
| $R^2$ | 0.103 | | 0.049 | |
| F | 2.31** | | 1.92* | |
| Wald $\chi^2$ | | 32.73*** | | 24.4*** |
| Method | OLS | FGLS | OLS | FGLS |

Significance levels *** 1% ** 5% *10%; robust standard error in brackets

infer from a unique independent variable that those parties are accountable to voters.

We may wonder about the third and fourth biggest parties of coalition governments. In the Appendix, Table 17 shows the statistical results.[31] The findings are the same as in the previous analysis. I cannot identify any relevant statistical relationship between electoral results and economic indicators. Only inflation is statistically significant for the third biggest parties. Therefore, economic performance does not explain the electoral results of these parties.

Finally, as above, we may assume that the electoral results of these parties are related. To this point, I have analyzed coalition parties as if they were individual actors. However, this is an strong assumption. For that reason, I have applied Seemingly Unrelated Regression (SURE). But the Breusch–Pagan test of independence reveals that the estimate using SURE is not correct: residuals between equation are not correlated. For that reason, these results may be biased and I do not use them as part of my empirical evidence.

---

[31]Models 1 and 2 deal with the electoral payoffs of the third biggest parties and models 3 and 4 analyze fourth biggest parties. I use Ordinary Least Squares in all models. I leave out the spatial and time dimensions. The Breusch and Pagan test and the F test reveal that those dimensions are statistically unnecessary.

**Table 14** Second biggest party in coalition. Analysis on electoral payoffs

| Variables | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Inflation | 0.003 | 0.003 | −0.002 | −0.0004 |
| | (0.007) | (0.008) | (0.006) | (0.007) |
| GDP per capita | 0.145 | 0.051 | 0.1 | 0.0004 |
| | (0.122) | (0.076) | (0.113) | (0.06) |
| Unemployment | 0.222 | 0.203 | 0.169 | 0.169 |
| | (0.216) | (0.133) | (0.195) | (0.112) |
| Left party | −2.274** | −1.796*** | −1.699* | −1.767*** |
| | (1.1) | (0.545) | (0.962) | (0.456) |
| Government expenditure | −0.036 | −0.051 | −0.006 | −0.036 |
| | (0.07) | (0.038) | (0.058) | (0.034) |
| Left * expenditure | 0.267*** | 0.26*** | 0.178** | 0.144*** |
| | (0.087) | (0.056) | (0.075) | (0.043) |
| $\lambda$ | 1.86** | 1.928*** | 1.256 | 1.425*** |
| | (0.923) | (0.388) | (0.831) | (0.292) |
| Intercept | −6.447** | −6.849*** | −4.776* | −4.742*** |
| | (2.731) | (1.208) | (2.498) | (0.879) |
| N | 66 | 66 | 66 | 66 |
| n | | 19 | | 19 |
| $R^2$ | 0.188 | | 0.144 | |
| F | 3.38*** | | 2.62** | |
| Wald $\chi^2$ | | 62.22*** | | 51.15*** |
| Method | OLS | FGLS | OLS | FGLS |

Significance levels *** 1% ** 5% *10%; robust standard error in brackets

In sum, accountability does not work if we classify parties taking into account their weight in the government. If accountability is a question of economic outputs and electoral results, people may not assign responsibilities to those coalition parties. Thus, the empirical evidence presented in this subsection confirms what literature says. Moreover, these outcomes fit my theoretical arguments. I argue that accountability is a question about tasks and not about size.

# 5 Conclusion

This chapter originated from a simple question: how do elections work when people face divided power? For a long time, philosophers and social scientist have concluded that this situation is a challenge for democracy. They argued that this institutional framework does not permit clarity of responsibility and, therefore, citizens may not be able to control politicians. Thus, if accountability is a key feature of democracy, divided power would not allow politicians to be responsible for their policies and performance. I argued that this hypothesis, termed by scholars as 'clarity of responsibility', was not developed theoretically and the empirical evidence was poor. Therefore, this chapter aimed to fill these academic gaps, focusing on coalition governments as an example of divided power.

I argued that in order to define accountability correctly it is very important to consider parties as the key actor. "Accountability is individual rather than collective" said Elster (Przeworski et al. 1999: 255). To put is another way, responsibility is a question about parties, not about governments. This assumption contrasts with studies to date. Scholars have dealt with the question of 'clarity of responsibility' by treating governments as single actors. But making this theoretical mistake, they do not do any intragovernmental analysis.

This chapter has resolved these problems. Thus, I have presented the causal mechanisms of accountability, using incumbent parties as unit of analysis. As I developed, coalition governments are as accountable as single party government. Hence, we cannot argue that multiparty cabinets are unaccountable. We observed that the state of the economy significantly affected the electoral outcomes of Prime Minister parties. Thus, these parties assume the responsibility for the incumbent performance. Nevertheless, I did not observe the same results for the biggest parties of the coalition. This confirms that accountability is a question about task, not one of size.

In sum, the literature was mistaken. The empirical evidence provided by this chapter corrects the hypothesis of 'clarity of responsibility' and concludes that multiparty cabinets are accountable and that voters may control them.

# Appendix

See Tables 15–17.

**Table 15** Selection equation

| Variables | Single-party governments |
|---|---|
| Proportional | −0.706*** |
| | (0.142) |
| Mixed | −1.22*** |
| | (0.226) |
| Multilevel | −0.689*** |
| | (0.154) |
| ENEP | −0.532*** |
| | (0.06) |
| Intercept | 2.279*** |
| | (0.237) |
| N | 689 |
| $R^2$ | 0.2485 |
| Wald $\chi^2$ | 158.55 |
| Method | Probit |

**Table 16** Ministry of Education and Ministry of Health. Analysis on electoral payoffs

| Variables | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Inflation | 0.09 | 0.068 | 0.042 | 0.018 | 0.78 | 0.669 |
| | (0.074) | (0.057) | (0.068) | (0.048) | (1.633) | (1.488) |
| GDP per capita | 0.091 | −0.028 | 0.028 | −0.091 | 1.032 | 0.781 |
| | (0.156) | (0.07) | (0.128) | (0.057) | (0.563) | (0.516) |
| Unemployment | −0.028 | −0.138 | 0.014 | −0.108 | 1.13 | 0.889 |
| | (0.212) | (0.127) | (0.173) | (0.113) | (1.448) | (1.285) |
| Left party | −1.013 | −1.664** | −1.008 | −1.216* | 1.055 | 0.69 |
| | (1.302) | (0.79) | (1.069) | (0.707) | (3.48) | (3.25) |
| Expenditures on education | −0.034 | −0.02 | −0.022 | −0.012 | | |
| | (0.029) | (0.013) | (0.026) | (0.009) | | |
| Expenditures on health | | | | | −0.091 | 0.034 |
| | | | | | (0.359) | (0.342) |
| Left* exp. on education | 0.024 | 0.002 | 0.022 | 0.002 | | |
| | (0.037) | (0.021) | (0.033) | (0.02) | | |
| Left* Exp. on health | | | | | −0.364 | −0.483 |
| | | | | | (0.283) | (0.274) |
| $\lambda$ | 1.615 | 2.048*** | 0.864 | 1.255** | −3.86 | −3.114 |
| | (1.263) | (0.704) | (1.071) | (0.638) | (2.63) | (2.561) |
| Intercept | −5.859 | −6.01*** | −3.753 | −3.912** | 2.518 | 2.115 |
| | (3.292) | (2.113) | (2.713) | (1.915) | (9.772) | (9.475) |
| N | 56 | 56 | 56 | 56 | 19 | 19 |
| n | | 15 | | 15 | | |
| $R^2$ | 0.096 | | 0.048 | | 0.453 | 0.36 |
| F | 1.58 | | 0.75 | | 8*** | 2.83* |
| Wald $\chi^2$ | | 19.37*** | | 14.32*** | | |
| Method | OLS | FGLS | OLS | FGLS | OLS | OLS |

Significance levels *** 1% ** 5% *10%; robust standard error in brackets

**Table 17** Third and fourth biggest party in coalition. Analysis on electoral payoffs

| Variables | Third biggest party | | Fourth biggest party | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| Inflation | −0.009* | −0.008* | 0.004 | 0.005 |
| | (0.005) | (0.004) | (0.012) | (0.009) |
| GDP per capita | 0.133 | 0.118 | 0.102 | −0.001 |
| | (0.101) | (0.081) | (0.197) | (0.138) |
| Unemployment | −0.124 | −0.14 | −0.336 | −0.158 |
| | (0.126) | (0.108) | (0.401) | (0.258) |
| Left party | 0.359 | −0.098 | 0.048 | 0.873 |
| | (0.875) | (0.676) | (1.882) | (1.203) |
| Expenditures | 0.068 | 0.064 | −0.031 | −0.052 |
| | (0.046) | (0.038) | (0.112) | (0.089) |
| Left* expenditures | 0.048 | 0.035 | 0.172 | 0.124 |
| | (0.068) | (0.052) | (0.162) | (0.103) |
| $\lambda$ | −0.978 | −0.824 | −0.39 | −1.666 |
| | (1.484) | (0.827) | (2.422) | (1.509) |
| Intercept | 1.827 | 1.435 | 1.425 | 4.902 |
| | (3.048) | (2.561) | (6.855) | (4.311) |
| N | 41 | 41 | 27 | 27 |
| $R^2$ | 0.079 | 0.103 | 0.069 | 0.109 |
| F | 1.93* | 1.45 | 0.23 | 0.56 |
| Method | OLS | OLS | OLS | OLS |

Significance levels *** 1% ** 5% *10%; robust standard error in brackets

# References

Adsera AC, Boix C, Payne M (2003) Are you being served? Political accountability and quality of government. J Law Econ Organ 19:445–490

Anchen CH, Bartels LM (2004) Musical chairs: pocket-book voting and the limits of democratic accountability. In: Annual Meeting of the American Political Science Association, Cambridge, IL

Anderson CJ (2000) Economic voting and political context: a comparative perspective. Elect Stud 19:151–170

Austen-Smith D (1992) Strategic models of talk in political decision making. Int Polit Sci Rev 13:45–58

Austen-Smith D, Banks J (1988) Elections, coalitions and legislative outcomes. Am Polit Sci Rev 82:405–422

Austen-Smith D, Banks J (1989) Electoral accountability and incumbency. In: Ordershook PC (ed) Models of strategic choice in politics. University of Michigan Press, Ann Arbor, MI

Baltagi BH (2005) Econometric analysis of panel data, 3rd edn. Wiley, London

Barreiro B (2007) Explaining electoral performance of incumbents in democracies. In: Maravall JM, Sánchez-Cuenca I (eds) Controlling governments: voters, institutions and accountability. Cambridge University Press, Cambridge

Barro RJ (1973) The control of politicians: an economic model. Public Choice 14:19–42

Bengtsson A (2004) Economic voting: the effect of political context, volatility and turnout on voters' assignment of responsibility. Eur J Polit Res 43:749–767

Berganza JC (2000) Two roles for elections: discipline the incumbent and selecting a competent candidate. Public Choice 105:165–193

Breen R (1996) Regression models. Censored, sample selected and truncated data. Sage, Thousand Oaks, CA

Caramani D, Flora P, Kraus F, Rothenbacher F (2000) Elections in Western Europe since 1815: electoral results by constituencies. Macmillan, London

Castilla EJ (1998) Análisis dinámico. Centro de Investigaciones Sociológicas, Madrid

Chapel HW, Keech WR (1985) A new view of political accountability for economic performance. Am Polit Sci Rev 79:10–27

Duch RM, Stevenson R (2005a) Context and the economic voting: a multilevel analysis. Polit Anal 13:387–409

Duch RM, Stevenson R (2005b) Assessing the magnitude of the economic vote over time and across nations. Elect Stud 25:528–547

Ferejohn J (1986) Incumbent performance and electoral control. Public Choice 50:5–25

Fiorina MP (1981) Retrospective voting in American national elections. Yale University Press, New Haven, CT

Fisher SD, Hobolt SB (2010) Coalition government and electoral accountability. Elect Stud 29:358–369

Golder SN (2006) Pre-electoral coalition formation in parliamentary democracies. Br J Polit Sci 36:193–212

Greene WH (2003) Econometric analysis. Prentice Hall, Upper Saddle River, NJ

Heckman JJ (1974) Shadow prices, market wages and labor supply. Econometrica 42:679–694

Heckman JJ (1979) Sample selection bias as a specification error. Econometrica 47:153–162

Hsiao C (2003) Analysis of panel data. Cambridge University Press, Cambridge

Kramer GH (1983) The ecological fallacy revisited: aggregated-versus individual level findings on economic and elections and sociotropic voting. Am Polit Sci Rev 77(1):92–111

Laffont J-J, Martimort D (2002) The theory of incentives. The principal-agent model. Princeton University Press, Princeton, NJ

Lewis-Beck MS (1986) Comparative economic voting: Britain, France, Germany and Italy. Am J Polit Sci 30(2):315–346

Lewis-Beck MS (1988) Economics and elections: the major western democracies. University of Michigan Press, Ann Arbor, MI

López-Nava K (2007) Democractic accountability for economic performance: information and institutional dynamics in economic voting across democracies, 1970–2005. Ph.D. thesis, Stanford University

Macho I, Pérez D (2005) Introducción a la Economía de la Información, 2nd edn. Ariel, Barcelona

Mackie TT, Rose R (1982) International almanac of electoral history. Facts on file. New York

Mackie TT, Rose R (1997) A decade of election results: updating the international almanac. Center for the Study of Public Policy, University of Strathclyde, Glasgow

Maravall JM (2007) Accountability and the survival of the governments. In: Boix C, Stokes S (eds) The Oxford handbook of comparative politics. Oxford University Press, Oxford

Maravall JM, Przeworski A (1999) Reacciones políticas a la economía. Revista Española de Investigaciones Sociológicas 87:11–52

Mershon C (1996) The cost of coalition: coalition theories and Italian governments. Am Polit Sci Rev 90:534–554

Mershon C (2002) The cost of coalition. Stanford University Press, Stanford, CA

Montabe J (1997) El Gobierno. In: Martínez MA (ed) Política y Gobierno en España. Tirant lo Blanch, Valencia

Nadeu R, Niemi RG, Yoshinaka A (2002) A cross-national analysis of economic voting: taking into account of the political context across time and nations. Elect Stud 21:403–423

Norpoth H (2001) Divided government and economic voting. J Polit 63:414–435

Norpoth H, Lewis-Beck MS, Lafay J-D (1991) Economic and politics: the calculus of support. The University of Michigan Press, Ann Arbor, MI

Peltzman S (1990) How efficient is the voting market? J Law Econ 33:27–63

Persson T, Tabellini G (2000) Political economics. Basic, New York

Pitkin HF (1967) The concept of representation. University of California Press, Berkeley, CA

Powell GB (2000) Elections as instruments of democracy: majoritarian and proportional visions. Yale University Press, New Haven, CT

Powell GB, Whitten GD (1993) A cross-national analysis of economic voting: taking into account of the political context. Am J Polit Sci 37:391–414

Przeworski A (2003) States and markets. A primer in political economy. Cambridge University Press, Cambridge

Przeworski A (2007) Is the science of comparative politics possible? In: Boix C, Stokes S (eds) Oxford Handbook of comparative politics. Oxford University Press, Oxford

Przeworski A, Stokes SC, Manin B (1999) Democracy, accountability and representation. Cambridge University Press, Cambridge

Royed FJ, Leyden KM, Borrelli SA (2000) Is 'clarity of responsibility' important for economic voting? Revisited Powell and Whitten's hypothesis. Br J Polit Sci 30:669–685

Sánchez-Cuenca I, Barreiro B (2000) Los efectos de la acción de gobierno en el voto durante la etapa socialista (1982–1996). Centro de Investigaciones Sociológicas, Madrid

Stokes S (1996) Public opinion and market reforms: the limits of economic voting. Comp Polit Stud 29:499–519

Strom K (1990) Minority governments and majority rules. Cambridge University Press, Cambridge

Strom K, Bergman T, Muller WC (2003) Delegation and accountability in parliamentary democracies. Oxford University Press, Oxford

Tucker JA (2006) Regional economic voting. Cambridge University Press, Cambridge

Whitten GD, Palmer HD (1999) Cross-national analyses of economic voting. Elect Stud 18(1):49–67

Wilkin S, Haller B, Norpoth H (1997) From Argentina to Zambia: a world-wide test of economic voting. Elect Stud 16(3):301–316

Woldendorp JH, Keman H, Budge I (1998) Party government in 20 democracies: an update (1990–1995). Eur J Polit Res 33:125–164

# Part II
# Democracy and Voting

# Empirical and Formal Models of the United States Presidential Elections in 2000 and 2004

**Norman Schofield, Christopher Claassen, Maria Gallego, and Ugur Ozdemir**

## 1 Introduction

The formal literature on two party electoral competition has typically been based on the assumption that parties or candidates adopt positions in order to win, and has inferred that parties will converge to the electoral median, under deterministic voting in one dimension (Downs 1957) or to the electoral mean in stochastic models.[1] An early empirical paper by Poole and Rosenthal (1984) provided some evidence that candidates in US presidential elections did not converge.

In the standard spatial model, only candidate *positions* matter to voters. However, as Stokes (1963, 1992) has emphasized, the non-policy evaluations, or *valences*, of candidates by the electorate are equally important. In empirical models, a party's valence is usually assumed to be independent of the party's position, and adds to the statistical significance of the model. In general, valence reflects the overall degree to which the party is perceived to have shown itself able to govern effectively in the past, or is likely to be able to govern well in the future (Penn 2009).

An extensive literature has developed over the last decade that considers deterministic or probabilistic voting models including valence or bias towards one or other of the candidates.[2]

---

[1]See the earlier work by Enelow and Hinich (1982, 1989), Erikson and Romero (1990) and more recent work by Duggan (2006), McKelvey and Patty (2006) and Patty et al. (2009).

[2]Ansolabehere and Snyder (2000), Groseclose (2001), Aragones and Palfrey (2002, 2005), Adams and Merrill (2002, 2005), Banks and Duggan (2005), Ashworth and Bueno de Mesquita (2009), Jessee (2009, 2010), Zakharov (2009), Serra (2010).

N. Schofield (✉), C. Claassen, and U. Ozdemir
Washington University in Saint Louis, St. Louis, MO, USA
e-mail: schofield.norman@gmail.com; classen@wustl.edu; uozdemir@wustl.edu

M. Gallego
Washington University in Saint Louis, St. Louis, MO, USA
and
Department of Economics, Wilfrid Laurier University, Waterloo, ON, Canada
e-mail: mgallego@wlu.ca

This chapter offers a general model of elections based on the assumption that valence can be measured in a number of ways. The first kind is a fixed or *exogenous valence*, which for a party $j$ is denoted $\lambda_j$. As in empirical work, we assume that $\lambda_j$ is held constant at the time of an election, and so is independent of the party's position. Exogenous valence can be estimated as the intercept term in a stochastic model. Earlier work (Schofield and Sened 2006) has shown that, in models involving exogenous valence, if the valence differences are sufficiently large, then vote maximizing parties will not converge.

Here we construct stochastic models of the 2000 and 2004 US presidential elections involving exogenous valence and show that the valence of the two candidates were similar enough so that the unique Nash equilibrium was one where both candidates converge to the electoral origin in order to maximize vote share.[3]

Recent empirical work by Clarke et al. (2009a, b) has analyzed recent US presidential elections and British general elections. These works have shown that valence, as measured by the perceptions of the character traits of the candidates, or of party leaders, is a key element of these elections.[4]

In the empirical analysis we show that a voter's perception of each candidate's traits has a very significant impact on the probability that the voter chooses one candidate or the other. The simulation of the spatial model, based on both position and character traits, allows us to estimate what we call *Local Nash equilibria (LNE)*[5] to the vote maximizing game.[6] The LNE of this traits model is slightly perturbed from the electoral origin, so that the two candidates are located at the same position, slightly to the right on the economic axis, and at a neutral position on the social axis. These equilibrium positions differ from the estimated positions of the two candidates.

In order to account for the discrepancy between the estimated positions and the positions obtained by equilibrium analysis, we introduce a different kind of valence known as *activist valence*. When candidate $j$ adopts a policy position $z_j$, in the policy space, $X$, then the activist valence of the party is denoted $\mu(z_j)$. Implicitly we adopt a model originally due to Aldrich (1983). In this model, activists provide crucial resources of time and money to their chosen candidate, and these resources are dependent on the candidate position.[7] The candidate can then use these

---

[3]The empirical analyses were based on the 2000 and 2004 American National Election Surveys (ANES).

[4]See also Clarke et al. (2005). Jesee (2009, 2010) has also examined partisan bias in the 2004 and 2008 Presidential elections.

[5]We focus on *local equilibria* because we consider that candidates will only be able to make small adjustments to their policy statements as the election nears.

[6]We focus on *vote maximizing* rather than maximizing the *probability of winning* because the former model is linear and would seem to more closely characterize the likely behavior of candidates adapting to electoral information obtained from polls and the like. As Patty (2002, 2007) has shown, these two classes of models differ in the equilibria.

[7]For convenience, it is assumed that $\mu(z_j)$ is only dependent on $z_j$, and not on $z_k$, $k \neq j$, but this is not a cucial assumption.

resources to enhance the candidate's image before the electorate, thus affecting the candidate's overall valence.

Moreover, because activist support is denominated in terms of time and money, it is reasonable to suppose that the activist function will exhibit decreasing returns. We point out that when these functions are sufficiently "concave" with respect to candidate positions,[8] then the activist vote maximizing model will exhibit a Nash equilibrium.

The difference we find between the estimated positions of the two candidates and those inferred to be equilibria from the full trait model gives us an estimate for the influence of activists.

Empirical analysis of the 2000 and 2004 US elections suggests that party activists tend to have more extreme policy positions than the typical voter. The problem for each candidate is that by choosing a position to maximize activist support, the candidate loses centrist voters and by choosing to be closer to centrist voters the candidate can loose activist support. The candidate must determine the trade-off between attracting resources from activists and appealing to the voters. This trade-off is captured by the "optimal marginal condition" that maximizes vote share. This is given as a (first order) *balance condition*.

Grossman and Helpman (1996), in their game theoretic model of activists, consider two distinct motives for interest groups:

> Contributors with an *electoral motive* intend to promote the electoral prospects of preferred candidates, [while] those with an *influence motive* aim to influence the politicians' policy pronouncements.

In our first activist model the term $\mu_j(z_j)$ influences every voter and thus contributes to the electoral motive for candidate $j$. In addition, the candidate must choose a position to balance electoral and activist support.

We argue that the influence of activists on the two candidates can be characterized in terms of activist gradients. For the two candidates, these gradients point into opposite quadrants of the policy space. We also obtained information from the American National Election Surveys on activists, namely those who contributed resources to one or other of the two parties. The mean positions of the two sets of party activists were shown to be compatible with our estimated party activist gradients.

Because each candidate is supported by multiple activists, we extend the activist model by considering a family of potential activists, $\{A_j\}$ for each candidate, $j$, where each $k \in A_j$ is endowed with a utility function, $U_k$, which depends on candidate $j$'s position $z_j$, and the preferred position of the activist. The resources allocated to $j$ by $k$ are denoted $R_{jk}(U_k(z_j))$. Let $\mu_{jk}(R_{jk}(U_k(z_j)))$ denote the effect that activist $k$ has on voters' utility. Note that the activist valence function for $j$ is the same for all voters. With multiple activists, the *total activist valence function* for agent $j$ is the linear combination

---

[8]We mean by this that the appropriate Hessians have negative eigenvalues of sufficient magnitude.

$$\mu_j(z_j) = \sum_{k \in A_j} \mu_{jk}(R_{jk}(U_k(z_j))).$$

Bargains between the activists supporting candidate $j$ then gives a *contract set* of activist support for candidate $j$, and this contract set can be used formally to determine the balance locus, or set of optimal positions for each candidate. This balance locus can then be used to analyze the pre-election contracts between each candidate and the family of activist support groups.

Consider now the situation where these contracts have been agreed, and each candidate is committed to a set of feasible contracts as outlined in Grossman and Helpman (1994, 1996, 2001). Suppose further that the activists have provided their resources. Then at the time of the election the effect of this support is incorporated into the empirical estimates of the various exogenous, sociodemographic and trait valences. Consequently, when we estimate these valences we also estimate the aggregate activist influence. The estimated positions of the candidates can then be regarded as incorporating policy preferences of the activists. Electoral models where candidates have policy positions, as proposed by Wittman (1977), Calvert (1985), Duggan and Fey (2005), and Duggan (2006) implicitly assume that candidates would be willing to accept defeat because of an adherence to particular policy positions.

We argue that it is more plausible that the estimated positions of the candidates are the result of maximizing candidate utility functions that balance the electoral consequences of position-taking with the necessity of obtaining activist resources to contest the election. This calculation requires an estimate of the degree to which these resources will influence the perceptions that the electorate has of the various valences associated with the model.

In the final version of the model we allow the activist valence function to be individual specific. The total resources available to candidate $j$ are now denoted $\mathbf{R}_j(z_j)$, and these may be allocated to individuals, with resource $m_{ij}$ targeted on voter, or "voter class", $i$ by candidate $j$. Since $m_{ij}$ will depend on $z_j$, we write this allocation as $m_{ij}(z_j)$, so the budget constraint is

$$\mathbf{R}_j(z_j) = \sum_{k \in A_j} R_{jk}(U_k(z_j))$$
$$= \sum_{i \in N} m_{ij}(z_j).$$

The optimization problem is now a more complex one, subject to this constraint. In actual fact candidates will generally not allocate resources to individuals *per se*, but to voter classes via media outlets in different regions, or "zip codes". Indeed, much of the action in political campaigns is concerned with the analysis of local data so as to determine how voters might be targeted in an optimal fashion. For example, the logit models in this chapter give sociodemographic analyses that would, in principle allow for the targeting of specific groups in the polity.

The general balance condition presented in the Technical Appendix specifies how these resources should be allocated throughout the polity.

A recent literature on elections has focussed on the effects of campaign expenditure on US election results (Coate 2004).[9] Herrera et al. (2008) suggest that electoral volatility forces candidates to spend more, while Ashworth and Bueno de Mesquita (2009) suppose that candidates buy valence so as to increase their election chances. Meirowitz (2008) notes that

> candidates and parties spending this money thought that it would influence the election outcome. Downsian models of competition cannot explain how candidates choose spending campaign levels or what factors influence these decisions.

Meirowitz proxies the choice of expenditure in terms of candidate choice of effort, but his model does not explicitly deal with the budget question.

Ansolabehere et al. (2003) provide an empirical analysis of Congressional and Presidential election campaign contributions up to 2000. They note that candidates, parties and organizations raised and spent about $3 billion in the 1999–2000 election cycle. However, the federal government at that time spent about $2 trillion, so the prize from influencing politics was of considerable value. They suggest that so little is spent in contributions relative to the possible gains because contributions are a consumption good, rather than an investment good. However, they do observe that the electoral motive is not insignificant: they suggest that the marginal impact of $100,000 spent in a House race is about 1% in vote share.

The essence of the model presented here is that it attempts to endogenize the question of the resource budget of candidates since the total resources used by candidates in seeking election victory come from the *implicit contracts* they can make with their supporting activists.[10] The activists must solve their own optimization problem by estimating the benefit they receive from the electoral and influence motives, in deciding what resources to make available to their chosen candidate.

Essentially there is an arms race between candidates over these resources due to a feedback mechanism between politics and economics. As the outcome of the election becomes more important, activists become increasingly aware that the resources they provide have become crucial to election victories, and they become more demanding of their chosen candidates. Because of the offer of resources, candidates are forced to move to more radical positions, and polarization in candidate opositions increases, even though there may be little change in the degree of polarization of the electorate.

In the conclusion we suggest that the results presented here lend some support to the activist model proposed by Miller and Schofield (2003, 2008) and elaborated in Schofield and Miller (2007). Changes in voter choice appear to result not only from changes in the distribution of electoral preferences, but from the shifts in electoral perceptions. In turn, these changes are the consequence of the shifting pattern of

---

[9]An earlier paper by Groseclose and Snyder (1996) looked at vote buying, but in the legislature.

[10]Snyder and Ting (2008) also consider the contracts between activists and candidates but assume that the policy space is one dimensional.

activist support for the candidates. Since the importance of electoral contributions has increased, this has enhanced the influence of activist groups.[11] The empirical and formal models presented here give a reason why electoral politics has become very polarized in the United States.[12] This polarization appears to have benefited the wealthy in society and may well account for the increase in inequality in income and wealth distribution that has occurred over the last decade. (Hacker and Pierson 2006, 2010; Pierson and Skocpol 2007).

As Miller and Schofield (2008) have argued, over the long run the coalition structure among activist groups for the parties will shift in response to exogenous shocks, leading to a shift in the activist coalitions. This may be the cause of the slow realignment that appears to have occurred over many decades in U.S. politics.[13]

In the next section we present the empirical methodology that was used, together with the computation method to find equilibria. Section 3 draws some conclusions from the analysis, including a number of inferences from the model relating to Madison's argument about the "probability of a fit choice" in the Republic. Section 4 is a Technical Appendix that gives the details of the spatial model that we deploy.

## 2 Methodology: Spatial Models of the 2000 and 2004 Elections

### 2.1 The 2000 Election

To construct a model of the 2000 election, we used survey data from the 2000 American National Election Study (ANES 2000). The survey is a nationally representative sample of the voting age population, with 1555 pre- and post-election respondents in 2000. The first step was to build up a map of the policy space and to assign each surveyed individual a two-dimensional ideal point on that space. Following Schofield et al. (2003), we constructed a two dimensional policy space based on economic and social issues. Exploratory factor analysis led to the ten survey items, reported in the Data Appendix from which two factors were extracted. The factor loadings per item are given in Table 1. Figure 1 gives a smoothing of the estimated voter distribution. Essentially, left on the economic ($x$) axis in this figure is pro-redistribution. The second social axis ($y$) is determined by attitudes to abortion and gays, so we interpret north on this axis to be in support of certain civil rights. Figure 2 gives a perspective plot of the electoral distribution.

Respondent's partisan choice was measured with the following question:

"Who do you think you will vote for in the election for President?"

---

[11]Indeed, Herrera et al. (2008) observe that spending by parties in federal campaigns went from 58 million dollars in 1976 to over 1 billion in 2004 in nominal terms.

[12]See the works by Fiorina et al. (2005), Fiorina and Abrams (2009) and McCarty et al. (2006) on polarization in the electorate and Layman et al. (2010) on polarization among activists. Schofield et al. (2011) gives similar results for the 2008 election.

[13]See also the earlier work by Sundquist (1973).

**Table 1** Factor loadings from the American national election survey, 2000

| Question | Social policy | Economic policy |
|---|---|---|
| 1. Economic problems | 0.02 | 0.32 |
| 2. Federal spending | 0.09 | 0.36 |
| 3. Equality | 0.21 | 0.50 |
| 4. African American | 0.15 | 0.46 |
| 5. Immigrants | 0.08 | 0.31 |
| 6. Liberal vs conservative | 0.38 | 0.33 |
| 7. Guns | 0.17 | 0.34 |
| 8. Abortion | 0.51 | 0.02 |
| 9. Gays | 0.65 | 0.18 |
| 10. Family | 0.44 | 0.24 |
| % Variance explained | 11.2 | 11.0 |
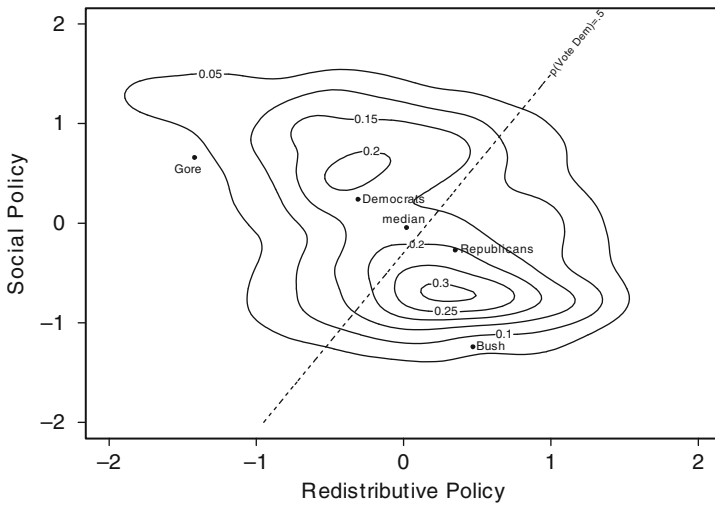


**Fig. 1** Contour plot of the voter distribution in 2000 with the equiprobable cleavge line

Activism was measured with the following question:

"During an election year people are often asked to make a contribution to support campaigns. Did you give money to an individual candidate running for public office?" OR

"Did you give money to a political party during this election year?"

An activist is thus operationally defined here as someone who claimed to donate money to either a candidate or a party. Activists were then coded as Republican or Democrat depending on the party or partisan affiliation of the candidate to which they gave money. Of the sample, 4.5% ($n = 70$) were Republican activists, and 2.9% ($n = 47$), Democrat activists, in 2000. As Table 2 shows, the mean Democratic partisan position was $\left(x_{dem}^{part}, y_{dem}^{part}\right) = (-0.31, 0, 24)$ with standard error (0.029, 0.03) while the activist mean was $\left(x_{dem}^{act}, y_{dem}^{act}\right) = (-0.54, 0.48)$ with standard error (0.10, 0.10). For the Republican partisans we find $\left(x_{rep}^{part}, y_{rep}^{part}\right) = (0.36, -0.27)$ with standard error (0.027, 0.03) while the activist mean was
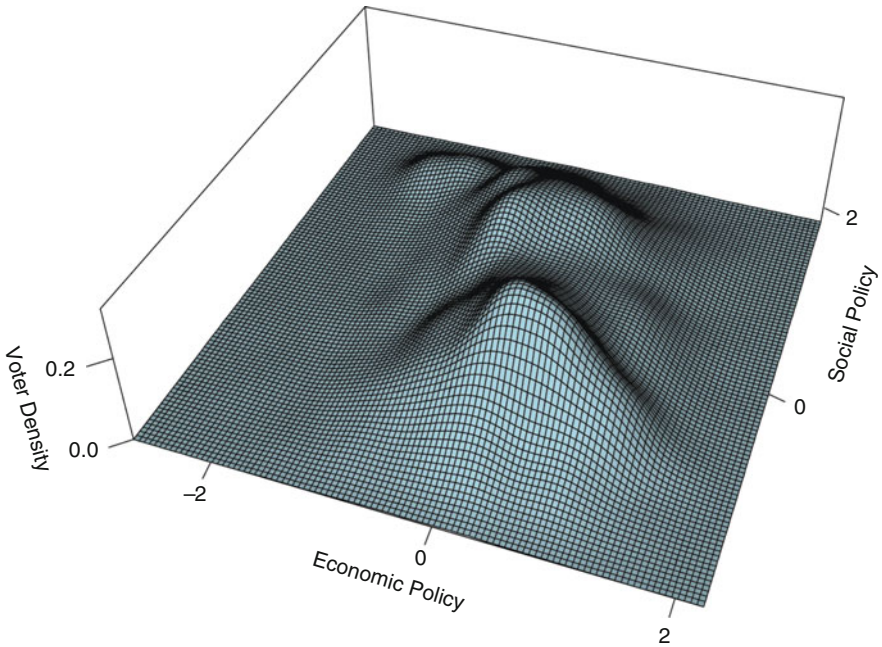
**Fig. 2** Perspective plot of the voter distribution in 2000

**Table 2** Descriptive data 2000

|               | Econ mean | Policy std. err. | C.I | Social mean | Policy std. err. | C.I | $n$ |
|---------------|-----------|------------------|-----|-------------|------------------|-----|-----|
| Activists     |           |                  |     |             |                  |     |     |
| Democrats     | −0.54     | 0.10             | [−0.74,−0.34] | 0.48 | 0.10 | [+0.28,+0.68] | 47 |
| Republicans   | 0.42      | 0.07             | [+0.28,+0.56] | −0.30 | 0.08 | [−0.46,−0.14] | 70 |
| Non-activists |           |                  |     |             |                  |     |     |
| Democrats     | −0.31     | 0.029            | [−0.34,−0.28] | 0.24 | 0.03 | [+0.18,+0.30] | 634 |
| Republicans   | 0.36      | 0.027            | [+0.3,+0.42] | −0.27 | 0.03 | [−0.33.−0.21] | 536 |

$\left( x_{rep}^{act}, y_{rep}^{act} \right) = (0.42, -0.30)$, with standard error $(0.07, 0.08)$. The 95% confidence intervals on these estimates give some weak evidence that the activist mean positions are more extreme than the partisan mean positions. Figure 3 shows the mean activist and voter positions, and standard error bars with the voter bars the smaller.

The positions of the major presidential candidates, Bush and Gore, in 2000 were estimated in a similar fashion to that of the sampled individuals. These estimated responses of the candidates are given in Table 3. Scores were assigned to each candidate for each of the constituent survey items based on press reports of their

**Fig. 3** Comparison of mean
partisan and activist positions
in 2000
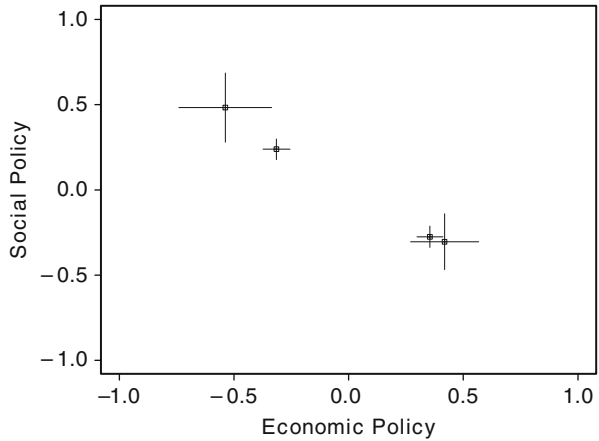


**Table 3** Gore and Bush
estimated responses, 2000

| Question | Bush | Gore |
|---|---|---|
| 1. Economic problems | 3 | 1 |
| 2. Federal spending | 3 | 2 |
| 3. Equality | 3 | 1 |
| 4. African American | 4 | 2 |
| 5. Immigrants | 3 | 3 |
| 6. Liberal vs conservative | 3 | 1 |
| 7. Guns | 2 | 1 |
| 8. Abortion | 3 | 2 |
| 9. Gays | 3 | 1 |
| 10. Family | 5 | 3 |
| Estimated position: economic policy | +0.47 | −1.42 |
| Estimated position: social policy | −1.24 | +0.66 |

campaign stances on various issues. The factor analysis was used to obtain estimated
policy positions $z_{gore} = (x_{Gore}, y_{Gore}) = (-1.42, 0.66)$, and $z_{Bush} = (x_{Bush}, y_{Bush})$
$= (0.47, -1.24)$. Notice that the candidate positions are more extreme on both axes
than the partisan and activist positions. In particular Gore's position $x_{Gore}$ is well
outside the confidence intervals of both partisans and activists on the economic
axis, while Bush's position $y_{Bush}$ lies outside the confidence intervals on the social
axis. This suggests that both candidates were influenced by more radical activists.

The contour plot of Figure 1 includes an estimated cleavage line dividing likely
Democrat candidate voters from Republican candidate voters. This partisan cleav-
age line was derived from a standard binomial logit model, designed to test the
effects of each policy dimension on vote choice. We do not report the full results of
the positional model here.

Because this logit model involves the preferred positions of voters, we refer to it
as a *pure positional model*. Note however that the position of each candidates is
implicit, rather than explicit in this model. Our estimates of the log-likelihood,
Akaike information criterion (AIC) and Bayesian information criterion (BIC) were

all quite acceptable, and all coefficients were significant with probability $< 0.01$. A voter $i$, with preferred position $(x_i, y_i)$ is estimated to vote Republican with probability

$$\rho_{rep} = \frac{[\exp(\lambda_r + bx_i + cy_i]}{1 + [\exp(\lambda_r + bx_i + cy_i]}. \tag{1}$$

The estimated coefficients in this model are $(\lambda_r, b, c) = (-0.24, 1.30, -0.70)$.

These empirical results suggest that economic policy (the *x-axis*) is a more salient dimension than social policy (the *y-axis*) in modeling vote choice for the Republican candidate.

According to this model, any voter with preferred point lying on the cleavage line has equal probability of picking one or other of the candidates. This cleavage line is given by the equation

$$y = 1.87x - 0.34. \tag{2}$$

Note that the cleavage line is very similarly positioned to the cleavage lines for the 1964 and 1980 elections, estimated by Schofield et al. (2003). Although this positional model has very significant coefficients, it takes the positions of Bush and Gore as exogenous, and so cannot be used to estimate the vote maximizing positions. We now present the formal stochastic model based on estimates of the candidate positions, and use this to explain the data just presented.

The pure spatial model, $\mathbb{M}(\lambda, \beta)$, is based on the voter utility assumption

$$u_{ij}(x_i, z_j) = \lambda_j - \beta \lVert x_i - z_j \rVert^2 + \varepsilon_j \tag{3}$$

$$= u_{ij}^*(x_i, z_j) + \varepsilon_j. \tag{4}$$

Here $u_{ij}^*(x_i, z_j)$ is the observable and $\{\varepsilon_j\}$ denote Type I extreme value errors, as discussed in Sect. 4.1. At a vector, $\mathbf{z}$, the probability that a voter $i$ chooses candidate $j$ is:

$$\rho_{ij}(\mathbf{z}) = \Pr[[u_{ij}(x_i, z_j) > u_{il}(x_i, z_l)], \text{ for all } l \neq j].$$

Table 4 presents the estimations of these spatial various models.[14]

We can compare these models using the differences in log likelihoods, as in Table 5.

---

[14]All models in Table 4 are given with Gore as the base, so the results give the estimations of the probability of voting for Bush.

**Table 4** Spatial logit models for USA 2000 (Base = Gore)

| Variable | (1) $\mathbb{M}(\boldsymbol{\lambda},\beta)$. Spatial | (2) $\mathbb{M}(\boldsymbol{\lambda},\boldsymbol{\alpha},\beta)$. Sp. and traits | (3) $\mathbb{M}(\boldsymbol{\lambda},\boldsymbol{\theta},\beta)$. Sp. and dem | (4) $\mathbb{M}(\boldsymbol{\lambda},\boldsymbol{\theta},\boldsymbol{\alpha},\beta)$. Full |
|---|---|---|---|---|
| Bush valence $\boldsymbol{\lambda}$ | −0.43*** | −0.69*** | −0.39 | 0.48 |
| | (5.05) | (5.64) | (0.95) | (0.72) |
| Spatial coeff. $\beta$ | 0.82*** | 0.35*** | 0.89*** | 0.38*** |
| | (14.9) | (3.69) | (14.8) | (3.80) |
| Bush trait | | 3.559*** | | 3.58*** |
| | | (13.84) | | (13.60) |
| Gore trait | | −3.22*** | | −3.15*** |
| | | (14.25) | | (13.64) |
| Age | | | −0.14** | −0.22** |
| | | | (2.33) | (2.17) |
| Gender (F) | | | −0.139 | −0.39 |
| | | | (1.00) | (1.41) |
| African American | | | −1.57*** | −1.45*** |
| | | | (5.85) | (3.67) |
| Hispanic | | | −0.27 | −0.23 |
| | | | (0.77) | (0.49) |
| Class | | | −0.20 | −0.12 |
| | | | (1.30) | (0.47) |
| Education | | | 0.18*** | 0.11 |
| | | | (3.60) | (1.32) |
| Income | | | 0.042** | −0.01 |
| | | | (3.6) | (0.32) |
| Observations | 1,238 | 1,238 | 1,238 | 1,238 |
| Log likelihood (LL) | −707.8 | −277.3 | −661.3 | −263.7 |
| AIC | 1,420 | 562.7 | 1,341 | 549.4 |
| BIC | 1,432 | 585.9 | 1,393 | 613.4 |

|$t - stat$| in parentheses
***prob < 0.001, **prob < 0.01, *prob < 0.05

**Table 5** Differences in LL for US model comparisons in 2000

| | $\mathbb{M}_2$ | JPT | JST | ST | T |
|---|---|---|---|---|---|
| $\mathbb{M}_1$ | JPT | na | 4* | 18*** | 25*** |
| | JST | −4 | na | 14 | 21*** |
| | ST | −18 | −14 | na | 7*** |
| | T | −25 | −21 | −7 | na |

*JPT* joint positional with traits, *JST* joint spatial with traits, *ST* pure spatial with traits, *T* pure traits
*** = highly significant

Note that the log likelihood of the pure spatial model given in Table 4(1) is −708, which we found to be very similar to the log likelihood of the pure positional model.

We use the equilibrium concept of local Nash equilibrium (LNE). This is simply a vector, $\mathbf{z}$, such that each candidates vote share, $V_j(\mathbf{z}) = \frac{1}{n}\sum_i \rho_{ij}(\mathbf{z})$, is locally maximized.

To determine whether the joint origin, $\mathbf{z}_0 = (z_{Gore}, z_{Bush}) = (0, 0)$ is an equilibrium for the pure spatial model, $\mathbb{M}(\boldsymbol{\lambda}, \beta)$, we need to examine the Hessians of the vote share functions.

The distribution of voter ideal points is characterized by *electoral covariance matrix*

$$\nabla_0 = \begin{bmatrix} 0.58 & -0.20 \\ -0.20 & 0.59 \end{bmatrix}.$$

The principal component of the electoral distribution is given by the vector $(1.0, -3.05)$ with variance 0.785, while the minor component is given by the orthogonal eigenvector $(1.0, 0.327)$ with variance 0.385. The correlation between these two factors is only $-0.344$.

Table 4(1) shows the intercept term $\lambda_{Bush}$, or exogenous valence for Bush in comparison to Gore, to be $-0.43$, while the $\beta$-coefficient is 0.82.

From the results in Sect. 4.1, it follows, according to the model $\mathbb{M}(\boldsymbol{\lambda}, \beta)$, that the probability that a generic voter, $i$, chooses Bush, when both Bush and Gore are at the electoral origin, $\mathbf{z}_0$, is:

$$\begin{aligned} \rho_{Bush} &= \frac{\exp[u^*_{ibush}(x_i, z_{bush})]}{\exp[u^*_{igore}(x_i, z_{gore})] + \exp[u^*_{ibush}(x_i, z_{bush})]} \\ &= \frac{\exp[-0.43]}{\exp[0] + \exp[-0.43]} \\ &= [1 + \exp(0.43)]^{-1} \\ &= [1 + 1.54]^{-1} = 0.40 \end{aligned}$$

Section 4.1 shows that the Hessian of Bush's vote share function at $\mathbf{z}_0$ is given by the characteristic matrix

$$\begin{aligned} C_{Bush} &= [2\beta(1 - 2\rho_{Bush})]\nabla_0 - I \\ &= [2 \times 0.82 \times 0.2 \times \nabla_0] - I \\ &= (0.33)\nabla_0 - I \\ &= \begin{bmatrix} 0.19 & -0.07 \\ -0.07 & 0.195 \end{bmatrix} - I = \begin{bmatrix} -0.81 & -0.06 \\ -0.06 & -0.80 \end{bmatrix} \end{aligned}$$

The determinant is positive, and the trace negative, so both eigenvalues are negative, and the joint origin is a LNE of the pure spatial model. The *convergence coefficient*, $c$, is defined to be

$$c = 2\beta(1 - 2\rho_1)\,trace\,(\nabla_0) = 0.37.$$

The valence theorem in Sect. 4.1 shows that a sufficient condition for convergence to $\mathbf{z}_0$ in the pure spatial model is the condition $c < 1$. Using the coefficients of the pure spatial model, simulation of vote maximizing behavior confirmed that the joint origin was a LNE for the 2000 Presidential election. We also included the third party candidates, Nader and Buchanan, in the estimation, but the estimates of their valences were so low that they had no impact on the Local Nash Equilibrium at

the joint origin. We also considered a model with different $\beta$ coefficients on the two axes, but found again that the joint origin was a LNE.

Note however that $z_{Gore} = (-1.42, 0.66)$, while $z_{Bush} = (0.47, -1.24)$, so these estimated positions did not locally maximize the two candidates' vote shares, contingent on the validity of the pure spatial model with exogenous valence.

We now extend the model by adding the sociodemographic variables and electoral perception of traits, based on the utility assumption

$$u_{ij}(x_i, z_j) = \lambda_j + (\theta_j \cdot \eta_i) + (\alpha_j \cdot \tau_i) - \beta \left\| x_i - z_j \right\|^2 + \varepsilon_j \tag{5}$$

$$= u_{ij}^*(x_i, z_j) + \varepsilon_j \tag{6}$$

Here $\boldsymbol{\theta} = \{(\theta_j \cdot \eta_i)\}$ refers to sociodemographic characteristics while $\boldsymbol{\alpha} = \{(\alpha_j \cdot \tau_i)\}$ refer to the electoral perception of traits. We also obtained the results of a pure traits model, denoted $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\alpha})$, and a pure sociodemographic model, $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta})$, respectively. These results showed that the pure sociodemographic model was statistically quite weak in comparison to the traits model.

Table 4 (models 2,3,4) gives the spatial models with traits, and sociodemographics, denoted $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\alpha}, \beta)$ and $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \beta)$, respectively, as well as the model with both sociodemographics and traits, $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$. The Table shows that a number of the sociodemographic coefficients were significantly different from zero in the models $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \beta)$ and $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, particularly those given by the categorical variable associated with African–American voters. Education is significant in $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \beta)$, but not when traits are included. Moreover, the difference between the log-likelihoods of this joint model, $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \beta)$, and the pure spatial model, $\mathbb{M}(\boldsymbol{\lambda}, \beta)$, was a significant +46.

Table 4 shows that the models with traits are far superior to the models without traits. This can be seen from the comparison of the log-likelihoods of the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\alpha}, \beta)$ against $\mathbb{M}(\boldsymbol{\lambda}, \beta)$ and $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$ against $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \beta)$. Note in particular that in the spatial model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\alpha}, \beta)$, with traits, both the spatial coefficient, $\beta$, and exogenous valence, $\boldsymbol{\lambda}$, are still significant. In the model, $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, the spatial coefficient remains significant, but the exogenous valence becomes insignificant. This suggests that the traits together with the sociodemographic variables provide a measure of candidate valence, but that the traits are weakly correlated with the sociodemographic variables.

For the model with traits, we found the difference between the maximum and minimum values of the Bush traits to be 4.5, while the difference for the Gore traits was 4.8. Since the coefficients on the Bush and Gore traits were 3.6 and 3.1 respectively, the magnitudes of these effects are highly significant in explaining voter behavior.[15]

---

[15]Table 4 shows that the pure traits model has an AIC of 574.8. This compares with an AIC of 664.3 for the trait model found by Clarke et al. (2009a).

We can justify the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$ by comparing it with the joint positional model. This positional model with traits, denoted JPT, has a very similar AIC to the model presented by Clarke et al. (2009a), and correctly classifies approximately 90% of the voter choice. The difference in log-likelihoods between this model and the joint spatial model with traits, denoted JST in Table 9, is only +4. We can infer that both models give very statistically significant estimations of voter response. Notice also that both the difference between the log-likelihood of the pure spatial model with traits, denoted ST in Table 5, and of JST over the pure traits model are +7 and +21, respectively.[16] We can infer that though there may be some correlation between voter perception and voter preference, the significance of the model with traits is greatly enhanced by using spatial characteristics and the sociodemographics. As suggested by the related work by Clarke et al. (2009a), the spatial and traits models complement one another.

However, simulation of the model, $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, showed that the unique local equilibrium was very close to the joint origin at

$$\mathbf{z}^{el} = \begin{bmatrix} & Bush & Gore \\ x & 0.027 & 0.027 \\ y & -0.02 & -0.02 \end{bmatrix}$$

At this LNE, we estimated the vote share for Bush to be 46% while the share for Gore was 54%. Although we find the traits model provides a statistically significant method of examining voter choice, it does not provide a satisfactory model of candidate positioning.

We now extend the model and assume voter utility is given by

$$u_{ij}(x_i, z_j) = \lambda_j + \mu_j(z_j) + (\theta_j \cdot \eta_i) + (\alpha_j \cdot \tau_i) - \beta \|x_i - z_j\|^2 + \varepsilon_j \tag{7}$$

$$= u_{ij}^*(x_i, z_j) + \varepsilon_j \tag{8}$$

where $\mu_j(z_j)$ is an activist term determined by candidate location. As Sect. 4.1 argues, the equilibrium, $\mathbf{z}^{el}$, of the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, provides an estimate of the weighted electoral means for the two candidates.

We now assume the estimated positions given in Fig. 1 comprise an LNE, $z^{\mathbb{V}*}$, of the model where each candidate has induced policy preferences from the supporting activists, as discussed in Sect. 4.3. Then the difference between the estimated positions and the weighted electoral means provides an estimate of the total activist pulls. Thus:

---

[16]The AIC results are similar. The AIC for the pure traits model is 575, which drops to 563 when $\beta$ is added and drops further to 549 with the addition of the sociodemographics.

$$[\mathbf{z}^{\mathbb{V}*} - \mathbf{z}^{el}] = \begin{bmatrix} & Bush & Gore \\ x & 0.47 & -1.42 \\ y & -1.24 & +0.66 \end{bmatrix} - \begin{bmatrix} & Bush & Gore \\ x & 0.027 & 0.027 \\ y & -0.02 & -0.02 \end{bmatrix}$$

$$= \begin{bmatrix} & Bush & Gore \\ x & +0.44 & -1.45 \\ y & -1.22 & +0.68 \end{bmatrix}.$$

From the Activist Theorem 2 of Sect. 4.3, each term in this gradient equation is given by the expression

$$[z_j^{\mathbb{V}*} - z_j^{el}] = \frac{n^* \delta_j}{2(1-\delta_j)\beta} \frac{d\mu_j}{dz_j}(z_j^{\mathbb{V}*}) \equiv \frac{n^* \delta_j}{2(1-\delta_j)\beta} \sum_{k \in A_j} a_k \frac{dU_k}{dz_j} \text{ for } j \in P.$$

Here $j = $ Bush, Gore, and the $\frac{dU_k}{dz_j}$ terms are activist gradients pointing towards the preferred positions of the various activists for the two parties. The terms $\delta_j$ are weighting parameters, with $\delta_j = 0$ corresponding to a pure vote maximizing strategy.

Now the mean activist positions for the two parties are

$$\begin{bmatrix} 2000 & Rep\ Act & Dem\ Act \\ x & +0.42 & -0.54 \\ y & -0.30 & +0.48 \end{bmatrix}$$

The gradient activist pulls and the mean activist positions are compatible if we assume that Republican activists strongly favor conservative social policies, while Democrat activists strongly favor liberal economic policies. As discussed above, Miller and Schofield (2003, 2008) have proposed a model of this kind, where each party has two classes of activists, economic and social. The set of bargains each set of party activists may make over the influence they exert on their parties is given by a one dimensional contract curve, as shown in Fig. 7 in Sect. 4.2, below. The set of possible optimal positions for each candidate is then given by a one dimensional balance locus. The actual optimal position will depend on the "eccentricity" of the utility functions of the activists, namely the trade off between activist marginal willingness to contribute and their demand for policy gains.

These estimates indicate that there is a tug of war between voters and activists over the location of the party candidates. The distribution of voters preferred points is concentrated in the electoral center, so the weighted electoral gradient of each candidate points towards the center, as is consistent with the standard spatial model. However, activists for the two parties are more concerned with social policy (for the Republicans), and economic policy (for the Democrats). There will be conflict between activists for each party as recent events have suggested. The overall effect draws the candidates into the opposing quadrants, as suggested by Fig. 1. We infer that activists with more radical policy preferences have a significant influence on the candidates. Indeed, comparison of the estimated vote share at $\mathbf{z}^{el}$ suggest that

activist contributions helped the Bush campaign by increasing his vote share from the estimated 46% to about 50%.

Here we have shown that the electoral model requires additional terms of the form $\{\mu_j\}$. We have argued that these terms are due to activists.[17] Determining the precise form requires a solution of a complex activist-candidate bargaining game as proposed by Grossman and Helpman (1994, 1996, 2001) and Baron (1994). For our purpose we have assumed that the candidate-activist contracts have been concluded prior to the election, and the empirical estimates of the various valences are the consequence of the shifts in positions away from the electoral equilibria that we have estimated.

Note that in the version of the model given in Sect. 4.4, where candidates target voters, the precise theoretical equilibrium positions will depend not only on the activist positions, but on the willingness of voters to be persuaded.

## 2.2   The Election of 2004

We repeated the above analysis for the 2004 election contest between Bush and Kerry, using the same set of responses from the ANES 2004, as for 2000.

The positions of the major presidential candidates, Bush and Kerry, in 2004 were estimated in a similar fashion to that of the sampled individuals. Scores were assigned to each candidate for each of the constituent survey items based on press reports of their campaign stances on various issues. Factor loadings and descriptive statistics are given in Tables 6 and 7. These estimates were used to obtain estimated policy positions for each candidate, as in Table 8.

Figure 4 gives the contour plot of the electoral distribution in the policy space in 2004 while Fig. 5 gives a perspective plot. Again, left on the economic ($x$) axis is

| | Question | Social policy | Economic policy |
|---|---|---|---|
| **Table 6** Factor loadings from the American national election survey, 2004 | 1. Economic problems | −0.02 | 0.39 |
| | 2. Federal spending | 0.04 | 0.39 |
| | 3. Equality | 0.30 | 0.43 |
| | 4. African American | 0.32 | 0.49 |
| | 5. Immigrants | 0.27 | 0.17 |
| | 6. Liberal vs conservative | 0.39 | 0.34 |
| | 7. Guns | 0.13 | 0.34 |
| | 8. Abortion | 0.57 | −0.05 |
| | 9. Gays | 0.60 | 0.11 |
| | 10. Family | 0.62 | 0.21 |
| | % Variance explained | 15.1 | 10.5 |

---

[17]As we show in Sect. 4.3, we can interpret these terms as policy preferences on the part of candidates, but induced from the policy preferences of activists.

**Table 7** Descriptive data 2004

|  | Econ mean | Policy std.err | 95% C.I | Social mean | Policy std.err | 95% C.I | n |
|---|---|---|---|---|---|---|---|
| Activists |  |  |  |  |  |  |  |
| Democrats | −0.49 | 0.07 | [−0.63, −0.55] | 0.75 | 0.14 | [0.47, 1.03] | 47 |
| Republicans | 0.55 | 0.06 | [0.43, 0.67] | −0.48 | 0.06 | [−0.6, −0.36] | 63 |
| Non-activists |  |  |  |  |  |  |  |
| Democrats | −0.33 | 0.03 | [−0.39, −0.27] | 0.37 | 0.04 | [0.29, 0.45] | 413 |
| Republicans | 0.30 | 0.03 | [0.24, 0.36] | −0.28 | 0.03 | [−0.34, −0.22] | 440 |

**Table 8** Kerry and Bush estimated responses, 2004

| Question | Bush | Kerry |
|---|---|---|
| 1. Economic problems | 3 | 1 |
| 2. Federal spending | 3 | 2 |
| 3. Equality | 3 | 1 |
| 4. African American | 4 | 2 |
| 5. Immigrants | 3 | 3 |
| 6. Liberal vs conservative | 3 | 1 |
| 7. Guns | 2 | 1 |
| 8. Abortion | 3 | 2 |
| 9. Gays | 3 | 1 |
| 10. Family | 5 | 3 |
| Estimated position: Social policy | −1.02 | 0.83 |
| Estimated position: Economic policy | 0.57 | −1.30 |
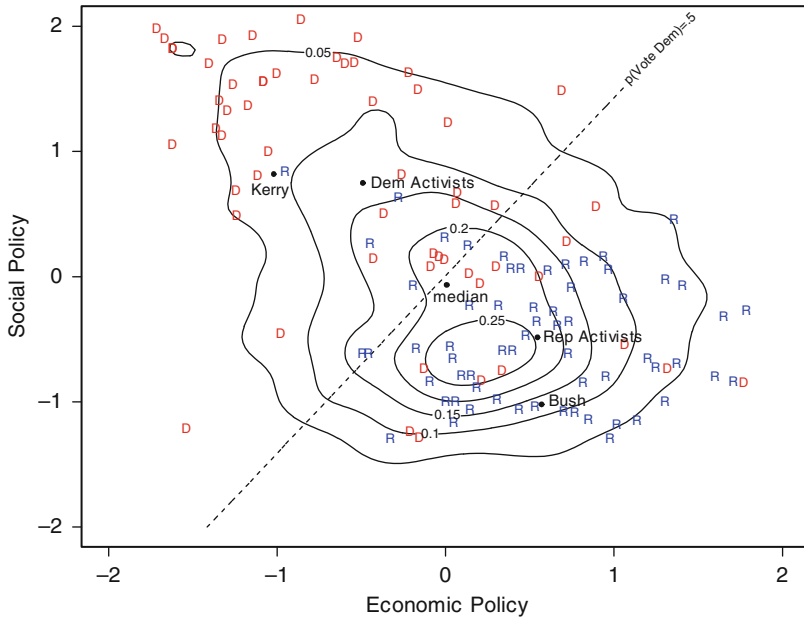


**Fig. 4** 2004 estimated voter distribution, and activist positions (Democrat activists are denoted D and Republican activists are denoted R)
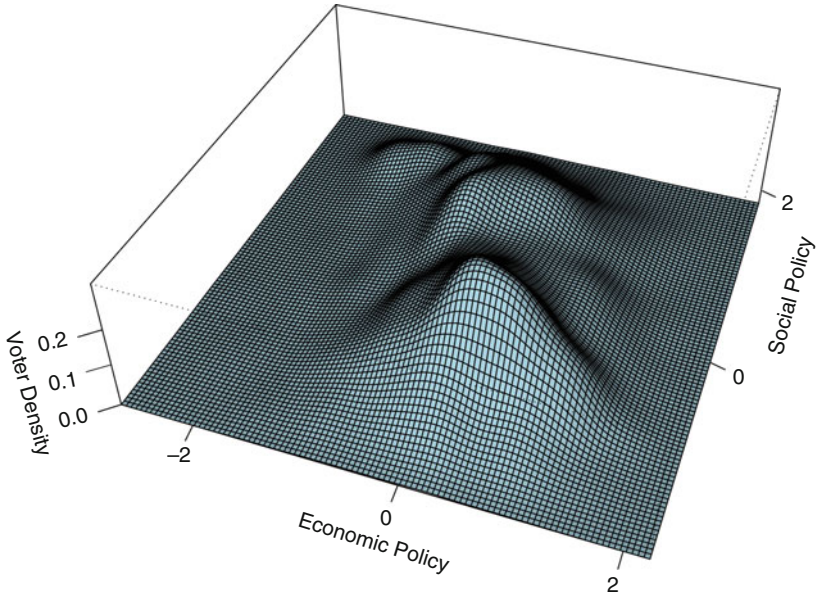
**Fig. 5** Pespective plot of the sample electorate in 2004
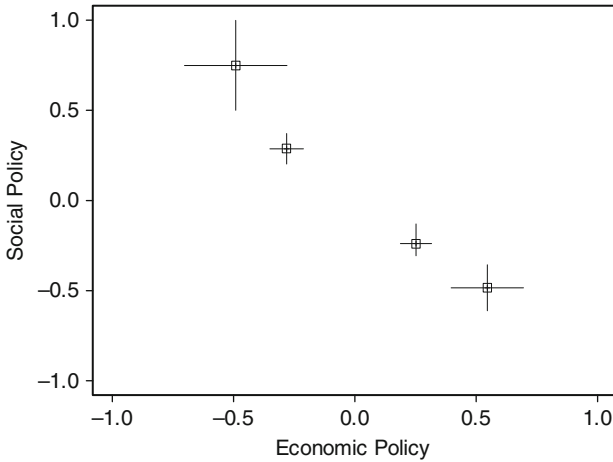


**Fig. 6** Comparison of mean partisan and activist positions in 2004

pro-redistribution, while north on the social ($y$) axis is pro-individual civil rights. Figure 6 gives the estimated mean positions of Democrat and Republican partisans and activists for 2004, together with the the error bars of the estimates. The mean partisan positions are characterized by smaller error bars. Figure 4 is also annotated with the estimated positions of Republican and Democrat activists, and the two party presidential candidates.

Note that the positions of the presidential candidates are given by

$$\mathbf{z}^* = \begin{bmatrix} candidate & Bush & Kerry \\ x & 0.57 & -1.30 \\ y & -1.02 & +0.83 \end{bmatrix}$$

Figure 4 also shows the estimated threshold dividing likely Democrat candidate voters from Republican candidate voters. Again this partisan cleavage line was derived from a binomial logit model, designed to test the effects of each policy dimension on vote choice. The results were very similar to those obtained for 2000 and are not reported here.

According to the positional model, a voter $i$, with preferred position $(x_i, y_i)$ is estimated to vote Republican with probability

$$\rho_{rep} = \frac{\exp(a + bx_i + cy_i)}{1 + \exp(a + bx_i + cy_i)} \tag{9}$$

where $(a, b, c) = (-0.20, 1.34, -0.93)$

That is, any voter with preferred point lying on the cleavage line has equal probability of picking one or other of the candidates. This cleavage line is given by the equation

$$y = 1.44x - 0.21. \tag{10}$$

which almost goes through the electoral origin. The effect of economic policy preferences is the stronger of the two dimensions in determining choice between the Democrat and Republican candidates.

The estimated position for Kerry is $z_{Kerry} = (-1.30, 0.83)$, while the Democrat activist and partisan mean positions are $z_{DEM}^{act} = (-0.49, 0.75)$ and $z_{DEM}^{part} = (-0.33, 0.37)$. Thus Kerry's position on the economic axis is a distance about $+0.8$ more extreme that the mean activist position on this axis.

Similarly, the estimated position for Bush is $z_{Bush} = (0.57, -1.02)$, while the Republican activist and partisan mean positions are $z_{REP}^{act} = (0.55, -0.48)$ and $z_{REP}^{part} = (0.30, -0.28)$. Thus Bush's position on the social axis is a distance about $+0.5$ more extreme that the mean activist position on this axis.

Note however, that there are Democratic and Republican activists located at far more extreme positions that the two candidates. It is consistent with the model of activists discussed in the Technical Appendix that more extreme activists will have a disproportionate effect on the candidate positions.

To use estimations of a spatial model, we consider the various logit models presented in Table 9.

The electoral covariance matrix obtained from the factor analysis is given by

$$\nabla_0 = \begin{bmatrix} 0.58 & -0.177 \\ -0.177 & 0.59 \end{bmatrix}.$$

**Table 9** Spatial logit models for USA in 2004 (Base = Kerry)

| Variable | (1) $\mathbb{M}(\boldsymbol{\lambda},\beta)$. Spatial | (2) $\mathbb{M}(\boldsymbol{\lambda},\boldsymbol{\alpha},\beta)$. Sp. and traits | (3) $\mathbb{M}(\boldsymbol{\lambda},\boldsymbol{\theta},\beta)$. Sp. and dem | (4) $\mathbb{M}(\boldsymbol{\lambda},\boldsymbol{\theta},\boldsymbol{\alpha},\beta)$ Full |
|---|---|---|---|---|
| Bush valence $\lambda$ | −0.43*** | −0.15 | −1.72*** | −0.670 |
|  | (5.05) | (1.00) | (3.50) | (0.70) |
| Spatial coeff. $\beta$ | 0.95*** | 0.47*** | 1.09*** | 0.475*** |
|  | (14.21) | (3.49) | (13.76) | (3.125) |
| Bush trait |  | 4.18*** |  | 4.22*** |
|  |  | (11.49) |  | (11.40) |
| Kerry trait |  | −4.20*** |  | −4.14*** |
|  |  | (11.58) |  | (11.13) |
| Age |  |  | −0.16** | 0.03 |
|  |  |  | (2.61) | (0.25) |
| Gender (F) |  |  | 0.08 | −0.38 |
|  |  |  | (0.44) | (1.18) |
| African American |  |  | −1.62*** | −1.13* |
|  |  |  | (6.11) | (2.30) |
| Hispanic |  |  | −0.26 | 0.14 |
|  |  |  | (0.75) | (1.75) |
| Class |  |  | 0.22 | 0.26 |
|  |  |  | (1.20) | (0.75) |
| Education |  |  | 0.15*** | 0.136 |
|  |  |  | (2.37) | (1.12) |
| Income |  |  | 0.056*** | 0.012 |
|  |  |  | (3.29) | (0.038) |
| Observations | 935 | 935 | 935 | 935 |
| Log likelihood (LL) | −501.7 | −145.3 | −448.2 | −137.8 |
| AIC | 1,007 | 298.5 | 914.4 | 297.7 |
| BIC | 1,018 | 320.7 | 964.2 | 358.6 |

|$t$ – $stat$| in parentheses
***prob < 0.001, **prob < 0.01, *prob < 0.05

The principal component of the electoral distribution is given by the vector $(1.0, -1.04)$ with variance 0.765, while the minor component is given by the orthogonal eigenvector $(1.0, 0.96)$ with variance 0.405. The total variance, $\sigma^2 = trace(\nabla_0)$ is 1.15, and the *electoral standard deviation* (esd) is $\sigma = 1.07$.

Table 9(1) shows the coefficients in 2004 for the pure spatial model to be

$$(\lambda_{Kerry}, \lambda_{Bush}, \beta) = (0, -0.43, 0.95).$$

According to the model $\mathbb{M}(\lambda, \beta)$, the probability that a voter chooses Bush, when both Bush and Kerry are at the electoral origin, $\mathbf{z}_0$, is

$$\rho_B = [1 + \exp(0.43)]^{-1} = [1 + 1.52]^{-1} = 0.40.$$

Then from the valence theorem presented in Sect. 4.1, the Hessian for Bush, when both candidates are at the origin, is given by:

| | JPT | JST | ST | T |
|---|---|---|---|---|
| **Table 10** Differences in LL for US model comparisons in 2004 | | | | |
| JPT | na | 2 | 5 | 12 |
| JST | −2 | na | 7 | 14 |
| ST | −5 | −7 | na | 7 |
| T | −12 | −14 | −7 | na |

*JPT* joint positional with traits, *JST* joint spatial with traits, *ST* spatial with traits, *T* pure traits

$$C_{Bush} = [2\beta(1 - 2\rho_{Bush})\nabla_0 = 2 \times 0.95 \times 0.2 \times \nabla_0 - I]$$
$$= (0.38)\nabla_0 - I$$
$$= (0.38)\begin{bmatrix} 0.53 & -0.18 \\ -0.18 & 0.66 \end{bmatrix} - I$$
$$= \begin{bmatrix} -0.8 & -0.06 \\ -0.06 & -0.75 \end{bmatrix}$$

It is obvious that both eigenvalues are negative. The convergence coefficient is $c = (0.38) \times (1.19) = 0.45$, and the joint origin is a LNE.

For the joint model given in Table 9(3), we cannot assert from first principles that $(z_{Bush}, z_{Kerry}) = (0, 0)$ satisfies the first order condition for a LNE. However, simulation of the model led us to infer that the joint origin was indeed a LNE of this model.

Although a number of the sociodemographic valences were significantly different from zero, we can infer that their magnitude is insufficient to perturb the equilibrium away from the joint origin.[18]

Note however that $z_{Kerry} = (-1.30, 0.83)$, while $z_{Bush} = (0.57, -1.02)$, so these estimated positions did not locally maximize the two candidates vote shares, contingent on the validity of the pure spatial model with exogenous valence.

We now extend the model by adding the electoral perception of traits.

It is also clear from Table 9 that the spatial models with traits (ST) and (JST) are far superior to all other spatial models without traits. Note in particular that in the joint spatial model with traits, the spatial coefficient is still significant, while the coefficient on the exogenous valence becomes insignificant. This is to be expected, as the traits are a substitute for the measure of candidate valence.

Table 10 gives the comparisons of the log-likelihoods between the joint positional (JPT) model with traits, and the spatial (ST) and joint spatial (JST) models with traits, as well as the pure traits model (T). There is only a minor difference of

---

[18]Notice that the difference between the loglikihoods of the joint spatial model and that of the pure spatial model is +53. It is possible that adding further demographic variables would change the LNE of the joint model. However, since the effect of sociodemographic variables is limited, it is unlikely that there would be any substantial effect on the LNE.

+2 in log-likelihood between JPT and JST, which we take as justification for the estimates of candidate positions in the stochastic model.[19] Note however that the AIC values for JST and ST are similar (297.7 versus 298.5). Finally, traits do not capture all the electoral characteristics. The AIC results are similar. The AIC for the pure traits model is 311, which drops to 298.5 when $\beta$ is added and drops slightly more to 297.7 with the addition of the sociodemographics. The log-likelihood differences are 7 and 14 respectively.

Simulation of the joint spatial model with traits showed that the joint origin was not an equilibrium, but the LNE was very close to the joint origin:

$$\mathbf{z}^{el} = \begin{bmatrix} & Bush & Kerry \\ x & 0.03 & 0.03 \\ y & -0.021 & -0.021 \end{bmatrix}.$$

We now assume the estimated positions comprise an LNE of the full activist model, so

$$\mathbf{z}^{\mathbb{V}*} = \begin{bmatrix} candidate & Bush & Kerry \\ x & 0.57 & -1.30 \\ y & -1.02 & +0.83 \end{bmatrix}$$

Since $\mathbf{z}^{el} = (z_{Bush}^{el}, z_{Kerry}^{el})$ is an LNE from the joint model, with no activist valence terms, we infer that $\mathbf{z}^{el}$ is the vector of weighted electoral means. Thus by the balance condition, as given in Sect. 4.1:

$$\mathbf{z}^{\mathbb{V}*} - \mathbf{z}^{el} = \begin{bmatrix} & Bush & Kerry \\ x & 0.57 & -1.30 \\ y & -1.02 & +0.83 \end{bmatrix} - \begin{bmatrix} & Bush & Kerry \\ x & 0.03 & 0.03 \\ y & -0.021 & -0.021 \end{bmatrix}$$
$$= \begin{bmatrix} & Bush & Kerry \\ x & 0.54 & -1.33 \\ y & -1.0 & +0.85 \end{bmatrix}.$$

The difference between $\mathbf{z}^{\mathbb{V}*}$ and $\mathbf{z}^{el}$ thus provides an estimate of the activist pull on the two candidates. In this election, we estimate that activists pull the two candidates into opposed quadrants of the policy space. The estimated distributions of activist positions for the two parties, in these two opposed quadrants (as given in Fig. 4) are compatible with this inference. These estimates indicate that the more extreme economic activists exerted significant pulls on both candidates in 2004, drawing them into the opposite quadrants.

---

[19]Clarke et al. (2009a) obtained an AIC of 239 for a composite version of the model here called JPT. However, they used many more sociodemographic variables. The value of 297.7 for the AIC of the spatial model, JST, suggests that it is a valid model of electoral behavior.

$$\begin{bmatrix} 2004 & \text{Rep Act} & \text{Dem Act} \\ x & 0.55 & -0.49 \\ y & -0.48 & +0.75 \end{bmatrix}.$$

As in the analysis for 2000, if we assume that the Democrat activists tend to be more concerned with liberal economic policy and Republican activists tend to be more concerned with conservative social policy, then we have an explanation for the candidate shifts from the estimated equilibrium.

## 3  Concluding Remarks and Recent Events

Valence, whether exogenous or based on electoral perceptions of character traits, is intended to model that component of voting which is determined by the judgements of the citizens. In this respect, the formal stochastic valence model provides a framework for interpreting Madison's argument in *Federalist X* over the nature of the choice of Chief Magistrate in the Republic (Madison [1787],1999). Schofield (2002) has suggested that Madison's argument may well have been influenced by Condorcet's work on the so-called "Jury Theorem" (Condorcet 1785, McLennan 1998). However, Madison's conclusion about the "probability of a fit choice" depended on assumption that electoral judgment would determine the political choice. The analysis presented here does indeed suggest that voters' judgements, as well as their policy preferences, strongly influence their political choice.

This chapter can be seen as a contribution to the development of a Madisonian conception of elections in representative democracies as methods of aggregation of both preferences and judgements. One inference from the work presented here does seem to belie Riker's arguments (1980, 1982) that there is no formal basis for populist democracy. Since voters' perceptions about candidate traits strongly influence their political decisions, the fundamental theoretical question is the manner by which these perceptions are formed.

The empirical and formal models presented here do suggest that these perceptions are the result of the influence of activist groups. Changes in voter choice appear to result not only from changes in the electoral distribution, but from the shifts in electoral perceptions. In turn, these changes are the result of the competition between the candidates over activist support. As we noted in the introduction, the importance of electoral contributions has increased, and this has enhanced the influence of activist groups.

While the analysis presented here has focused on a presidential election, it can, in principle, be applied to congressional elections as well. In this case, instead of dealing with cooperation between activist groups for a single party, we could model competition between activist groups over candidate choice for a party. Recent events over the election for the New York 23rd congressional district show how contentious this competition can be. On November 1, 2009, conservative pressure

forced the centrist Republican candidate, Dede Scozzafava, to drop out of the primary race and endorse the Democrat candidate, Bill Owens. On November 3, Owens won the election in what had been a Republican district since 1872.

In the mid term 2010 election cycle total campaign spending was about $4 billion, with Republican spending somewhat higher than that of the Democrats. The extremely high level of expenditure (especially for a midterm election) is particularly interesting because of the increasing degree of polarization mentioned in the Introduction. In this election the Democrats lost 63 seats in the House, leading to a Republican majority of 242–192. In the Senate the Democrats lost 6 seats but retained a majority of 53 (with 2 pro-Democrat Independents) to 47. President Obama was eventually able to strike a deal with the Republicans in December, 2010, to extend unemployment benefits and implement a 1-year payroll-tax cut for most workers. He was forced to accept the Republican demand for a continuation of Bush tax cuts even for the very wealthy.[20] The House speaker, Nancy Pelosi of California, accused Republicans of forcing Democrats "to pay a king's ransom in order to help the middle class". The bill passed the Senate on 15 December by 81–19, and at midnight on 16 December, 139 House Democrats voted with 138 House Republicans for the bill, against 112 Democrats and 36 Republicans. Obama immediately signed the bill.

After this initial compromise, the logjam seemed to have broken when Congress, on December 21, did approve a temporary spending bill up until March 2011.

However, on December 18, the "Dream Act" Bill to allow illegal immigrant students to become citizens failed on a Senate vote of 55–41. The Senate did vote 65–31 to repeal the "Don't Ask, Don't Tell" legislation, making it possible for gays to serve openly in the military. The House had previously approved this repeal by 250–175.

On December 20, the Senate voted 59–37 to reject an amendment to the new arms control treaty, "New Start," with Russia. The amendment would have killed the treaty because any change to the text would have required the United States and Russia to renegotiate the treaty. On December 22, the Senate voted 71–26 for the treaty. This treaty was seen as the most tangible foreign policy achievement of President Obama. Thirteen Republicans joined a unanimous Democratic caucus to vote in favor, exceeding the two-thirds majority required by the Constitution.

The Senate also voted for a $4.3 billion bill to cover medical costs for rescue workers after the 2001 terrorist attack. The House immediately voted for the bill 206–60, and it was sent to President Obama to sign into law. Congress also passed a defense authorization bill covering costs for Afghanistan and Iraq.[21]

As Obama said:

> I think it's fair to say that this has been the most productive post-election period we've had in decades, and it comes on the heels of the most productive 2 years that we've had in generations. If there's any lesson to draw from these past few weeks, it's that we are not

---

[20]This deal between the two opposed coalitions will have a serious effect on the overall U.S. debt.

[21]The bill did make it more difficult to transfer detainees from Guantánamo.

doomed to endless gridlock. We've shown in the wake of the November elections that we have the capacity not only to make progress, but to make progress together.

One of the first moves by the House in the new 112th Congress was to vote, on January 19, 2011, to repeal the Health Care Bill by a margin of 245–189. However, this repeal cannot pass the Democrat majority in the Senate.

A general inference from the model presented here is that the earlier debate about whether elections are chaotic, or are institutionally stabilized, needs to be recast. It may well be that social choice in the two party system, in the presence of activist conflict, need be neither chaotic nor equilibrating. Instead, as Miller and Schofield ([2008](#)) have argued, these activist conflicts appear to have induced a transformation of both parties in the United States. Over the long run, these continuing transformations induce a slow realignment of the political configuration.

## 4 Technical Appendix

### 4.1 *The Stochastic Model of Elections*

The electoral model that we deploy is an extension of the multiparty stochastic model of McKelvey and Patty ([2006](#)), modified by inducing asymmetries in terms of valence. The justification for developing the model in this way is the empirical evidence that valence is a natural way to model the judgements made by voters of party leaders or candidates for office.[22] There are a number of possible choices for the appropriate model for multiparty competition. The simplest one, which is used here, is that the utility function for the agent $j$ is proportional to the anticipated vote share, $V_j$, of the party in the election.[23]

With this assumption, we can examine the conditions on the parameters of the stochastic model which are necessary for the existence of a pure strategy Nash equilibrium (PNE). Because the vote share functions are differentiable, we use calculus techniques to obtain conditions for positions to be locally optimal. Thus we examine what we call *local pure strategy Nash equilibria* (LNE). From the definitions of these equilibria it follows that a PNE must be a LNE, but not conversely.

The stochastic electoral model utilizes socio-demographic variables and voter perceptions of character traits. For this model we assume that voter $i$ utility is given by the expression

---

[22]We can use the model either for party leaders or candidates for office, as in the United States. In the following we shall use the term *agents* to mean either one.

[23]For refining the model, and for empirical analysis, we could adapt the model so that parties choose positions to maximize their seat shares, relative to a given constituency structure. We adopt the simplifying vote share assumption in order to present the essential structure of the formal model.

$$u_{ij}(x_i, z_j) = \lambda_j + \mu_j(z_j) + (\theta_j \cdot \eta_i) + (\alpha_j \cdot \tau_i) - \beta||x_i - z_j||^2 + \varepsilon_{ij} \qquad (11)$$

$$= u_{ij}^*(x_i, z_j) + \varepsilon_j \qquad (12)$$

Here $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_p)$ is the exogenous valence vector which we assume satisfies the ranking condition $\lambda_p \geq \lambda_{p-1} \geq \cdots \geq \lambda_2 \geq \lambda_1$. The agents are labelled $(1, \ldots, p)$ and $\lambda_j$ is the exogenous valence of agent or candidate $j$. The points $\{x_i : i \in N\}$ are the preferred policies of the voters, in the compact Euclidean space $X$, of finite dimension $w$. The vector $\mathbf{z} = \{z_j : j \in P\}$ gives the positions of the agents in the same space. The term

$$||x_i - z_j|| = \left[\sum\nolimits_{t=1}^{w} (x_{it} - z_{jt})^2\right]^{1/2}$$

is simply the Euclidean distance between $x_i$ and $z_j$. We assume the error vector $\boldsymbol{\varepsilon}$ is distributed by the type I extreme value (Gumbel) distribution, as in empirical conditional logit estimation. This assumption means that all $\{\varepsilon_j\}$ are iid, so we write $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_j, \ldots, \varepsilon_p)$. The variance of $\varepsilon_j$ is fixed at $\frac{\pi^2}{6}$. As a result, the spatial parameter, $\beta$, has dimension $L^{-2}$, where $L$ is whatever unit of measurement is used in $X$.

In empirical models, the valence vector $\boldsymbol{\lambda}$ is given by the intercept term for each agent in the model. The symbol $\boldsymbol{\theta}$ denotes a set of $k$-vectors $\{\theta_j : j \in P\}$ representing the effect of the $k$ different sociodemographic parameters (class, domicile, education, income, religious orientation, etc.) on voting for agent $j$ while $\eta_i$ is a $k$-vector denoting the $i$th individual's relevant "sociodemographic" characteristics. The compositions $\{(\theta_j \cdot \eta_i)\}$ are scalar products, called the *sociodemographic valences* for $j$.

The terms $\{(\alpha_j \cdot \tau_i)\}$ are scalars giving voter $i$'s perceptions and beliefs. These can include perceptions of the character traits of agent $j$, or beliefs about the state of the economy, etc. We let $\boldsymbol{\alpha} = (\alpha_p, \ldots, \alpha_1)$. A *trait score* can be obtained by factor analysis from a set of survey questions asking respondents about the traits of the agent, including 'moral', 'caring', 'knowledgable', 'strong', 'honest', 'intelligent', etc. The perception of traits can be augmented with voter perception of the state of the economy, etc. in order to examine how anticipated changes in the economy affect each agent's electoral support.[24]

The terms $\{\mu_j : j \in P\}$ are the *activist valence functions*. The full model including activists is denoted $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$.

Partial models are:

1. Pure *sociodemographic*, denoted $\mathbb{M}(\boldsymbol{\lambda}, \theta)$, with only exogenous valence and sociodemographic variables
2. *Pure spatial*, denoted $\mathbb{M}(\boldsymbol{\lambda}, \beta)$, with only exogenous valence and $\beta$

---

[24]See Clarke et al. (2009a) for the same empirical procedure.

3. *Joint spatial*, denoted $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \beta)$, with exogenous valence, sociodemographic variables and $\beta$

4. *Spatial sociodemographic model with traits*, denoted $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, without the activist components

In all models, the probability that voter $i$ chooses candidate $j$, when party positions are given by $\mathbf{z}$ is:

$$\rho_{ij}(\mathbf{z}) = \Pr\left[\left[u_{ij}(x_i, z_j) > u_{il}(x_i, z_l)\right], \text{for all } l \neq j\right].$$

A (strict) local Nash equilibrium (LNE) for a model $\mathbb{M}$ is a vector, $\mathbf{z}$, such that each candidate, $j$, chooses $z_j$ to locally (strictly)[25] maximize the expected vote share

$$V_j(\mathbf{z}) = \frac{1}{n}\Sigma_i\rho_{ij}(\mathbf{z}), \text{ subject to } \mathbf{z}_{-j} = (z_1, ..z_{j-1}, z_{j+1}, ..z_p).$$

In these models, political candidates cannot know precisely how each voter will choose at the vector $\mathbf{z}$. The stochastic component as described by the vector $\boldsymbol{\varepsilon}$ is one way of modeling the degree of risk or uncertainty in the candidates' calculations. Implicitly we assume that they can use polling information and the like to obtain an approximation to this stochastic model in a neighborhood of the initial candidate locations. For this reason we focus on LNE. Note however, that as candidates adjust positions in response to information in search of equilibrium then the empirical model may become increasingly inaccurate.

A strict Nash equilibrium (PNE) for a model $\mathbb{M}$ is a vector $\mathbf{z}$ which globally strictly maximizes $V_j(\mathbf{z})$. Obviously if $\mathbf{z}$ is not a LNE then it cannot be a PNE. Indeed there may not exist a PNE.

It follows from Train (2003) that, for the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\mu}, \beta)$, the probability, $\rho_{ij}(\mathbf{z})$, that voter $i$, with ideal point, $x_i$, picks $j$ at the vector, $\mathbf{z}$, of candidate positions is given by

$$\rho_{ij}(\mathbf{z}) = \left[1 + \sum\nolimits_{k \neq j}\left[\exp(f_{kj})\right]\right]^{-1}$$
$$\text{where } f_{kj} = u_{ik}^*(x_i, z_k) - u_{ij}^*(x_i, z_j).$$
$$\text{Thus } \frac{d\rho_{ij}(\mathbf{z})}{dz_j} = -\left[1 + \sum\nolimits_{k \neq j}\left[\exp(f_{kj})\right]\right]^{-2}\frac{d}{dz_j}\left[\sum\nolimits_{k \neq j}\left[\exp(f_{kj})\right]\right]$$
$$= \{2\beta(x_i - z_j) + \frac{d\mu_j}{dz_j}(z_j)\}[\rho_{ij} - \rho_{ij}^2].$$

We use this gradient equation to show that the first order condition for $\mathbf{z}^*$ to be a LNE is given by the following balance equations:

---

[25]We keep to strict equilibria to avoid non-generic problems when one eigenvalue is zero.

$$0 = \frac{dV_j(z)}{dz_j} = \frac{1}{n} \sum_{i \in N} \frac{d\rho_{ij}}{dz_j}$$

$$= \frac{1}{n} \sum_{i \in N} [\rho_{ij} - \rho_{ij}^2] \left\{ 2\beta(x_i - z_j) + \frac{d\mu_j}{dz_j}(z_j) \right\}.$$

$$\text{Hence } z_j^* = \sum_i \frac{[\rho_{ij} - \rho_{ij}^2]x_i}{\sum_{k \in N} [\rho_{kj} - \rho_{kj}^2]} + \frac{1}{2\beta} \frac{d\mu_j}{dz_j}(z_j),$$

$$\text{or } z_j^* = \frac{1}{2\beta} \frac{d\mu_j}{dz_j}(z_j) + \sum_{i=1}^{n} \varpi_{ij} x_i.$$

$$\text{This can be written } 0 = \left[ z_j^{el} - z_j^* \right] + \frac{1}{2\beta} \frac{d\mu_j}{dz_j}(z_j)$$

$$\text{where } z_j^{el} = \sum_{i=1}^{n} \varpi_{ij} x_i.$$

Here $z_j^{el}$ is the *weighted electoral mean* of agent $j$. Because this model is linear, it is possible to modify these weights to take account of the differential importance of voters in different constituencies.[26]

A similar analysis holds for the full model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$. We can therefore write the first order balance condition at an equilibrium, $\mathbf{z}^* = (z_1^*, .., z_j^* .. z_p^*)$, as a set of *gradient balance conditions*

$$\frac{dE_j^*}{dz_j}(z_j^*) + \frac{1}{2\beta} \frac{d\mu_j}{dz_j}(z_j^*) = 0. \tag{13}$$

The first term in this equation is the *centripetal marginal electoral pull* for agent $j$, defined at $z_j$

$$\frac{dE_j^*}{dz_j}(z_j) = \left[ z_j^{el} - z_j \right]$$

The second gradient term, $\frac{d\mu_j}{dz_j}(z_j)$ is the *centrifugal marginal activist pull for $j$*, at $z_j$.

Writing $\mathbf{z}^{el} = \left( z_1^{el}, z_2^{el}, .. z_p^{el} \right)$, and $\frac{d\boldsymbol{\mu}}{d\mathbf{z}}(\mathbf{z}^*) = \left( .., \frac{d\mu_j}{dz_j}(z_j^*), .. \right)$, then in vector notation, the first order condition can be written

$$\left[ \mathbf{z}^* - \mathbf{z}^{el} \right] = \frac{1}{2\beta} \frac{d\boldsymbol{\mu}}{d\mathbf{z}}(\mathbf{z}^*)$$

---

[26]For example, presidential candidates may attempt to maximize total electoral votes, so voters can be weighted by the relative electoral college seats of the state they reside in.

The vector $\mathbf{z}^*$ is said to be a *balance solution* if this vector equation is satisfied.

To determine the LNE for the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$ it is of course necessary to consider the Hessians $\frac{dV_j^2(\mathbf{z})}{dz_j^2}$. These will involve the second order terms $\frac{d^2\mu_j}{dz_j^2}$. Schofield (2006) proved the following Activist Theorem. The theorem suggests that there will be natural conditions under which the second order terms $\frac{d^2\mu_j}{dz_j^2}$ will be negative definite. Indeed if the eigenvalues are negative and of sufficiently large modulus, then we may expect the existence of PNE.

**Activist Theorem 1.** *Consider the electoral model* $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$

(i) *The first order condition for* $\mathbf{z}^*$ *to be an LNE is that it is a balance solution.*
(ii) *If all activist valence functions are sufficiently concave,*[27] *then a balance solution will be a PNE.* □

For the pure spatial model, $\mathbb{M}(\boldsymbol{\lambda}, \beta)$, it is clear that when the candidate positions are identical, then $\rho_{kj} = \rho_j$, is independent of the voter suffix $k$. Thus all $\varpi_{ij} = \frac{1}{n}$ gives the first order condition for a LNE. By a change of coordinates, it follows that $\mathbf{z}_0 = (0, \ldots 0)$ is a candidate for a LNE.[28] Note however that this argument does not follow for the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, and generically $\mathbf{z}^{el} = (z_1^{el}, z_2^{el}, ..z_p^{el}) \neq (0, \ldots 0)$.

Since the valence functions are constant in the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, the marginal effects, $\frac{d\mu_j}{dz_j}$, will be zero. However, since the weights in the weighted electoral mean for each candidate will vary from one individual to another, it is necessary to simulate the model to determine the LNE $\mathbf{z}^{el} = \left(z_1^{el}, z_2^{el}, ..z_p^{el}\right)$.[29] Notice also that the marginal vote effect, $\frac{d\rho_{ij}}{dz_j}$, for a voter with $\rho_{ij}(\mathbf{z}) \simeq 1$ will be close to zero. Thus in searching for LNE, each candidate will seek voters with $\rho_{ij}(\mathbf{z}) < 1$.

For the pure spatial model, $\mathbb{M}(\lambda, \beta)$, we have shown that $\mathbf{z}_0$ satisfies the first order condition for LNE. The necessary and sufficient second order condition for LNE at $\mathbf{z}_0$ in the pure spatial model, $\mathbb{M}(\lambda, \beta)$, is determined as follows. When all candidates are at the electoral origin, and agent 1 is, by definition, the lowest valence candidate, then the probability that a generic voter picks candidate 1 is given:

$$\rho_1 = \left[1 + \sum_{k \neq 1}\left[\exp(\lambda_k)\right]\right]^{-1} \tag{14}$$

To compute the Hessian of candidate 1, we proceed as follows:

---

[27] By this we mean that the eigenvalues of the activist functions are negative and of sufficient magnitude everywhere. That is to say, there exists $\alpha < 0$, such that all eigenvalues $< \alpha$ is sufficient to guarantee existence of a PNE.

[28] It is worth observing that if we use just distance rather than distance squared then the first order condition is essentially a counting rule, giving a median position of the candidates as equilibrium.

[29] This can be done using the gradient equation given below.

$$\frac{dV_1^2(\mathbf{z})}{dz_1^2} = \frac{d}{dz_1} \frac{1}{n} \sum_{i \in N} 2\beta(z_1 - x_i)\left[\rho_{i1}^2 - \rho_{i1}\right]$$

$$= \frac{2\beta}{n} \sum_{i \in N} \left\{ [\rho_{i1}^2 - \rho_{i1}] \frac{d}{dz_1}(z_1 - x_i) - (1 - 2\rho_{i1}) \frac{d\rho_{i1}}{dz_1}.(z_1 - x_i) \right\}$$

$$= \frac{2\beta}{n} \sum_{i \in N} \left[\rho_{i1} - \rho_{i1}^2\right][2\beta(1 - 2\rho_{i1})(x_i - z_1).(x_i - z_1) - I]$$

where $I$ is the $w$ by $w$ identity matrix, and we use (.) to denote scalar product. Again, when all candidates are at the origin then $\rho_{i1} = \rho_1$ is independent of $i$. Moreover,

$$\frac{1}{n} \sum_{i \in N} (x_i).(x_i) = \nabla_0 \tag{15}$$

is the $w$ by $w$ covariance matrix of the distribution of voter ideal points, taken about the electoral origin. Then setting $z_1 = 0$ in the above equation, we see that the Hessian of the vote share function of candidate 1 is given by

$$2\beta\left[\rho_1 - \rho_1^2\right][2\beta(1 - 2\rho_1)\nabla_0 - I] \tag{16}$$

Since $\left[\rho_1 - \rho_1^2\right] > 0, \beta > 0$, this Hessian can be identified with the $w$ by $w$ characteristic matrix for candidate 1, given by:

$$C_1 = 2\beta(1 - 2\rho_1)\nabla_0 - I, \tag{17}$$

Following Schofield (2007), this shows that the necessary and sufficient second order condition for an LNE at $\mathbf{z}_0$ is that $C_1$ has negative eigenvalues.[30]

These second order conditions can be interpreted in terms of the trace and determinant of $C_1$. Schofield (2007) also shows that a necessary condition for $\mathbf{z}_0 = (0, \ldots 0)$ to be an LNE is that a convergence coefficient, c, defined by

$$c = 2\beta(1 - 2\rho_1)\sigma^2$$

satisfies the critical convergence condition, $c < w$. Here $\sigma^2 = trace(\nabla_0)$ is the sum of the variance terms on all axes. We state this as the Valence Theorem.

**The Valence Theorem.**
(i) *The joint origin $\mathbf{z_0}$ satisfies the first order condition to be a LNE for the model* $\mathbb{M}(\lambda, \beta)$.

---

[30]Strictly speaking, the condition is that the eigenvalues are non-positive. To avoid the degenerate case with a zero eigenvalue, we focus on a strict local equilibrium associated with negative eigenvalues of the Hessian.

(ii) *The necessary and sufficient second order condition for LNE at $z_0$ is that $C_1$ has negative eigenvalues.*[31]

(iii) *A necessary condition for $z_0$ to be a LNE for the model $\mathbb{M}(\lambda, \beta)$ is that $c(\boldsymbol{\lambda}, \beta) < w$.*

(iv) *A sufficient condition for convergence to $z_0$ in the two dimensional case is that $c < 1$.* □

Note that $\beta$ has dimension $L^{-2}$, while $\sigma^2$ has dimension $L^2$ so $c$ is dimensionless, and is therefore independent of the units of measurement.

In the two dimensional case, Schofield (2007) also shows that a *sufficient* condition for negative eigenvalues, and thus for convergence to $z_0$, is that $c < 1$. Note that when $c \ll 1$, and the eigenvalues are both negative and of sufficient magnitude then we might expect that the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$ will have an equilibrium $(z_1^{el}, z_2^{el}, ..z_p^{el}) \simeq \mathbf{0}$, that is close to the joint origin. We found this for the two elections discussed here.[32]

When the necessary condition fails, then the lowest valence candidate has a best response that diverges from the origin. In this case there is no guarantee of existence of a PNE.

## 4.2 Extension to the Case with Multiple Activist Groups

We adapt the model presented in Schofield and Cataife (2007), where there are multiple activist groups for each party.

1. For each party candidate, $j$, let $\{A_j\}$ be a family of potential activists, where each $k \in A_j$ is endowed with a utility function, $U_k$, which is a function of the position $z_j$. The resources allocated to $j$ by $k$ are denoted $R_{jk}(U_k(z_j))$. The total activist valence function for candidate $j$ is the linear combination

$$\mu_j(z_j) = \sum_{k \in A_j} \mu_{jk}\big(R_{jk}\big(U_k(z_j)\big)\big). \tag{18}$$

where $\{\mu_{jk}\}$ are functions of the contributions $\{R_{jk}(U_k(z_j))\}$, and each $\mu_{jk}$ is a concave function of $R_{jk}$.

2. Assume the gradients of the valence functions for $j$ are given by

$$\frac{d\mu_{jk}}{dz_j} = a_k^* \frac{dR_{jk}}{dz_j} = a_k^* a_k^{**} \frac{dU_k}{dz_j} \tag{19}$$

where the coefficients $a_k^*, a_k^{**}$ are all differentiable functions of $z_j$ and $> 0$.

---

[31]In the usual way, the condition for an LNE is that the eigenvalues are negative semidefinite.

[32]Indeed, we found the same result for the 2008 election reported in Schofield et al. (2011).

3. Under these assumptions, the first order equation $\frac{d\mu_j}{dz_j} = 0$ becomes

$$\frac{d\mu_j}{dz_j} = \sum_{k \in A_j} \frac{d}{dz_j} \left[ \mu_{jk} \left( R_{jk} \left( U_k(z_j) \right) \right) \right] \tag{20}$$

$$= \sum_{k \in A_j} \left( a_k^{**} a_k^* \right) \frac{dU_k}{dz_j} = 0. \tag{21}$$

The *Contract Set* generated by the family $\{A_j\}$ is the locus of points satisfying the gradient equation

$$\sum_{k \in A_j} a_k \frac{dU_k}{dz_j} = 0, \text{ where } \sum_{k \in A_j} a_k = 1 \text{ and all } a_k > 0. \tag{22}$$

Here we normalize by setting $a_k = \frac{a_k^{**} a_k^*}{\sum_{m \in A_j} a_m^{**} a_m^*}$.

The *Balance Locus* for the candidate $j$, defined by the family, $\{A_j\}$, is the solution to the first-order gradient equation

$$\left[ z_j^{el} - z_j^* \right] + \frac{1}{2\beta} \left[ \sum_{k \in A_j} a_k \frac{dU_k}{dz_j} \right] = 0. \tag{23}$$

The simplest case, discussed in Schofield and Cataife (2007) is in two dimensions, where each candidate has two activist groups. In this case, the contract curve for each candidate's supporters will, generically, be a one-dimensional arc. Miller and Schofield (2003) also supposed that the activist utility functions were ellipsoidal, mirroring differing saliences on the two axes. In this case the contract curves would be *catenaries*, and the balance locus would be a one dimensional arc. The balance solution for each candidate naturally depends on the position(s) of opposed candidate(s), and on the coefficients, as indicated above, of the various activists. The determination of the balance solution can be obtained by computing the vote share Hessian along the balance locus.

Figure 7 gives an illustration of the equilibrium balance solution for a Republican candidate, denoted $z_1^*(z_2)$ in the figure. Here $z_2$ denotes the position taken by the Democrat candidate.

Since the activist valence function for candidate $j$ depends on the resources contributed by the various activist groups to this candidate, we may expect the marginal effect of these resources to exhibit diminishing returns. Thus the activist valence functions can be expected to be concave in the activist resources, so that the Hessian of the overall activist valence, $\mu_j$, can be expected to have negative eigenvalues. When the activist functions are sufficiently concave (in the sense that the Hessians have negative eigenvalues of sufficiently large modulus) then we may infer not only that the LNE will exist, but that they will be PNE.
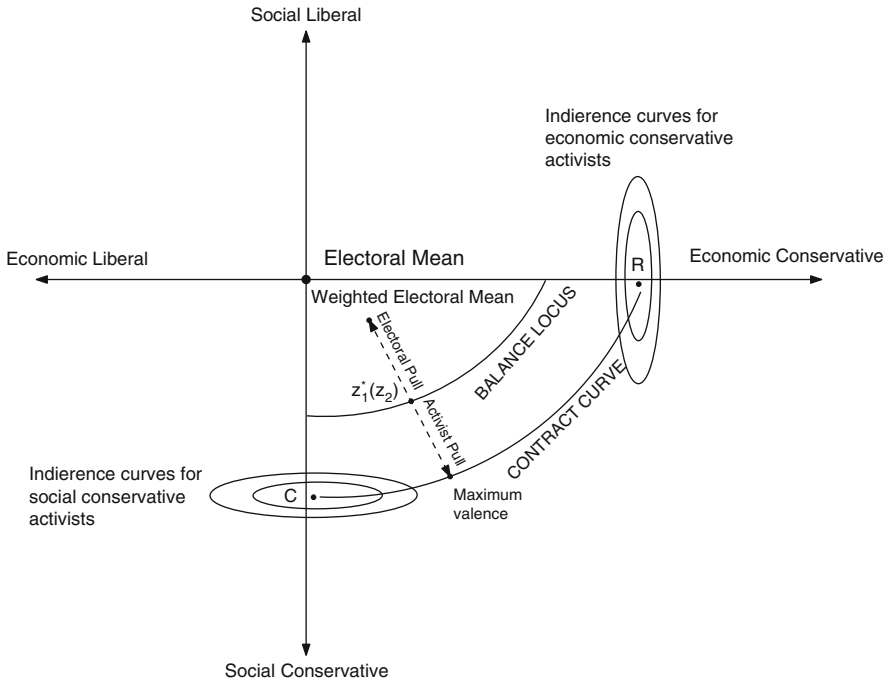
**Fig. 7** Optimal Republican position

## 4.3 Policy Preferences of Candidates

If we associate the utilities $\{U_k\}$ with leaders of the activist groups for the parties, then the combination

$$\sum_{k \in A_j} a_k \frac{dU_k}{dz_j}$$

may be interpreted as the marginal utility of the candidate of party $j$, induced by the activist support.

To see this suppose that each candidate were to maximize the function $\mathbb{V}$, given by

$$\mathbb{V}_j(\mathbf{z}) = \delta_j \mu_j(z_j) + \frac{1 - \delta_j}{n} \sum_i \rho_{ij}(\mathbf{z})$$

where $\mu_j$ is no longer an activist function, but a policy determined component of the candidate's utility function, while $\delta_j \in [0, 1]$ is the weight given to the policy preference. This model has been proposed by Wittman (1977), Calvert (1985), Duggan and Fey (2005) and Peress (2010).

If we let $z_j^{\mathbb{V}*}$ be the solution with these policy preferences, then the solutions for $\{z_j^{\mathbb{V}*}\}$ will depend on $j$, and so $\rho_{ij}$ will depend on voters' identity, i.e., will depend on $\{x_i \in X\}_{i \in N}$. Thus $\rho_{ij}$ cannot be written as $\rho_j$. The first order condition becomes

$$\frac{d\mathbb{V}_j(\mathbf{z})}{dz_j} = \delta_j \frac{d\mu_j}{dz_j}(z_j) + \frac{1 - \delta_j}{n} \sum_{i=1}^{n} 2\beta(x_i - z_j)\left[\rho_{ij} - \rho_{ij}^2\right] = 0$$

$$\Rightarrow z_j^{\mathbb{V}*}(1 - \delta_j) \sum_{k \in N}\left[\rho_{kj} - \rho_{kj}^2\right] = \frac{n\delta_j}{2\beta}\frac{d\mu_j}{dz_j} + (1 - \delta_j)\sum_{i=1}^{n}\left[\rho_{ij} - \rho_{ij}^2\right]x_i.$$

so

$$z_j^{\mathbb{V}*} = \frac{n^*\delta_j}{2(1 - \delta_j)\beta}\frac{d\mu_j}{dz_j} + \sum_{i \in N}[\bar{\omega}_{ij}]x_i = \frac{n^*\delta_j}{2(1 - \delta_j)\beta}\frac{d\mu_j}{dz_j} + z_j^{el},$$

where

$$n^* = \frac{n}{\sum_{k \in N}\left[\rho_{kj} - \rho_{kj}^2\right]}.$$

The new "balance equation" becomes

$$\left[z_j^{el} - z_j^{\mathbb{V}*}\right] + \frac{n^*\delta_j}{2(1 - \delta_j)\beta}\frac{d\mu_j}{dz_j}(z_j^*) = 0.$$

Here $\frac{d\mu_j}{dz_j}(z_j^*)$ is a gradient at $z_j^*$ pointing towards the policy preferred position of the candidate.

Suppose now that all agents have contracted with their various activists groups, as given above by the multiple activist model, as above. Suppose further that the activists have provided resources which have been deployed to influence voters. If we now estimate the spatial sociodemographic model with traits, $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, at the time of the election, then the effect of these resources will be incorporated in the parameters of the model. Simulation of this model will give $\mathbf{z}^{el}$. Suppose further that the agent is *committed* to the contract with the activists, so that the agent's equilibrium position, $z_j^{\mathbb{V}*}$, is that which is obtained from the model where the agent adopts a policy position induced from this contract. Comparing the above equation with the multiple activist model, we can make the identification

$$\left[z_j^{\mathbb{V}*} - z_j^{el}\right] = \frac{n^*\delta_j}{2(1 - \delta_j)\beta}\frac{d\mu_j}{dz_j}(z_j^{\mathbb{V}*}) \equiv \frac{n^*\delta_j}{2(1 - \delta_j)\beta}\sum_{k \in A_j}a_k\frac{dU_k}{dz_j} \text{ for } j \in P.$$

The weighted electoral mean $z_j^{el}$ can be obtained for the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$ using simulation. This equation implies that the agent's marginal policy preference can be identified with a combination of the marginal preferences of the party

activists. To solve this equation in detail requires first solving the contract game between activists and agents, as outlined in Grossman and Helpman (1994, 1996, 2001). For our purposes it is sufficient to use this reduced form, as we are interested in the difference $\left[z_j^{\mathbb{V}*} - z_j^{el}\right]$ between the estimated policy position, $z_j^{\mathbb{V}*}$ of agent $j$ and its weighted electoral mean, $z_j^{el}$.

We now present these results as a Theorem.

**Activist Theorem 2.** *Suppose each candidate, $j$, is committed to a contract with a family of activists $\{A_j\}$ with utility functions $\{U_k : k \in A_j\}$. Let $z_j^{el}$ be the estimated equilibrium position according to the spatial sociodemographic model, $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, at the time of the election. Then the influence of the activists is given by the set of equations*

$$\left[z_j^{\mathbb{V}*} - z_j^{el}\right] = \frac{n^*\delta_j}{2(1-\delta_j)\beta}\frac{d\mu_j}{dz_j}(z_j^{\mathbb{V}*}) \equiv \frac{n^*\delta_j}{2(1-\delta_j)\beta}\sum_{k \in A_j} a_k \frac{dU_k}{dz_j} \text{ for } j \in P. \qquad \square$$

The advantage of this version of the result is that while the activist resources affect the voter probabilities, these are already included in the estimation of the model and the estimated weighted means $\{z_j^{el}\}$. Thus the effect of activist support is subsumed in the empirical estimates of the exogenous valences and the additional sociodemographics and trait valences. While this does not allow us to solve for the nature of the contracts, it does give an estimate of the nature of the contracts between agents and activists. A further advantage of the model is that it provides a rationale for agents to act as though they had policy preferences, while at the same time choosing policy options that are aimed at maximizing vote share.

## 4.4   Extension of the Activist Model: Targeting Voters

As before we let $\{A_j\}$ be the family of activist supporters for $j$ and now write

$$\mathbf{R}_j(z_j) = \sum_{k \in A_j} R_{jk}\big(U_k(z_j)\big). \tag{24}$$

for the total resources obtained by agent $j$ from the various activist groups in $\{A_j\}$. Here, we again assume the activist utilities are functions of $z_j$. These resources are denominated in terms of time (multiplied by the wage rate for labor) or money, so we can take the units to be currency, dollars say.

These resources are used to target the individual voters and the voter utility function is now

$$\begin{aligned} u_{ij}(x_i, z_j) &= \lambda_j + \mu_i(m_{ij}) + (\theta_j \cdot \eta_i) + (\alpha_j \cdot \tau_i) - \beta||x_i - z_j||^2 + \varepsilon_j \\ &= u_{ij}^*(x_i, z_j) + \varepsilon_j. \end{aligned}$$

Here $\mu_i(m_{ij})$ is the valence effect of the expenditure of resources, $(m_{ij})$ on the targeting of voter $i$, by agent $j$. We assume that the greater the resources $m_{ij}$ spent on persuading voter $i$, the greater the implicit valence associated with candidate $j$, so $\frac{d\mu_i(m_{ij})}{dm_j} > 0$. We may also assume decreasing returns: $\frac{d^2\mu_i(m_{ij})}{dm_j^2} < 0$. Obviously we can partition the voters into different categories, in terms of their sociodemographic valences. Note that different agents may target the same voter or group of voters.

We assume that for each $j$ the budget constraint is satisfied:

$$\mathbf{R}_j(z_j) = \sum_{k \in A_j} R_{jk}\left(U_k(z_j)\right) = \sum_{i \in N} m_{ij} \tag{25}$$

We further assume that $j$ solves the optimization problem that we now construct. Since $\mathbf{R}_j(z_j)$ determines the budget constraint for $j$, we can write $m_{ij} \equiv m_{ij}(z_j)$, so

$$\mu_i(m_{ij}) \equiv \mu_i(m_{ij}(z_j)) \equiv \mu_{ij}(z_j).$$

We shall also assume that the solution to the optimization problem is smooth, in the sense that $\mu_{ij}(\text{-})$ is a differentiable function of $z_j$.

Then just as above, the first order condition gives a more general balance condition as follows:

$$0 = \frac{dV_j(\mathbf{z})}{dz_j} = \frac{1}{n} \sum_{i \in N} \frac{d\rho_{ij}}{dz_j}$$

$$= \frac{1}{n} \sum_{i \in N} \left[\rho_{ij} - \rho_{ij}^2\right] \left\{2\beta(x_i - z_j) + \frac{d\mu_{ij}}{dz_j}(z_j)\right\}.$$

So $z_j \sum_{i \in N} \left[\rho_{ij} - \rho_{ij}^2\right] = \sum_{i \in N} \left[\rho_{ij} - \rho_{ij}^2\right] \left\{x_i + \frac{1}{2\beta}\frac{d\mu_{ij}}{dz_j}(z_j)\right\}.$

Hence $z_j^* = \dfrac{\sum_i \left[\left[\rho_{ij} - \rho_{ij}^2\right]\left[x_i + \frac{1}{2\beta}\frac{d\mu_{ij}}{dz_j}(z_j)\right]\right]}{\sum_{k \in N}\left[\rho_{kj} - \rho_{kj}^2\right]}$ and $z_j^* = \sum_{i=1}^{n} \bar{\omega}_{ij}(x_i + \gamma_i)$ where $\gamma_i = \frac{1}{2\beta}\frac{d\mu_{ij}}{dz_j}(z_j)$

and $\bar{\omega}_{ij} = \dfrac{\left[\rho_{ij} - \rho_{ij}^2\right]}{\sum_{k \in N}\left[\rho_{kj} - \rho_{kj}^2\right]}.$

This can be written $\left[z_j^* - z_j^{el}\right] = \sum_{i=1}^{n} \bar{\omega}_{ij}\gamma_i$ where $z_j^{el} = \sum_{i=1}^{n} \bar{\omega}_{ij}x_i$.

When $\frac{d\mu_{ij}}{dz_j}(z_j) = \frac{d\mu_i}{dz_j}(z_j)$ this reduces to the previous result.

The difference now is that instead of there being a single *centrifugal marginal activist pull* $\frac{1}{2\beta}\frac{d\mu_i}{dz_j}(z_j)$ *there is an aggregate activist pull*

$$\sum_{i=1}^{n} \bar{\omega}_{ij}\gamma_i = \frac{1}{2\beta} \sum_{i=1}^{n} \frac{\left[\rho_{ij} - \rho_{ij}^2\right]}{\sum_{k \in N}\left[\rho_{kj} - \rho_{kj}^2\right]} \frac{d\mu_{ij}}{dz_j}(z_j)$$

determined by the budget constraint.

Notice that the first order condition depends on the marginal terms, $\frac{d\mu_{ij}}{dz_j}(z_j)$, associated with policy positions, and these will depend on the marginal valence effects $\frac{d\mu_i(m_{ij})}{dm_j}$. Although these valence effects can be assumed to exhibit decreasing returns, these will vary across different classes of voters. The plausibility of existence of Nash equilibria turns on whether the induced second order terms $\frac{d^2\mu_{ij}}{dz_j^2}(z_j)$ have negative eigenvalues. The assumption of negative eigenvalues would give a version of the activist theorem.

Note also that if $\rho_{ij}$ is close to 0 or 1, then $\bar{\omega}_{ij}$ will be close to 0, so the optimal calculation will be complex, though in principle solvable. It is plausible the candidate should expend resources on pivotal voters for whom $\rho_{ij}$ is close to $1/2$.[33]

## 4.5 *Endogenizing Activist Support*

To sketch an outline of a general model to endogenize activist support, we first let the electoral mapping

$$\boldsymbol{\rho} : X^p \times \mathbb{B}^{n\times p} \to [0,1]^{n\times p}$$

specify the voter probabilities in terms of candidate positions in $X^p$ and the voter distribution, in $\mathbb{B}^{n\times p}$, of resources $\{m_{ij}\}$ to all voters.[34] In principle $\boldsymbol{\rho}$ could be obtained by empirical analysis. We assume that $\boldsymbol{\rho}$ is common knowledge to agents and activists.

We then let

$$\mathbf{V} = V_1 \times .. \times V_p : X^p \times \mathbb{B}^{n\times p} \to [0,1]^p$$

be the *agent profile function*, mapping agent positions and voter distributions to vote shares, as given by the above models. Indeed, for a more general model we could consider multiparty systems where agents form beliefs about coalition behavior, as suggested in Schofield and Sened (2006). In this case the mapping would be

$$\mathbf{V} = V_1 \times .. \times V_p : X^p \times \mathbb{B}^{n\times p} \to \mathbb{R}^p.$$

We let the **k** activists have preferences over the positions taken by the $p$ political agents and agent vote shares, so the *activist profile function* is a map

---

[33]Stokes (2005) make a somewhat similar inference, discussing clientist models of politics, where $m_{ij}$ is simply a monetary bribe to $i$. Obviously the marginal benefit to a poor voter is greater than to a wealthy voter, under the usual assumption of decreasing marginal utility for money. Dal Bo (2007) also considers a model of bribery but does not consider income effect *per se*.

[34]It is reasonable to assume that the resource distributions lie in a compact ball, namely $\mathbb{B}^{n\times p}$.

$$\mathbf{U} : X^p \times [0, 1]^p \to \mathbb{R}^k.$$

It is reasonable to suppose that both **V** and **U** are differentiable. We now regard the activists as principals who choose offers to make to the political agents. This offer can be regarded as a mapping

$$\mathbf{U}^* : X^p \to \mathbb{B}^p.$$

which specifies the provision of activist resources to each of the agents. Note that we assume that these principals are assumed to make inferences about how the agents will respond to the offer mapping, on the basis of common knowledge about the electoral mapping, $\boldsymbol{\rho}$.

The agents in turn choose a best response to $\mathbf{U}^*$. We seek an equilibrium to a game form which may be written

$$\begin{aligned}
\mathbf{U}^* \otimes \mathbf{V} \quad &: \quad X^p \to X^p \times \mathbb{B}^p \to X^p \times \mathbb{B}^{n \times p} \to \mathbb{R}^k \times [0, 1]^p. \\
&: \quad (\mathbf{z}) \to (\mathbf{z}, \mathbf{U}^*(\mathbf{z})) \to (\mathbf{z}, \mathbf{m}) \to ((\mathbf{U}(\mathbf{z}, \mathbf{V}((\mathbf{z}, \mathbf{m})), \mathbf{V}((\mathbf{z}, \mathbf{m}))
\end{aligned}$$

On the basis of the offer mapping, $\mathbf{U}^*$, the agents choose a position vector **z** and a distribution matrix, $m \in \mathbb{B}^{n \times p}$, such that $(\mathbf{z}, \mathbf{m})$ is a LNE for the agent profile function, **V**, subject to the constraint that **m** is compatible with the offer $\mathbf{U}^*(\mathbf{z})$.

This is an extremely complex dynamical game, and we do not attempt to explore the full ramifications of this model here.[35] Notice that the game form just presented attempts to endogenize activist choices based on an asumption of differentiability. It is quite possible that, in actual applications of the model, the activist offer mapping may be non differentiable, as activists may switch allegiance from one agent or party to another.[36]

Earlier results of Schofield (1978) and McKelvey (1979) had suggested chaos could be generic in electoral models. The model proposed here does not exhibit chaos. However, the application of a simpler version of this model (Schofield et al. 2003) to the historical development of the U.S. political economy suggests that the equilibria of the model are subject to both "circumferential" and "radial" transformations over time, as activists switch support, and candidates move nearer or further away from the origin.

---

[35]See Coram (2010) and Duggan and Kalandrakis (2011) for dynamical versions of bargaining models. Acemoglu and Robinson (2008) also develop a model based on Markov Perfect Equilibrium where the elite, the activists, have different preferences for the public good, in $X$ and contribute to the de facto power, or political strength, of the political leader. However, they do not assume competing political leaders.

[36]The "matching" model proposed by Jackson and Watts (2010) embeds the Nash equilibrium within a coalition game, and would allow the principals to switch from one agent coalition to another.

# Data Appendix

Question Wordings for the American National Election Surveys, for 2000 and 2004.

1. We need a strong government to handle today's complex economic problems [1]; or the free market can handle these problems without government being involved [3].
2. Should federal spending on welfare programs be increased [1], decreased [3], or kept about the same? [2].
3. This country would be better if we worried less about how equal people are. Do you agree strongly [5], agree somewhat [4], neither agree nor disagree [3], disagree somewhat [2], or disagree strongly [1] with this statement?
4. Many minorities overcame prejudice and worked their way up. African Americans should do the same without any special favors. Do you agree strongly [5], agree somewhat [4], neither agree nor disagree [3], disagree somewhat [2], or disagree strongly [1] with this statement?
5. Do you think the number of immigrants from foreign countries who are permitted to come to the United States to live should be increased a lot [1], increased a little [2], left the same as it is now [3], decreased a little [4], or decreased a lot [5]?
6. We hear a lot of talk these days about liberals and conservatives. Here is a three-point scale on which the political views that people might hold are arranged from extremely liberal to extremely conservative. liberal = [1], moderate = [2], all conservative = [3].
7. Do you think the federal government should make it more difficult [1] for people to buy a gun than it is now, make it easier [3] for people, or keep the rules the same [2].
8. Which one of the opinions on this page best agrees with your view [on abortion]? By law, abortion should never be permitted [3]; The law should permit abortion only in case of rape, incest, or when the woman's life is in danger [2]; The law should permit abortion for reasons other than rape, incest, or danger to the woman's life, but only after the need for the abortion has been clearly established [2]; By law, a woman should always be able to obtain an abortion as a matter of personal choice [1].
9. Do you think gay or lesbian couples, in other words, homosexual couples, should be legally permitted to adopt children? Yes [1], No [3].

10. This country would have many fewer problems if there were more emphasis on traditional family ties. Do you agree strongly [5], agree somewhat [4], neither agree nor disagree [3], disagree somewhat [2], or disagree strongly [1] with this statement?

# References

Acemoglu D, Robinson J (2008) Persistence of power, elites, and institutions. Am Econ Rev 98:267–293

Adams J, Merrill S III (2002) Centrifugal incentives in multi-candidate elections. J Theor Polit 14:275–300

Adams J, Merrill S III (2005) Policy seeking parties in a parliamentary democracy with proportional representation: a valence-uncertainty model. Br J Polit Sci 39:539–558

Aldrich J (1983) A Downsian spatial model with party activists. Am Polit Sci Rev 77:974–990

Ansolabehere S, Snyder J (2000) Valence politics and equilibrium in spatial election models. Public Choice 103:327–336

Ansolabehere S, de Figueiredo JM, Snyder J (2003) Why is there so little money in U.S. politics? J Econ Perspect 17:105–130

Aragones E, Palfrey T (2002) Mixed equilibrium in a Downsian model with a favored candidate. J Econ Theory 103:131–161

Aragones E, Palfrey T (2005) Spatial competition between two candidates of different quality: the effects of candidate ideology and private information. In: Austen-Smith D, Duggan J (eds) Social choice and strategic decisions: eassays in honor of Jeffrey Banks. Springer, Berlin

Ashworth S, Bueno de Mesquita E (2009) Elections with platform and valence competition. Games Econ Behav 67:191–216

Banks J, Duggan J (2005) Probabilistic voting in the spatial model of elections: the theory of office motivated candidates. In: Austen-Smith D, Duggan J (eds) Social choice and strategic decisions: essays in honor of Jeffrey Banks. Springer, Berlin

Baron DP (1994) Electoral competition with informed and uninformed voters. Am Polit Sci Rev 88:33–47

Calvert RL (1985) Robustness of the multidimensional voting model: candidates, motivations, uncertainty and convergence. Am J Polit Sci 29:69–85

Clarke HD, Sanders D, Stewart M, Whiteley P (2005) Political choice in Britain. Oxford University Press, Oxford

Clarke HD, Kornberg A, Scotto T (2009a) Making political choices. Toronto University Press, Toronto

Clarke HD, Sanders D, Stewart M, Whiteley P (2009b) Performance politics and the British voter. Cambridge University Press, Cambridge

Coate S (2004) Political competition with campaign contributions and informative advertising. J Eur Econ Assoc 2:772–804

Condorcet N (1785) Essai sur l'application de l'analyse a la probabilite des decisions rendus a la pluralité des voix. Imprimerie Royale, Paris

Coram A (2010) Resource spending over time in an electoral competition. Elect Stud 29:497–508

Dal Bo E (2007) Bribing voters. Am J Polit Sci 51:789–803

Downs A (1957) An economic theory of democracy. Harper and Row, New York

Duggan J (2006) Candidate objectives and electoral equilibrium. In: Weingast BR, Wittman DA (eds) The Oxford handbook of political economy. Oxford University Press, Oxford

Duggan J, Fey M (2005) Electoral competition with policy-motivated candidates. Games Econ Behav 51:490–522

Duggan J, Kalandrakis T (2011) A Newton collocation method for solving dynamic bargaining games. Soc Choice Welfare 36:611–651

Enelow JM, Hinich M (1982) Nonspatial candidate characteristics and electoral competition. J Polit 44:115–131

Enelow JM, Hinich M (1989) The location of American presidential candidates. Math Comput Model 12:417–435

Erikson RS, Romero DW (1990) Candidate equilibrium and the behavioral model of the vote. Am Polit Sci Rev 84:1103–1126

Fiorina M, Abrams SJ (2009) Disconnect: the breakdown of representation in American politics. University of Oklahoma Press, Norman

Fiorina M, Abrams SJ, Pope JC (2005) Culture war?: the myth of a polarized America. Pearson Longman, New York

Groseclose T (2001) A model of candidate location when one candidate has a valence advantage. Am J Polit Sci 45:862–886

Groseclose T, Snyder J (1996) Buying supermajorities. Am Polit Sci Rev 90:303–315

Grossman GM, Helpman E (1994) Protection for Sale. Am Econ Rev 84:833–850

Grossman GM, Helpman E (1996) Electoral Competition and Special Interest Politics. Rev Econ Stud 63:265–286

Grossman GM, Helpman E (2001) Special interest groups. MIT Press, Cambridge

Hacker JS, Pierson P (2006) Off center: the republican revolution and the erosion of American democracy. Yale University Press, New Haven

Hacker JS, Pierson P (2010) Winner-take-all politics: how washington made the rich richer – and turned its back on the middle class. Simon and Schuster, New York

Herrera H, Levine D, Martinelli C (2008) Policy platforms, campaign spending and voter participation. J Public Econ 92:501–513

Jackson M, Watts A (2010) Social games: matching and the play of finitely repeated games. Games Econ Behav 70:170–191

Jessee SA (2009) Spatial voting in the 2004 presidential election. Am Polit Sci Rev 103:59–81

Jessee SA (2010) Partisan bias, political information and spatial voting in the 2008 presidential election. J Polit 72:327–340

Layman GC et al (2010) Activists and conflict extension in American party politics. Am Polit Sci Rev 104:324–346

Madison J ([1787],1999) The Federalist No. 10. In: Rakove J (ed) James Madison: writings. The Library of America, New York

McCarty N, Poole K, Rosenthal H (2006) Polarized America. MIT Press, Cambridge

McKelvey RD (1979) General conditions for global intransitivities in formal voting models. Econometrica 47:1085–1112

McKelvey RD, Patty J (2006) A theory of voting in large elections. Games Econ Behav 57:155–180

McLennan A (1998) Consequences of the Condorcet jury theorem for beneficial information aggregation by rational agents. Am Political Sci Rev 92:413–418

Meirowitz A (2008) Electoral contests, incumbency advantages and campaign finance. J Polit 70:681–699

Miller G, Schofield N (2003) Activists and partisan realignment in the U.S. Am Polit Sci Rev 97:245–260

Miller G, Schofield N (2008) The transformation of the republican and democratic coalitions in the U.S. Perspect Polit 6:433–450

Patty J (2002) Equivalence of objectives in two-candidate elections. Public Choice 112:151–166

Patty J (2007) Generic difference of expected vote share maximization and probability of victory maximization in simple plurality elections with probabilistic voters. Soc Choice Welfare 29:149–173

Patty J, Snyder JM, Ting M (2009) Two's company, three's an equilibrium: strategic voting and multicandidate elections. Quart J Polit Sci 4:251–278

Penn E (2009) A model of far-sighted voting. Am J Polit Sci 53:36–54

Peress M (2010) The spatial model with non-policy factors: a theory of policy motivated candidates. Soc Choice Welfare 34:265–294

Pierson P, Skocpol T (eds) (2007) The transformation of American politics: activist government and the rise of conservatism. Princeton University Press, Princeton

Poole K, Rosenthal H (1984) U.S. presidential elections 1968–1980: a spatial analysis. Am J Polit Sci 28:283–312

Riker WH (1980) Implications from the disequilibrium of majority rule for the study of institutions. Am Polit Sci Rev 74:432–446

Riker WH (1982) Liberalism against populism. Freeman, San Francisco

Schofield N (1978) Instability of simple dynamic games. Rev Econ Stud 45:575–594

Schofield N (2002) Evolution of the constitution. Br J Polit Sci 32:1–20

Schofield N (2006) Equilibria in the spatial stochastic model with party activists. Rev Econ Des 10:183–203

Schofield N (2007) The mean voter theorem: necessary and sufficient conditions for convergent equilibrium. Rev Econ Stud 74:965–980

Schofield N, Cataife G (2007) A model of political competition with activists with an application to the elections of 1989 and 1995 in Argentina. Math Soc Sci 53:213–231

Schofield N, Miller G (2007) Elections and activist coalitions in the U.S. Am J Polit Sci 51:518–531

Schofield N, Sened I (2006) Multiparty democracy: elections and legislative politics. Cambridge University Press, Cambridge

Schofield N, Miller G, Martin A (2003) Critical elections and political realignments in the U.S.: 1860–2000. Pol Stud 51:217–240

Schofield N, Claassen C, Ozdemir U, Zakharov AV (2011) Estimating the Effects of activists in two-party and multi-party systems: comparing the United States and Israel. Soc Choice Welfare 36:483–518

Serra G (2010) Polarization of what? A model of elections with endogenous valence. J Polit 72:426–437

Snyder J, Ting M (2008) Interest groups and the electoral control of politicians. J Public Econ 92:482–500

Stokes D (1963) Spatial models and party competition. Am Polit Sci Rev 57:368–377

Stokes D (1992) Valence politics. In: Kavanagh D (ed) Electoral politics. Clarendon, Oxford

Stokes SC (2005) Perverse accountability: a formal model of machine politics with evidence from Argentina. Am Polit Sci Rev 99:315–325

Sundquist JL (1973) Dynamics of the party system: alignment and realignment of political parties in the United States. The Brookings Institution, Washington, DC

Train K (2003) Discrete choice methods for simulation. Cambridge University Press, Cambridge

Wittman D (1977) Candidates with policy preferences: a dynamic model. J Eco Theory 14:180–189

Zakharov AV (2009) A model of candidate location with endogenous valence. Public Choice 138:347–366

# Modelling Elections in Post-Communist Regimes: Voter Perceptions, Political Leaders and Activists

**Norman Schofield, JeeSeon Jeon, Marina Muskhelishvili, Ugur Ozdemir, and Margit Tavits**

## 1 Introduction

Recent work has argued that institutional characteristics of political systems, such as presidentialism vs. parliamentarianism, or majoritarianism vs. proportionality, will have significant effects on the stability of government and the nature of redistributive politics.[1] These arguments have been based on cross country empirical analyses and relatively simple one dimensional spatial models. The formal underpinning of these models has often been based on the assumption that parties or candidates adopted positions in order to win. This assumption leads to the inference that parties will converge to the electoral median (under deterministic voting in one dimension, as in Downs 1957; Riker and Ordeshook 1973) or to the electoral mean in stochastic models (See the discussion in Schofield et al., 2011a.).

These various spatial models treat vote choice as a function of voters' policy preferences only. Yet, in almost every polity we witness electoral or policy outcomes that are difficult to explain in terms of the pure spatial model. An example would be the apparent increase in "polarization" even in mature democracies such as the United Kingdom and the United States. There has also been evidence of political fragmentation in established democracies such as the Netherlands and Belgium, as well as in Post-Communist East European countries.[2] These observations

---

[1]Bawn and Rosenbluth (2005), Persson and Tabellini (2000, 2003), Dow (2001, 2011), Ezrow (2010, 2011).

[2]See Markowski and Tucker (2010a) and de Vries and Edwards (2009) on "extremist" Euroskeptic parties.

N. Schofield (✉), J. Jeon, U. Ozdemir, and M. Tavits
Center in Political Economy, Washington University in Saint Louis, MO 63130
e-mail: schofield.norman@gmail.com; jsjeon@wustl.edu; uozdemir@wustl.edu; tavits@wustl.edu

M. Muskhelishvili
Center for Strategic Studies, 19, I.Chavchavadze ave. Tbilisi 0179, Georgia
e-mail: marinmus@yahoo.com

suggest that the pure spatial model is missing something fundamental: centrist political equilibrium cannot be a defining property of electoral politics.

The analysis presented here suggests that the differences in political configurations may result from the very different incentives that activist groups face in the various institutional environments. In this paper we contrast electoral models for a proportional electoral system used for parliamentary elections in Poland with the winner take-all-electoral system used for presidential elections in Georgia and Azerbaijan.

To estimate voter positions we use the national election surveys and carry out a factor analysis of this responses. This allows us to construct, for each election, a policy space, $X$. Using the voter information, we can infer a preferred policy point, $x_i$, in $X$, for each voter.

There are a number of ways to estimate party or candidate positions. One can use expert estimates, as in Benoit and Laver (2006), or analysis or party manifestos (Budge et al. 1987, 2001; Ezrow 2010). A difficulty with these methods is to guarantee that the space of voter preferences and the space of party locations is the same.[3] Here we estimate party positions using the notion of *partisan constituencies*. The idea here is that party leaders can fairly easily obtain information about the policy positions of their supporters, and each can respond by advocating policies that are close to the mean of the preferences of their respective supporters. This satisfies what Huber and Powell (1994) call ideological congruence between citizens and policy makers. For each party or candidate, we estimate the position by taking the mean of the positions of the voters who actually chose that party or candidate.

On the other hand, the standard Downsian (1957) model of political competition is that of "opportunistic," office seeking parties. Each voter is assumed to choose the party whose policy position is closest while parties are assumed to maneuver so as to gain as many votes as possible.

To estimate such opportunistic behavior we model the relationship between electoral response and party positions using a mixed logit stochastic model. On the basis of such an empirical electoral model, we then use the results of a general formal model to determine how changes in party position effect election results. It is then natural to seek the existence of "Nash equilibria" in the empirical model – a set of party positions from which no party may deviate to gain advantage in terms of its vote share. Since the "utility functions" of parties are, in fact unknown, it is possible to use "counter factual experiments" to make inferences about the political game. That is, after modelling the relationship between party positions and election outcome (for a given electoral distribution), we may make assumptions about the utility functions of leaders and examine the Nash equilibria under these assumptions, to determine whether the Nash equilibria so determined correspond to the actual positions of the parties or candidates.

The technique we use to compare elections in different polities is a formal stochastic model of elections that emphasizes the importance of *valence*. As discussed in Schofield et al. (2011a), the standard Downsian spatial model is based on the assumption that it is only party *positions* that matter to voters. In the models we

---

[3]In Schofield et al. (2011a) we estimated the candidate positions using the same survey questionnaire.

employ we utilize the notion of valence of candidates or party leaders. By valence we mean the electoral perception of the *quality* of candidates (Stokes 1963, 1992). Based on the empirical work presented here, we argue that neither the Downsian convergence result nor the "social chaos theorem" (Riker 1980) gives a complete picture of elections. Both position and valence matter in a fundamental way. We then use this model to suggest that the nature of the electoral system influences the calculations of the leaders of the activist groups who provide the resources that are critical for political success.

We consider an empirical stochastic model, denoted $\mathbb{M}(\boldsymbol{\lambda}, \beta)$, where $\boldsymbol{\lambda}$ is the vector of party valences, and $\beta$ is the spatial parameter. As shown in Schofield et al. (2011a), there exist a "convergence coefficient", denoted $c(\boldsymbol{\lambda}, \beta)$, defined by $(\boldsymbol{\lambda}, \beta)$ and the covariance matrix of the voter preferred positions. A previous theorem (Schofield 2007) asserts that if the dimensionless coefficient, $c(\boldsymbol{\lambda}, \beta)$, exceeds 2, then according to the pure spatial model, under any vote maximizing Nash equilibrium, all parties should diverge away from the electoral origin.

To illustrate this result, we examine a sequence of elections in the multiparty polity of Poland for 1997, 2001 and 2005. In these three election models, the $\beta$ coefficients are all highly significant and take values about 1.5. Indeed the convergence coefficients are calculated to lie in the range [5.92, 6.82]. Moreover, the Hessian of the lowest valence party at the joint origin is shown to have both eigenvalues positive in each election. This implies that the origin is a vote *minimizing* position for such a party. As a consequence we infer that any Nash equilibrium under the vote maximizing spatial model is one where *all* parties *diverge* from the origin.[4] We verified this inference by simulating these models to determine the equilibria in the spatial models, with and without sociodemographic variables, and confirmed their divergent nature. Our estimates of party positions as well as the equilibria suggested that the political system in Poland in these years could be seen to be quite chaotic.

As a second example we modelled the 2008 presidential election in Georgia. This post-communist polity has a very powerful president, Mikheil Saakashvili, and a fragmented opposition. It may be regarded as a partial democracy or "anocracy".[5] By "anocracy" we mean a polity with some democratic methods in place, but where the media are weak and opposition groups find it difficult to coordinate. As in Poland, the spatial model for Georgia gives two dimensions: westernization and democracy. A positive value in the *West* dimension is taken to mean a strong anti-western attitude. The democracy dimension is defined by voters' judgement about current democratic environment in Georgia. Larger values in the democracy dimension are associated with negative judgement about the current state of democratic institutions in Georgia, coupled with a demand for a greater democracy. Our analysis obtained a $\beta$ coefficient of 0.78 but a high value of $c(\boldsymbol{\lambda}, \beta) = 2.39$,

---

[4]Similar results have been obtained for Israel and Turkey (Schofield et al. (2011b, e).

[5]See Epstein et al. (2006), Gandhi and Vreeland (2004), Vreeland (2008), Regan and Bell (2010), Fjelde (2010) for discussion of stability in an anocracy.

implying that the small, low valence parties should diverge from the origin. Our equilibrium analysis indicates that Saakashvili, with very high valence, should locate near the electoral origin. Our estimate of his policy was that his position was very pro-west and opposed to further democratization.

The third example is the election in Azerbaijan in November 2010, where President Ilham Aliyev's ruling Yeni Azerbaijan Party obtained a majority of 72 out of 125 seats. Other independent candidates, aligned with the government, received 38 seats, and 10 small opposition or quasi-opposition parties won the remaining 13 seats. The survey obtained by the Institute of Strategic Studies of the Caucasus allowed us to infer that the policy space was uni-dimensional. Our analysis obtained a $\beta$ coefficient of 1.34, and indicated that that the Hessian of the low valence opposition had a single positive eigenvalue, implying divergence away from the origin by all parties. This model is only one dimensional, so the result is not quite compatible with the analysis of Georgia. However, if the model were two-dimensional, and symmetric in the sense that voter variances were identical on both axes, then the convergence coefficient would be $c(\lambda, \beta) = 2.89$, very similar to the result for Georgia.

In contrast to these three examples, we have shown in Schofield et al. (2011a) that the unique vote maximizing equilibrium for both the 2000 and 2004 elections in the U.S. had both candidates adopting positions at, or very close to the electoral origin.[6] The considerable difference between Nash equilibria in elections in Poland, Georgia and Azerbaijan, in contrast to the analysis presented for the United States in Schofield et al. (2011a) suggests that that the difference may be due to a very different logic governing the influence of activist groups in these different polities.

Based on a comparison of estimated and simulated equilibrium positions for the three elections in Poland, we argue that the difference between the estimated party positions and the equilibrium positions is much less dramatic than in the U.S., suggesting that the influence of activists is less pronounced in Poland than in the United States. We infer that this is because under the proportional electoral system of Poland small activist groups can still expect to influence policy outcomes, through party membership of coalition government. Thus there is little tendency for activist groups to coalesce under this method of proportional electoral rule. Consequently, political fragmentation will be maintained.

In Georgia and Azerbaijan, on the other hand, because the presidential election is a winner-take-all system, based therefore on plurality rule, coalitions are not the norm. The media play a crucial role in enhancing candidates' valence. Opposition groups, with restricted access to the media, find it very difficult to combine resources behind a single opposition leader. This tends to preserve the dominance of the presidential party coalition.

In the conclusion we discuss results on other polities with proportional electoral systems such as Israel and Turkey, and comment that the models for recent elections in these polities also have high convergence coefficients in the range [4.0, 6.0].

---

[6]A similar result has been obtained for the 2008 U.S. Presidential election (Schofield et al. (2011b).

In contrast, parliamentary polities with fairly majoritarian electoral systems, like Canada and the United Kingdom, have convergence coefficients in the medium range of [0.84, 1.94]. Other work (Schofield and Zakharov 2010) has also obtained a value for the convergence coefficient for the 2007 Duma election in Russia of 1.7. For the anocratic polities of Georgia and Azerbaijan the convergent coefficient, or its analogue, lies in the range [2.3, 3.0]. These values are quite different from the low values for $c(\lambda, \beta)$ obtained for the United States which we found to lie in the range [0.4, 1.1]. We suggest that the convergence coefficient of a polity is a theoretically useful way of classifying the fundamental properties of the electoral system.

## 2  Elections in Poland 1997–2005

### 2.1  Background

Poland held regular elections in 1997, 2001, and 2005. For all of these elections Poland used an open-list proportional representation (OLPR) electoral system with a threshold of 5% nationwide vote for parties and 8% for electoral coalitions. The rules of the 1997 elections were slightly different from the ones used since 2001: the number of districts was larger (52 compared to 41) and in addition to districts there was a 69-seat national list. In 1997 and since 2005 votes are translated into seats by the D'Hondt method rather than the more proportional modified Saint-Laguë method used in 2001.

The party system in Poland is relatively unstable – in each election new parties emerge and some existing ones die, and the vote shares fluctuate considerably for those parties that manage to survive multiple elections. Table 1 lists, by election

**Table 1** Seats in Polish Sejm elections

| Party | 1997 (%) | 2001 (%) | 2005 (%) |
|---|---|---|---|
| Democratic Left Alliance (SLD) | 164 (35.6) | 200[a] (43.4[a]) | 55 (12.0) |
| Polish People's Party (PSL) | 27 (5.8) | 42 (9.1) | 25 (5.4) |
| Freedom Union (UW) | 60 (13.0) | 0 | |
| Solidarity Election Action (AWS) | 201 (43.6) | 0 | |
| Labor Party (UP) | 0 | 16[a] (3.5[a]) | |
| Union of Political Realism (UPR) | 0 | | |
| Movement for Reconstruction of Poland (ROP) | 6 (1.3) | | |
| Self Defense, Samoobrona (SO) | | 53 (11.5) | 56 (12.1) |
| Law and Justice (PiS) | | 44 (9.5) | 155 (33.7) |
| Civic Platform (PO) | | 65 (14.1) | 133 (29.0) |
| League of Polish Families (LPR) | | 38 (8.2) | 34 (7.4) |
| Democratic Party (DEM) | | | 0 |
| Social Democracy of Poland (SDP) | | | 0 |
| German minority | 2 | 2 (0.4) | 2 (0.4) |
| Total | 460 | 460 | 460 |

[a]Coalition of SLD with UP

**Table 2** Vote shares in Polish Sejm elections

| Party | 1997 | 2001 | 2005 |
|---|---|---|---|
| Democratic Left Alliance (SLD) | 27.1 | 41.0[a] | 11.3 |
| Polish People's Party (PSL) | 7.3 | 9.0 | 7.0 |
| Freedom Union (UW) | 13.4 | 3.1 | |
| Solidarity Election Action (AWS) | 33.8 | 5.6 | |
| Labor Party (UP) | 4.7 | | |
| Union of Political Realism (UPR) | 2.0 | | |
| Movement for Reconstruction of Poland (ROP) | 5.6 | | |
| Self Defense (SO) | | 10.2 | 11.4 |
| Law and Justice (PiS) | | 9.5 | 27.0 |
| Civic Platform (PO) | | 12.7 | 24.1 |
| League of Polish Families (LPR) | | 7.9 | 8.0 |
| Democratic Party (DEM) | | | 2.5 |
| Social Democracy of Poland (SDP) | | | 3.9 |

[a]Coalition of SLD with UP

year, the names of the parties and their seat shares while Table 2 gives their vote shares. Usually about five or six parties win seats in the Sejm (lower house).

The main political parties during the time period under consideration include the following. The left-wing ex-communist Democratic Left Alliance (SLD) and the agrarian Polish Peoples' Party (PSL), both of which have participated in all three elections considered here and been the most frequent governing parties in the post-communist period. In 1997 Solidarity Election Action (AWS) and the Freedom Union (UW) were also important players. Both parties had grown out of the Solidarity movement. AWS combined various mostly right wing and Christian groups under one label, while UW was formed based on the liberal wing of Solidarity. After the 2001 election, Civic Platform (PO), Law and Justice (PiS), League of Polish Families (LPR), and Self-Defense (SO) emerged as significant new parties. The first three parties were formed on the ruins of AWS and UW. PO combines the liberals from both parties, while PiS represents the conservatives. LPR's ideology combines nationalism with Catholic fundamentalism and the party is sometimes considered a far-right entity. SO is a leader-centered agrarian party that is left-wing on economic policy but very right-wing religious on values. Both LPR and SO did not survive as significant political players and are no longer represented in the Polish Sejm.

Existing literature suggests that the two main axis of Polish electoral politics along which both voters and parties align are the economic dimension and social values dimension (Kitschelt et al. 1999; Fidrmuk 2000a, b; Powers and Cox 1997; Tavits and Letki 2009; Tucker 2006; Owen and Tucker 2010; Markowski 2006). This has remained true for the entire post-communist era. The first dimension encompasses issues related to economic transition and economic performance such as the speed and nature of privatization, reducing unemployment, and increasing social security. The social values' dimension includes attitudes towards communist past, the role of church in politics, moral issues, and nationalism (Grzymala-Busse 2002; Szczerbiak 1998). Over the years, these social issues have gained increasing prominence in political rhetoric and as determinants of

vote choice (Markowski and Tucker 2010a, b). The relevance of social issues is further underlined by the significant influence of the Catholic church on Polish party politics (Markowski 2006) and the high salience of the divide between the anti-communists and ex-communists.

## 2.2   The Elections

We analyzed the three Polish elections based on data from the respective Polish National Election Studies (PNES). These are surveys of the adult population conducted after each national parliamentary election. We were able to use responses from samples of sizes 660, 657 and 1,095, respectively for the pure spatial models. The dependent variable in our analyses is the respondent's vote choice. We use the spatial distance between parties and voters, and voters' socio-demographic characteristics to explain this vote choice. See Appendix 1 for the question wordings.

The PNES includes a battery of questions asking respondents' position on various issues. We identified issues pertaining to economic policy and social values and performed factor analysis to confirm the existence of the two dimensions in the data and obtain factor scores for each dimension. The following items loaded on the two dimensions (the items used depend on what was available in a given survey).

Economic dimension (all years): privatization vs. state ownership of enterprises, fighting unemployment vs. keeping inflation and government expenditure under control, proportional vs. flat income tax, support vs. opposition to state subsidies to agriculture, state vs. individual social responsibility.

Social values dimension: separation of church and state vs. influence of church over politics (1997, 2001, 2005), complete decommunization vs. equal rights for former nomenclature (1997, 2001), abortion rights regardless of situation vs. no such rights regardless of situation (1997, 2005).[7]

The factor loadings for the two dimensions are given in Appendix 2.

We adopted the notion of *partisan constituencies* and estimated party positions on these dimensions by taking the average of the positions of the voters for each party. In an alternative analysis, we obtained the information on the placement of political parties from Benoit and Laver (2006), which uses expert surveys to place parties on a variety of issues. The results of these alternative analyses were substantively similar to the ones presented here. However, the Benoit and Laver data were collected after the 2001 elections only. Using these placements to identify party positions in 1997 and 2005 may not be accurate because party positions may have changed. We therefore decided to use the more time-sensitive measures obtained from the PNES for the final analyses presented here.

---

[7]Respondent's opinion on each of these issues was recorded on an 11-point scale with the first option given scored as zero and the second option scored as ten. See Appendix 1 for the exact question wording.

Figures 1–3 display the estimate of the density contours of the electoral distribution of voter bliss points for each election year, as well as the estimated party positions.[8] Figures 4–6 give estimated Nash equilibria for these elections.

These party positions are given below.

$$\mathbf{Z}^*_{1997} = \begin{bmatrix} Party & SLD & PSL & UW & AWS & UP & UPR & ROP \\ x & 0.03 & -0.35 & 0.52 & 0.005 & 0.29 & 1.81 & 0.15 \\ y & -0.72 & -0.35 & -0.1 & 0.72 & -0.15 & -0.15 & 0.75 \end{bmatrix}$$
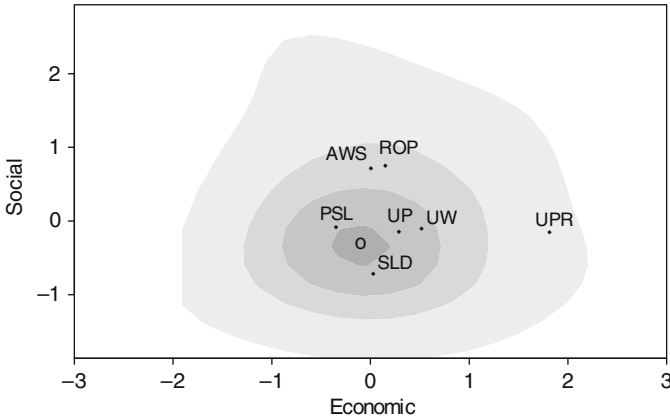


**Fig. 1** Voter distribution and party positions in Poland in 1997
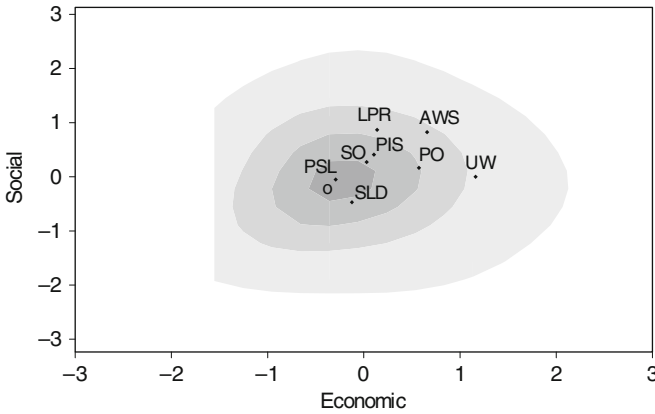


**Fig. 2** Estimated party positions in Poland in 2001

---

[8]For 2001, the positions of the LPR PO, PSL, SLD and UW are almost identical to those estimated by Benoit and Laver (2006), thus providing some justification for our method of estimating party positions.
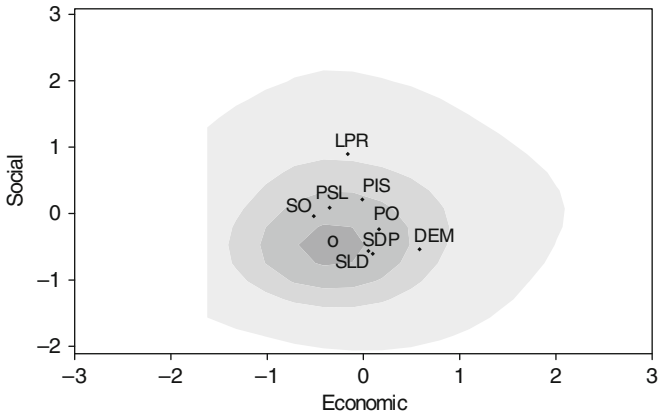
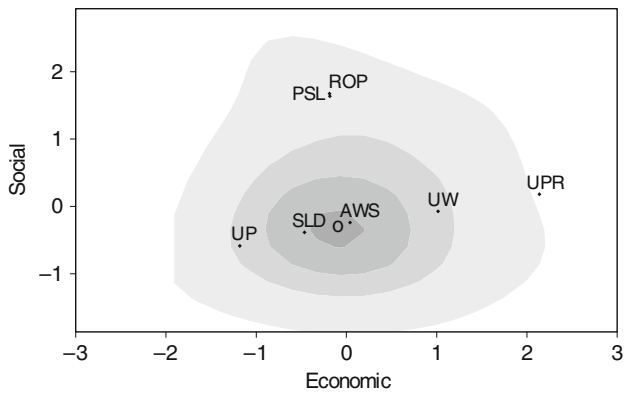**Fig. 3** Estimated party positions in Poland in 2005



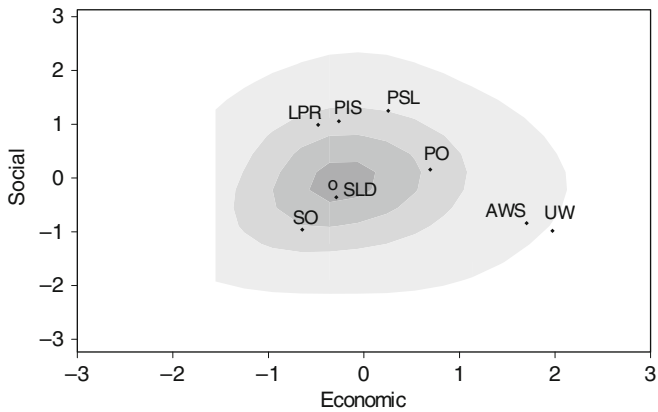**Fig. 4** Equilibrium positions under the joint model in Poland in 1997



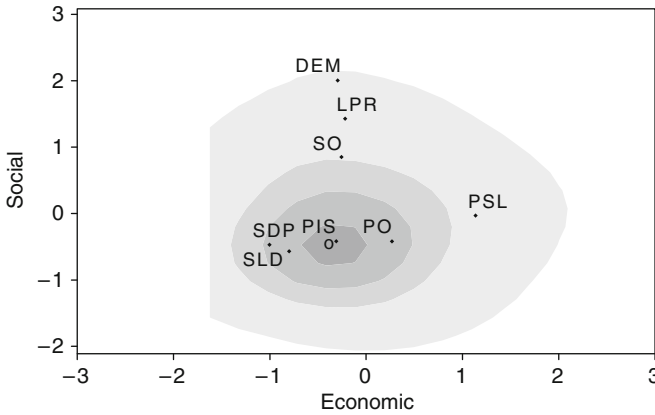**Fig. 5** Equilibrium positions under the joint model in Poland in 2001

**Fig. 6** Equilibrium positions under the joint model in Poland in 2005

In 1997, Solidarity Electoral Action (AWS), with 201 seats and based on the Solidarity trade union, formed a coalition with the Freedom Union (UW), a party on the right, supporting classical liberalism, with 60 seats. Together the coalition controlled 261 seats, out of 460. The election was a major setback for the Democratic Left Alliance (SLD) and the Polish People's Party (PSL) which were forced out of government.

$$\mathbf{Z}^*_{2001} = \begin{bmatrix} Party & SLD,\, UP & PSL & UW & AWS & SO & PiS & PO & LPR \\ x & -0.12 & -0.29 & 1.16 & 0.66 & 0.03 & 0.11 & 0.57 & 0.14 \\ y & -0.47 & -0.05 & 0.002 & 0.83 & 0.27 & 0.41 & 0.17 & 0.87 \end{bmatrix}$$

In the 2001 election, the coalition of SLD and UP won 216 of the 460 seats, and was able to form a government with the support of the Polish People's Party (PSL), with 42 seats, thus controlling 258 seats in all. The former ruling parties, the Solidarity Electoral Action (AWS) and the Freedom Union (UW) only gained about 10% of the vote but no seats. In its place several new parties emerged, including the center right LPR, SO, and PiS, and the further right PO. Figures 1 and 2 suggest that the AWS fractured into five factions, a small remnant AWS, and these four new parties.

$$\mathbf{Z}^*_{2005} = \begin{bmatrix} Party & SLD & PSL & DEM & SDP & SO & PiS & PO & LPR \\ x & 0.05 & -0.35 & 0.58 & 0.10 & -0.52 & -0.01 & 0.16 & -0.16 \\ y & -0.56 & 0.09 & -0.54 & -0.61 & -0.04 & 0.20 & -0.23 & 0.90 \end{bmatrix}$$

After 2003 a variety of factors combined to bring about a collapse of support for the government of the SLD-UP-PSL coalition. Discontent with high unemployment, government spending cuts (especially on health, education and welfare) and privatization was compounded by a series of corruption scandals, leading to the resignation of the Prime Minister Leszek Miller in May 2004, who was succeeded by Marek Belka.
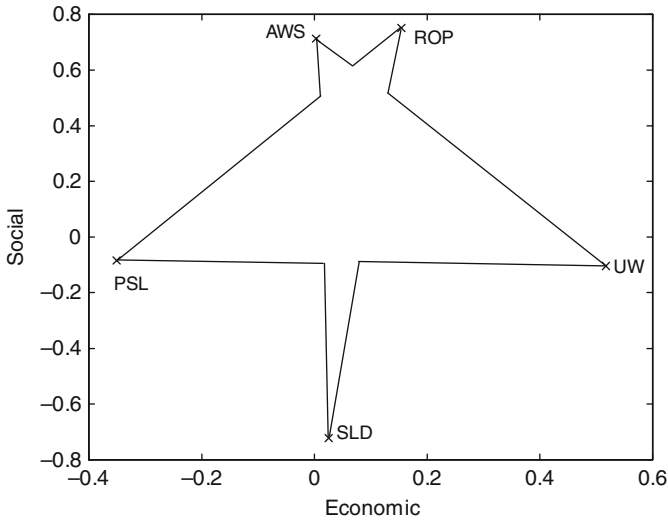
**Fig. 7** Estimate of the heart in 1997 in Poland

The parties running in the 2005 election were similar to those running in 2001, with the addition of SDP (a left wing splinter group from the SLD), and the right wing Democratic Party (DEM). Figure 7 suggests that the DEM was formed from the Freedom Union (UW), the moribund Solidarity Electoral Action (AWS) and some right wing SLD dissidents. Both these new parties failed to win seats, though they took about 6% of the vote.

The two larger center right parties, Law and Justice (PiS) and Civic Platform (PO), did much better in 2005, gaining over 60% of the vote and 288 seats. They had splintered off from the anti-communist Solidarity movement but differed on issues such as the budget and taxation. Law and Justice, with 155 seats, had a policy of tax breaks and state aid for the poor, and pledged to uphold traditional family and Christian values, while being suspicious of economic liberalism. The Civic Platform, with 133 seats, supported free market forces and wanted to introduce a flat 15% rate for income tax, corporation tax and VAT. It promised to move faster on deregulation and privatisation, in order to adopt the euro as soon as possible.

Negotiations between PiS and PO about forming the new government collapsed in late October, precipitated by disagreement over who would be speaker of the Sejm. The PiS leader, Jaroslaw Kaczyński, declined the opportunity to become Prime Minister so as not to prejudice the chances of his twin brother, Lech Kaczyński, in the presidential election.[9] On 1 November, 2005, the PiS announced a minority government, with 155 seats, led by Kazimierz Marcinkiewicz as the Prime Minister.

---

[9] Lech Kaczyński became President after that election, but died in the airplane crash on April 10, 2010, on his way to Russia to commemorate the Katyn massacre of Polish officers in 1940.

A major stumbling block against the PiS forming a coalition with the PO was the insistence by the PO that it receive the Interior portfolio, if it were to enter a coalition government with the PiS, to prevent one party from controlling all three of the "power" ministries (Security, Justice and Interior), thus the police and security services. The PO also opposed a "tactical alliance" between the PiS and Samoobrona, who shared eurosceptic and populists sentiments, although differing on economic policy. The election campaign, in which both of these center-right parties had competed mainly against each other, rather than with parties on the left, accentuated differences and created an antagonistic relationship between the two parties.

The PiS minority government depended on the support of the radical Samoobrona (SO), with 56 seats, and the conservative League of Polish Families (LPR), with 34 seats. On 5 May 2006 PiS formed a coalition government with Samoobrona and LPR, controlling 245 seats. In July 2006, Marcinkiewicz tendered his resignation, because of disagreements with the PiS party leader, Jaroslaw Kaczyński. Kaczyński then formed a new minority government and was sworn in on July 14, 2006, finally becoming prime minister. His party, Law and Justice, was defeated in the November 2007 election and Donald Franciszek Tusk, co-founder and chairman of Civic Platform, became Prime Minister.[10]

Figure 3 indicates the policy differences that existed between the PiS and the more left-wing Samoobrona, SO, and the centrist LPR on the one hand, and the more right-wing party, the PO, on the other.

As the tables on election results illustrate, the electoral system in Poland is highly proportional, though the SLD gained a higher seat share than vote share in 1997 and 2001.

Tables 3–5 give the party valences for three pure spatial logit models (one for each election year) based on the estimated positions of the parties. We also estimated pure sociodemographic models and joint models, based on the spatial model and including sociodemographic variables. For the sociodemographic variables we chose age in years, regular monthly income, former communist party membership, and religiosity (believer vs. atheist or agnostic). This choice follows previous literature that identifies these demographics as important determinants of vote choice and party preference (Markowski 2006; Wade et al. 1995).[11] Table 6 gives the comparison of the log likelihoods for these models for 1997. Indeed, the loglikelihoods for the joint models were superior to the pure spatial and sociodemographic models for all years. For all spatial models, in Tables 3–5, the $\beta$-coefficient is highly significant (at the 0.001 level). The high valence values were also significant in the pure spatial and joint models. Only a few of the sociodemographic variables were found to be significant.

---

[10]After President Lech Kaczyński's death in the plane crash in April, 2010, his brother, Jaroslaw Kaczyński, ran against acting president Bronislaw Komorowski in the presidential election on 20 June. Kaczyński received 36.46% of votes in the first round, while Komorowski received 41.54%. In the second round, Kaczyński was defeated with 47% of the vote to Komorowski's 53%.

[11]See Schofield et al. (2010) for full details of these joint models.

**Table 3** Poland 1997 pure spatial model (Base = ROP)

| Variable | Party | Coefficient | Std. error | |t-Value| |
|---|---|---|---|---|
| Spatial $\beta$ | | 1.739*** | 0.116 | 15.04 |
| Valence $\lambda$ | UP | −0.558 | 0.262 | 2.13 |
| | UW | 0.731*** | 0.199 | 3.66 |
| | AWS | 1.921*** | 0.174 | 11.046 |
| | SLD | 1.419*** | 0.19 | 7.47 |
| | PSL | 0.073 | 0.222 | 0.328 |
| | UPR | −2.348*** | 0.501 | 4.685 |
| $n = 660$ | LL = −855 | AIC = 1,725 | | |

*LL* loglikelihood
*Prob < 0.05; **prob < 0.01; ***prob < 0.001

**Table 4** Poland 2001 pure spatial model (Base = LPR)

| Variable | Party | Coefficient | Std. error | |t-Value| |
|---|---|---|---|---|
| Spatial $\beta$ | | 1.48*** | 0.118 | 12.61 |
| Valence $\lambda$ | SLD | 1.99*** | 0.174 | 11.41 |
| | AWS | −0.37 | 0.248 | 1.49 |
| | UW | −1.00*** | 0.308 | 3.24 |
| | SO | 0.41* | 0.202 | 2.04 |
| | PIS | 0.43* | 0.200 | 2.16 |
| | PSL | 0.09 | 0.218 | 0.41 |
| | PO | 0.80*** | 0.192 | 4.19 |
| $n = 657$ | LL = −1,004 | AIC = 2,024 | | |

*LL* log likelihood
*Prob < 0.05; **prob < 0.01; ***prob < 0.001

**Table 5** Poland 2005 pure spatial model (Base = LPR)

| Variable | Party | Coefficient | Std. error | |t-Value| |
|---|---|---|---|---|
| Spatial $\beta$ | | 1.55*** | 0.115 | 13.41 |
| Valence $\lambda$ | SO | 0.82*** | 0.161 | 5.09 |
| | DEM | −1.04*** | 0.260 | 4.01 |
| | SDP | −0.34 | 0.205 | 1.66 |
| | PIS | 1.95*** | 0.146 | 13.40 |
| | SLD | 0.47** | 0.172 | 2.72 |
| | PO | 1.50*** | 0.152 | 9.88 |
| | PSL | −0.17 | 0.196 | 0.85 |
| $n = 1,095$ | LL = −1,766 | AIC = 3,549 | | |

*LL* log likelihood
*Prob < 0.5; **prob < 0.01; ***prob < 0.001

**Table 6** Comparisons of LL for Poland in 1997

| | $\mathbb{M}_2$ | Joint | Spatial | Socio-dem. |
|---|---|---|---|---|
| $\mathbb{M}_1$ | Joint | na | 34 | 629 |
| | Spatial | −34 | na | 595 |
| | Socio-dem. | −595 | −629 | na |

Table 3 shows that the estimates for the pure spatial model in 1997 were:

$$(\lambda_{UPR}, \lambda_{UP}, \lambda_{ROP}, \lambda_{PSL}, \lambda_{UW}, \lambda_{SLD}, \lambda_{AWS}; \beta)$$
$$= (-2.3, -0.56, 0.0, 0.07, 0.73, 1.4, 1.92; 1.74)$$

The covariance matrix is:

$$\nabla_0 = \begin{bmatrix} 1.0 & 0.0 \\ 0.0 & 1.0 \end{bmatrix}.$$

As in Schofield et al. (2011a), the probability, $\rho_{UPR}$, that a voter chooses the lowest valence party, when all parties are at the joint origin, is given by the model $\mathbb{M}(\boldsymbol{\lambda}, \beta)$ as

$$\rho_{UPR} \simeq \frac{1}{1 + e^{1.92+2.3} + e^{1.4+2.3}}$$
$$= \frac{1}{1 + 66 + 40} \simeq 0.01$$

Thus $2\beta (1 - 2\rho_{UPR}) = 2 \times 1.74 \times 0.98 = 3.41$

$$\text{and } C_{UPR} = (3.41) \begin{bmatrix} 1.0 & 0.0 \\ 0.0 & 1.0 \end{bmatrix} - I$$
$$= \begin{bmatrix} 2.41 & 0.0 \\ 0.0 & 2.41 \end{bmatrix},$$

$$\text{so } c = 3.41 \times 2 = 6.82.$$

Using the necessary condition for $c$ for convergence from the valence theorem in Schofield et al. (2011a), we infer that all parties diverge in equilibrium. Similar results for the elections of 2001 and 2005 show divergence for these pure spatial models.

In 2001, we find $\beta = 1.482$, so $c \simeq 5.92$, and in 2005, $\beta = 1.548$, so $c \simeq 6.192$. See Tables 4 and 5.

Computation, using a MATLAB simulation program, showed the vote maximizing local equilibrium for 1997 to be the vector

$$\mathbf{Z}^{el}_{1997} = \begin{bmatrix} Party & SLD & PSL & UW & AWS & UP & UPR & ROP \\ x & -0.47 & -0.11 & 1.01 & 0.04 & -1.18 & 2.14 & -0.12 \\ y & -0.39 & 1.61 & -0.07 & -0.24 & -0.59 & 0.18 & 1.64 \end{bmatrix},$$

as shown in Fig. 4. Figures 5 and 6 give the equilibria in 2001 and 2005.[12]

All parties, in equilibrium, scatter away from the electoral origin. Note that in 1997, the two high valence parties, the AWS and the SLD, have equilibrium

---

[12]Because the Hessians have positive eigenvalues, the party preference correspondences are not convex valued, so no general argument can be used to assert existence of pure strategy Nash equilibria (PNE). If a PNE were to exist it would coincide with one of the LNE.

positions very close to the electoral origin. Similarly, in 2001 only the highest valence party, the SLD, and in 2005, only the highest valence party, the PIS, have equilibrium positions that are located at, or very close to, the electoral origin. The significant drop in the valence of the AWS between 1997 and 2001 should have forced it even further from the origin than the position that it did indeed adopt. A robust inference from these figures is that parties do not locate themselves at positions that maximise the vote shares, as estimated by the joint spatial model. We suggest that parties' positions are effectively decided by small activist groups whose preferred positions are adopted by the parties. For example, when the AWS fragmented in 2001, new parties like the PiS, SO, PO and LPR adopted positions in the upper right quadrant of the policy space. When the UW disappeared in 2005, its place was taken by the DEM, whose position was controlled by an activist faction that had controlled the UW. These observations are consistent with the hypothesis that the activist groups supporting the AWS and the UW fragmented in 2001, and this led to the creation of these new parties.

We can see the nature of bargaining over coalition governments in these three elections by constructing the "median lines" between pairs of parties that pivot between majority coalitions, as in Figs. 7–9. When these medians do not intersect, then they bound a finite, star shaped set known as the "heart". Schofield (1999) has suggested that each election heart gives a heuristic estimate of the set of possible coalition policy outcomes.

For example, note that the coalition government of AWS, and the small party, the UW, in 1997 can be represented by the upper right median in Fig. 7.

The coalition of the SLD and the small party, the PSL, in 2001, can be represented by the median line on the lower left in Fig. 8.

Finally, the complex negotiations involving the PiS and the small parties, the SO and LPR, against the PO in 2005 all refer to the triangular heart bounded by these party positions in the upper left of Fig. 9. If we are correct in our inference that the break-up of the AWS activist group led to the creation of the smaller SO, PiS and LPR parties, we may infer that the minority PiS government, supported by the SO and LPR provided policy benefits of some kind for the activist groups supporting these parties.[13] It is interesting to note that according to the spatial model, the PiS could have located itself at the electoral origin, in which case it would have been a core party, in the sense of Laver and Schofield (1990). To do so however, it would have had to change its policy position by moving "south" on the policy axis. Notice that the three coalitions that formed after these three elections were all minimal winning (Riker 1962) although the one in 2005 was a minority coalition with support. Obviously coalition formation in a fragmented polity is made very complex by the configuration of party positions.

These figures suggest that even small parties can hope to belong to government. It follows that activist groups supporting these parties can aspire to influence government policy. We hypothesize that such activist groups have little incentive

---

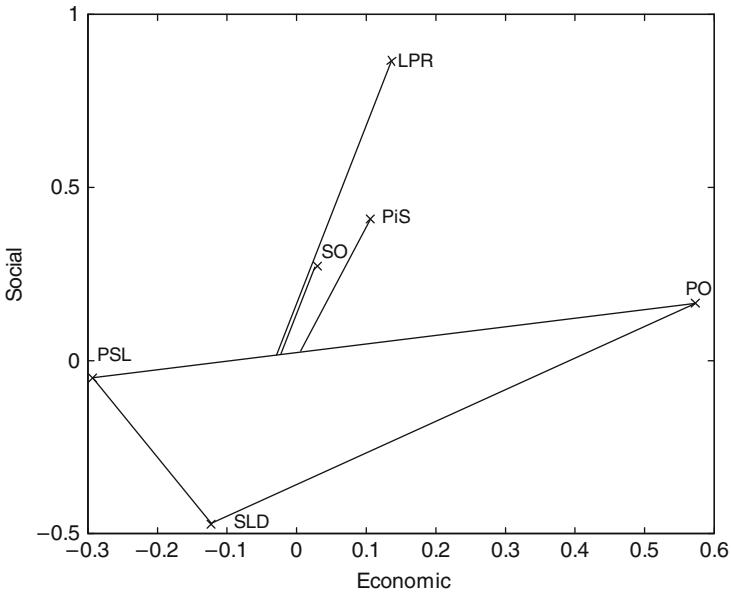[13]We may refer to the logic of these choices as "hunting the heart".

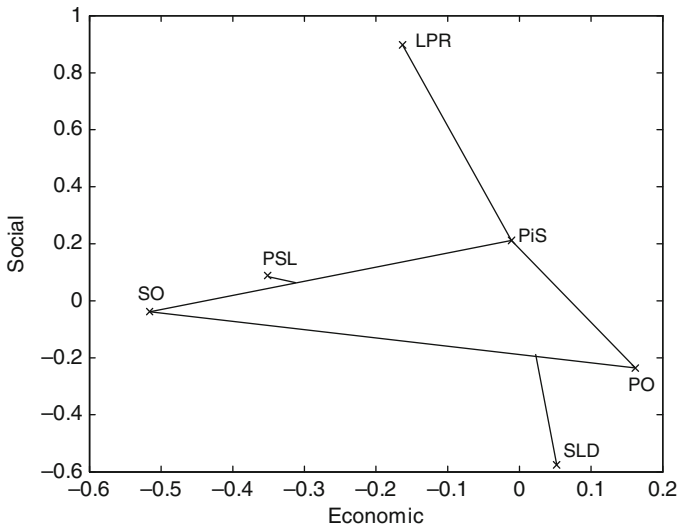**Fig. 8** Estimate of the heart in 2001 in Poland



**Fig. 9** Estimate of the heart in 2005 in Poland

to coalesce in a highly proportional electoral system. Indeed, some of these activist groups may have every incentive to fragment. The logic of such maneuvering would seem to involve both analysis of the stochastic model, in order to gauge

electoral response, coupled with coalition bargaining theory to make sense of the formation of government. In the next two sections we consider elections in the anocracies of Georgia and Azerbaijan where fragmentation is much less pronounced because of the dominance of the president's party.

## 3 Georgian Presidential Election of 2008

### 3.1 Background

By the time of the dissolution of Soviet Union in the late 1980s there were sharp political tensions in the Caucasus. The sharpest and the most violent division was the Nagorno-Karabakh separatist war between Armenia and Azerbaijan, which lasted from 1988 to 1994 and cost many hundreds of thousands of causalities. In Georgia the National Independence Movement was leading the country towards independence, while separatist movements within Abkhaz and Ossetian ethnic minorities triggered violent conflicts in the regions, which later on developed into civil wars.

Nation building and territorial conflicts were only part of the complicated political agenda of the region. Liberation from the Soviet rule induced a deep institutional shock that encompassed all spheres of the political system. Countries of the region had to reform almost all aspects of social activity as the Soviet model of social arrangement collapsed. As the crisis was systemic and the new arrangements could not evolve from the old one, it required the creation of a new paradigm. One was provided by the logic of neoliberal globalization and "democratization".

Besides the challenges of nation building, and the transformation of the political and economic systems, the societies of the region experienced a culture shock. All aspects of culture, including knowledge and symbols, patterns and norms of social arrangement, values and perceptions started to change dramatically. A majoritarian democracy, with political competition through free multiparty elections, was considered to be the main institution through which all these controversies could be governed. Elections in Georgia were therefore viewed not just a matter of elite competition, but instead were required to legitimate the shift of power and to stabilize mass beliefs.

From the time of Perestroika to the present, Georgia has experienced three major changes of government, each of which was preceded by mass mobilization and unrest.

The first was the shift of power from the Communist party to the Round Table – Free Georgia block (headed by Zviad Gamsakhurdia) in 1990.

The second was the shift of power from Gamsakhurdia to Eduard Shevardnadze, through the interim government of 1992. After the first post-Soviet Georgian constitution established a presidential democratic republic, Shevardnadze was elected as a president in November 1995, with 70% of the vote. He won a second term in April 2000.

In 2003 Shevardnadze resigned under the pressure of mass protests, and in the third shift of the November 2003 "Rose Revolution" Mikheil Saakashvili, leader of the United National Movement Party, took 96% of the vote, becaming president on 25 January 2004.

Each of these transfers of power was radical in a sense that it changed not only the ruling elite, but also the dominant trend of political development.

National liberation stances were dominant after the politics of Glasnost and Perestroika allowed for the political involvement of the population. These stances dominated the Supreme Council elections of 1990, where Gamsakhurdia defeated the Communist Party. In 1991, Gamsakhurdia declared independence for Georgia, but he failed, however, to incorporate the agenda of liberal and democratic trans-formation and to gain support from the ethnic minorities as well as from the democratic opposition.

As a result, the regime was confronted with a new wave of protests. In January 1992, a coup d'état forced Gamsakhurdia to flee from Georgia, and Shevardnadze was invited back to the country from Moscow, in order to halt the collapse into total civil war. Shevardnadze was appointed acting chairman of the Georgian State Council in March 1992, and was elected as the head of state in the first post-Soviet multiparty election.

By late 1993, struggles over the issues of Abkhazian and Ossetian separatism developed into a fully-fledged civil war. In 1993, Georgian troops were defeated in their attempt to restore control over the breakaway regions, "Ethnic cleansing" caused more than 200,000 Georgians to flee from the Abkhaz and Tskhinvali territories. By 1995, however, the period of civil war was over.

The constitution of 1995, as well as the basic economic reforms of 1994–1996 (including the introduction of a national currency, privatization, and structural adjustment in line with the Washington consensus) together established the fundamental framework for social, political and economic activities. However, there remained a serious gap between formal arrangements and de facto practices.

Despite the declared pro-democratic and pro-western stance of the Shevard-nadze regime, this was a hybrid system that existed until the end of his rule in 2003. On the one hand, Shevardnadze did not restrict freedom of society but allowed the emergence of new political and economic relations. On the other hand, he would not accept major changes within the state and government structures. The greater the demand for change, the more conservative he tended to become. As a result, corruption penetrated all spheres of life and distrust deepened against the state institutions.

The almost unanimous discontent with the conservative, weak and corrupt executive power of the regime overshadowed all other possible political divisions, and unified the opposition to Shevardnadze. The agenda of further democratization became dominant, promoted by the oppositional TV Rustavi2, which supported the "reformers" among the ruling elite – Zurab Jvania and Mikheil Saakashvili. The people eventually mobilized against Shevardnadze, and the revolution of 2003

forced him to resign. Saakashvili became the unchallenged leader of the mass protest movement, taking 96% of the vote for president, and becoming president on 25 January 2004.

Welt (2010) comments that

> Georgia's Rose Revolution stemmed from Georgians' discontent with an ineffective, criminalized, and corrupt ruling regime. Georgia's ruling regime was not only unpopular before the 2003 election, but also weak.

This time the country found new leadership, composed of a young energetic generation of risk-taking activists who opted for a quick political changes. Slow, piecemeal and negotiations-based decision-making, typical for the democratic process, contradicted their perception of themselves as a vanguard of pro-western development. Rule of law, civil and political rights, together with constitutional checks and balances, were supposed to be the norm, but in fact were subject to manipulation and were sometimes clearly violated.

> For the leaders of the revolution, for the National Movement, democracy was important, as much as democracy was the identity marker of becoming part of the West. In this sense, democracy was an external attribute, a self-declared ideology that aligned Georgia with the West, rather than a certain political practice concerning the organization of the political sphere through competitive elections, and other internal attributes of democratic performance (Cheterian 2008).

The change of the constitution in 2004, a decrease in the freedom of the media, as well as cases of the redistribution of property and other violations of the law, marked a growing gap between the pro-western stance of governmental policies and the de facto concentration of power in the hands of a small elite who seemed above the law (See also Anable 2006; Broers 2005).

The incompatibility of the pro-western orientation and non-democratic practices split society into two poles. The government promoted its agenda of externally oriented policies, including integration into NATO, arguing that this required strong leadership. The opposition insisted on the agenda of democracy and rule of law, demanding greater equality.

The split of public opinion into two poles could be interpreted as a normal political struggle between those who supported a "Western integration" agenda against those who opted for "democracy and rule of law," were it not for the illiberal environment in which the split occurred. Moreover, this split induced a change in attitude towards the U.S.

> At one time, pro-American feeling was nearly universal in Georgia. This has begun to somewhat change-as manifested by protests in front of the U.S. Embassy and increasing charges levied by the opposition that the United States has chosen to support Saakashvili rather than democracy (Mitchell 2008).

Each of these two poles had the support of different media outlets, particularly TV channels. Saakashvili controlled Rustavi2, formerly for the opposition, but by this time pro-government. The opposition initially depended on Imedi, owned by Patarkatsishvili.

Television is the main source of political information and opinion formation in Georgia, as almost everywhere. Even in a very liberal and apolitical environment, television, by its very nature, is an agenda-setting institution: it sequences, frames and contextualizes information. When this medium is not free, as in Georgia, then this tool may be used in a very goal-oriented way, creating a biased picture of political reality.

The two opposed TV channels, Rustavi2 and Imedi, had two very different views of politics. By the Fall of 2007, the governing elite and the leaders of the opposition appeared on their own channels, and seemed to ignore each other. The resulting split within society became extremely polarized.

> There are two realities in Georgia today – one seen by Saakashvili supporters and the other by the opposition and more apolitical members of society" (Sumbadze 2009).

This split in society, in which two versions of possible development existed simultaneously but separately, was a novelty for Georgia, and dominated the election of 5 January 2008. A series of anti-government demonstrations had led to clashes between police and demonstrators in the streets of Tbilisi on 7 November, 2007, and a declaration of a state of emergency. The oppositional TV channel Imedi was closed and its equipment partly destroyed by the police. These events led to harsh criticism of the Saakashvili government by the Human Rights Watch for using "excessive" force against protesters. The International Crisis Group warned of growing authoritarianism.

## 3.2   The Election in 2008

The presidential election on 5 January 2008 gave Saakashvili 53.5% of the vote, as shown in Table 7. Muskhelishvili et al. (2009) commented that the election result

> created suspicion, since cases of stuffing ballots . . . were registered in many precincts. . .
> Being unable to either change the regime or improve its quality through elections the
> opposition movement gradually lost momentum. The main opposition parties refused to
> consider these results legitimate. Because. . . a large share of society welcomed this refusal

**Table 7** Georgian presidential election 2008

| Candidate | Party | Vote share |
|---|---|---|
| Saakashvili | United National Movement | 53.5 |
| Gachechiladze | Opposition coalition | 25.7 |
| Patarkatsishvili | Media tycoon | 7.1 |
| Natelashvili | Georgian Labour Party | 6.5 |
| Gamkrelidze | New Right | 4.0 |
| Maisashvili | Party of the Future | 0.7 |
| Sarishvili-Chanturia | Hope party | 0.2 |
| Repeated ballots | | 1.7 |
| Invalid ballots | | 0.6 |
| Total | | 100 |

> by participating in mass post-electoral protest demonstrations, the political crisis of 2007 was not resolved by the [Presidential and Parliamentary] elections of 2008.

In August 2008, a series of clashes between Georgian and South Ossetian forces resulted in Saakashvili ordering an attack on the town of Tskhinvali. In response, the Russian army invaded South Ossetia, followed later by the invasion of other parts of Georgia. Eventally there was a ceasefire agreement, and on 26 August the Russian president, Dmitry Medvedev, signed a decree recognizing Abkhazia and South Ossetia as independent states. On August 29, 2008, in response to Russia's recognition of Abkhazia and South Ossetia, Georgia broke off diplomatic relations with Russia.

Opposition against Saakashvili intensified in 2009, when there were mass demonstrations against him. The next presidential election is planned for 2013. In preparation, on October 15, 2010, the Parliament approved, by 112 to 5, a constitutional amendment that increased the power of the prime minister over that of the president. It was thought that this was a device to allow Saakashvili to take on the role of prime minister in 2013, just as Putin had done in Russia.[14]

We used a sample survey to construct a formal election model in an attempt to understand the nature of politics in Georgia. Table 8 gives the survey vote shares for the candidates, while Table 9 gives the factor model, based on the survey questions, given in the Appendix 3.

The first factor dimension, *West*, is strongly related with the respondents' attitude toward the US, EU and NATO. Those who have favorable opinion toward the United States, European Union and NATO have smaller values in this

**Table 8** Sample vote shares among the four candidates in Georgia

| Candidate | Vote | % |
|---|---|---|
| Saakashvili | 252 | 63.2 |
| Gachechiladze | 85 | 21.3 |
| Patarkatsishvili | 39 | 9.8 |
| Natelashvili | 23 | 5.8 |
| Total | 399 | 100 |

**Table 9** Factor loadings

| ($n = 399$) | West | Dem |
|---|---|---|
| 1.General direction | 0.12 | 0.77 |
| 2.Democracy | 0.15 | 0.85 |
| 3.Next election fair | 0.20 | 0.66 |
| 4.Opinion USA | 0.63 | 0.26 |
| 5.Opinion EU | 0.78 | |
| 6.Opinion NATO | 0.91 | 0.15 |
| % variance | 0.32 | 0.30 |
| Cumulative % variance | 0.32 | 0.62 |

[14]See Bunce and Wolchik (2010) for a general discussion of the wave of democratic change that has occurred in the last 20 years in post-Soviet countries, sometimes leading from autocracy to democracy and then back again. See also Muskhelishvili (2010).

dimension. Thus, larger value in the *West* dimension means stronger antiwestern attitude. The second dimension, *Democracy*, is defined by respondents' judgement about current democratic environment in Georgia. Larger values in the *Democracy* dimension are associated with negative judgement about the current state of democratic institutions in Georgia, and a demand for a greater democracy.

The electoral covariance matrix is:

$$\nabla_0 = \begin{bmatrix} & Democracy & West \\ Democracy & 0.83 & 0.05 \\ West & 0.05 & 0.87 \end{bmatrix}$$

The voter distribution is displayed in Fig. 10. The points (S, G, P, N) represents estimated candidate positions, corresponding to Saakashvili (S), Gachechiladze (G), Patarkatsishvili (P), Natelashvili (N). Since there was no other information that can be used to estimate party position we used the mean value of the factor scores of those voters who voted for each candidate. Figure 11 gives the actual voter positions by candidate.

The estimated party positions were:

$$\mathbf{Z}^* = \begin{bmatrix} & S & G & P & N \\ Democracy & -0.43 & 0.86 & 0.53 & 0.67 \\ West & -0.11 & 0.00 & 0.48 & 0.41 \end{bmatrix}$$

Since the three opposition candidates are supported by voters who have similar negative judgments about democracy in Georgia, Fig. 10 takes the democracy axis as the *x*-axis and attitudes to the west as the *y*-axis. The pure spatial model in Table 10 gives the following:



**Fig. 10** Estimated party positions and voter distribution in Georgia in 2008
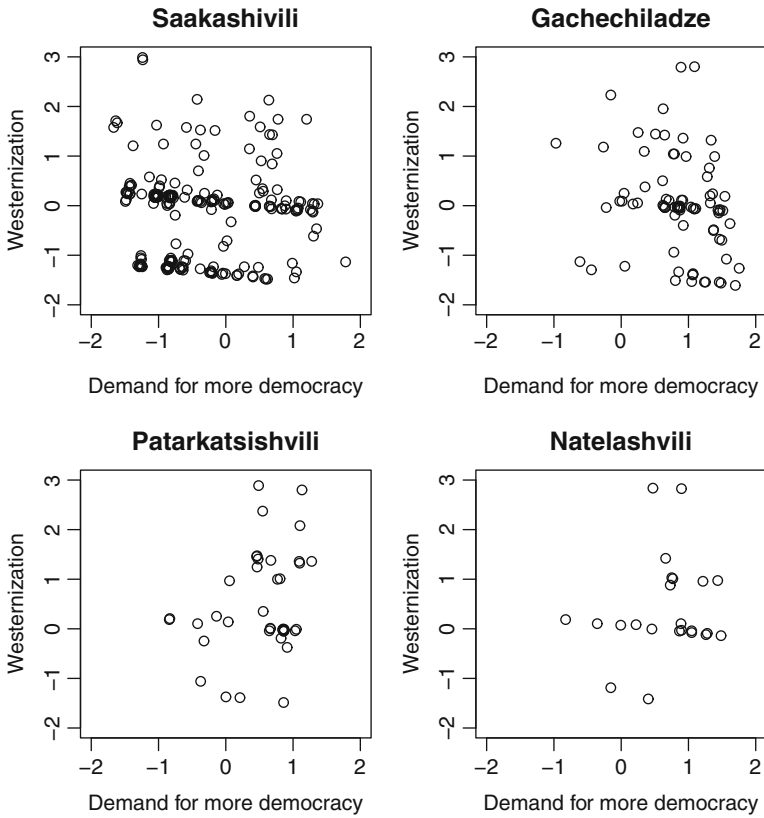
**Fig. 11** Voter positions by candidate choice in Georgia in 2008

**Table 10** Pure spatial model for Georgia (Natelashvili as baseline)

| Variable | Coeff. | Std. error | |t| Value |
|---|---|---|---|
| Spatial $\beta$ | 0.78*** | 0.07 | 11.15 |
| Valence $\lambda_S$ | 2.48*** | 0.24 | 10.41 |
| Valence $\lambda_G$ | 1.34*** | 0.24 | 5.59 |
| Valence $\lambda_P$ | 0.51 | 0.26 | 1.94 |
| $n$ | 388 | | |
| Log likelihood | $-305.97$ | | |
| ***$Prob < 0:001$ | | | |

$$\lambda_S = 2.48, \ \lambda_G = 1.34, \ \lambda_P = 0.51, \ \lambda_N \equiv 0.0$$
$$\beta = 0.78.$$

Given these coefficients, the probability that a typical voter chooses Natelashvili when all parties locate at the origin is:

$$\rho_N \quad = \quad \frac{\exp[\lambda_N]}{\sum\limits_{k=1}^{4} \exp[\lambda_j]} = \frac{e^0}{e^0 + e^{0.51} + e^{1.34} + e^{2.48}} \simeq 0.05,$$

and $(\rho_s, \rho_G, \rho_P, \rho_N) = (0.65, 0.21, 0.09, 0.05)$.

Since $2\beta(1 - 2\rho_N) = 2 \times 0.78 \times 0.9 = 1.4$, we use the formula (from the valence theorem in Schofield et al. 2011a) to obtain the characteristic matrix of Natelashvili:

$$C_N = (1.4) \begin{bmatrix} 0.83 & 0.05 \\ 0.05 & 0.87 \end{bmatrix} - I = \begin{bmatrix} 1.17 & 0.07 \\ 0.07 & 1.22 \end{bmatrix} - I$$

$$= \begin{bmatrix} 0.17 & 0.07 \\ 0.07 & 0.22 \end{bmatrix}.$$

Both eigenvalues are positive and

$$c \equiv c(\boldsymbol{\lambda}, \beta) = 1.4 \times 1.7 = 2.39.$$

Thus the joint origin is a minimum for Natelashvili.

Appendix 4 gives the results of the spatial sociodemographic model. Only gender has a statistically significant effect, with women in favor of Saakashvili. Age, education, and financial situation are not significant.

To estimate local Nash equilibrium, we stimulated the model by estimating each candidates best response to the given positions in Fig. 10, obtaining

$$\begin{bmatrix} & S & G & P & N \\ Democracy & 0.26 & 0.44 & 0.42 & 0.40 \\ West & 0.08 & 0.01 & 0.65 & 1.06 \end{bmatrix}.$$

Reiterating this procedure, staring with Saakashvili, and taking the best response in turn of each candidate until no party can increase vote share further, we obtain an estimate for the local Nash equilibrium:

$$\mathbf{Z}^{el} = \begin{bmatrix} & S & G & P & N \\ Democracy & -0.01 & 0.08 & -0.52 & 0.38 \\ West & -0.03 & -0.15 & -0.23 & 1.00 \end{bmatrix}.$$

Figure 12 gives the estimated equilibrium positions.

As expected, the high valence candidate, Saakashvili, has an equilibrium position very near the origin, followed by Gachechiladze, followed by Patarkatsishevili, with Natelashvili furthest away. The difference between these two estimates is:

**Fig. 12** Estimated equilibrium positions in Georgia in 2008

$$\mathbf{Z}^* - \mathbf{Z}^{el} = \begin{bmatrix} & S & G & P & N \\ Democracy & -0.43 & 0.86 & 0.53 & 0.67 \\ West & -0.11 & 0.00 & 0.48 & 0.41 \end{bmatrix}$$

$$- \begin{bmatrix} & S & G & P & N \\ Democracy & -0.01 & 0.08 & -0.52 & 0.38 \\ West & -0.03 & -0.15 & -0.23 & 1.00 \end{bmatrix}$$

$$= \begin{bmatrix} & S & G & P & N \\ Democracy & -0.42 & 0.78 & 1.05 & 0.29 \\ West & -0.8 & -0.05 & 0.71 & -0.59 \end{bmatrix}.$$

We infer that activists pull Saakashvili to the lower left while the other candidates respond to their activists in demanding more democracy.

## 4 The Election in Azerbaijan in 2010

In the 2010 election in Azerbaijanm, 2,500 candidates filed application to run in the election, but only 690 were given permission by the electoral commission.

The parties that competed in the election were: Yeni Azerbaijan Party (the governing party), Civic Solidarity Party, Motherland Party, and Musavat.

Many national and foreign experts expect no major improvement in the conduct of these elections. No elections after 1992 has been fully in accordance with national and international democratic standards. So far Azerbaijan has been

convicted twice of election fraud during the 2005 parliamentary elections by the European Court of Human Rights in Strasbourg. More cases are expected to be decided soon. The pre-election atmosphere was tense with the media complaining of pressure and non-transparent financial transactions of state officials.

The opposition alleged irregularities and Musavat declared that the election was illegitimate. It also asserted that the West did not criticize the regime because of Azerbaijan's geostrategic location. President Aliyev, however, rejected the criticisms claiming the election "conformed to European standards".

President Ilham Aliyev's ruling Yeni Azerbaijan Party obtained a majority of 72 out of 125 seats. Nominally independent candidates, who were aligned with the government, received 38 seats, and 10 small opposition or quasi-opposition parties got the remaining 13 seats. Civic Solidarity retained its 3 seats, and Ana Vaten kept the 2 seats that they had in the previous legislature; the Democratic Reform party, Great Creation, the Movement for National Rebirth, Umid, Civic Unity, Civic Welfare, Adalet (Justice), and the Popular Front of United Azerbaijan, most of which were represented in the previous parliament, won one seat a piece. For the first time, not a single candidate from the main right-wing opposition Azerbaijan Popular Front (AXCP) or Musavat was elected.

The Central Election Commission said turnout was 50.1%, out of a total 4.9 million people eligible to vote. Opposition leaders suggested the low turnout was due to candidate disqualifications by the CEC, and consequent discouragements to vote after their choice of candidate was excluded.

Table 11 gives the election results and the Appendix 5 gives the survey questions.

Our analysis relies on the pre-election surveys conducted by the International Center for Social Research (ICSR), Baku, Azerbaijan. The survey data include questionnaires about respondents' evaluation on the democratic situation, political

**Table 11** Summary of the 7 November 2010 Azerbaijan election results

| Party | Votes | Seats |
|---|---|---|
| Yeni Azerbaijan Party (YAP) | 1,104,528 (45.8%) | 72 |
| Civic Solidarity Party (VHP) | 37,994 (1.6%) | 3 |
| Motherland Party (AVP) | 32,935 (1.4%) | 2 |
| Equality Party (MP) | 42,551 (1.8%) | – |
| Azerbaijani Popular Front Party (AXCP) | 31,068 (1.3%) | – |
| Independents | 1,160,053 (48.2%) | 48 |
| Of which supported government | | (38) |
| Opposition[a] | | (10) |
| Total turnout (50.1%) | 2,409,129 | 125 |

[a]Opposition Parties and seats
1-Democratic Reforms party
1-Great Creation
1-The Movement for National Rebirth
1-Umid
1-Civic Welfare
1-Adalet (Justice)
1-The Popular Front of United Azerbaijan

**Table 12** Factor loadings for Azerbaijan

|  | Demand for democracy |
| --- | --- |
| Q2 Democratic satisfaction | 0.844 |
| Q3A Democratic improvement | 0.771 |
| Q3B Free opinion | 0.761 |
| Q6.1 Trust Parliament | 0.717 |
| Q6.2 Trust Government | 0.656 |
| Q6.3 Trust President | 0.883 |
| Q6.5 Trust elections | 0.742 |
| Q10.1 Political inactiveness | 0.709 |
| Q29 Free election | 0.774 |
| % var | 0.584 |
| $n$ | 149 |

institutions, and economic situation in Azerbaijan, as well as voting intention. The number of respondents in the original dataset is 1,002. The final number of observation used in this analysis was 149 for three reasons. First, a large number of respondents (636) are abstainers (those who answered that they would not vote). Thus there is no available information on their party preference. Second, among the remainder are 138 who were independent voters (those who answered that they would vote for independent candidates) and 53 who reported that they intended to vote for the parties other than YAP, VHP, AVP, AXCP and MP. Among the remaining 173 cases, only 160 had completed the factor analysis questions. The number of each party's voters are (YAP, VHP, AVP, AXCP-MP) = (113, 7, 4, 36).[15] For the parties VHP and AVP, the estimation of party positions was very sensitive to inclusion or exclusion of one respondent. We therefore used only a small subset of voters (149) who completed the factor analysis questions and intended to vote for YAP or AXCP-MP.

Table 12 gives the one-dimensional factor model. Larger values of the resultant factor score were associated with negative evaluation of the current democratic state in Azerbaijan. Specifically, the respondents with larger values tended to be dissatisfied with the current Azerbaijani democracy, did not think that free opinion is allowed, had a low degree of trust in key national political institutions, and expected that the 2010 parliamentary election would be undemocratic. This dimension is called "Demand for democracy". Figure 13 displays the distribution of respondents along the dimension. The electoral variance is 0.93. Figure 13 also shows the estimated party positions (where party positions were estimated using the mean of the party voters' positions. The party positions were estimated to be

$$(\text{YAP, AXCP} - \text{MP}) = (-0.47, 1.48).$$

We considered voters who evaluated themselves as a supporter of a party as activists. The activists means for the two parties are located at $(-0.63, 1.57)$. The

---

[15]Because of the survey design, AXCP and MP were not differentiated and are regarded as one party block. See question wording in Appendix 5 for vote choice.
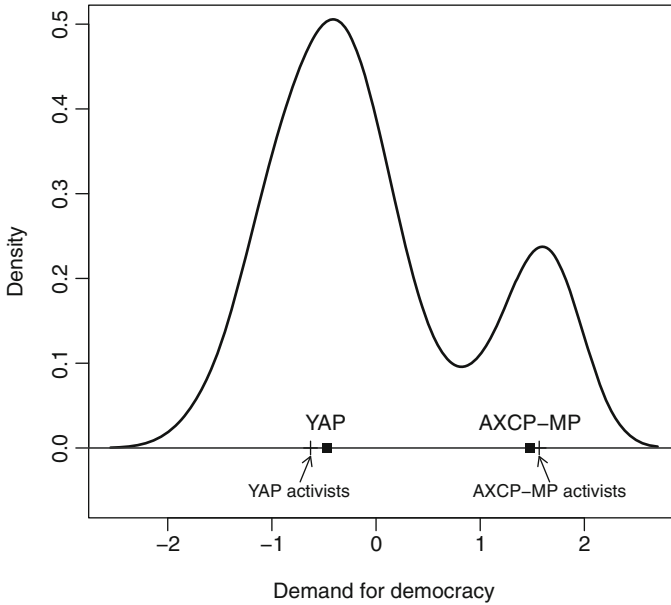
**Fig. 13** Voter distribution and activist positions in Azerbaijan in 2010

number of activists for YAP and AXCP-MP is 48 and 19, respectively. Mean activist positions are also shown in Fig. 13.

Table 13(i) presents the pure spatial binomial logit model while Table 13(ii) gives the spatial sociodemographic model.

In the first model, $\beta = 1.34$ and $(\lambda_{YAP}, \lambda_{AXCP-MP}) = (1.30, 0)$. None of the sociodemographic variables are statistically significant.[16]

Then, $(\rho_{yap}, \rho_{axcp-mp}) = (0.79, 0.21)$[17] and,

$$\begin{aligned} C_{axcp-mp} &= 2\beta(1 - 2\rho_{axcp-mp}) \cdot \text{variance} - 1 \\ &= 2 \cdot (1.34) \cdot (1 - 2 \cdot 0.21) \cdot 0.93 - 1 \\ &= 0.45. \end{aligned}$$

Since the single eigenvalue is positive, we expect divergence away from the origin by all parties for the pure spatial model. As before, we infer that the activists pull the two parties further away from the origin. This model is only one dimensional, so the result is not quite compatible with the analysis of Georgia. However, if the model were two-dimensional, and symmetric in the sense that voter variances

---

[16]The variable 'city' is a binary variable indicating whether the respondent resides in city area or not. The category 1,2 and 3 in the question 'type of location' are coded as city, and 4 and 5 are coded as non-city residents.

[17]Among the two parties, the sample voteshare is (0.76, 0.24).

**Table 13** Pure spatial and sociodemographic models for Azerbaijan (baseline AXCP-MP)

|  | (i) Coeff. (|t-value|) | (ii) Coeff. (|t-value|) |
|---|---|---|
| Distance | 1.34*** | 1.65*** |
|  | (4.62) | (3.38) |
| $\lambda_{YAP}$ | 1.30* | –4.57 |
|  | (2.14) | (0.99) |
| City |  | 1.40 |
|  |  | (0.94) |
| Gender (female) |  | –0.65 |
|  |  | (0.4) |
| Age |  | –0.14 |
|  |  | (0.15) |
| Education |  | 0.65 |
|  |  | (1.01) |
| Financial situation |  | 0.90 |
|  |  | (1.08) |
| $n$ | 149 | 149 |
| Log likelihood | –11.48 | –10.02 |
| McFadden $R^2$ | 0.86 | 0.88 |

$Prob < 0.05$, ***$prob < 0.001$

were 0.93 on each axis, then the convergence coefficient would be $c = 2.89$, very similar to the result for Georgia.

# 5 Concluding Remarks

The discussion of elections in Schofield et al. (2011a) of this volume and in this chapter suggests the electoral models are very different in a majoritarian political system such the United States and one based on a proportional electoral system such as Poland. The illustration of the Georgian election in 2008 and the election in Azerbaijan in 2010 suggests that presidential systems in these two post-Communist polities lie midway between the plurality polities and the proportional polities.

As we have seen in Schofield et al. (2011a), the convergence coefficients for the United States elections in 2000 and 2004 were only 0.37 and 0.45, respectively. According to our model, this implies that the electoral effect dominates, so that the candidates should converge to the electoral origin.

In contrast, the empirical analyses presented here show that the convergence coefficient for the 1997, 2001 and 2005 elections in Poland were 6.82, 5.92 and 6.19 respectively. Related work has shown that the convergence coefficients were 5.94 for the 2002 election in Turkey (Schofield et al. 2011d) and 3.98 for the 1996 election in Israel (Schofield et al. 2011b). In these polities with electoral systems based on proportional representation (PR), the convergence coefficients are very high because the spatial coefficient ($\beta$) and the total variance in the electoral covariance matrix are both large. As a result, under PR, the pure electoral motive

is sufficient to pull parties away from the center. We suggest that in the United States, the activist effect dominates over the electoral effect, and activist groups therefore exert a considerable influence on candidate positions. In proportional representative systems, this activist influence can be much weaker.

A standard way of estimating political fragmentation is in terms of the *effective number of party vote strength (env)* or *effective number of party seat strength (ens)*.[18] For example, in Poland in 1997 the *env* increased from about 5.5 in 1997 to 7.7 in 2005, while the *ens* increased from 3.1 to 5.0. In Israel in 1996 the *env* and *ens* were both about 7.0, and in Turkey in 2002 the *env* was about 7.5. The *env* and *ens* are convenient measures, intended to capture the nature of the distribution of electoral preferences and how these are turned into political configurations. We propose that the convergence coefficient is a theoretically consistent way of classifying the degree of political fragmentation, based as it is on the underlying political preferences and political response. These estimates for fragmented polities suggest that high convergence coefficients are associated with high estimates of the *env* and *ens*. Consider the following examples of polities with different electoral systems.

In the winner take-all presidential elections of 2000 and 2004 in the United States, the *env* was about 2.0 and the *ens* can be taken to be 1.0, corresponding to the low convergence coefficients of 0.37 and 0.45 (See Table 14). The convergence coefficients are very low because the two major parties in the US have very similar electoral support.[19]

Canada has a Parliamentary polity with a plurality electoral system, giving two large parties, the Conservatives and Liberals. However, small parties, the Bloc Québécois and the New Democratic Party, can survive because of regionalism, so its electoral system is not as majoritarian as the United States. In the elections of 2004 and 2008, the *env* was about 4.0 while the *ens* increased from about 3.1 to 3.5. Schofield et al. (2011d) found that the convergence coefficient for Canada in the 2004 election was 1.94. This estimate is greater than that of the U.S. but less than that of fragmented polities such as Poland, Israel or Turkey.

Similarly, the United Kingdom has two large parties, Labour and Conservative, and three small parties, Liberal Democrats, Scottish Nationalists and Plaid Cymru, as well as small factional parties from Northern Ireland. The results of Schofield et al. (2011c) give convergence coefficients of 0.84 for the 2005 election and 0.98 for the 2010 election in Britain. The difference between Canada and the Britain was the lower $\beta$ in the election in the Britain. The *env* for this election in Britain was about 2.7, while the *ens* was about 2.5, indicating that the electoral system is more majoritarian than that of Canada.

---

[18]Fragmentation can be identified with the *effective number* (Laakso and Taagepera 1979). That is, let $H_v$ (the Herfindahl index) be the sum of the squares of the relative vote shares and $env = H_v^{-1}$ be the *effective number of party vote strength*. In the same way we can define *ens* as the effective number of party seat strength using shares of seats.

[19]We could of course measure *ens* in terms of party strength in Congress, giving a value close to 2.0.

**Table 14** Convergence coefficients and fragmentation

| Variable | Country | | |
|---|---|---|---|
| | US | Britain | Canada |
| Conv. Coeff. | [0.40,1.1] (2000–2008) | [0.84,0.98] (2005–2010) | 1.94 (2004) |
| Political system | Pres.[a] PL.[b] | Parl.[a] PL.[b] | Parl.[a] PL.[b] |
| env | 2.0 | 3.2 (1997) | 4.0 (2004) |
| env | | 2.7 (2005) | 4.1 (2008) |
| ens | 1.0 | 2.2 (1997) | 3.1 (2004) |
| ens | | 2.5 (2005) | 3.5 (2008) |
| | Russia | Georgia | Azerbaijan |
| Conv. Coeff. | 1.7 (2007) | 2.4 (2008) | 2.89[c] (2010) |
| Political system | Anoc Pres.[d] PL.[b] | Anoc Pres.[d] PL.[b] | Anoc Pres.[d] PL.[b] |
| env | 2.3 | 2.9 (2008) | 2.27 |
| ens | 2.0 | 1.0 (2008) | 1.3 |
| | Israel | Turkey | Poland |
| Conv. Coeff. | 3.98 (1996) | 5.94 (2002) | 6.82 (1997) |
| Political system | Frag.[e] PR[b] | Frag.[e];PR[b], cut off | Frag.[e] PR[b] |
| env | 6.5 (1996) | 7.7 (1999) | 5.5 (1997) |
| env | 10.0 (2009) | 4.0 (2007) | 7.7 (2005) |
| ens | 6.5 (1996) | 5.0 (1999) | 3.1 (1997) |
| ens | 10.0 (2009) | 2.3 (2007) | 5.0 (2005) |

[a]*Parl* parliamentary, *Pres.* presidential

[b]*PL* plurality; *PR* proportional representation

[c]Convergence coefficient modified for two dim

[d]*Anoc.Pres* Anocratic presidential

[e]*Frag.* fragmented

These observations suggest a variation of the Duverger (1954) and Riker (1953) hypotheses regarding the difference between plurality and proportional electoral rule. We hypothesize that in an election based on proportional electoral methods, if the convergence coefficient derived from the spatial model is high, then there will be very little motivation for interest groups to coalesce. Consequently, the fragmentation of interest groups will lead to a degree of fragmentation in the polity. Without a dominant centrally located party, there may be coalitional instability resulting from a fragmented polity and a complex configuration of parties.

Indeed, we hypothesize that the difference between these various polities can be summed up as follows.

Under democratic proportional electoral methods, the convergence coefficient will tend to be large (of order >4.0). Bargaining to create winning coalitions occurs *after* the election, and there need be no strong tendency forcing activist groups to coalesce, in order to concentrate their influence. Indeed, there can exist incentives for activist groups to fragment If activist groups respond to this impulse, then activist fragmentation will result in party fragmentation. Parties can be scattered throughout the policy space. Activist groups, linked to small parties, may aspire to affect policy outcomes, by gaining access to the governing coalition. This is indicated by the observation that the bargaining domain in the legislature (the heart) will depend on the location of small parties. Party strengths will fluctuate

in response to exogenous shocks, and the structure of the heart will be affected by these changes. We conjecture that activist groups will attempt to maneuver the party, partly with a view to gaining votes, but more importantly, to be positioned in the heart.

Under the strong version of plurality rule, as in the United States, the convergence coefficient will be low (in the range 0.4 to 1.0). If interest groups do not form a coalition *before* the election, then they will have little impact on political outcomes. Consequently, small, third parties cannot obtain representation. Unlike the situation in a polity based on proportional rule, an activist group linked to a small party in a plurality polity has little expectation of influencing government policy. Thus activist groups face "increasing returns to size". In the United States, presidential candidates must balance the centripetal electoral effect against the centrifugal activist effect, and plurality rule induces what is essentially a two party system, through this effect on activist groups. Although the two party configuration may be in equilibrium at any time, the tension within the activist coalitions can induce a slow transformation of party positions, and thus political realignment.

In Parliamentary systems based on plurality rule, such as Britain and Canada, the convergence coefficient will tend to take low to intermediate values (between 0.8 and 2.0). Large and small parties can co-exist, since small parties can depend on regional support. The influence of activist groups will depend on the degree of regional orientation of these parties.

There is a very large literature on category of "partial democracies" or "anocracies"[20] These exhibit mixed characteristics of both democratic and autocratic regimes. The Russian polity in 2007 had a single dominant party, United Russia, with 64% of the vote and 70% of the seats, and two smaller parties with representation in the Duma. There were also a number of parties with very small vote share and no seats. The degree of majoritarianism can be inferred from the *env* of 2.3 and *ens* of 2.0. The convergence coefficient for that election was estimated to be 1.7 (Schofield and Zakharov 2010).

The empirical analysis of the 2008 presidential election in Georgia that we have presented here has found a convergence coefficient of about 2.4. Georgia is similar to Russia in the sense that the party supporting the president is dominant, with 53.5% of the vote, while the opposition parties are fragmented, giving an *env* of 2.94. Because the presidential election is winner-take-all, we take the *ens* to be 1.0. Azerbaijan is an even more extreme case. The electoral system is very majoritarian, and the dominant party controls almost all resources, taking about 46% of the vote and 58% of the seats, or 88% when its support coalition is included. It is difficult to give meaningful estimates of the *env and ens* for Azerbaijan, because of the support given to the

---

[20]See Carothers (2002) for the difficulty of transition from an anocracy to a full democracy.

dominant party, but Table 14 presents values of 2.27 for the *env* and 1.3 for the *ens*. The analogue of the convergence coefficient we have taken to be about 2.8.

In these "anocratic" Presidential systems, such as Georgia and Azerbaijan, that we have considered here, as well as in Russia, small opposition parties can exist but their supporting activist groups will find it difficult to coalesce because they cannot obtain support through the media. The opposition parties thus find it almost impossible to present an united front against the regime. In contrast, since the president has control over much of the media and can offer political bribes to his supporters, the pro-regime activist groups will coalesce in support, and his valence will remain high.

We have seen in this essay how even when democratic elections are in place, political leaders can gain overwhelming power by the control of the media, and through the resources provided by pro-regime activists. Oppositional groups as a result have little opportunity to gain sufficient valence, or electoral esteem to in order to offer attractive alternatives to the political leader.

We suggest that the convergence coefficient for such polities will tend to lie the intermediate range. Table 14 suggests that the convergence coefficient in various polities does indeed provide a method of classifying the nature of political competition.

# Appendix 1: Question Wording for Poland

These question wordings are based on the 2001 PNES. We have also indicated any noteworthy differences in question wording for the other years.

**Vote Choice**

"For which party or coalition candidate did you vote in the Sejm elections?"

The issue positions of voters

"A variety of solutions and policies aimed at solving the above mentioned issues are conceivable. On subsequent CARDS we present opposite solutions to each issue. Please read them carefully and tell me, where would you place your own opinions and stances. In doing so, please use the 11-point scale, where: 0 – means

full acceptance of the statement (solution) proposed on the left side of the CARD, 10 – means full acceptance of the statement (solution) – on the right side, 5 – means that you favor solutions lying in between both opposite ones, and the remaining scale points indicate different levels of acceptance of each of those opposite statements".

**Economic Dimension**

(1) Privatization

    00) State owned enterprises should be privatized quickly; the inefficient ones should be liquidated

    10) Enterprises should remain state property and their modernization financed from the state budget

(2) Unemployment

    00) Fighting unemployment should be an absolute policy priority of the government, even if it leads to higher spending and inflation

    10) Many other – more important than unemployment – issues should be governmental priority, i.e. balanced budget, fighting inflation, etc.

(3) Income tax

    00) The higher one's income, the higher the percentage it should be taxed

    10) Everyone should be taxed the same percentage of his/her income, irrespectively of the income level

(4) Subsidies to agriculture

    00) Agriculture should receive subsidies from the budget, otherwise many farms will go bankrupt

    10) Agriculture should not receive subsidies from the budget, because no single social group should live at the expense of society

(5) State vs. individual responsibility for social welfare

    00) The state should grant its citizens the widest possible social safety net, i.e. health care, social welfare, free education, etc.

    10) Citizens should take care and responsibility of their health, self-help, children's education, etc on their own

**Social Values Dimension**

(6) Church and state

    00) The Church should be completely separated from the state and should not interfere with politics

    10) The Church should exert influence over politics and state policies

(7) Decommunization

    00) Individuals occupying high positions under communism ('nomenclatura') should now be forbidden to perform responsible state functions

    10) These individuals ('nomenclatura') should have the same rights as all others in competing for public offices and state positions

(8) Abortion

    00) Women should have abortion right regardless of situation

    10) Abortion should not be allowed regardless of situation

We reversed the coding on Privatization and Decommunization so that (00) could be regarded as a more left wing, or pro-communist response.

We used factor analysis to obtain the positions of voters on the economic and social values dimension.

### Sociodemographics

For the sociodemographic variables we used the responses to the following questions.

(1)   Income

"What was your average monthly income last year?"
The measure is recorded in Polish zloty.

(2) Age

"Your year of birth. . ."
We subtracted respondent's year of birth from the year of election to obtain respondent's age in years.

(3) Communist party membership

"Did you ever happen to be a member of PZRP, ZSL, or SD?"
1. Yes
2. No

The 2005 survey had an additional option (3) "Was too young." We collapsed this with "no" in order to maintain a dichotomous measure.

The 2005 survey asked about membership in PZRP only and not in the other two communist regime satellite parties. The 1997 survey asked about membership in each of the ex-communist parties separately. We only used the information about former PZRP membership because this was the main communist party whereas the others were satellites that cooperated with the regime.

(4) Religion

"How would you describe your attitude towards religion? Are you: (1) atheist (2) agnostic (3) believer (4) devout believer".
We collapsed the first two and last two categories to obtain a dichotomous measure of 1 = religious, 0 = not religious.

## Appendix 2: Factor Loadings for Poland

**Table A1a** Factor loadings from the Polish National Election Survey, 1997

| Question | 1. Economic | 2. Social |
|---|---|---|
| 1. Privatization | **0.45** | 0.003 |
| 2. Unemployment | **0.70** | −0.07 |
| 3. Income tax | **0.53** | −0.04 |
| 4. Subsidies | **0.65** | −0.17 |
| 5. Social welfare | **0.76** | 0.02 |
| 6. Church and state | 0.07 | **0.80** |
| 7. Decommunization | −0.01 | **0.52** |
| 8. Abortion | 0.14 | **0.80** |
| Eigenvalues | 2.00 | 1.59 |

**Table A1b** Factor loadings from the Polish National Election Survey, 2001

| Question | 1. Economic | 2. Social |
|---|---|---|
| 1. Privatization | **0.537** | 0.266 |
| 2. Unemployment | **0.656** | −0.133 |
| 3.Income tax | **0.555** | −0.225 |
| 4. Subsidies | **0.695** | −0.166 |
| 5. Social welfare | **0.737** | −0.176 |
| 6. Church and state | 0.31 | **0.538** |
| 7. Decommunization | 0.186 | **0.795** |
| Eigenvalues | 2.185 | 1.119 |

**Table A1c** Factor loadings from the Polish National Election Survey, 2005

| Question | 1. Economic | 2. Social |
|---|---|---|
| 1. Privatization | **0.59** | −0.070 |
| 2. Unemployment | **0.69** | 0.03 |
| 3. Income tax | **0.58** | −0.14 |
| 4. Subsidies | **0.61** | −0.30 |
| 5.Social welfare | **0.74** | −0.03 |
| 6. Church and state | 0.28 | **0.75** |
| 8. Abortion | 0.12 | **0.80** |
| Eigenvalues | 2.115 | 1.315 |

## Appendix 3: Question Wording for Georgia

**Data:** Post-election surveys conducted by GORBI-GALLUP International from March 19 through April 3, 2008. In the original dataset $n = 1,000$. Among the respondents, 745 answered that they cast a vote on the election day. In the case of listwise deletion of missing data, the number of observation is $n = 399$. Those 399 voters (1) cast a vote; (2) to one of the four candidates who got more than 5% of the vote; and (3) answered all the questions used in the factor analysis.

[**Vote Choice**]

Please tell me which candidate did you vote for during the presidential elections on the 5th of January 2008? 1 Levan Gachechiladze; 2 Badri Patarkatsishevili; 3 Davit Gamkrelidze; 4 Shalva Natelashvili; 5 Mikheil Saakashvili; 6 Gia Maisashvili; 7 Irina Sarishvili; 8 Against all; 9 NA (recoded) 1 Saakashvili, 2 Gachechiladze, 3 Patarkatsishevili, 4 Natelashvili, NA:NA

[**Questions Used in Factor Analysis**]

(1) In your opinion, are things in Georgia generally going in the right direction or the wrong direction?

   1 Right direction; 2 Wrong direction; 9 DK/NA

(2) In general would you say that currently democracy works in Georgia very well, rather well, rather poorly, very poorly?
   1 very well, 2 rather well, 3 DK, 4 rather poorly, 5 very poorly, 9 NA.

(3) Tell me your overall opinion of USA.
   1 very favorable; 2 somewhat favorable; 3 somewhat unfavorable; 4 very unfavorable; 9 NA

(4) Tell me your overall opinion of EU.
   1 very favorable; 2 somewhat favorable; 3 somewhat unfavorable; 4 very unfavorable; 9 NA

(5) Tell me your overall opinion of NATO.
   1 very favorable; 2 somewhat favorable; 3 somewhat unfavorable; 4 very unfavorable; 9 NA

(6) How much confidence do you have that upcoming parliamentary elections will be transparent and fair?
   1 great deal of confidence; 2 fair amount of confidence; 3 no much confidence; 4 no confidence at all; 9 NA

[**Questions Considered but not Included in the Factor Analysis**]

The question regarding Iraq was loaded heavily ($> 0.5$) in the democratic dimension, but it was not included because it did not seem to be directly related with democratic attitude. The factor loadings of other questions were mostly around 0.1.

As you know, the plebiscite was conducted during the presidential elections held on the 5th of January. Did you vote for or against that the next parliamentary elections should be held in spring 2008?
1 Yes; 2 No; 9 NA

Did you vote fore or against that Georgia should pursue integration into NATO?
1 Yes; 2 No; 9 NA

To what extent do you approve the Georgian government's decision to send its armed forces to Iraq?
1 Fully approve; 2 approve; 3 Neither approve nor disapprove; 4 disapprove; 5 totally disapprove; 9 NA

Generally democracy is the best system of government for governing the country comparing with other systems.

Georgia should leave the CIS.

Our opposition is in alliance with the National Movement.

Usage of military methods in order to regain Georgian territorial integrity are approved.

1 strongly agree; 2 somewhat agree; 3 somewhat disagree; 4 strongly disagree; 9 DK; 0 NA

Tell me your overall opinion of Russia. 1 very favorable; 2 somewhat favorable; 3 somewhat unfavorable; 4 very unfavorable; 99 NA

How much confidence do you have that upcoming parliamentary elections will be transparent and fair?

1 great deal of confidence; 2 fair amount of confidence; 3 no much confidence; 4 no confidence at all; 9 NA (recoded NA:NA)

Please tell me to what extent do you agree or disagree with the following statement: If I could I would go back to Shevardnadze's Georgia.

1 Strongly agree; 2 somewhat agree; 3 neither agree nor disagree; 4 disagree; 5 strongly disagree; 9 NA

[**Sociodemographic Variables**]

(SD1) gender

male = 1, female 2

(SD2) Age

1 18–24: 2 25–30: 3 31–39: 4 40–50: 5 51–60: 6 60+

(SD3) education

1 pre-primary: 2 primary: 3 incomplete general secondary, vocational: 4 complete specialized secondary: 5 complete general secondary: 6 incomplete higher: 7 PHD, post graduate courses

(SD4) financial situation

1 no money for food, 2 not for clothing, 3 not for expensive things, 4 expensive things, 5 whatever we want, 9 NA

(SD5) region

1 Tbilisi; 2 Kakheti; 3 Shida Kartli; 4 Kvemo Kartli; 5 Samtskhe-Javakheti; 6 Adjara; 7 Guria; 8 Samegrelo; 9 Imereti/Racha/Svaneti; 10 Mtskheta-Tianeti

# Appendix 4: Spatial Sociodemographic Model for Georgia (Natelashvili as Baseline)

|  | Variable | Coeff. | Std. error | $|t|$ Value |
|---|---|---|---|---|
|  | Spatial $\beta$ | 0.82*** | 0.07 | 11.16 |
| Saakashvili | $\lambda_S$ | 1.75 | 1.35 | 1.29 |
|  | Gender (female) | 0.99* | 0.49 | 2.01 |
|  | Age | 0.16 | 0.16 | 0.95 |
|  | Education | −0.21 | 0.17 | 1.25 |

| | Variable | Coeff. | Std. error | \|t\| Value |
|---|---|---|---|---|
| | Financial situation | 0.40 | 0.34 | 1.17 |
| Gachechiladze | $\lambda_G$ | 0.27 | 1.39 | 0.19 |
| | Gender (female) | 0.72 | 0.50 | 1.45 |
| | Age | 0.06 | 0.17 | 0.35 |
| | Education | −0.15 | 0.17 | 0.87 |
| | Financial situation | 0.66 | 0.35 | 1.89 |
| Patarkatsishevili | $\lambda_P$ | 0.94 | 1.49 | 0.63 |
| | Gender (female) | 1.04 | 0.55 | 1.88 |
| | Age | −0.09 | 0.18 | 0.49 |
| | Education | −0.25 | 0.19 | 1.30 |
| | Financial situation | 0.36 | 0.38 | 0.94 |
| | $n$ | 399 | | |
| | Log likelihood | −298.23 | | |

*$Prob < 0.05$ ***$prob < 0.001$

# Appendix 5: Question Wording for the Azerbaijan Election

## *Survey Items*

[**Vote Choice**]

[*Q*23] Are you going to vote for the candidate from political party/block or for the independent candidate?

1. Candidate from political party/block; 2. Independent candidate; 77, 88, 99. NA

[*Q*24] Here is the list of political parties and blocks, which will run for coming parliamentary elections on 7 November, 2010. Please tell me, which of them you would vote for?

1. Yes, for sure; 2. Very likely; 3. Likely; 4. Indifferent; 5. Not likely; 6. No, for sure; 77. NA; 88. Don't know/hard to say; 99. Refusal

## A. Blocks

1. AXCP-MUSAVAT; 2. KARABAKH (UMID, ADP, AYDINLAR); 3. INSAN NAMINA (VIP, ALP); 4. ISLAHAT (BQP, BAXCP, ADALAT); 5. DEMOKRA-TIYA (VHP, ADIP)

## B. Political Parties

1. KXCP; 2. YAP; 3. ALDP; 4. SOCIAL DEMOKRAT; 5. DADP; 6. ANA VATAN; 7. MILLI DEMOKRAT; 8. MMP; 9. AMIP

[**Activist**]

[*Q*14] Some people think of themselves as usually being a supporter of one political party rather than another. Do you usually think of yourself as being a supporter of one particular party or not?

1. Yes (name); 2. No; 3. It is difficult to answer; 4. Refusal

[**Survey Items Used for Factor Analysis: Demand for Democracy**]

[*Q*2] Are you satisfied with the current state of democracy in Azerbaijan?
1. Fully satisfied; 2. Partially satisfied; 3. Neither satisfied nor dissatisfied; 4. Partially dissatisfied; 5. Completely dissatisfied; 88. Don't know/hard to say; 99. Refusal

[*Q*3] Would you agree with the following two statements?

[*A*]. Azerbaijan is more democratic now than it was 10 years ago.

[*B*]. People in Azerbaijan are free to express their opinions and concerns.

1. Strongly agree; 2. Agree; 3. Disagree; 4. Strongly disagree; 88. Don't know/ hard to say; 99. Refusal

[*Q*6] What is the degree of your confidence towards the following institutions?

(1) Parliament (Milli Mejlis)
(2) Government (Cabinet of Ministers)
(3) President of the country
(4) Elections on different levels

1. High; 2. Average; 3. Low; 88. Don't know/hard to say; 99. Refusal

[*Q*10.1] As is known, many people in our country are not politically active. To what extent do you agree or disagree with the following statements about the reason for this?

(1) Lack of freedom and Democracy

1. Fully disagree; 2. To some extent disagree; 3. Neither agree, neither disagree; 4. To some extent agree; 5. Fully agree; 88. Don't know/hard to say; 99. Refusal

[*Q*29] Do you believe that forthcoming parliamentary elections in Azerbaijan will be really democratic (free, open, transparent and fair)?

1. Yes; 2. No; 88. Don't know/hard to say; 99. Refusal

[**Demographics**]

Type of location: 1. Capital city; 2. Large city; 3. Small city; 4. Village; 5. Camp for IDPs

[*Q*31] Gender: 1. male; 2. female

[*Q*32] Age group: 1. 18–24; 2. 25–34; 3. 35–44; 4. 45–54; 5. 55–64; 6. 65+

[*Q*35] Education: 1. Without any education; 2. Primary school; 3. Incomplete secondary; 4. Complete secondary; 5. Secondary technical; 6. Incomplete higher; 7. Higher

[*Q*44] Household economic situation: Pick the phrase which best describes the economic situation in your family

1. There is not enough money even for food, we have to go into debt or get help from relatives or friends
2. There is enough money for food, but we have difficulty buying clothes

3. There is enough money for food and clothes, but expensive durable goods such as TV or refrigerator are a problem for us
4. We can buy durable goods from time to time, but the purchase really expensive things, such as an automobile, home, or a trip abroad, are beyond our means
5. Nowadays we can afford many things – an automobile, home, foreign travel – in a word, we do not deny ourselves anything

88. Don't know/hard to say
99. Refuse

# References

Anable D (2006) The role of Georgia's media-and western aid-in the rose revolution. Harvard Int J Press/Polit 11:7–43

Bawn K, Rosenbluth F (2005) Short versus long coalitions: electoral accountability and the size of the public sector. Am J Polit Sci 50:251–265

Benoit K, Laver M (2006) Party policy in modern democracies. Routledge, London

Broers L (2005) After the 'revolution': civil society and the challenges of consolidating democracy in Georgia. Central Asian Surv 24:333–350

Budge I, Robertson D, Hearl DJ (eds) (1987) Ideology, strategy, and party change: spatial analyses of post-war election programmes in 19 democracies. Cambridge University Press, Cambridge

Budge I, Klingemann H-D, Volkens A, Bara J, Tannenbaum E (eds) (2001) Mapping policy preferences-estimates for parties, electors, and governments 1945–1998. Oxford University Press, Oxford

Bunce V, Wolchik S (2010) The regional tradition. In: Bunce V, Wolchik S (eds) Democracy and authoritarianism in the post Communist world. Cambridge University Press, Cambridge

Carothers T (2002) The end of the transition paradigm. J Democracy 13:5–21

Cheterian V (2008) Georgia's rose revolution: change or repetition? tension between state-building and modernization projects. Nationalities Papers 36:689–712

De Vries CE, Edwards EE (2009) Taking Europe to its extremes. Party Polit 15:5–28

Dow JK (2001) A comparative spatial analysis of majoritarian and proportional elections. Elect Stud 20:109–125

Dow JK (2011) Party-system extremism in majoritarian and proportional electoral systems. Br J Polit Sci 41:341–361

Downs A (1957) An economic theory of democracy. Harper and Row, New York

Duverger M (1954) Political parties: their organization and activity in the modern state. Wiley, New York

Epstein D, Bates R, Goldstone J, Kristensen I, O'Halloran S (2006) Democratic transitions. Am J Polit Sci 50:551–568

Ezrow L (2010) Linking citizens and parties. Oxford University Press, Oxford

Ezrow L (2011) Reply to Dow: party positions, votes and the mediating role of electoral systems. Br J Polit Sci 41:448–452

Fidrmuk J (2000a) Economics of voting in post-communist countries. Elect Stud 19:199–217

Fidrmuk J (200b) Political support for reforms: economics of voting in transition countries. Eur Econ Rev 44:1491–1513

Fjelde H (2010) Generals, dictators and kings. Conflict Manage Peace Res 27:195–218

Gandhi J, Vreeland J (2004) Political institutions and Civil War: unpacking anocracy. In: Working paper, Emory University

Grzymala-Busse A (2002) Redeeming the communist past. Cambridge University Press, Cambridge

Huber J, Bingham Powell G Jr (1994) Congruence between citizens and policy makers in two visions of liberal democracy. World Polit 1:73–111

Kitschelt H, Mansfeldova Z, Markowski R, Toka G (1999) Post-communist party systems. Cambridge University Press, Cambridge

Laakso M, Taagepera R (1979) Effective number of parties: a measure with applications to west Europe. Comp Polit Sci 12:3–27

Laver M, Schofield N (1990) Multiparty government: the politics of coalition in Europe. Oxford University Press, Oxford. Reprinted 1998. University of Michigan Press, Ann Arbor

Markowski R (2006) Polish elections of 2005: pure chaos or restructuring of the party system. West Eur Polit 29:814–832

Markowski R, Tucker J (2010a) Euroskepticism and the emergence of political parties in Poland. Party Polit 1–26

Markowski R, Tucker J (2010b) Subjective vs. objective proximity in Poland: new directions for the empirical study of political representation. In: Working paper, New York University

Mitchell L (2008) Democracy bound. Natl Interest 95:70–76

Muskhelishvili M (2010) Georgia in a new wave of transformation: the Caucasus & globalization. J Polit Econ Stud 4:35–41

Muskhelishvili M et al (2009) Georgia's ongoing struggle for a better future continued: democracy promotion through civil society development. Democratization 16:682–708

Owen A, Tucker JA (2010) Past is still present: micro-level comparisons of conventional vs transitional economic voting in three polish elections. Elect Stud 29:25–39

Persson T, Tabellini G (2000) Political economics. MIT Press, Cambridge, MA

Persson T, Tabellini G (2003) The economic effect of constitutions. MIT Press, Cambridge MA

Powers DV, Cox JH (1997) Echoes from the past: the relationship between satisfaction with economic reforms and voting behavior in Poland. Am Polit Sci Rev 91:617–33

Regan PM, Bell SR (2010) Changing lanes or stuck in the middle: why are anocracies more prone to civil wars? Polit Res Quart 63:747–759

Riker WH (1953) Democracy in the United States. Macmillan, New York

Riker WH (1962) The theory of political coalitions. Yale University Press, New Haven, CT

Riker WH (1980) Implications from the disequilibrium of majority rule for the study of institutions. Am Polit Sci Rev 74:432–46

Riker WH, Ordeshook PC (1973) An introduction to positive political theory. Prentice-Hall, Englewood Cliffs, NJ

Schofield N (1999) The heart and the uncovered set. J Econ Suppl 8:79–113

Schofield N (2007) The mean voter theorem: necessary and sufficient conditions for convergent equilibrium. Rev Econ Stud 74:965–980

Schofield N, Zakharov AV (2010) A stochastic model of the 2007 Russian Duma election. Public Choice 142:177–194

Schofield N, Ozdemir U, Tavits M (2010) Elections in Poland. In: Working paper, Washington University in Saint Louis

Schofield N, Claassen C, Gallego M, Ozdemir U (2011a) Empirical and formal models of the United States presidential elections in 2000 and 2004. Chapter 9 in this volume

Schofield N, Claassen C, Ozdemir U, Zakharov AV (2011b) Estimating the effects of activists in two-party and multi-party systems: comparing the United States and Israel. Soc Choice Welf 36:483–518

Schofield N, Gallego M, Jeon J (2011c) "Leadership" in the elections in Great Britain 2005 and 2010. Electoral Studies 31:in press

Schofield N, Gallego M, Jeon J, Ozdemir U (2011d) Modelling elections in Canada. In: Working papers, Washington University in Saint Louis

Schofield N, Gallego M, Ozdemir U, Zakharov AV (2011e) Competition for popular support: a valence model of elections in Turkey. Soc Choice Welf 36:451–482

Stokes D (1963) Spatial models and party competition. Am Polit Sci Rev 57:368–377

Stokes D (1992) Valence politics. In: Kavanagh D (ed) Electoral politics. Clarendon, Oxford

Sumbadze N (2009) Saakashvili in the public eye: what public opinion polls tell us. Central Asian Surv 28:185–197

Szczerbiak A (1998) Electoral politics in Poland: the parliamentary elections of 1997. J Communist Stud Trans Polit 14:58–83

Tavits M, Letki N (2009) When left is right: party ideology and policy in post-communist Europe. Am Polit Sci Rev 103:555–569

Tucker JA (2006) Regional economic voting: Russia, Poland, Hungary, Slovakia and the Czech Republic. Cambridge University Press, Cambridge

Vreeland J (2008) The effect of political regime on Civil War. J Conflict Res 52:401–425

Wade L, Lavelle P, Groth AJ (1995) Searching for voting patterns in post-communist Poland's Sejm elections. Communist Post-Communist Stud 28:411–425

Welt C (2010) Georgia's Rose Revolution. In: Bunce V, Wolchik S (eds) Democracy and authoritarianism in the post Communist world. Cambridge University Press, Cambridge

# Electoral Systems and Party Responsiveness

**Lawrence Ezrow**

Whether parties respond to the mean voter position or to their core supporters is a question that is at the core of understanding how political representation occurs.[1] Do political parties respond to the ideological shifts of their supporters or to those of the mean voter (or to neither)? Previous theoretical and empirical research highlights the importance of the mean or median voter's policy preference as the starting point for democratic representation (Downs 1957; Powell 2000; Huber and Powell 1994; McDonald and Budge 2005; Stimson et al. 1995; Erikson et al. 2002; Adams et al. 2004, 2006). An alternative and equally compelling vision of policy representation emphasizes the policy preference of the *mean party supporter* in explaining party-citizen linkages (Dalton 1985; Wessels 1999; Weissberg 1978). The first model of political representation is referred to as the *general electorate model*, and the second model as the *partisan constituency model*.

With respect to the general electorate model and the partisan constituency model, the following questions are addressed: First, are shifts in the preference of the mean voter position in the general electorate accompanied by roughly corresponding policy shifts of the parties in a given party system? Alternatively, are shifts in the preferences of the party's supporters accompanied by roughly similar shifts in the party's position? Finally, are these citizen-party linkages mediated by the electoral system in which parties compete?

The empirical analyses examine political parties in 15 Western European democracies from 1973 to 2003.[2] The results reported below support the following three conclusions. First, in systems that feature proportional electoral systems, parties are systematically responsive to the mean voter position. Second, parties in disproportional systems do *not* respond systematically to the preferences of the

---

[1]This chapter is based on Ezrow et al. (2011) and Ezrow (2010: Chapter 6). I thank Catherine de Vries, Marco Steenbergen, and Erica Edwards for their ideas, and allowing me to use them here.

[2]The following fifteen countries are included: Austria, Belgium, Denmark, Finland, France, Germany, Great Britain, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, and Sweden.

L. Ezrow
Department of Government, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK
e-mail: ezrow@essex.ac.uk

mean voter position. Finally, neither system induces parties to respond to their supporters.[3] (On first glance, these conclusions may seem alarming for disproportional systems in which parties are neither responsive to their supporters or to the mean voter position. However, as I discuss later, this finding can easily be explained if party strategies vary within these systems, perhaps due to valence considerations).

These conclusions, however, come with several caveats. First, due to measurement issues (discussed below), the empirical analyses are limited to 15 party systems in Western democracies. While the scope of the study thereby covers a significant portion of the population we are attempting to understand (i.e. stable and industrialized democracies), it nevertheless warrants caution about extrapolating these conclusions to political systems outside of the study.

Second, the fluidity of more elegant two-dimensional spatial mappings in a smaller number of countries (see, e.g., Schofield 1997; Dow 2001) has been sacrificed for unidimensional measurements of ideology in order to widen the geographical scope of this study. Nevertheless, the analysis of Left–Right policy responsiveness is still illuminating. With respect to this issue, Ian Budge and Michael McDonald comment that "while the issues involved in Left–Right divisions do not cover the whole spectrum of democratic politics, few would deny they are at the centre of them" (Budge and McDonald 2006, p. 453). There is complementary research that supports these authors' remarks, and suggests that the Left-Right dimension captures an important and meaningful component of political competition across the national settings and time period that are under review here (see, e.g., Powell 2000; Huber and Powell 1994; Huber 1989; McDonald and Budge 2005).

The third caveat is that while it is assumed that parties *respond* to public preferences, an equally plausible alternative is that parties shape or 'cue' the preferences of the electorate (see Steenbergen et al. (2007)). The measurement instruments do not allow parsing out the direction of causality in the empirical analyses. However a similar approach to Steenbergen et al. (2007) is employed to address endogeneity (discussed later), and the results from these analyses suggest that the assumption that parties respond to shifts in public preferences is a reasonable one. Moreover, this endogeneity issue cannot detract from the finding that political parties competing in proportional systems are more responsive to the mean voter position than are parties in disproportional systems. Under either causal scenario, to the extent that patterns are uncovered between the policy preferences of parties and citizens, these findings contribute to our understanding of party competition in Western European democracies.

These above limitations notwithstanding, the result that parties across all Western European democracies, regardless of electoral system, tend to respond to the mean voter position (and not their supporters) has important implications for

---

[3]Following Ezrow et al. (2011), I also report that mainstream parties (i.e. parties belonging to the Social Democratic, Conservative, Christian Democratic, or Liberal party families) tend to respond to shifts in the mean voter position as opposed to the policy shifts of their supporters, and that the opposite holds for niche parties.

political representation and for our understanding of electoral system effects. The study relates to political representation and specifically the model of dynamic representation developed by Stimson et al. 1995 (see also Erikson et al. 2002), which identifies party responsiveness to shifts in public opinion as a key component to political representation.[4] Similarly, I assess the empirical validity of a partisan constituency model that is based on several influential studies which emphasize the importance of party-constituency agreement (see Dalton 1985; Weissberg 1978; Wessels 1999).[5]

Finally, the study contributes to our understanding of electoral system effects. There are several prominent representation studies that articulate persuasive arguments, which are summarized in the next section, that suggest that electoral systems affect which groups of citizens to which parties respond. I present empirical evidence that examines these claims, and conclude that parties in PR systems tend to respond to the mean voter position. Furthermore, parties are not responsive to their supporters across systems.

## 1   Electoral Systems and Party Responsiveness

Do electoral systems help determine whether parties respond to the mean voter position or to their core supporters? Theory leads to conflicting answers. One reasonable expectation is that:

> H1a: Parties in disproportional systems are responsive to the mean voter position, and parties in proportional systems are responsive to their supporters.

The first argument that underlies this expectation is that less proportional voting systems – like plurality systems – plausibly motivate political parties to emphasize vote-seeking objectives, so that the parties competing in these systems can be expected to be more responsive to changes in the mean voter position. The reason that disproportional systems plausibly promote vote-seeking behavior by parties is because of the well-known fact that such voting systems tend to "punish" smaller parties and reward larger parties when national vote returns are converted into parliamentary seats (Cox 1997, see also Taagepera and Shugart 1989). Disproportional systems, for this reason, motivate parties to place a premium on gaining

---

[4]I note that mean voter representation and dynamic representation do differ conceptually. While the former concerns the party-citizen linkage and "giving voice" to electors, the latter refers to responsiveness of governing institutions in terms of policy outputs.

[5]Indeed this study reinforces the findings of Ezrow et al. (2011) that highlight the applicability of the partisan-constituency model to the policy shifts of niche parties. Furthermore, the results corroborate and expand upon the conclusions of recent studies by Adams et al. (2006) and Meguid (2005, 2008) who present theoretical and empirical arguments suggesting that spatial theories of electoral competition (Downs 1957; Enelow and Hinich 1984) should account for the party families competing in elections.

substantial vote shares, so that they can be among the large parties that benefit from this effect.

Jay Dow (2001, 2011) offers a related argument that in disproportional systems the major parties may reasonably aspire to win a single-party parliamentary majority, which gives these major parties added incentives to maximize votes. For instance, the plurality-based postwar elections held in Britain and New Zealand (the latter country featured plurality until its switch to PR in 1996) returned single-party parliamentary majorities in over 80% of the cases.[6] Dow (2001) argues that this "winner-take-all" feature of disproportional, plurality-based elections motivates political parties to be highly responsive to voters' policy preferences.

By contrast the lower effective seat thresholds associated with highly proportional voting systems plausibly motivate the parties in these systems to emphasize policy objectives and to thereby be more ideologically "rigid" in the face of mean voter shifts, because these parties are assured of at least some parliamentary representation when they are confident that their vote shares will exceed the relatively low national thresholds that are necessary to obtain legislative seats in highly proportional systems. Thus, parties in proportional systems are freer to respond to their core supporters' ideological preferences because they do not have to compete for marginal voters at the 'center' in order to gain representation in the legislature.

An additional set of considerations is that parties in PR systems – which should be smaller, on average, because there are more of them – should have more information about their supporters, and, furthermore, these parties are more flexible in terms of responding to shifts in their supporters' positions. In contrast, the large parties associated with electoral competition in disproportional systems should have more difficulty collecting information about their supporters, and their correspondingly larger organizational structures should make it more difficult to respond to their supporters' ideological shifts.[7]

While the above considerations suggest that the general electorate model applies to disproportional systems and that the partisan constituency model applies to proportional systems, there are another set of theoretical considerations which suggest the opposite holds; specifically, that:

> H1b: Parties in disproportional systems are responsive to their supporters, and parties in proportional systems are responsive to the mean voter position.

Arguments on this side of the ledger refer to research on valence, party activists, and coalitions. The strategic implications of "valence" dimensions of party evaluation (i.e. dimensions related to voters' impressions of party elites' competence, honesty, or charisma) suggest that "valence-disadvantaged" parties in disproportional systems would not be oriented towards the mean voter position; instead, they

---

[6]15 out of 17 postwar British elections have returned single-party parliamentary majorities, while in New Zealand each of the postwar elections held under plurality through the mid-1990 s returned parliamentary majorities.

[7]I thank Gary Marks for raising several points in this paragraph.

have electoral incentives to differentiate themselves on policy grounds: if these parties present centrist policies that are similar to those advocated by valence-advantaged parties, then voters will choose based on the valence dimension – that is, they will choose parties that have superior valence images (Schofield and Sened 2005, 2006; see also MacDonald and Rabinowitz 1998). To the extent that Schofield's argument captures real-world parties' electoral strategies, we should not expect all vote-seeking parties to appeal to the mean voter position. Even if plurality systems motivate parties to attach greater weight to vote-seeking, this will not in turn imply plurality elections motivate policy convergence.[8]

While the arguments in the paragraph above suggest that parties in plurality (or disproportional) systems are not always motivated to respond to the mean voter position, these arguments however do not suggest that parties would be responsive to their core supporters. Miller and Schofield (2003), building on Aldrich (1983a, b, 1995), develop a second, related, motivation for vote-seeking parties which revolves around strategic incentives related to *party activists*. The Miller–Schofield argument is that parties can enhance their vote shares by appealing to party activists who provide scarce campaign resources (i.e. time and money). Specifically, the authors argue that parties can use the added campaign resources they acquire via their policy appeals to activists to enhance their images along valence dimensions such as competence and integrity – and that this in turn will increase the parties' electoral support among rank-and-file voters. If campaigns are more important in elections in disproportional systems, given the above considerations on valence and activists, we might expect the partisan constituency model to apply to disproportional systems.

By contrast, party responsiveness to the mean voter position in proportional systems may not be so surprising if parties are concerned with maximizing their likelihood of being included in the governing coalition. If this is the case then appealing to the center may be a viable strategy. Schofield et al. (1998) examine Dutch and German elections and determine that parties try to put themselves in good positions for the post-coalition negotiations. This entails presenting policies that are acceptable to potential coalition partners, which may provide incentives for policy moderation. If proportional systems motivate parties to present policies that are acceptable to coalition partners then these parties may well present centrist positions. This finding is also in line with that of the other prominent coalition scholars who present theoretical and empirical results that support the claim that in proportional systems, gaining membership in the governing coalition is closely linked with centrist positioning (Axelrod 1970; Laver and Shepsle 1996; Powell 2000; Huber and Powell 1994). When facing the decision to respond to core supporters or the mean voter, parties may choose the latter option with the

---

[8]Adams and Merrill (1999, 2000) present an alternative argument that voters' partisan loyalties can motivate vote-seeking parties to diverge from the center, in the direction of the policies favored by the members of their partisan constituencies (see also Adams et al. 2005).

expectation that this will enhance the possibility of joining the governing coalition after an election.

There are of course, additional possibilities: that is, that neither model of party responsiveness applies uniformly to either type of electoral system. The arguments above on party positioning incentives might counterbalance leading to the unremarkable yet plausible conclusion that sometimes parties respond to their core supporters; sometimes to the mean voter position; sometimes to both (if they shift in the same direction); and sometimes to neither – and that electoral systems do not apply incentives uniformly to the parties that compete within them.

## 2   Data and Measurement

### 2.1   Measuring Parties' Policy Positions and Public Policy Preferences

Each hypothesis posits that the changes in the voters' ideological preferences are somehow linked to parties' policy positions. Thus, to test this proposition it is necessary to develop longitudinal, cross-national measures of parties' policy programs as well as measures of voters' policy preferences.

To measure party policy positions over time, I use estimates that are reported by Budge et al. (2001) from the Comparative Manifesto Project (CMP). These data are comprised of party manifestos from the main political parties in 25 democracies in the postwar period and provide the only longitudinal and cross-national estimates of party policies. The analytical payoff of the CMP data is that it allows party positions to be mapped over the entire time period and in all of the countries under investigation.[9] Moreover, as the content of these manifestos is often the result of intense intra-party debate, the CMP estimates should be reliable and accurate statements about parties' positions at the time of elections. Indeed, these measures are generally consistent with those from other party positioning studies, such as those based upon expert placements, citizen perceptions of parties' positions, and parliamentary voting analyses. This provides additional confidence in the longitudinal and cross-national reliability of these estimates (see Hearl 2001; McDonald and Mendes 2001; Laver et al. 2003).

While the methods used by the CMP to map party policy positions based on election programs are described at length elsewhere, I briefly review them here.[10] Under the CMP framework, policy preferences are characterized by systematic examination of party stances on policies based on content analysis of election

---

[9]In *Mapping Policy Preferences II* the CMP updates their estimates of parties' policy positions through 2003 and expands the number of countries for which they place parties (Klingemann et al. 2006).

[10]For a more thorough description of the coding process, see Appendix 2 in Budge et al. (2001).

programmes (Budge et al. 2001). Individual coders isolate "quasi-sentences" in a party's manifesto and pair them with policy categories (e.g. education, defense, law and order, morality, etc.) using a pre-established, common classification scheme. The classification scheme is made up of 56 categories and the percentages of each category provide the basis for estimating the policy priorities of a party. The Left–Right ideological scores for parties' manifestos range from $-100$ (extreme left) to $+100$ (extreme right).

The measure of public opinion is based on Eurobarometer surveys dating from 1973 (the first year that the Left–Right self-placement item appears on the Eurobarometer survey)[11] until 2002 (the last year for which the "vote intention" item discussed below appears on the survey). In these surveys approximately 2,000 respondents in each country in each year are asked to place themselves on a 1–10 Left–Right ideological scale.[12]

Finally, to estimate the policy position of the mean party supporter, the "vote intention" question on the Eurobarometer surveys is used in combination with the Left–Right self-placement data described above. Specifically, the question asks respondents the following: "If there were a 'general election' tomorrow, which party would you support?" The mean party supporter is calculated as the mean Left–Right self-placement for all respondents that indicated that they would support the party in the upcoming parliamentary elections.[13] Appendix 1 presents the countries, parties, inter-election periods, party family designations, and the mean Left–Right party supporter positions that are used in the empirical analyses.[14]

## 2.2 Model Specification for the Hypotheses

In order to test the Hypotheses 1a and 1b, I specify the following multivariate regression model (referred to as the core model specification):

---

[11]For the public opinion data, the *Mannheim Eurobarometer Trend File, 1970–2002* (Schmitt and Scholz 2005) was relied upon, which has compiled the Eurobarometers for the time period under investigation.

[12]Specifically, the Eurobarometer surveys ask, "In political matters, people talk of 'the left' and 'the right.' How would you place your views on this scale?"

[13]The mean party supporter estimates are based on at least 50 responses to the Left–Right self-placement item in each country in each year. I note that country-year observations were based on a relatively large number (approximately 2,000) of respondents so that only a few parties did not reach this criterion for inclusion. In addition, only parties that were observed in at least three successive elections are included in the empirical analyses.

[14]Addressing the research question requires aggregating individual-level observations up to party- and country-levels of analysis. Thus while the statistical analyses are based on 309 party-level observations, it should be clarified that these aggregated observations are based on slightly over 800,000 individual responses for the time period and countries under consideration.

Change in party position (t)

$$= B_1 + B_2[\textit{Meanshift} - \textit{allvoters}(t)]$$
$$+ B_3[\textit{Meanshift} - \textit{partysupporters}(t)]$$
$$+ B_4[\textit{Disproportionality} \times \textit{Meanshift} - \textit{allvoters}(t)] \qquad (1)$$
$$+ B_5[\textit{Disproportionality} \times \textit{Meanshift} - \textit{partysupporters}(t)]$$
$$+ B_6[\textit{Disproportionality}]$$
$$+ B_7[\textit{Changeinpartyposition}(t-1)]$$

where,

*Change in party position (t)* = the change in a party's Left–Right policy position in the current election compared to its position in the previous election (election t−1), based on the CMP data.

*Change in party position (t−1)* = the difference in the CMP Left−Right estimates of a party's policy position between election t−1 and election t−2.

*Mean shift − all voters (t)* = the change in the mean Left–Right self-placement score of *all* respondents in a country between the year of the current election and the year of the previous election (election t−1), based on the Eurobarometer data.

*Mean shift − party supporters (t)* = the change in the mean Left–Right self-placements for all of the respondents who indicated that they would vote for the party in the upcoming national election, between the year of the current election and the year of the previous election.

*Disproportionality (t)* = 1 if a country is categorized as disproportional, based on the Gallagher (1991) Index of Disproportionality is $\sqrt{1/2 \sum (vi - si)^2}$, where $v_i$ an $s_i$ are the vote shares and subsequent seat shares for party $i$; and 0 if a country is categorized as relatively proportional.[15]

The dependent variable [*change in party position (t)*] represents the inter-election shift in parties' Left–Right policies. The variable is constructed so that positive scores indicate that the parties' policies are moving "rightward" between elections and negative scores denote "leftward" party shifts. The key independent variables [*mean shift − all voters*] and [*mean shift − party supporters*] are similarly constructed.

Recall that the Hypotheses 1a and 1b state that party-citizen linkages are mediated by the electoral system. To test this proposition, I include two interaction variables in the core model specification, [*disproportionality × mean shift − all voters*] and [*disproportionality × mean shift − party supporters*], which interact public opinion shifts and mean party supporter shifts with the dummy variable

---

[15]There is a relatively large gap of four points between the groupings proportional and disproportional. The dummy variable is used to enhance the interpretation of the results. I note that the results remain unchanged when the continuous variable is employed.

[*niche*]. The [*disproportionality*] variable is based on the widely used Gallagher's Index of Disproportionality (Lijphart 1999; Dow 2011).[16]

The interaction terms allow us to estimate differences in the degree to which public opinion or party supporters influence parties' policy positions conditional on the electoral system. Let us first consider the effects of the variables [*mean shift − all voters*] or [*mean shift − party supporters*] on the policy shifts of parties. For parties in proportional electoral systems, the dummy variable [*disproportionality*] equals zero, and the coefficients $B_2$ and $B_3$ on the variables [*mean shift − all voters*] and [*mean shift − party supporters*] estimate the effects of public opinion shifts and party supporter shifts on mainstream parties' policy shifts. If parties are generally responsive to shifts in public opinion and to their supporters, coefficients $B_2$ and $B_3$ will be positive and statistically significant.

The effect of public opinion shifts and party supporter shifts on the policy shifts in disproportional systems are measured in the instances where [*disproportionality*] equals one. The effects of changes in public opinion on niche parties' policy programs will be captured by the sum of the coefficients $B_2$ and $B_4$ on the variables [*mean shift − all voters*] and [*disproportionality × mean shift − all voters*] in (1). Similarly, the sum of the coefficients $B_3$ and $B_5$ on variables [*mean shift − party supporters*] and [*disproportionality × mean shift − party supporters*] will estimate the influence of changes in the mean Left–Right position of party supporters on parties' policy shifts in disproportional systems.

Two additional variables in the core model specification are included. First, a lagged version of the dependent variable [*change in party position (t−1)*] is included, which measures the party's policy shift between election t−2 and election t−1. The lagged dependent variable addresses autocorrelation (discussed further below). Additionally, the [*change in party position (t−1)*] variable addresses policy alternation, a possibility raised by Budge (1994) and Adams (2001), that party elites may have electoral incentives to move their party's position in the opposite direction from their shift in the previous election. Policy alternation, according to Budge, is a rational response by party leaders to placate different wings within the party (see also Budge et al. 2010). Adams emphasizes the existence of non-policy related factors such as the party identification of voters that would explain similar zigzag patterns of party movement.[17] Under either scenario, the direction of parties' policy shifts in the previous election might influence party leaders' Left–Right

---

[16]Based on this measure the countries categorized as proportional systems are Austria, Belgium, Denmark, Finland, Germany, Ireland, Italy, Luxembourg, The Netherlands, Portugal and Sweden. Countries categorized as disproportional systems are the United Kingdom, Spain, France, and Greece.

[17]There are two additional considerations which would also explain party policy alternation. Burt (1997) proposes a random ideologies model that explains policy alternations by assuming a random selection of three successive party ideologies from a random probability distribution. Measurement error in the CMP estimates of parties' Left–Right positions is a fourth factor that would explain policy alternation. To the extent that parties' "true" positions do not vary over time, and to the extent that the CMP estimates contain measurement error, the estimates will shift in the pattern predicted by Burt's random ideologies model.

strategies in the current election. The [*disproportionality*] variable is also included on its own in the model specification to ensure that the effects of the interaction terms are measured accurately (Braumoeller 2004).[18]


## 2.3   *Evaluating the Electoral System Effects Hypotheses*

The hypotheses are evaluated using time-series cross-sectional data from 15 Western European democracies over the period 1973–2002. One possible concern is the existence of unobserved differences between countries or parties: estimating a simple regression on the pooled data containing these unobserved differences could lead to erroneous inferences (Hsiao 2003; Green et al. 2001). In Table 1 estimates are reported for the core model specification that controls for country-specific effects.[19] The results indicate that unobserved differences between countries are not driving the major findings.

The parameter estimates for the core model specification are presented in Column 1 of Table 1. The coefficient estimate on the [*Change in party position (t−1)*] variable is negative and statistically significant, which is consistent with the theoretical arguments of Budge (1994); see also (Budge et al. 2010) and Adams (2001) that parties tend to shift their positions in the opposite direction from their shifts in the previous inter-election period. With respect to the key hypotheses, the parameter estimates suggest that mainstream parties respond to shifts in the mean voter position: specifically, the parameter estimate on the [*mean shift − all voters (t)*] variable is positive and statistically significant (+11.55). Furthermore, the magnitude of this estimate suggests that the effect is *substantively* significant: the coefficient indicates that when the mean Left–Right self-placement of respondents in a country shifts by a unit along the 1–10 Eurobarometer Left–Right scale during an inter-election period then mainstream parties' Left–Right positions tend to shift 11.55 units in the same direction along the 200-point CMP Left–Right scale. Accordingly, the evidence supports that finding that shifts in parties' Left–Right policy positions systematically respond to shifts in the mean voter position.

Table 1 also reports the parameter estimates for the [*mean shift − party supporters*] variable, which are close to zero (+0.41) and statistically insignificant. Thus the evidence does not support a finding that parties in proportional systems are systematically responsive to their supporters.

With respect to the expectations raised above, if there is evidence that parties are relatively more sensitive to shifts in the mean voter position in disproportional

---

[18]One of the central implications of the study by Braumoeller (2004) is that properly estimating the effects of interaction terms involves including the constitutive (i.e. lower-order) terms in the model specification.

[19]Parameters were also estimated for a model with party-specific effects, and the substantive results remained unchanged. Also, the [*disproportionality*] variable does not vary within country, and so this term naturally drops out of the model specification.

**Table 1** Explaining parties' policy shifts

|  | Basic (1) | Type of party (2) | Past election results (3) | Ideology (4) | Full (5) |
|---|---|---|---|---|---|
| Mean shift − all voters (t) | 11.55** | 13.62*** | 10.69** | 11.54** | 12.81** |
|  | (4.85) | (5.07) | (4.84) | (4.86) | (5.06) |
| Mean shift − party supporters (t) | 0.41 | −1.48 | 0.82 | 0.46 | −1.01 |
|  | (1.67) | (1.85) | (1.68) | (1.68) | (1.86) |
| Disproportionality × mean shift − all voters (t) | −11.42 | −11.74 | −9.82 | −11.47 | −10.10 |
|  | (7.23) | (7.24) | (7.20) | (7.24) | (7.22) |
| Disproportionality × mean shift − party supporters (t) | −3.81 | −4.15 | −4.57 | −3.78 | −4.89 |
|  | (5.13) | (5.12) | (5.12) | (5.13) | (5.12) |
| Niche × mean shift − all voters |  | −4.93 |  |  | −5.40 |
|  |  | (7.87) |  |  | (7.83) |
| Niche × mean shift − party supporters |  | 8.84** |  |  | 8.58** |
|  |  | (3.84) |  |  | (3.82) |
| Niche |  | −0.98 |  |  | −1.07 |
|  |  | (1.73) |  |  | (1.84) |
| Change in party position (t−1) × vote change (t−1) |  |  | 0.024** |  | 0.02* |
|  |  |  | (0.012) |  | (0.01) |
| Vote change (t−1) |  |  | −0.23* |  | −0.23* |
|  |  |  | (0.14) |  | (0.14) |
| Ideology |  |  |  | 0.36 | 0.10 |
|  |  |  |  | (0.75) | (0.80) |
| Change in party position (t−1) | −0.43*** | −0.43*** | −0.44*** | −0.43*** | −0.43*** |
|  | (0.05) | (0.05) | (0.05) | (0.05) | (0.05) |
| Intercept | 1.91*** | 2.06*** | 1.91*** | 1.98*** | 1.95*** |
|  | (0.66) | (0.74) | (0.66) | (0.67) | (0.74) |
| N | 309 | 309 | 309 | 309 | 309 |
| $R^2$ | 0.19 | 0.20 | 0.19 | 0.19 | 0.21 |

*Notes*: $*p < 0.10$, $**p < 0.05$, $***p < 0.01$, two-tailed tests. Standard errors are in parentheses. The dependent variable is the change in a party's Left–Right policy position, based on the codings of parties' policy programmes that are reported in the CD-ROM in Budge et al. (2001) and Klingemann et al. (2006). The model is estimated with country-specific intercepts, and disproportionality (t) drops out of the model because the estimates do not vary by country

systems than in proportional systems, we would expect the coefficient on the [*disproportionality* × *mean shift* − *all voters*] variable to be positive and statistically significant. Alternatively, if the coefficient on the [*disproportionality* × *mean shift* − *party supporters*] variable is positive and statistically significant, this would indicate that parties in less proportional systems are more responsive to their supporters than parties in proportional systems.

The results are quite striking. The coefficient on the [*disproportionality* × *mean shift* − *all voters*] variable is approaching significance ($p = 0.115$), and it is substantively large (B = −11.42). This suggests that the evidence does not support the hypothesis that parties are systematically responsive to the mean voter position in disproportional systems (the conditional parameter estimates on the [*mean shift* − *all voters (t)*] variable for disproportional systems are: B2 + B4 = 0.13; s.e. = 5.54, $p = 0.98$). Furthermore, the coefficient on the [*disproportionality* × *mean shift* − *party supporters*] variable is negative and insignificant. The conditional

**Table 2** Summary of findings

| Electoral system | Model of party responsiveness | |
|---|---|---|
| | Mean voter | Partisan-constituency |
| Proportional | ✓ | X |
| Disproportional | X | X |

parameter estimates for the [mean shift − party supporters] variable do not support a finding that parties in disproportional systems are responsive to their supporters ($B3 + B5 = -3.40$; s.e. $= 4.84$, $p = 0.48$).

To summarize the findings briefly, as depicted in Table 2, parties display responsiveness to the mean voter position in proportional electoral systems. However, the evidence does not support a similar finding for parties in disproportional systems; namely, there is no evidence that parties systematically respond to the mean voter position in these systems. Finally, there is no evidence to suggest that parties under either system systematically respond to their supporters.

## 2.4 Sensitivity Analyses

The possibility of serially correlated errors within countries is also addressed. Given the structure of the data, the causal processes which generate the change in the policy position of a party at time $t$ could also be operating during the prior inter-election period $t-1$. This concern is addressed by including the lagged version of the dependent variable, [*change in party position (t−1)*] in the core specification given in (1) (see Beck and Katz 1995, 1996).[20]

Columns 2–5 in Table 1 report parameter estimates for pooled data analyses that control for additional factors that plausibly influence parties' policy positions, including the type of party; effects of past election results; party system convergence; and a full specification that accounts for all of the factors in the analysis.

### 2.4.1 Type of Party

Column 2 reports estimates for a *Type of Party* model, which accounts for the type of party, 'niche' or 'mainstream', competing in the election. Ezrow et al. (2011) have argued that mainstream parties (i.e., parties belonging to the Social Democratic, Conservative, Christian Democratic, or Liberal party families) tend to adjust their Left–Right positions in response to shifts in the mean voter position, but appear unresponsive to the policy shifts of their supporters. Additionally, they have

---

[20]The coefficient on the variable [*change in party position (t−1)*] is negative and statistically significant, which suggests that parties tend to shift policy in the opposite direction from their previous policy shift. This result is consistent with conclusions reported by Budge (1994) and Adams (2001).

noted that the opposite pattern is true for niche parties (i.e., parties belonging to the Communist, Nationalist, and Green party families), and that these parties are highly sensitive to shifts in the position of their mean supporter, while they do not respond systematically to the mean voter in the general electorate. The parameter estimates in Column 2 support these authors' findings that mainstream parties are responsive to the mean voter and that niche parties are responsive to their supporters. Moreover, the parameter estimates continue to support the substantive conclusions.

### 2.4.2 Past Election Results

Column 3 reports estimates for a *Past Election Results* model, which is identical to the basic model *except* that it controls for the possibility that parties adjust their Left–Right positions in response to the outcome of the previous election. Specifically, building on Budge's (1994) empirical finding that parties tend to shift their policies in the same direction as the last time if they gained votes at the previous election, and in the opposite direction if they lost votes (see also Adams et al. 2004; Somer-Topcu 2009; Budge et al. 2010), a variable [*vote change (t−1)*] is incorporated that denotes the party's vote gain or loss at the previous election, and the variable [*vote change (t−1) × change in party position (t−1)*] that interacts the vote change variable with the party's Left–Right shift at the previous election. A positive coefficient estimate on this interactive variable will indicate that parties tend to shift their positions in the same direction as their previous policy shift if they gained votes at the previous election, and in the opposite direction if they lost votes. The parameter estimate on this variable that is reported in column 3 is indeed positive and statistically significant, which supports Budge's arguments. More importantly, the parameter estimates continue to support the substantive conclusion that parties tend to be responsive to the mean voter position across electoral systems.

### 2.4.3 Party System Convergence

Previous studies by Adams and Somer-Topcu (2009), Ezrow (2007), and Keman and Pennings (2006) report results suggesting that parties tend to moderate their Left–Right positions over time, i.e. left-wing parties tend to shift to the right while right-wing parties shift leftward. To evaluate this hypothesis, a model was estimated that was identical to the basic model *except* that a [*party ideology*] variable was incorporated that was scored at +1 for left-wing parties, −1 for right-wing parties, and zero for centrist parties.[21] Column 4 of Table 1 reports the parameter

---

[21]Parties are defined as left-wing if the CMP classified the party as being a member of the Social Democratic party family, while right-wing parties were those that the CMP classified as belonging to the Conservative or Christian Democratic party families. Parties were defined as centrist if they were classified as members of the Liberal party family. The parties' family designations are reported in Appendix 1.

estimates for this party moderation model. The estimated coefficient on the [*party ideology*] variable is positive but is not statistically significant. The inclusion of this variable does not alter the substantive conclusions.

### 2.4.4 A Fully-Specified Model

Column 5 in Table 3 reports the parameter estimates for a *Fully-Specified Model*, which controls for type of party; past election results; and party system convergence. The coefficient estimates for this model continue to support the substantive conclusions.

### 2.4.5 Collinearity

The possibility of collinearity between public opinion and supporter positions was also considered. If it were the case that these variables are highly collinear then parsing out their effects would be difficult. This might be a problem especially for mainstream parties, where one could argue that supporters may be a more representative cross-section of the public than is the case for niche parties. There is modest evidence of this: the correlation between public opinion changes and changes in the positions of party supporters is 0.20 ($p < 0.01$). However, for mainstream parties the correlation is not so high as to create severe collinearity. In quite a few cases (40.9%) changes in public opinion are in the opposite direction to those among party supporters. Moreover, parameter estimates for a *Public Opinion* model and a *Party Supporter* model, where the effects of *only* changes in mean voter; and *only* changes in supporter positions were estimated, and in each case, the results remain unchanged.[22]

## 3 Discussion and Conclusion

## 3.1 *Summary of Findings*

How parties represent the policy preferences of citizens is a crucial aspect of political representation. In spite of the importance of understanding these linkages, there has been very little systematic cross-national empirical examination of the *dynamic* relationships that exist between parties and electorates. This chapter hurdles some of the macro-level observational barriers that are required to analyze

---

[22]These estimates are available from the author upon request. Additionally, Appendix 2 addresses two variants of endogeneity, 'cueing' and treating the data as a panel, and concludes that these do not affect the substantive conclusions.

theories at the country- and party- levels. In so doing, three findings have been identified: the first is that changes in the mean voter position cause corresponding shifts in parties' policy positions in proportional electoral systems.

The second major finding is that there is evidence that electoral systems mediate citizen-party linkages. The evidence suggests that proportional systems display mean voter representation, and disproportional systems do not. Several scholars have argued that adopting some form of proportional representation may be desirable (Lijphart 1999; see also Ezrow 2010), and this chapter suggests that they outperform disproportional systems in terms of mean voter representation. This finding, that parties are highly responsive to the mean voter position in PR systems, comports well with scholars that have emphasized post-election coalition negotiations as a factor in explaining party positioning strategies in systems featuring PR (Schofield et al. 1998; Axelrod 1970; Laver and Shepsle 1996).

The third major finding is that – apart from niche parties – there is no evidence to suggest that parties are responsive to shifts in their supporters' positions in political systems featuring either proportional or disproportional electoral systems. The central implication of this result is that the citizen-party linkage is particularly useful for understanding the policy shifts of niche parties in Western Europe, but that the remaining parties in these systems do not display a similar pattern of responsiveness. The result that niche parties respond to their supporters corroborates and extends the research that emphasizes the type of party (Adams et al. 2006; Meguid 2005, 2008; Ezrow 2008; see also Calvo and Hellwig 2011) in spatial analyses of elections and political representation.[23]

## 3.2  Mixed Party Strategies in Disproportional Systems

While the finding that parties in proportional systems respond to the mean voter position can be interpreted in a fairly straightforward fashion by referring to the work of prominent coalition theorists (e.g. Axelrod 1970; Laver and Shepsle 1996), the finding that parties neither respond to the mean voter nor their supporters in disproportional systems requires additional explanation. It has been posited that in disproportional systems, votes are more important due to the 'mechanical effect' which hurts small parties and favors large parties when national vote shares are translated into seat shares in the lower house (see Cox 1990; Dow 2001). On the surface, this would lead to the expectation that parties in disproportional systems would cater to the center of the voter distribution, in a Downsian fashion, to maximize votes. The empirical findings that are reported in this chapter obviously do not support this explanation.

---

[23]Furthermore, these conclusions support the perspective of Laver (2005) and Fowler and Laver (2008), that it is worthwhile to model competition between sets of parties that employ different decision rules.

Even if votes are more important for parties in disproportional systems, this would not necessarily translate into all parties adopting strategies that are highly responsive to the center of the voter distribution. Indeed the central implication of the Groseclose (2001) formal study of valence in two-party systems (i.e. systems that tend to rate highly in terms of disproportionality) is that when there are valence inequalities this should lead the valence-advantaged party to locate at the center – but that the valence disadvantaged party should locate in a distinctly non-centrist position. The reason is that the valence-disadvantaged party must try to distinguish itself in policy terms in order to have any chance of winning the election. Under this scenario, it makes sense for one party to follow the mean voter model, and another to conceivably follow the party supporter model to enhance its valence.

## 3.3  Looking Forward

Although there is no evidence suggesting that parties in disproportional systems are uniformly behaving as the mean voter model or the partisan-constituency model predict, this does not preclude different parties from pursuing different strategies. That is, it may be that valence-advantaged parties are responding to the mean voter, and valence-disadvantaged parties to their supporters. Identifying which parties are responding to which voters is one area for future study.

This discussion raises several additional questions for future research. Western Europe does not have a great deal of variation in terms of disproportionality. Thus, another path is to explore the validity of the general electorate model and the partisan constituency model in additional disproportional settings such as the United States and Australia.

While the evidence suggests that there are indeed direct linkages between voter preferences and the policy positions that are on offer by parties in a political system, the explanations put forth in this chapter are only tentative. Consequently, a comprehensive explanation requires *contextual* analyses of Western European parties: namely, of parties' organizational structures, of party elites, as well as of rank-and-file party supporters (see, e.g., Kitschelt 1988). An analysis of *why* different parties are apparently receiving different signals from different segments of the electorate, though outside the scope of this study, is necessary in order to reach a better understanding of how changes occur to the policy choices that political parties present to the electorate.

The results of this analysis, nonetheless, are relevant to our understanding of the democratic process. This chapter demonstrates the existence of linkages between the policy preferences of citizens and parties primarily through the mean voter position.

# Appendix 1: List of Countries, Inter-Election Periods, Parties, Party Families and Mean Left–Right Party Supporter Positions Included in the Empirical Analyses

| Country Inter-election period | Party | Party family | Mean left–right party supporter position |
|---|---|---|---|
| *Austria* 1995–1999; 1999–2003 | Austrian Peoples' Party (ÖVP) | Conservative | 5.84 |
| | League of the Independents, later named Freedom Movement (VdU/FPÖ) | Liberal | 6.63 |
| | Social Democratic Party (SPÖ) | Social democratic | 4.33 |
| | Green Alternative (GA) | Green | 4.79 |
| *Belgium* 1974–1977; 1977–1978; 1979–1981; 1985–1987; 1987–1991; 1991–1995; 1995–1999 | Christian Social Party (PSC) | Christian democratic | 6.43 |
| | Christian People's Party (CVP) | Christian democratic | 6.89 |
| | Liberal Reformation Party (PRL) | Liberal | 6.47 |
| | Liberal Reformation Party-Francophone Democratic Front (PRL-FDF) | Liberal | 6.22 |
| | Flemish Liberals and Democrats (VLD) | Liberal | 6.37 |
| | Francophone Socialist Party (PS) | Social democratic | 3.94 |
| | Flemish Socialist Party (SP) | Social democratic | 4.34 |
| | AGALEV | Green | 4.45 |
| | ECOLA | Green | 4.58 |
| | Flemish Bloc (VB) | Nationalist | 6.36 |
| *Denmark* 1977–1979; 1979–1981; 1981–1984; 1984–1987; 1987–1988; 1988–1990; 1990–1994; 1994–1998; 1998–2001 | Conservative People's Party (KF) | Conservative | 7.33 |
| | Radical Party (RV) | Liberal | 5.42 |
| | Liberals (V) | Liberal | 6.86 |
| | Social Democratic Party (SD) | Social democratic | 4.95 |
| | Center Democrats (CD) | Social democratic | 6.42 |
| | Socialist People's Party (SF) | Communist | 3.48 |
| | Progress Party (FP) | Nationalist | 7.11 |
| *Finland* 1995–1999; 1999–2003 | National Rally (KOK) | Conservative | 7.91 |
| | Finnish Center (KESK) | Liberal | 6.46 |
| | Finnish Social Democrats (SSDP) | Social democratic | 4.48 |
| | Left Wing Alliance (VL) | Communist | 3.01 |
| | Green Union (VL) | Green | 5.23 |
| *France* 1978–1981; 1981–1986; | Gaullists | Conservative | 7.14 |
| | Rally for the Republic (RPR) | Conservative | 7.18 |

*(continued)*

| Country Inter-election period | Party | Party family | Mean left–right party supporter position |
|---|---|---|---|
| 1986–1988; 1988–1993; 1993–1997; 1997–2002 | Union for French Democracy (UDF) | Conservative | 6.43 |
| | Socialist Party (PS) | Social democratic | 3.66 |
| | French Communist Party (PCF) | Communist | 2.59 |
| *Germany* 1976–1980; 1980–1983; 1983–1987; 1987–1990; 1990–1994; 1994–1998; 1998–2002 | Christian Democratic Party/ Christian Social Union (CDU/CSU) | Christian democratic | 6.43 |
| | Free Democratic Party (FDP) | Liberal | 5.86 |
| | Social Democratic Party (SDP) | Social democratic | 4.40 |
| | Party of German Socialism (PDS) | Communist | 4.23 |
| *Greece* 1981–1985; 1985–1989 (June); 1989–1989 (Nov); 1989–1990; 1990–1993; 1993–1996; 1996–2000 | New Democracy (ND) | Christian democratic | 8.14 |
| | Panhellenic Socialist Movement (PASOK) | Social democratic | 4.59 |
| | Communist Party of Greece (KKE) | Communist | 2.12 |
| | Progressive Left Coalition (SAP) | Communist | 2.69 |
| *Ireland* 1977–1981; 1981–1982 (Feb); 1982–1982 (Nov); 1982–1987; 1987–1989; 1989–1992; 1992–1997; 1997–2002 | Fianna Fail | Conservative | 6.56 |
| | Fine Gail | Christian democratic | 6.37 |
| | Progressive Democrats (PD) | Liberal | 6.15 |
| | Labour Party (LP) | Social democratic | 4.88 |
| *Italy* 1976–1979; 1979–1983; 1987–1992; 1992–1994; 1994–1996; 1996–2001 | Italian Social Movement (AN) | Nationalist | 8.26 |
| | Northern League (LN) | Nationalist | 6.10 |
| | Go Italy (FI) | Conservative | 7.01 |
| | Italian People's Party (PPI) | Christian democratic | 5.70 |
| | Republican Party (PRI) | Liberal | 4.95 |
| | Italian Democratic Socialist Party (PSDI) | Social democratic | 4.50 |
| | Socialist Party (PSI) | Social democratic | 3.62 |
| | Newly Founded Communists (RC) | Communist | 2.06 |
| | Democrats of the Left (DS) | Communist | 2.48 |
| *Luxembourg* 1979–1984; 1984–1989; 1989–1994; 1994–1999 | Christian Social People's Party (PCS/CSV) | Christian democratic | 6.82 |
| | Patriotic and Democratic Group (PD/DP) | Liberal | 5.76 |
| | Socialist Workers' Party (POSL/LSAP) | Social democratic | 4.20 |
| | Communist Party (PCL/KPL) | Communist | 3.14 |
| *Netherlands* 1977–1981; 1981–1982; 1982–1986; 1986–1989; 1989–1994; 1994–1998; 1998–2002 | Christian Democratic Appeal (CDA) | Christian democratic | 6.67 |
| | People's Party for Freedom and Democracy (VVD) | Liberal | 6.84 |
| | Labour Party (PvdA) | Social democratic | 3.76 |

(*continued*)

| Country Inter-election period | Party | Party family | Mean left–right party supporter position |
|---|---|---|---|
| | Democrats 66 (D'66) | Social democratic | 4.75 |
| | Green Left (GL) | Green | 3.34 |
| *Portugal* 1987–1991; 1991–1995; 1995–1999 | Center Social Democrats (CDS/PP) | Conservative | 7.37 |
| | Popular Democratic Party (PPD/PSD) | Social democratic | 6.95 |
| | Portuguese Socialist Party (PSP) | Social democratic | 4.63 |
| | Unified Democratic Coalition, (CDU) | Communist | 2.64 |
| *Spain* 1986–1989; 1989–1993; 1993–1996; 1996–2000 | Popular Alliance (AP/PP) | Conservative | 7.29 |
| | Convergence and Union (CiU) | Conservative | 5.33 |
| | Spanish Socialist Workers' Party (PSOE) | Social democratic | 3.54 |
| | Communist Party (IU) | Communist | 2.74 |
| *Sweden* 1994–1998; 1998–2002 | Moderate Coalition Party (MSP) | Conservative | 7.80 |
| | Christian Democratic Community Party (KdS) | Christian democratic | 6.48 |
| | People's Party (FP) | Liberal | 6.23 |
| | Center Party (CP) | Liberal | 5.90 |
| | Social Democratic Labour Party (SdsP) | Social democratic | 4.18 |
| | Communist Party (VP) | Communist | 2.90 |
| | Green Party | Green | 4.63 |
| *United Kingdom* 1979–1983; 1983–1987; 1987–1992; 1992–1997; 1997–2001 | Conservative Party | Conservative | 6.97 |
| | Liberal Democrats (LD) | Liberal | 5.39 |
| | Labour Party | Social democratic | 4.26 |

*Notes*: Parties are observed in at least three successive elections. The mean Left-Right party supporter position is calculated as the average of the mean party supporter positions for all of the elections in which the party is included in the empirical analysis

# Appendix 2: Addressing Two Types of Endogeneity

The first variant of endogeneity concerns the possible cueing of public opinion or supporters by parties. Although this study concentrates on the ideological *linkages* between parties and citizens rather than the direction in which ideological preferences are transmitted between these groups, cueing is nevertheless relevant. Indeed, it may be that citizens respond to parties, rather than the reverse, in which case the coefficients reported may be biased and inconsistent. Past studies that have explicitly addressed this issue of causality have found that any cue-giving effects by parties tend to be weaker than the corresponding cues that voters transmit to parties (see Carrubba 2001; Steenbergen et al. 2007). Nonetheless,

a Durbin-Wu-Hausman test was performed in which public opinion and supporter positions were modeled as functions of their lagged values, computed the residuals, and entered these residuals into a model of parties' policy positions. If public opinion and supporter positions were endogenous with respect to party positions, then these positions would be reflected in the residuals; namely, the residuals would exert a significant effect on parties' Left–Right positions. There was no evidence of this for mainstream parties. For niche parties there was evidence that the public opinion residual is significant. For mainstream parties, the effect of the public opinion residuals yields $p = 0.941$, while the corresponding $p$-value for the supporter residuals is 0.261. For niche parties, the $p$-values are 0.376 for the supporter residual and 0.029 for the public opinion residual. The upshot is that the estimates reported in the tables are not marred by endogeneity. With the possible exception of the effect of public opinion on niche parties, these estimates are consistent.

The second type of endogeneity that could be present in the empirical analyses is that the results could inherently be tilted towards supporting the partisan constituency model of political representation. To the extent that this type of endogeneity is a problem, it should bias the statistical analyses in favor of the finding that parties respond to their supporters. For example, if a party shifts away from a segment of its constituency it should lose support from some of these voters. This policy shift should also produce a subsequent increase in support for the party by virtue of its moving towards a new set of voters. If this 'party switching' process were taking place, we would observe a pattern in which the mean Left–Right supporter position would shift in tandem with the party's policy shift even if none of the voters are actually shifting their ideological positions. Since this should bias the empirical findings towards a finding that *all* parties are responsive to their supporters, it cannot account for the observation that only niche parties are responding to their supporters. Given this unavoidable feature of the empirical analyses, it actually further strengthens the finding that mainstream parties respond to the mean voter position. Thus to the extent that this second type of endogeneity is a "problem," it actually strengthens the substantive conclusion that mainstream parties are *not* disproportionately responsive to their supporters, since the statistical analyses may be biased in the opposite direction.

# References

Adams J (2001) Party competition and responsible party government: a theory of spatial competition based upon insights from behavioral voting research. University of Michigan Press, Ann Arbor

Adams J, Merrill S (1999) Party policy equilibrium for alternative spatial voting models: an application to the Norwegian storting. Eur J Polit Res 36:235–255

Adams J, Merrill S (2000) Spatial models of candidate competition and the 1988 French presidential election: are presidential candidates vote-maximizers? J Polit 62:729–56

Adams J, Somer-Topcu Z (2009) Moderate now, win votes later: the electoral consequences of parties' policy shifts in twenty-five postwar democracies. J Polit 71:678–692

Adams J, Clark M, Ezrow L, Glasgow G (2004) Understanding change and stability in party ideologies: do parties respond to public opinion or to past election results? Br J Polit Sci 34:589–610

Adams J, Merrill S, Grofman B (2005) A unified theory of party competition: a cross-national analysis integrating spatial and behavioral factors. Cambridge University Press, Cambridge

Adams J, Clark M, Ezrow L, Glasgow G (2006) Are Niche parties fundamentally different from mainstream parties?: the causes and electoral consequences of Western European parties' policy shifts, 1976–1998. Am J Polit Sci 50:513–529

Aldrich J (1983a) A Downsian spatial model with party activists. Am Polit Sci Rev 77:974–990

Aldrich J (1983b) A spatial model with party activists: implications for electoral dynamics. Public Choice 41:63–100

Aldrich J (1995) Why parties? University of Chicago Press, Chicago

Axelrod R (1970) Conflict of interest. Markham, Chicago

Beck N, Katz J (1995) What to do (and not to do) with time-series cross-section data. Am Polit Sci Rev 89:634–647

Beck N, Katz J (1996) Nuisance vs. substance: specifying and estimating time-series-cross-section models. Polit Anal 6:1–36

Braumoeller B (2004) Hypothesis testing and multiplicative interaction terms. Int Organ 58:807–820

Budge I (1994) A new theory of party competition: uncertainty, ideology, and policy equilibria viewed comparatively and temporally. Br J Polit Sci 24:443–467

Budge I, McDonald M (2006) Choices parties define: policy alternatives in representative elections – 17 countries 1945–1998. Party Polit 12:451–466

Budge I, Klingemann H-D, Volkens A, Bara J, Tanenbaum E (2001) Mapping policy preferences: estimates for parties, electors, and governments 1945–1998. Oxford University Press, Oxford

Budge I, Ezrow L, McDonald M (2010) Ideology, party factionalism and policy change: an integrated dynamic theory. Br J Polit Sci 40:781–804

Burt G (1997) Party policy: decision rule or chance? a note on budge's new spatial theory of party competition. Br J Polit Sci 27:647–58

Calvo E, Hellwig T (2011) Centripetal and centrifugal incentives under different electoral systems. Am J Polit Sci 55:27–41

Carrubba C (2001) The electoral connection in European union politics. J Polit 63:141–58

Cox G (1990) Centripetal and centrifugal incentives in electoral systems. Am J Polit Sci 34:905–935

Cox G (1997) Making votes count. Cambridge University Press, Cambridge

Dalton R (1985) Political parties and political representation. Comp Polit Stud 17:267–299

Dow JK (2001) A comparative spatial analysis of majoritarian and proportional elections. Elect Stud 20:109–125

Dow JK (2011) Party system extremism in majoritarian and proportional electoral systems. Br J Polit Sci 41:341–361

Downs A (1957) An economic theory of democracy. Harper, New York

Enelow J, Hinich M (1984) The spatial theory of voting. Cambridge University Press, Cambridge

Erikson R, MacKuen M, Stimson J (2002) The macro polity. Cambridge University Press, Cambridge

Ezrow L (2007) The variance matters: how party systems represent the preferences of voters. J Polit 69:182–192

Ezrow L (2008) Research note: on the inverse relationship between votes and proximity for niche parties. Eur J Polit Res 47:206–220

Ezrow L (2010) Linking citizens and parties. Oxford University Press, Oxford

Ezrow L, De Vries CE, Steenbergen M, Edwards EE (2011) Mean voter representation and partisan constituency representation: do parties respond to the mean voter position or to their supporters? Party Polit 17:275–301

Fowler J, Laver M (2008) A tournament of party decision rules. J Conflict Resolut 52:68–92

Gallagher M (1991) Proportionality, disproportionality and electoral systems. Elect Stud 10:33–51

Green DP, Kim SY, Yoon DH (2001) Dirty pool. Int Organ 55:441–468

Groseclose T (2001) A model of candidate location when one candidate has a valence advantage. Am J Polit Sci 45:862–86

Hearl D (2001) Checking the party policy estimates: reliability. In: Budge I, Klingemann H-D, Volkens A, Bara J, Tanenbaum E (eds) Mapping policy preferences: estimates for parties, electors, and governments 1945–1998. Oxford University Press, Oxford

Hsiao C (2003) Analysis of panel data, 2nd edn. Cambridge University Press, Cambridge

Huber J (1989) Values and partisanship in left-right orientations: measuring ideology. Eur J Polit Res 17:599–621

Huber J, Powell GB (1994) Congruence between citizens and policymakers in two visions of liberal democracy. World Polit 46:291–326

Keman H, Pennings P (2006) Competition and coalescence in European party systems: social democracy and Christian democracy moving into the 21st century. Swiss Polit Sci Rev 12:95–126

Kitschelt H (1988) Organization and strategy of Belgian and West German ecology parties: a new dynamic of party politics in Western Europe? Comp Polit 20:127–154

Klingemann H-D, Volkens A, Bara J, Budge I, McDonald M (2006) Mapping policy preferences II: estimates for parties, electors and governments in central and Eastern Europe, European Union and OECD 1990–2003. Oxford University Press, Oxford

Laver M (2005) Policy and the dynamics of political competition. Am Polit Sci Rev 99:263–281

Laver M, Shepsle K (1996) Making and breaking governments. Cambridge University Press, Cambridge

Laver M, Benoit K, Garry J (2003) Extracting policy positions from political texts using words as data. Am Polit Sci Rev 97:311–331

Lijphart A (1999) patterns of democracy: government forms and performance in thirty-six countries. Yale University Press, New Haven

Macdonald SE, Rabinowitz G (1998) Solving the paradox of nonconvergence: valence, position, and direction in democratic politics. Elect Stud 17:281–300

McDonald M, Budge I (2005) Elections, parties, and democracy: conferring the median mandate. Oxford University Press, Oxford

McDonald M, Mendes S (2001) Checking the party policy estimates: convergent validity. In: Budge I, Klingemann H-D, Volkens A, Bara J, Tanenbaum E (eds) Mapping policy preferences: estimates for parties, electors, and governments 1945–1998. Oxford University Press, Oxford

Meguid B (2005) Competition between unequals: the role of mainstream party strategy and niche party success. Am Polit Sci Rev 99:347–360

Meguid B (2008) Party competition between unequals. Cambridge University Press, Cambridge

Miller G, Schofield N (2003) Activists and partisan realignment in the United States. Am Polit Sci Rev 97:245–260

Powell GB (2000) Elections as instruments of democracy. Yale University Press, New Haven

Schmitt H, Scholz E (2005) The Mannheim eurobarometer trend file, 1970–2002 [Computer file]. Prepared by Zentralarchiv fur Empirische Sozialforschung. ICPSR04357-v1

Schofield N (1997) A comparison of majoritarian and proportional electoral systems based on spatial modeling and "Rational" politicians. Presented at the conference on constitutional issues in modern democracies, Messina, Italy, 25–27 Sept 1997

Schofield N, Sened I (2005) Modeling the interaction of parties, activists, and voters: why is the political center so empty? Eur J Polit Res 44:355–390

Schofield N, Sened I (2006) Multiparty democracy: parties, elections, and legislative politics. Cambridge University Press, Cambridge

Schofield N, Martin A, Quinn K, Nixon D (1998) Multiparty electoral competition in the Netherlands and Germany: a model based on multinomial probit. Public Choice 97:257–93

Somer-Topcu Z (2009) Timely decisions: the effects of past national elections on party policy change. J Polit 71:238–248

Steenbergen M, Edwards EE, De Vries CE (2007) Who's cueing whom?: Mass-elite linkages and the future of European integration. Eur Un Pol 8:13–35

Stimson J, MacKuen M, Erikson R (1995) Dynamic representation. Am Polit Sci Rev 89:543–565

Taagepera R, Shugart M (1989) Seats and votes: the effects and determinants of electoral systems. Yale University Press, New Haven

Weissberg R (1978) Collective vs. dyadic representation in congress. Am Polit Sci Rev 72: 535–547

Wessels B (1999) System characteristics matter, empirical evidence from ten representation studies. In: Miller W, Homberg S, Pierce R (eds) Policy representation in western democracies. Oxford University Press, Oxford

# Electoral Institutions and Political Corruption: Ballot Structure, Electoral Formula, and Graft

**Daniel Max Kselman**

## 1 Introduction

How do electoral institutions affect the incidence of political corruption? In contrast to previous research, this paper suggests that the most important institutional parameter for reducing political graft is its *ballot structure*, and in particular whether it permits citizens to cast intra-party candidate votes which affect the organizational allocation of legislative seats. Sections 2–5 present a game theoretic model from which this core hypothesis emerges. I then test the prediction with cross-national data on electoral institutions and political corruption. Taken together, the paper's results provide a strong counter-argument to the notion that majoritarian institutions generate better governance than their proportional representation counterparts.

Legislative electoral institutions occupy an important explanatory position in contemporary political science. Most studies of formal electoral rules examine the distinction between majoritarian (MAJ) and proportional (PR) *electoral formulae*. The most common MAJ formula is the well-known First-Past-The-Post (FPTP) system used, among elsewhere, in Great Britain, Canada, and the United States. Recent research has also examined a system's *ballot structure*, i.e., the formal rules governing how citizens cast their votes. Of particular interest have been: (a) *preferential* rank-ordering systems which allow voters to identify not only their most-preferred candidate but also their second-most-preferred, third-most preferred, etc; and (b) *open-list* systems in which voters may simultaneously express support for political party organizations *and* particular candidates within these organizations' electoral lists.

One stream of research on the consequences of electoral institutions investigates their effect on a country's *party system*. Early work by Duverger (1959) on the relationship between electoral formulae and party-system fragmentation has since been formalized and qualified by Riker (1982), Palfrey (1989), and Cox (1994, 1997). In a distinct set of papers Cox (1987, 1990) investigates the consequences of electoral formulae and preferential rank-scoring systems for party system

D.M. Kselman
Center for Advanced Studies in the Social Sciences, Juan March Institute, C/Castelló 77, 28006 Madrid, Spain
e-mail: dkselman@march.es

polarization.[1] In turn, both the number of parties and their systemic polarization comprise important mechanisms in studies of PR's consequences for *ethnic conflict* (Lijphart 1977; Horowitz 1985; Reilly 2001; Fraenkel and Grofman 2004).

Another body of work investigates the consequences of electoral institutions for *economic policy*. A series of articles addresses the relative merits of PR as opposed to MAJ systems in generating socio-economic redistribution (Austen-Smith 2000; Iversen and Soskice 2006; Long Jusko 2009). As well, Persson and Tabellini (2000, 2003) argue that PR competition favors the production of "public good" policies applicable to society as a whole, whereas MAJ competition favors the production of more decentralized "club good" policies targeted to individual geographic constituencies.[2] Finally, Myerson (1993a) argues that pure plurality rule should generate economic policies more narrowly targeted to exclusive social minorities than those in preferential systems such as the Alternative Vote, the Borda Count, and the Negative Plurality Vote.

As is clear from the preceding paragraphs and footnotes, most work on the consequences of electoral institutions for party systems, ethnic conflict, and economic policy has emphasized the causal effects of electoral formulae and preferential rank ordering systems, to the exclusion of open-list systems which permit intra-party candidate voting. While early work on the relationship between electoral rules and *political corruption* similarly emphasized the impact of MAJ vs. PR electoral formulae (Myerson 1993b; Lijphart 1999; Persson and Tabellini 2000), more recent contributions have distinguished between open-list PR systems (OLPR) in which voters may simultaneously express support for a political party and a particular candidate within that party's electoral list; and closed-list PR systems (CLPR) in which voters may not express preferences for specific candidates within a party's list (Persson et al. 2003; Persson and Tabellini 2003; Kunicova and Rose-Ackerman 2005).

A growing consensus in this literature claims that "...proportional representation (PR) systems are more susceptible to corrupt political rent-seeking than plurality systems."[3] More particularly, the claim is that FPTP elections generate stronger ties of personal accountability between legislators and constituents than *both* OLPR and CLPR systems, which in turn makes FPTP particularly effective at constraining political corruption (Gingerich 2009 offers a more nuanced evaluation of OLPR). These same papers argue that OLPR produces levels of legislative accountability and political corruption *intermediate* to high accountability FPTP systems and low accountability CLPR systems. Like FPTP systems, OLPR contains direct candidate voting which is likely to strengthen legislative accountability; but like CLPR, legislative accountability in OLPR systems is diluted by the existence of party lists. Since

---

[1]Among other results he establishes that multi-candidate plurality rule elections generate non-centrist policy platforms; identifies a set of preferential rank-scoring rules that generate centrist equilibria in multi-candidate contests; and argues that PR systems with large district magnitudes should generate fairly polarized party competition.

[2]See Milesi-Ferretti et al. (2002) for a similar argument (albeit conceptually reversed: they refer to geographically targeted policies as "public goods").

[3]Kunicova and Rose-Ackerman 2005, p. 573.

OLPR shares elements of both MAJ and CLPR systems, the above authors hypothesize that it ought to generate levels of rent-seeking in between those of high accountability MAJ systems and low accountability CLPR systems.

## 2 The Basic Model

Contrary to this growing consensus, I now develop a game theoretic model which suggests that open-list systems' unique combination of intra- and inter-party competition should generate lower levels of corruption and higher levels of legislative accountability than *both* FPTP systems *and* closed-list electoral systems. Also counter to received wisdom, the model identifies conditions under which CLPR systems might themselves actually outperform FPTP systems, highlighting partisanship patterns in the electorate, the size of multi-member districts, and political parties' nomination procedures as crucial intervening variables. The model's primary strategic actors will be *incumbent legislators*, who must allocate scarce resources in the pursuit of distinct, and often mutually exclusive, political and material goals (Fenno 1978). This paper emphasizes two such behavioral options:

1. Devoting effort to the provision of goods and services for *district constituents* (e.g. pork projects, civil service jobs, ombudsman services, etc)
2. Devoting effort to the pursuit of one's own *personal enrichment* (e.g. through bribery, embezzlement, and other forms of *political corruption*)

In FPTP systems, individual incumbents are clearly affiliated with individual electoral districts, and their particularistic efforts on behalf of constituents will, by definition, be targeted to voters in their particular district. On the other hand, in *multi-member district* systems, it is generally impossible for incumbent legislators to develop particularistic relationships with *all* voters in these substantially larger electoral districts.[4] To model the process by which legislators in multi-member districts develop particularistic relationships, I will assume (**Assumption 1**) that individual incumbents each have *regions* (a.k.a. *bailiwicks*) within larger districts which they *may* target with particularistic goods and services. As well, I will assume that no two incumbents target *the same* region with their particularistic efforts (**Assumption 2**).[5]

In order to operationalize these assumptions, let $N$ represent the number of *incumbents* in a country's national Legislature, $D$ represent the number of *districts*

---

[4]For example, research on countries as varied as Ireland (Martin 2010), Brazil (Ames 1995), Argentina (Szwarcberg 2009), Columbia (Ingall and Crisp 2001), and Turkey (Kselman 2009) suggests that individual incumbents from multi-member districts target their particularistic efforts to well-defined regional or municipal strongholds (*aka* "bailiwicks") within larger electoral districts.

[5]Although a simplification, the case studies cited in ftn 4 suggest that incumbent legislators in multi-member district systems are adept at collectively carving out individual constituencies within which their jurisdiction is unchallenged; and that incumbents" primary competition for voters" particularistic allegiances often comes not from fellow legislators, but from *non-incumbent challengers*. As such, as a simplification the assumption is not only useful but also justifiable.

in the same country's electoral system, and $J$ represent the number of distinct *regions* contained within a country's territorial jurisdiction. Without loss of generality let $N = J$. Individual electoral districts will be denoted with the marker $d \in \{1, 2, .., D\}$, and the indicator $M_d \geq 1$ will represent the number of legislators which district $d$ sends to the national Legislator, i.e., its district *magnitude*. To model FPTP, I will assume that electoral districts and regions are *coterminous*, i.e. that each region is encompassed by a single electoral district which sends a single representative to the national Legislature (such that $D = N = J$).[6] On the other hand, in multi-member-district systems individual regions are regrouped into larger, multi-regional electoral districts, each of which sends multiple representatives to the Legislature. To implement Assumptions 1 and 2, assume that incumbents in multi-member-district systems each have the option of exerting particularistic effort on behalf of a single region within a larger electoral district, and no two incumbents devote particularistic effort to the same region. Denote individual regions within a multi-member district's boundaries as $j \in \{1, 2, ..., M_d\}$.

The current paper examines a game in which each of the $N$ incumbent legislators is affiliated with one of *two* political parties $P \in \{A, B\}$.[7] For an incumbent affiliated with party $P$, district $d$, and region $j$, denote $f_{j,d}^P$ as the level of effort devoted to securing regional voters' particularistic interests; and $c_{j,d}^P$ as the level of effort devoted to securing one's own material enrichment. All legislative incumbents from party $P$ are endowed with a fixed amount of effort $E^P$ which they divide exhaustively between $f_{j,d}^P$ and $c_{j,d}^P$, implying the effort constraint $f_{j,d}^P + c_{j,d}^P = E^P$. Define $\mathbf{F_d} = \{f_{1,d}, f_{2,d}, ..., f_{M_d,d}\}$ as a *strategy vector* containing the constituency effort allocations of all incumbents from a district $d$.[8] In single-member district systems $\mathbf{F_d} = \{f_{1,d}\}$.

Legislators will allocate this effort in the pursuit of two objectives: the desire for *re-election* and the desire for personal *material wealth*. Consider the following utility function for an incumbent affiliated with party $P$, district $d$, and region $j$:

$$U_{j,d}^P = c_{j,d}^P + [\pi_{j,d}^P(\mathbf{F_d}) \cdot \beta_{j,d}^P], \tag{1}$$

where $\pi_{j,d}^P(\cdot)$ represents the legislator's probability of gaining re-election and $\beta_{j,d}^P$ represents the benefit he or she associates with re-election. Legislators' utility thus increases linearly with effort devoted to the acquisition of material wealth. As demonstrated in the vote share expressions below, effort devoted to $f_{j,d}^P$ provides indirect and contingent benefits by increasing legislators' support among citizens in their affiliated regions, which may in turn increase $\pi_{j,d}^P(\cdot)$. The argument contains the strategy vector $\mathbf{F_d}$ rather than the individual choice $f_{j,d}^P$, since in multi-member

---

[6]Throughout I will assume that electoral districts do not "bisect" regions, i.e. the entirety of an individual region's territory is located in the same electoral district.

[7]As demonstrated elsewhere (Kselman 2010), the forthcoming results persist, and at times emerge even *more strongly*, in a theoretical environment where the number of parties is greater then two.

[8]I omit superscripts in strategy vectors for notational ease.

district systems re-election probabilities will depend not only one's own effort allocation, but also those of fellow incumbents from the same district.

In the game's first stage, all $N$ legislators will simultaneously allocate their fixed amount of effort $E^P$ between $f^P_{j,d}$ and $c^P_{j,d}$. In the second stage an election is held and citizens choose between parties $A$ and $B$. Appendix 1 develops an explicit model of voter choice to derive parties' *regional vote shares* given some allocation $f^P_{j,d}$ and $c^P_{j,d}$ by the affiliated regional incumbent; here I simply reproduce the key results derived therein. For a region whose incumbent is affiliated with party $P$, define $\ell^P_{j,d} \in [0, 1]$ as the percentage of regional *party loyalists*: voters whose "partisan bias" (see Appendix 1) towards the party of their regional incumbent is strong enough that they will support said party *even if* their regional incumbent chooses $f^P_{j,d} = 0$ and $c^P_{j,d} = E^P$. In turn, we can express a party $P$'s vote share in any region where the affiliated incumbent is a member of party $P$ as follows (Lemma 1 from Appendix 1):

$$V^P_{j,d}(f^P_{j,d}) = \begin{cases} (f^P_{j,d} + \ell^P_{j,d}) & \text{if } f^P_{j,d} < 1 - \ell^P_{j,d} \\ 1 & \text{if } f^P_{j,d} \geq 1 - \ell^P_{j,d} \end{cases} \qquad (2)$$

Naturally, party $P$'s vote share in region $j$ increases with $f^P_{j,d}$, the regionally targeted efforts of its legislative incumbent. Similarly, for equal allocations of $f^P_{j,d}$, the vote share $V^P_{j,d}(\cdot)$ will be higher in regions more where voters are more pre-disposed to party $P$ (i.e., regions with *higher* values of $\ell^P_{j,d}$). Finally, any constituency effort at or above the level $f^P_{j,d} = 1 - \ell^P_{j,d}$ yields a vote share of $V^P_{j,d}(\cdot) = 1$. Without loss of generality I will assume that voters cannot abstain, such that party $A$'s vote percentage in regions where the incumbent is from $B$ will simply be $[1 - V^B_{j,d}(\cdot)]$, and party $B$'s vote percentage in regions whose incumbent is from $A$ will be $[1 - V^A_{j,d}(\cdot)]$.

Incumbents must thus allocate their single unit of effort between $f^P_{j,d}$ and $c^P_{j,d}$ in the game's first stage so as to maximize their utility, taking into account the resulting vote outcomes in the game's subsequent electoral stage. In FPTP systems, incumbents can choose an optimal mix of $f^P_{j,d}$ and $c^P_{j,d}$ *decision-theoretically*: $\mathbf{F_d} = \{f_{1,d}\}$ is a one-dimensional vector, and effort allocations in one district have no bearing on other incumbents' likelihood of gaining re-election. Since only two parties compete, by the definition of plurality rule incumbents must secure *just over half* of their district's votes to gain re-election with certainty (ties are broken randomly). Define $\hat{f}^P_{1,d}$ as the *critical level* of constituency effort an incumbent from district $d$ must exert so as to win with certainty. From (2) above it is straight-forward to derive this critical level:

$$\hat{f}^P_{1,d} = \begin{cases} \dfrac{1}{2} - \ell^P_{1,d} + \varepsilon & \text{if } \ell^P_{1,d} \leq \dfrac{1}{2} \\ 0 & \text{if } \ell^P_{1,d} > \dfrac{1}{2} \end{cases}, \qquad (3)$$

where $\varepsilon \to 0$ represents the infinitesimal effort increment needed to make $V_{1,d}^P(\cdot) > \frac{1}{2}$ (Appendix 1 addresses the open-set problem associated with infinitesimal moves). Now, define $f_{1,d}^{P*}$ as the utility-maximizing choice of the incumbent from district $d$. Returning to expression (1) above we see that, trivially, if $E^P < \hat{f}_{1,d}^P$ then $f_{1,d}^{P*} = 0$, i.e., an incumbent without the resources necessary to gain re-election will simply shirk. Proposition 1 identifies $f_{1,d}^{P*}$ when $E^P \geq \hat{f}_{1,d}^P$:

**Proposition 1:** *In FPTP elections, if $E^P \geq \hat{f}_{1,d}^P$ then:*

$$f_{1,d}^{P*} = \begin{cases} \hat{f}_{1,d}^P & \text{if} \quad \beta_{1,d}^P > \hat{f}_{1,d}^P \\ 0 & \text{if} \quad \beta_{1,d}^P < \hat{f}_{1,d}^P \\ \hat{f}_{1,d}^P \text{ or } 0 & \text{if} \quad \beta_{1,d}^P = \hat{f}_{1,d}^P \end{cases}.$$

Not surprisingly, the lower the benefit associated with re-election, and the higher its costs in terms of regional effort, the more likely incumbents will choose maximal personal enrichment ($f_{1,d}^{P*} = 0$) as opposed to re-election ($f_{1,d}^{P*} = \hat{f}_{j,d}^P$).[9]

## 3  Legislative Equilibrium Under Closed-List Proportional Systems

In PR systems electoral districts send $M_d > 1$ incumbents to the national Legislature, and $\mathbf{F_d}$ is a multi-dimensional strategy vector. Incumbents' probability of winning $\pi_{j,d}^P(\mathbf{F_d})$ in multi-member systems will be determined not only by their own effort allocation $f_{j,d}^P$, but also by the effort allocations of the remaining $(M_d - 1)$ incumbents from their particular district. Thus, solving for optimal legislative effort allocations in PR systems becomes a *game theoretic* problem. Define $\bar{A}_d$ and $\bar{B}_d$ as the number of *current incumbents* from district $d$ who are affiliated with parties $A$ and $B$ respectively. Given a set of effort allocations $\mathbf{F_d}$ by district-incumbents in the game's first stage, we can derive $\mathbf{v_d^A}(\mathbf{F_d})$, party $A$'s percentage of district $d$'s total votes in the game's subsequent electoral stage:

---

[9]To prove this result, first note that the utility incumbents receive from choosing $\hat{f}_{1,d}^P$ is $(E^P - \hat{f}_{1,d}^P) + \beta_{1,d}^P$, i.e., they receive with certainty the fixed benefit associated with re-election and devote any surplus effort to pursuing personal material interests (such that $c_{1,d}^P = E^P - \hat{f}_{1,d}^P$). Were incumbent legislators to forgo re-election, the optimal allocation would be to choose $c_{1,d}^P = E^P$. A straight-forward utility comparison tells us that $(E^P - \hat{f}_{1,d}^P) + \beta_{1,d}^P$ is greater than $E^P$ as long as $\beta_{1,d}^P > \hat{f}_{1,d}^P$. Incumbents for whom $\beta_{1,d}^P < \hat{f}_{1,d}^P$ will thus prefer to devote all of their effort to personal enrichment. In this case, the district's seat transfers to a non-incumbent from the opposing party, who is assigned no strategic move in the game. Finally, if $\beta_{1,d}^P = \hat{f}_{1,d}^P$ the incumbent from $d$ will be indifferent between $f_{1,j}^{P*} = \hat{f}_{1,d}^P$ and choosing $f_{1,j}^{P*} = 0$.

$$\mathbf{v_d^A}(\mathbf{F_d}) = \frac{\sum_{\bar{A}_d} V_{j,d}^A(\cdot) + \sum_{\bar{B}_d} [1 - V_{j,d}^B(\cdot)]}{M_d}. \tag{4}$$

Recalling (2) above, the first term in (4)'s numerator represents the summation of party $A$'s vote shares in regions whose incumbent is from party $A$; and the second term represents the summation party $A$'s vote shares in regions whose incumbent is from party $B$. These additive terms must then be divided by the district's magnitude to generate an aggregate vote percentage. Party $B$'s district-level vote share can be expressed similarly, and is equal to $\mathbf{v_d^B} = (1 - \mathbf{v_d^A})$.

I employ a *simple quota* and *largest remainder rule* to model the process by which these district level vote shares are translated into legislative seats. Define $q = 1/M_d$ as the electoral quota needed to earn an individual legislative seat, and consider a district of magnitude $M_d = 10$, such that $q = 10\%$. As an example, if party $A$ secures a district-level vote share of $\mathbf{v_d^A} = 58\%$ party $B$ receives $\mathbf{v_d^B} = 42\%$, then $A$'s vote share contains 5 full quotas and $B$'s vote share contains 4 full quotas, implying that in a first allocation parties $A$ and $B$ will receive 5 and 4 seats respectively. As for the final seat, it will go to $A$ because her *remainder* of 8%, the vote share left over after her 5 quotas are subtracted from $\mathbf{v_d^A}$, is larger than $B$'s remainder of 2%. In a final tally $A$ will thus win 6 seats and $B$ will win 4. If parties have identical remainders of 5%, the final seat is allocated with an unbiased coin-flip.

At election time both parties present a *list* of $M_d$ candidates to a district's electorate. Among these candidates are the parties' legislative incumbents from district $d$ and a set of *non-incumbent candidates*, who are assigned no strategic move in the game. Party $A$'s ($B$'s) list thus contains $\bar{A}_d$ ($\bar{B}_d$) incumbents and $M_d - \bar{A}_d$ ($M_d - \bar{B}_d$) non-incumbents. Define $\mathbf{x_d^P}$ as the number of seats won by party $P$ in district $d$ during the game's election. After the election, these $\mathbf{x_d^P}$ seats are subsequently allocated to the *top $\mathbf{x_d^P}$ candidates on party $P$'s electoral list*. In CLPR systems, a candidate's position on his or her party's electoral list is *fixed prior to the general election*. Individual parties in CLPR systems may employ any number of organizational mechanisms to fix candidate list positions prior to a general election. In the text I examine the CLPR game in its simplest form by making the following assumption: incumbents from both parties $A$ and $B$ occupy *higher list positions than their parties' non-incumbents* (**Assumption 3**). This assumption implies no restriction as to which of a party's incumbents is $1^{st}$ on the list, which is $2^{nd}$, and so on; it stipulates only that incumbents have more favorable positions than non-incumbents. I now solve the following *district-level* game[10]:

1. In a first stage incumbent list positions in district $d$ are fixed, such that incumbents occupy higher list positions than non-incumbents.
2. In a second stage all incumbents from $d$ simultaneously choose $f_{j,d}^P$ and $c_{j,d}^P$.

---

[10]Note that Nash Equilibrium effort allocations by the $M_d$ incumbents in district $d$ do not depend on the effort allocation decisions of the $(N - M_d)$ incumbents from regions outside of $d$.

3. A third-stage election is held in which parties $A$ and $B$ receive vote percentages $\mathbf{v_d^A}$ and $\mathbf{v_d^B}$ respectively.
4. Parties $A$ and $B$ are allocated $\mathbf{x_d^A}$ and $\mathbf{x_d^B}$ seats respectively via the quota-remainder rule presented above.
5. And these seats go to $\mathbf{x_d^P}$ the candidates with party $P$'s highest list positions.

Under Assumption 3 it is straight-forward to show that, given any district-level vote outcomes $\mathbf{v_d^A}$ and $\mathbf{v_d^B}$ in the game's electoral stage, at least one of the two parties will have *all of its district-level incumbents re-elected.*[11] Define $\mathbf{F_d^o} = \{0, 0, ..., 0\}$ as the *full-shirking* strategy vector, that at which all incumbents from district $d$ choose $c_{j,d}^P = E^P$ and devote *no* effort to constituent interests; and let $P_d^+(P_d^-)$ denote the party from district $d$ whose incumbents are (are not) all re-elected when $\mathbf{F_d^o}$ is played.

Recalling the utility function specified in (1) above, we know that no incumbent who secures re-election when the full-shirking vector $\mathbf{F_d^o}$ is played has any incentive to alter his or her effort allocation: they secure re-election despite having chosen $c_{j,d}^P = E^P$, and any deviation would represent a needless transfer of effort away from the pursuit of personal enrichment. This disincentive to constituency effort applies to all incumbent candidates from $P_d^+$, but only to the top $S_d$ incumbents candidates on $P_d^-$'s electoral list, where $S_d$ denotes the number of "safe seats" won by $P_d^-$ at the full-shirking vector $\mathbf{F_d^o}$. What about the decision facing an incumbent from $P_d^-$ at list position $(S_d + 1)$, i.e., the incumbent with the highest list position not to receive one of the party's safe seats, and who is thus "next in line" for re-election? This *marginal candidate* may have a unilateral incentive to defect from $\mathbf{F_d^o}$, if by doing so he or she can push party $P_d^-$'s district-level vote share $\mathbf{v_d^{P-}}$ high enough to secure $P_d^-$ the additional seat needed for his or her re-election. Define $\hat{f}_{m,d}^{P-}$ as the *critical effort level* of the marginal candidate must exert in her region so as to secure re-election, and $f_{m,d}^{P-*}$ as the marginal candidate's equilibrium choice. As well, define $\mathbf{F_d^*} = \{f_{1,d}^*, f_{2,d}^*, ..., f_{M_d,d}^*\}$ as a *Nash Equilibrium* strategy vector of constituency effort allocations for incumbents from district $d$. I now present the following Theorem, proven in Appendix 2:

**Theorem 1:** *If incumbents occupy higher list positions than non-incumbents, then the unique district-level Nash Equilibrium to the above-specified CLPR game is characterized by the following strategy vector:*

$$\mathbf{F_d^*} = \{0, 0, ..., f_{m,d}^{P-*}, ..., 0, 0\},$$

*where $f_{m,d}^{P-*} \in \{\hat{f}_{m,d}^{P-}, 0\}$ depends on the marginal candidate's utility for re-election $\beta_{j,m}^{P-}$, her effort capacity $E^{P-}$, and percentage of party loyalists in her region $\ell_{m,d}^{P-}$.*

---

[11] For example, consider a district $d$ in which parties $A$ and $B$ hold $\bar{A}_d = 7$ and $\bar{B}_d = 3$ current seats. If an election is held in which $A$ and $B$ win $\mathbf{x_d^A} = 4$ and $\mathbf{x_d^B} = 6$ seats, then all of $B$'s incumbents (and 3 of its non-incumbents) secure re-election, while only 4 of $A$'s incumbents are re-elected.

Put simply, in equilibrium *at most one* of the district's incumbent legislators ever devotes *any* effort to securing the interests of regional constituents; and in many cases the CLPR Nash Equilibrium is the full-shirking vector itself (i.e., $\mathbf{F_d^*} = \mathbf{F_d^o}$).[12] This is due to the fact that incumbents with favorable list positions can *free-ride* on the regional vote-seeking efforts of fellow incumbents with lower list-positions. The proof in Appendix 2 demonstrates that, at any vector other than $\mathbf{F_d^*}$ in Theorem 1, either incumbents high on the list will defect so as to free-ride on their co-partisans' mobilizing efforts; or incumbents low on the list will defect to avoid having their efforts appropriated by those with higher list positions.[13]

# 4 Legislative Equilibrium Under Open-List Proportional Systems

In OLPR systems voters must simultaneously express support for a political party *and* for a particular candidate from within that political party's list. To capture this mechanism in our present formal context, $V_{j,d}^P(\cdot)$ will therefore represent not only the percentage of voters from region $j$ who contribute to party $P$'s district-level total, but also the percentage of regional voters who cast individual *candidate votes* in support of the region's incumbent. Similarly, votes cast against the regional incumbent's party will, by definition, serve as candidate votes for some candidate on the opposing party's list. Recall that both parties field a full slate of $M_d$ candidates on their electoral lists in district $d$, a slate which includes both legislative incumbents and non-incumbent candidates. In keeping with the analysis above, we will assume that in each region whose incumbent is from party $A$ ($B$), there exists a non-incumbent from party $B$ ($A$) who amasses the candidate votes of dissatisfied voters (**Assumption 4**). For example, in a region whose incumbent is from party $A$, $V_{j,d}^A(\cdot)$ will represent the percentage of candidate votes received by the regional incumbent; and $[1 - V_{j,d}^A(\cdot)]$ will represent the percentage of candidate votes accrued by one of party $B$'s non-incumbent candidates.

If party $P$ wins some number $\mathbf{x_d^P}$ seats in district $d$ as a result of the game's election, these seats go to the $\mathbf{x_d^P}$ candidates with the *highest candidate vote scores*.

---

[12]Kselman (2010) generalizes this result to *any* situation in which candidates' list positions are independent of incumbents' effort allocations decisions: regardless of the relative placement of incumbent and non-incumbent candidates, in a game with exogenously determined list positions *at most one of a district's incumbents* devotes positive effort to constituency service.

[13]In keeping with the tenor of recent research on intra-party dynamics in specific CLPR systems (e.g. Szwarcberg 2009 on Argentina), Kselman (2010) analyzes a distinct list formation mechanism, assuming that the incumbent from party $P$ who devotes the highest level of effort to $f_{j,d}^P$ receives the party's highest list position, the incumbent from party $P$ who devotes the second-highest level of effort to $f_{j,d}^P$ receives the party's second-highest list position, and so on. The results are discussed in the following Section.

I will assume that candidate vote ties between 2 or more incumbent candidates in a district $d$ are decided randomly and without bias (**Assumption 5**); and that candidate vote ties between incumbent and non-incumbent candidates are decided in favor of incumbents (**Assumption 6**). The latter assumption is purely expository, and the game is generalizable to situations in which ties between incumbents and non-incumbents are also broken randomly. I now derive Nash Equilibrium outcomes for the following district-level OLPR game:

1. In a first stage all incumbents from $d$ simultaneously choose $f_{j,d}^P$ and $c_{j,d}^P$.
2. A second-stage election is held in which incumbent candidates receive a candidate vote share of $V_{j,d}^P(\cdot)$ in their respective regions, and parties $A$ and $B$ receive district-level vote shears $\mathbf{v_d^A}$ and $\mathbf{v_d^B}$ respectively.
3. Parties $A$ and $B$ are allocated $\mathbf{x_d^A}$ and $\mathbf{x_d^B}$ seats respectively via the quota-remainder rule presented above.
4. And these seats go to the $\mathbf{x_d^P}$ candidates with $P$'s highest candidate vote totals.

This section and Appendix 3 investigate a stylized game in which the number of party loyalists $\ell_d \in [0,1]$ is identical across all regions in a particular electoral district $d$ (**Assumption 7**); and in which both $\beta_{j,d}^P \geq 1$ for all incumbents and $E^P \geq 1$ for both parties (**Assumptions 8** and **9**). Kselman (2010) extends the model to situations in which $\ell_{j,d}^P$ varies across both regions and parties, in which $\beta_{j,d}^P < 1$ for some or all incumbents, and in which $E^P < 1$ for one or both parties.

A district's partisanship is influential in defining its incumbents' Nash Equilibrium effort allocations. Proposition 2 in Appendix 3 shows that, when the percentage of loyalists $\ell_d$ is especially high, incumbent legislators can count on enough support from their party's loyal partisans to eliminate the need for constituency service, and the full-shirking vector $\mathbf{F_d^o}$ is the district-level Nash Equilibrium ($\mathbf{F_d^*} = \mathbf{F_d^o}$). However, at lower levels of party loyalty, incumbents desirous of re-election must devote effort to $f_{j,d}^P$ so as to *offset the candidate vote totals of non-incumbent candidates*. For example, consider a district $d$ in which parties $A$ and $B$ have $\bar{A}_d = 7$ and $\bar{B}_d = 3$ current incumbents respectively, and let district $d$'s loyalty level be $\ell_d = \frac{1}{4}$. If $\mathbf{F_d^o}$ is played, party $A$'s district-level vote share is equal to $\mathbf{v_d^A} = [\frac{1}{4} \cdot (7) + \frac{3}{4} \cdot (3)]/10 = 40\%$, and party $B$'s vote share is $\mathbf{v_d^B} = 60\%$. By the quota-remainder rule $A$ and $B$ thus each receive $\mathbf{x_d^A} = 4$ and $\mathbf{x_d^B} = 6$ seats.

Furthermore, these legislative seats are allocated almost exclusively to non-incumbent rather than incumbent candidates. To see this, note that at $\mathbf{F_d^o}$ all incumbents choose $f_{j,d}^P = 0$ and receive only $V_{j,d}^P(\cdot) = \frac{1}{4}$ candidate votes, while all non-incumbents receive $[1 - V_{j,d}^P(\cdot)] = \frac{3}{4}$ candidate votes. Since non-incumbents receive more candidate votes than incumbents, in the game's final stage party $B$'s 6 seats will be allocated to 6 of its 7 non-incumbent candidates, and party $A$'s 4 seats will be allocated to its 3 non-incumbents and 1 of its incumbent candidates. Given the above assumption that ties between 2 or more incumbents with identical candidate vote totals are randomly decided without bias, when $\mathbf{F_d^o}$ is played, each of $A$'s 7 incumbents has a probability $\pi_{j,d}^A = \frac{1}{7}$ of gaining the single seat allocated to an incumbent candidate.

The full-shirking vector $\mathbf{F_d^0}$ will thus not be a Nash Equilibrium. For example, any one of $A$'s 7 incumbents could then choose $f_{j,d}^A = \varepsilon$ ($\varepsilon \to 0$), increase her candidate vote total to just above that received by her fellow incumbents, and gain this individual seat with certainty (i.e., make $\pi_{j,d}^A(\cdot) = 1$). In turn, another of $A$'s incumbents could choose $f_{j,d}^A = \varepsilon'$ ($\varepsilon' > \varepsilon$) and gain the seat with certainty. But then a 3$^{rd}$ incumbent could do the same, and so on. Thus, party $A$'s 7 incumbents will jockey among themselves over the single legislative seat not allocated to $A$'s non-incumbent candidates. In addition to increasing one's own candidate vote total, this jockeying has two important strategic effects: (a) it *increases* $A$'s district vote share $\mathbf{v_j^A}$; and (b) it *decreases* the number of preference votes $[1 - V_{j,d}^A(\cdot)]$ received by the party $B$'s non-incumbent candidate in the same region. Thus, while an incumbent's quest for candidate votes emerges for purely competitive reasons, it also has certain 'positive externalities' for fellow incumbents from both parties.

I refer the interested reader to Appendix 3, which derives these dynamics' equilibrium consequences for any status quo incumbency pattern and any partisanship level. So as to present intuitively the Nash Equilibrium properties of OLPR competition, Fig. 1 uses the illustrative example of a district in which parties $A$ and $B$ have $\bar{A}_d = 7$ and $\bar{B}_d = 3$ current incumbents.

Figure 1's x-axis plots values of $\ell_d \in [0, 1]$ in descending order from left to right. The explicitly marked values of $\ell_d$ represent points at which the OLPR game's equilibrium properties change. As already noted, at particularly high values of party loyalty ($\ell_d \geq \frac{7}{8}$) the OLPR Nash Equilibrium is simply the full-shirking vector ($\mathbf{F_d^*} = \mathbf{F_d^0}$): at this outcome, parties $A$ and $B$ win $\mathbf{x_d^A} = 7$ and $\mathbf{x_d^B} = 3$ seats respectively by the quota remainder rule, and these seats are allocated to the parties' respective incumbent candidates, since their candidate vote totals outpace those of non-incumbent candidates.

At intermediate values of party loyalty ($\frac{7}{8} > \ell_d > \frac{7}{20}$), in equilibrium *all* minority party incumbents (i.e., those from party $B$) continue to choose $f_{j,d}^{B*} = 0$, but *all* majority party incumbents (i.e., those from party $A$) choose $f_{j,d}^{A*} = \hat{f}_d^A$, whose value
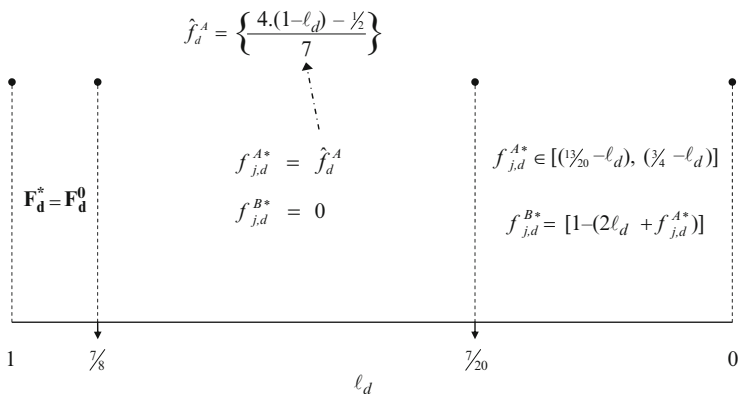


$$\hat{f}_d^A = \left\{ \frac{4 \cdot (1 - \ell_d) - \frac{1}{2}}{7} \right\}$$

$\mathbf{F_d^*} = \mathbf{F_d^0}$

$f_{j,d}^{A*} = \hat{f}_d^A$

$f_{j,d}^{B*} = 0$

$f_{j,d}^{A*} \in [(\frac{13}{20} - \ell_d), (\frac{3}{4} - \ell_d)]$

$f_{j,d}^{B*} = [1 - (2\ell_d + f_{j,d}^{A*})]$

1          $\frac{7}{8}$                                                    $\frac{7}{20}$                    0

$\ell_d$

**Fig. 1** OLPR Nash equilibria

is defined in the figure itself; and once again all incumbents from both parties are re-elected.[14] At lower levels of party loyalty ($\ell_d < \frac{7}{20}$) the Nash Equilibrium to the OLPR game is no longer unique, although the possible equilibrium outcomes occupy a narrowly defined range of incumbents' action spaces: majority party incumbents may choose $f_{j,d}^{A*}$ from the range defined in the figure itself, and minority party incumbents choose a corresponding value $f_{j,d}^{B*} = [1 - (2\ell_d + f_{j,d}^{A*})]$. I label these outcomes *Mutually-Assured Re-election Nash Equilibria*. In these equilibria, the parties once again win $\mathbf{x_d^A} = 7$ and $\mathbf{x_d^B} = 3$ seats; and incumbents' candidate vote shares *perfectly balance* the candidate vote shares of incumbents from the opposing party (i.e., $V_{j,d}^A = 1 - V_{j,d}^B$), such that all incumbents receive exactly as many candidate votes as their party's non-incumbent candidates, and are thus all re-elected. These particular equilibrium properties are a direct result of the pseudo-cooperative dynamics noted above.[15]

## 5 Institutional Comparative Statics

The theoretical results in the preceding sections apply to individual electoral districts. I will employ a two-stage process to specify empirical hypotheses about the aggregate prevalence of constituency service ($f_{j,d}^P$) and legislative corruption ($c_{j,d}^P$) generated by a particular electoral system across an *entire Legislature*. First, I employ Proposition 1, Theorem 1, and Theorem 2 to derive equilibrium outcomes in all of a system's individual electoral districts, which are a function of both the electoral system in place and status quo partisanship levels. Second, I aggregate these district-level effort allocations across all incumbents in the Legislature. Denote electoral institutions with the marker $I \in \{FPTP, CLPR, OLPR\}$, and define *total constituency service* as:

$$T^*(I, \ell) = \sum_{P \in \{A,B\}} f_{j,d}^{P*}. \tag{5}$$

---

[14]The effort level $\hat{f}_d^A$ represents the level of effort which, when chosen by all incumbents from $A$, pushes this party's vote share *just high enough* to win back $\mathbf{x_d^A} = 7$ seats. Put otherwise, when all incumbents from $A$ choose $\hat{f}_d^A$, they *split evenly* the cost, in terms of constituency effort, of re-electing the entire party.

[15]The result that, regardless of the district's level of part loyalty, all incumbents are re-elected in the OLPR game is an artifact of **Assumptions 8** and **9**, that both $\beta_{j,d}^P \geq 1$ for all incumbents and $E^P \geq 1$ for both parties. When either incumbents' utility for re-election $\beta_{j,d}^P$ or the effort capacity of *majority party* incumbents (in this case $E^A$) drop below a certain threshold, OLPR Nash Equilibria retain parallel properties, but some subset of incumbents will no longer gain re-election in equilibrium. As well, at unusually low values of $\beta_{j,d}^P$ the OLPR game may under certain circumstances generate instability. Section V addresses the consequences of this instability for the model's comparative static implications.

This is simply the sum of all constituency effort exerted, in equilibrium, by incumbents from both parties (for a given institution and partisanship SQ). The higher the level of $T^*(\cdot)$, the larger the ratio of constituency service to legislative corruption generated in a particular legislative context. The studies cited in Sect. 1 above (Persson and Tabellini 2003; Kunicova and Rose-Ackerman 2005) argue that FPTP elections should generate stronger ties of legislative accountability and lower levels of political corruption than *both* OLPR and CLPR systems; and that OLPR systems should occupy an intermediate position between FPTP systems and CLPR systems. The statistic $T^*(\cdot)$ allows us to assess this hypothesis, i.e., to identify the aggregate ratio of constituency service to legislative corruption which arises under alternative electoral institutions.

Although I will generate $T^*(\cdot)$ in a variety of exogenous contexts, to make the analysis tractable I will impose a series of parametric constraints, none of which restricts the results' generality. Firstly, the following computations assume that levels of party loyalty are identical across all of a country's $J = N$ regions; define this uniform level of partisanship as $\ell \in [0, 1]$. They also assume that $\beta_{j,d}^P \geq 1$ for all incumbents and that $E^P \geq 1$ for both parties.[16] Finally, in situations where OLPR Nash Equilibria are not unique, I analyze (without loss of generality) the equilibrium whose district-level constituency effort represents the *mean* of all possible equilibria (see Appendix 4). Consider a generic Legislature of $N = 200$ seats. In FPTP systems the 200 incumbents represent single-member districts. In PR systems, begin with a case in which these 200 seats are divided into $D = 20$ distinct regions, each with a magnitude of $M_d = 10$, with an incumbency status quo in which party $A$ has a slim 103-to-97 legislative majority. Appendix 4 contains a more detailed description of the parties' district-by-district incumbency status; and presents in detail the two-step process by which $T^*(\cdot)$ is calculated. Figure 2 plots values of $T^*(\cdot)$ for all three institutions at all possible values of $\ell \in [0, 1]$, which move in descending order from left to right on the figure's x-axis.

The explicitly marked values of $\ell$ represent key points of inflection on at least one of the $T^*(\cdot)$ plots. The first thing to note is that, aside from situations of unusually high party loyalty ($\ell > \frac{11}{12}$), *OLPR always generates higher levels of constituency service and lower levels of legislative corruption service than both FPTP and CLPR*. Put otherwise, the hypothesis which motivated past research on the relationship between electoral rules and corruption does not obtain in this theoretical context: OLPR is not an intermediate system, but rather outperforms both CLPR and FPTP in constraining legislative graft.

Secondly, note that the relationship between $T^*(CLPR)$ and $T^*(FPTP)$ varies according to levels of party loyalty. For $\ell > \frac{1}{2}$, CLPR may at times generate slightly

---

[16]As noted in ftn 15, at unusually low values of $\beta_{j,d}^P$ the OLPR game may under certain circumstances have no NE. At these extremely low values of $\beta_{j,d}^P$, neither FPTP nor CLPR systems generate *any* constituency service (i.e. at these low re-election utilities $T^*(FPTP) = T^*(CLPR) = 0$). As such, at these very low values of $\beta_{j,d}^P$, the fact that OLPR competition does not generate a stable outcome in fact makes it *more* constituency oriented then either FPTP or CLPR systems, which generate stable outcomes characterized by the categorical absence of constituency service.
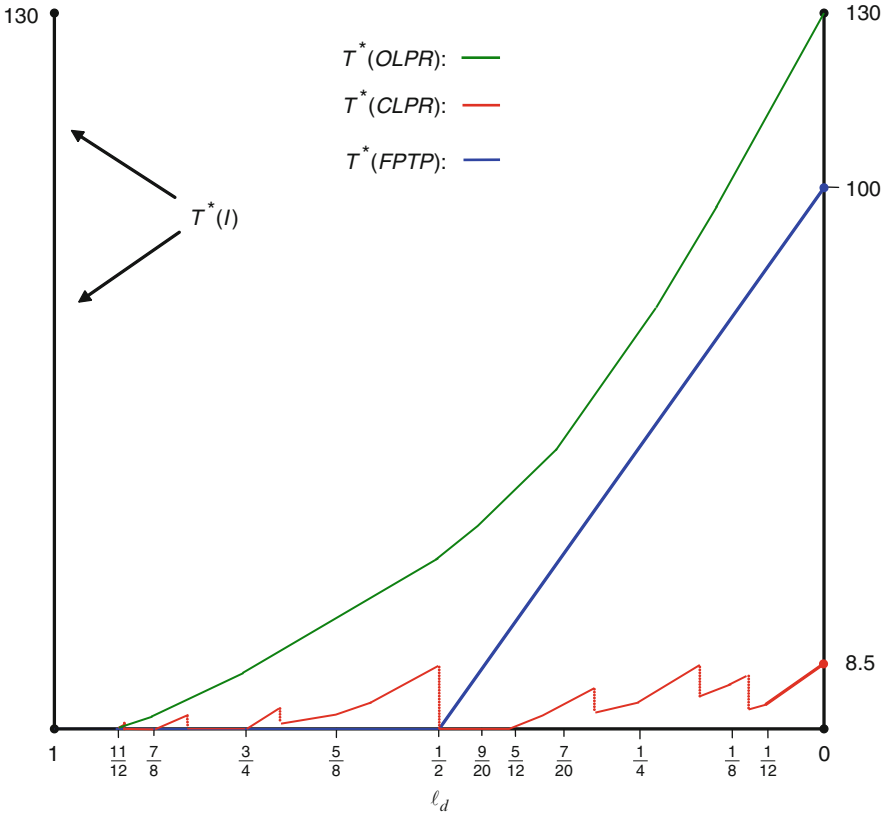
**Fig. 2** Aggregate constituency effort for FPTP and PR with 20 districts of magnitude 10

higher levels of constituency service than FPTP, although neither institution generates much constituency service to speak of. Once loyalty levels move below $\ell < \frac{1}{2}$, FPTP very quickly outpaces CLPR in generating particularistic effort. As such, and in contrast to the fairly uniform condemnation of CLPR as a suboptimal institution in the above-reviewed research, here we see that CLPR performs little differently from FPTP systems in electorates with high levels of partisan of partisan stability. The distinction between CLPR and FPTP systems only emerges at lower levels of partisan bias, i.e., contexts in which voter choice is more "elastic" to constituency effort.

Do these same relationships hold in distinct exogenous circumstances? Figure 3 again examines a Legislature of size $N = 200$, and in which party $A$ again has a slim legislative majority, but this time assumes that PR systems are composed of 40 districts of size $M_d = 3$ and 40 districts of size $M_d = 2$, such that the average district magnitude is significantly lower than that of the previous simulation. District-by-district incumbency details and derivations are again contained in Appendix 4.
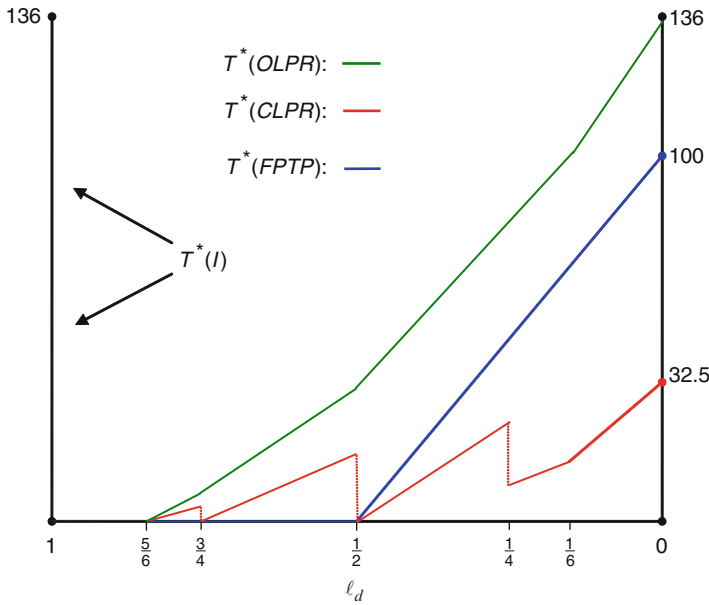
**Fig. 3** Aggregate constituency effort for FPTP and PR with 20 districts of magnitude 2 and 20 districts of magnitude 3

Once again, and in contrast to previous theoretical sentiments, except at the very highest levels of partisanship OLPR outpaces both CLPR and FPTP in generating constituency service and constraining political corruption. As well, smaller districts have the effect of *amplifying* the distinction between CLPR and FPTP when $\ell > \frac{1}{2}$ and *dulling* this distinction when $\ell < \frac{1}{2}$. At higher levels of party loyalty CLPR now *significantly* outperforms FPTP in generating constituency service, while at lower levels CLPR no longer lags as far behind FPTP as in the previous simulation.[17]

As demonstrated in Appendix 4, these institutional comparative statics do not depend on the legislative status quo: regardless of the district-by-district incumbency breakdown, OLPR outperforms both FPTP and CLPR in constraining legislative graft and generating constituency service at all but the highest levels of electoral partisanship (in which case $T^*(I) = 0$ for all three institutions). Furthermore, and again regardless the district-by-district incumbency breakdown, at higher (lower) levels of partisanship CLPR (FPTP) generates greater aggregate constituency

---

[17]I've analyzed a distinct mechanism by which list order is determined in CLPR systems, assuming that the incumbent from party $P$ who devotes the highest level of effort to $f_{j,d}^P$ receives the party's highest list position, the incumbent from party $P$ who devotes the second-highest level of effort to $f_{j,d}^P$ receives the party's second-highest list position, and so on (Kselman 2010). In this scenario, OLPR once again outperforms its counterparts all but the highest levels of partisanship. As well, under this distinct list formation assumption the circumstances in which CLPR outperforms FPTP are significantly expanded.

service than FPTP (CLPR). To summarize, the model's implications for the relationship between electoral institutions and legislative graft are quite distinct form those in the above-reviewed pieces. The following section subjects these distinct arguments to empirical analysis.

## 6   Electoral Formula, Ballot Structure, Political Corruption

The primary dependent variable employed in Persson and Tabellini (2003), Persson et al. (2003), and Kunicova and Rose-Ackerman (2005) is the *Control of Corruption Index*, which itself is one of six *Governance Indicators* compiled by World Bank researchers over the past 15 years.[18] Evidence from all three studies comes from cross-national investigation of data from 1997–1998, facilitating both statistical replication and the parsimonious evaluation of competing hypotheses. Persson and Tabellini (2003) and Persson et al. (2003) contain largely identical empirical analyses; henceforth I confine myself to the former so as to avoid redundancy. The authors use three institutional measures to study the relationship between a country's electoral formula, its ballot structure, and political corruption: MAJ, PIND, and PINDO. MAJ captures whether or not a country uses some form of *plurality rule* to elect its legislators, such that $MAJ = 1$ in plurality rule systems while $MAJ = 0$ in PR systems: far and away the most common form of plurality rule is FPTP in single-member districts. PIND captures the percentage of a country's legislators who are elected as individual candidates *independent of party lists*, such that $PIND = 1$ in FPTP systems with single-member districts and $PIND = 0$ in pure PR systems and other party list systems (PIND may assume values between 0 and 1 in countries which use a mix of party lists and direct candidate voting).[19] PINDO is a variant of PIND which accounts for the fact that legislators in OLPR systems occupy party lists, but are also the recipients of individually targeted candidate votes. It captures the percentage of legislators in a particular country which *are not* elected using *closed-party lists*, such that [PINDO $= 1$] for both pure FPTP and pure OLPR systems, while [PINDO $= 0$] for pure CLPR systems.

I begin by replicating the results found in Table 7.1 on pages 192–193 of Persson and Tabellini (2003). The dependent variable is GRAFT, a transformed measure of the World Bank's corruption index for which higher values indicate a greater presence of corruption, with a mean of 4.14 and standard deviation of 1.89 (ranging from a low of 0.74 in Denmark to a max of 6.92 in Paraguay). The statistical model is *weighted-least-squares*, where all regressions are weighted by the *inverse*

---

[18]This oft-used index aggregates into a single measure information from over 30 distinct public opinion and professional surveys which ask respondents for their subjective evaluations of a particular country's experience with political corruption. Treisman (2007) contains a detailed account of the strengths and weaknesses of this and other data sources on political corruption.

[19]As the authors readily admit, the variables MAJ and PIND are highly, though not perfectly, correlated ($r = 0.926$) due to the fact that by far the most common plurality rule system is FPTP in single-member districts, where legislators are *by definition* not elected on party lists.

*standard deviation* of the surveys which enter into the original index, to control for the fact that some countries generate higher levels of subjective uncertainty than others. Figure 4 contains the measurement specifics for GRAFT, MAJ, PIND, PINDO, and all control variables employed in the forthcoming analysis.

Control variables come directly from the publically available data set used in Persson and Tabellini's original analysis. For reasons of space I opt not to discuss results regarding these controls, and focus instead on the institutional measures of primary interest.

Columns 1 and 2 of Table 1 replicate the findings which motivate Persson and Tabellini's most basic empirical conclusions.[20] The results in column 1 come from a regression which includes both MAJ and PIND. Both coefficients are negative, which conforms to the authors' theoretical expectations: the direct legislative accountability associated with plurality rule should *reduce* corruption, while the muted accountability associated with party list competition should *increase* corruption. However only PIND attains statistical significance, due probably to the two measures' multi-colinearity ($r = 0.926$; see ftn 11). Column 2 presents results from a regression with replaces PIND with PINDO, whose correlation with MAJ is lower ($r = 0.680$); once again both coefficients are negative, but in this case only MAJ attains statistical significance, i.e., embedding open-list considerations into the variable PIND dilutes its statistical effect, and makes MAJ the most robust predictor. These results motivate Persson andTabellini's conclusion, quoted above, that plurality rule systems without party lists outperform *both* OLPR and CLPR in constraining corruption.

The first thing to note about these regressions is that they both contain the variable MAGN, which captures a country's *inverse district magnitude*, i.e., its number of electoral districts divided by its total number of legislative seats, such that [MAGN = 1] in pure single-member district systems and [MAGN < 1] in systems with at least one multi-member district. Not surprisingly, this variable is itself highly correlated with both MAJ ($r = 0.886$) and PIND ($r = 0.928$). Thus, the regression in column 1 contains three institutional measures correlated with one another at roughly $r = 0.9$, which makes the substantive interpretation of statistical coefficients a challenge. Column 3 contains the results of a regression identical to column 1 save for the exclusion of MAGN; without the inclusion of this highly multicolinear variable neither PIND nor MAJ attains statistical significance, and the sign on the former becomes positive.[21]

---

[20]Despite repeated attempts using all possible combinations of relevant control variables, I was unable to generate the exact weighted-least squares coefficients presented 7.1 of Persson and Tabellini (2003). As such, I settled on the set of control variables used by the authors themselves. The results presented here are nearly identical in terms of both substantive size and statistical significance to the original results.

[21]The nearly non-existent correlations between GRAFT and both PIND ($r = 0.038$) and MAJ ($r = 0.059$) lend support to the suspicion that these variables' statistical significance in columns 1 and 2 is largely a result of the institutional measures' multi-colinearity.

* All of the data in the following table comes directly from the publicly available dataset which accompanies Persson and Tabellini (2003), and which is available at http://people.su.se/~tpers/. All of the individual variable coding descriptions come directly from the data Appendix in Persson and Tabellini (2003).

-GRAFT: point estimate of 'Graft', the sixth cluster of Kaufman et al.'s *Governance Indicators* focusing on perceptions of corruption, with a possible range of 0-to-10, where lower values correspond to better outcomes.

-PIND: continuous measure of ballot structure, defined as [ 1 – (List Seats/Total Seats)], where the second term represents the percentage of legislators elected on party lists divided by the total number of seats in the Legislature. As such, PIND measures the percentage of legislators elected independent of party lists.

-PINDO: continuous measure of ballot structure, defined as [ 1 – (Closed List Seats/Total Seats)], where the second term represents the percentage of legislators elected on *closed party lists* divided by the total number of seats in the Legislature.

-MAJ: dummy variable which equals 1 for country's whose lower house is elected by plurality rule, and equals 0 otherwise.

-MAGN: inverse district magnitude, defined as the number of electoral districts inside a particular country divided by the number of seats in the country's Legislature.

-PRES: dummy variable equal to 1 in presidential regimes and 0 otherwise. Only regimes where the confidence of the Assembly is not necessary for the executive are excluded among presidential regimes.

-FEDERAL: dummy variable equal to 1 if the country has a federal political structure and 0 otherwise.

-GASTIL: average of indices for civil liberties and political rights, where each index is measured on a one-to-seven scale with one representing the highest degree of freedom and 7 the lowest.

-AGE: age of democracy, defined as (2000 – DEM_AGE)/200 and varying between 0 and 1, where the US is the world's oldest democracy with a value of AGE = 1. DEM_AGE is coded as the first year of democratic rule, corresponding to the first year of an uninterrupted string of positive yearly values on the variable POLITY until the end of the sample.

-COL_UK: dummy variable equal to 1 if the country was a former British colony and 0 otherwise.

-PROT80: percentage of the population in each country professing the Protestant religion in 1980.

-CATHO80: percentage of the population belonging to the Roman Catholic Church in 1980.

-CONFU: dummy variable equal to 1 if the majority of a country's population is Confucian/Buddhist/Zen, and equal to 0 otherwise.

-AVELF: index of ethno-linguistic fractionalization, approximating the level of ethnic and linguistic fragmentation within a country, ranging from 0 (homogeneous) to 1 (strongly fractionalized) an comprising an average of five different indices.

-LPOP: natural log of total population.

-EDUGER: total enrollment in primary and secondary education, as a percentage of the relevant age group in the population.

-LYP: natural log of per capita GDP, where real GDP is defined as per capita GDP in constant dollars expressed in international prices (base year 1985).

-TRADE: sum of exports and imports of goods and services measured as a share of GDP.

-LAAM: regional dummy variable equal to 1 if the country is in Latin America, Central America, or the Caribbean, and equal to 0 otherwise.

-OECD: dummy variable equal to 1 for all countries which were members of the OECD before 1993, and 0 for all other countries (except for Turkey, which is assigned a value of 0 despite having been a member nation prior to 1993).

**Fig. 4** Data from Persson and Tabellini (2003)

The second noteworthy aspect of this analysis is that the measures employed to capture a country's electoral formula and ballot structure group together systems with very different strategic properties. For example, the variable MAJ groups

**Table 1** Weighted-least-squares analysis of GRAFT ($N = 78$)

| | | | | | | |
|---|---|---|---|---|---|---|
| PIND | −1.783** (0.894) | | 0.601 (0.633) | 0.096 (0.655) | −0.442 (0.419) | −0.719 (0.467) |
| PINDO | −0.563 (0.592) | −0.385 (0.279) | | | | |
| MAJ | −1.208** (0.456) | −1.208** (0.456) | −0.857 (0.639) | −0.473 (0.643) | | |
| MAGN | 2.762*** (0.787) | 1.679*** (0.526) | | | | |
| OLPR | | | −0.508* (0.298) | −0.597** (0.271) | −0.638** (0.264) | −1.208** (0.506) |
| FPTP | | | | | 0.140 (0.412) | −0.153 (0.466) |
| CLPR | | | | | | −0.656 (0.498) |
| PRES | −0.699*** (0.278) | −0.608*** (0.284) | −0.508* (0.298)... | −0.408 (0.292) | −0.398 (0.294) | −0.397 (0.292) |
| FEDERAL | 0.185 (0.302) | 0.179 (0.308) | 0.189 (0.330) | 0.278 (0.322) | 0.340 (0.315) | 0.252 (0.321) |
| GASTIL | 0.076 (0.164) | 0.083 (0.167) | 0.107 (0.179) | 0.038 (0.176) | 0.028 (0.177) | 0.059 (0.178) |
| AGE | −0.210 (0.628) | −0.029 (0.629) | −0.008 (0.682) | −0.183 (0.666) | −0.292 (0.664) | −0.277 (0.660) |
| COL_UK | −0.776*** (0.285) | −0.591** (0.283) | −0.472 (0.296) | −0.567* (0.290) | −0.631** (0.279) | −0.701** (0.282) |
| PROT80 | −0.010** (0.005) | −0.009* (0.005) | −0.007 (0.005) | −0.007 (0.005) | −0.007 (0.005) | −0.007 (0.005) |
| CATHO80 | 0.003 (0.004) | 0.004 (0.004) | 0.004 (0.004) | 0.003 (0.004) | 0.003 (0.004) | 0.003 (0.004) |
| CONFU | 1.417*** (0.522) | 0.994** (0.463) | 0.559 (0.504) | 0.662 (0.491) | 0.795 (0.478) | 0.693 (0.482) |
| AVELF | 1.012* (0.564) | 0.661 (0.559) | 0.419 (0.587) | 0.383 (0.569) | 0.378 (0.573) | 0.450 (0.572) |
| LPOP | 0.026 (0.114) | 0.014 (0.116) | 0.043 (0.135) | 0.036 (0.121) | 0.023 (0.122) | 0.030 (0.121) |
| EDUGER | −0.012 (0.008) | −0.011 (0.008) | −0.001 (0.009) | −0.011 (0.008) | −0.011 (0.008) | −0.012 (0.008) |
| LYP | −0.859*** (0.234) | −0.958*** (0.231) | −1.026*** (0.250) | −1.016*** (0.290) | −0.984*** (0.242) | −0.930*** (0.244) |
| TRADE | −0.004 (0.003) | −0.004 (0.003) | −0.004 (0.003) | −0.004 (0.003) | −0.005 (0.003) | −0.005 (0.003) |
| OECD | −1.155*** (0.418) | −1.124*** (0.425) | −1.039** (0.455) | −1.150** (0.444) | −1.186*** (0.442) | −1.156** (0.440) |
| LAAM | 0.976*** (0.396) | 0.765** (0.416) | 0.748** (0.426) | 0.597 (0.419) | 0.555 (0.416) | 0.680 (0.424) |
| CONS | 12.56*** | 13.44*** | 14.04*** | 14.44*** | 14.35*** | 14.47*** |

Standard errors in parentheses; *($p > 0.10$); **($p > 0.05$); ***($p > 0.01$)

together FPTP systems with the Alternative Vote used in Australia and the Bloc Vote used in Mauritius and Thailand.[22] As well, the variable PIND regroups countries which use CLPR, OLPR, and a variety of hybrid list systems including aforementioned Bloc Vote and the Single-Transferable-Vote used in Ireland and Malta. Finally, the variable PINDO regroups FPTP systems and OLPR systems. The theoretical results derived in Sects. 2–5 demonstrate the potential hazards of this systemic conflation. I have thus created the variables FPTP, OLPR, CLPR, and HYBRID, each of which measures the percentage of a country's legislators elected under the relevant system (HYBRID groups together systems such as the Alternative Vote, the Bloc Vote, and the Single-Transferable Vote).[23] Appendix 4 presents all countries' individual values on these variables, along with coding rules used in their creation. In most cases a country's value on these variables is either 0 or 1; the few countries which used mixed systems have fractional values on two of these measures. For the latter cases, define a country's *predominant system* as the system used to elect a majority of its legislators.[24] In the 84 country dataset, 30 countries are completely or predominantly FPTP, 33 are CLPR, 14 are OLPR, and 6 are HYBRID. As a starting point consider Fig. 5, which presents the mean values of GRAFT among all countries of a particular predominant system.

This simple mean comparison defies quite strikingly the aforementioned conventional wisdom: mean levels of corruption are much lower in OLPR systems than in any other system, and are lower than those found in FPTP systems by nearly $\frac{3}{4}$ of a standard deviation on the GRAFT scale (3.17 as opposed to 4.56). In fact, FPTP systems, lauded in previous studies, register a higher mean level of corruption than *any* category!

Of course, simple mean comparisons are often misleading. Columns 4, 5, and 6 from Table 1 introduce the measures OLPR, FPTP, and CLPR into a more rigorous statistical setting. As my goal is a comparative evaluation of hypotheses pertaining to electoral formulae and ballot structure, I exclude MAGN in these regressions to

---

[22]As well, Chile is coded as a plurality rule system despite the fact that it has, since 1988, used the equivalent of OLPR in two-member districts (with d'Hondt as the effective formula) to elect its legislatures. Similarly, South Korea is coded as *non-plurality-rule* despite the fact that it has been a predominantly FPTP system since its transition to democracy. South Korea elects 245 legislators in single-member FPTP districts, and another 54 in a national-level upper-tier. However, the seat allocations in this small upper-tier are based on one's success in the FPTP districts (i.e. the more votes one gets in FPTP districts the more seats one receives in the upper-tier), such that parties' electoral calculations are driven nearly completely by single-member-district calculations.

[23]In keeping with the dependant variable's time point (1997–1998), these variables represent the system used in a particular country during the years 1994–1997.

[24]In almost all cases the coding of a country's predominant system is straight-forward. For example, Poland and Switzerland register scores of 0.85 (0.15) and 0.975 (0.025) respectively on the variable OLPR (FPTP), such that the predominant rule is clearly OLPR although both countries contain a small number of single-member-districts. Russia is the unique case in which the system is perfectly divided (both CLPR = 0.5 and FPTP = 0.5), and is thus not included in Figure 1's mean tally.
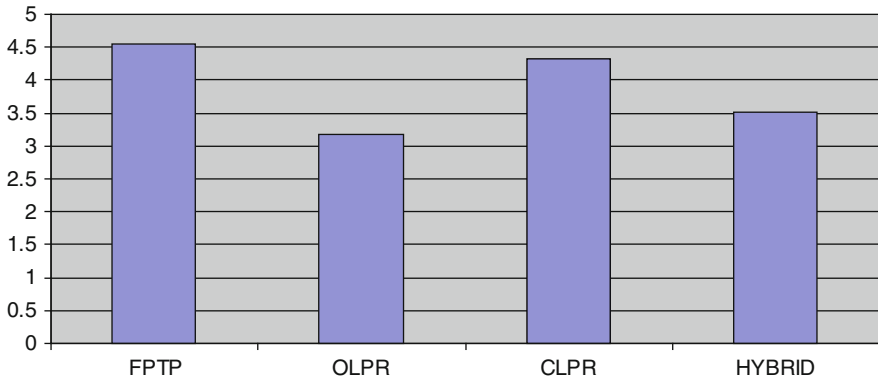
**Fig. 5** Mean comparison of graft

mitigate problems of multi-colinearity.[25] When introduced into a weighted-least-squares regression along with PIND and MAJ (column 4), only OLPR has a significant and reductive effect on a country's overall level of political corruption. Column 5 replaces the variable MAJ with the variable FPTP. Once again, neither FPTP nor PIND has a statistically significant effect on GRAFT, while OLPR has a significant reductive effect. Finally, column 6 introduces CLPR into the mix, with a similar qualitative result: only OLPR has a significant reductive effect on corruption. In this last column, this effect becomes substantively stronger, such that moving from a system where OLPR = 0 to a system where OLPR = 1 reduces corruption by two-thirds of a standard-deviation in the GRAFT measure.

These results also differ substantially from those uncovered in Kunicova and Rose-Ackerman (2005). Although for reasons of space I do not conduct a full replication of their analysis, a number of things can be said about these differences. Firstly, I reclassify a number of cases, coding for example both Chile and Poland as OLPR (the authors code them as plurality-rule and CLPR respectively), and coding both Hungary and Guatemala as CLPR systems (rather than plurality-rule systems). However, perhaps the most important difference between their analysis and that presented here is the sample size: their sample contains a non-negligible number of countries which are excluded from Persson and Tabellini's analysis due to their lack of democratic credentials (e.g. Jordan, Kazakhstan, Kuwait, Sierra Leone, Zimbabwe, Yemen, etc.). Thus, beyond the measurement issues noted above, behind our contradictory results lies an unresolved question, which exceeds my current scope, as to the consequences of electoral institutions in semi-democracies and/or non-democracies. To summarize, these empirical findings cast doubt on *both* the notion that FPTP systems generate lower levels corruption than PR systems, *and*

---

[25]I have run all regressions with both MAGN and PINDO included. PINDO has no effect on any of the following results. The inclusion of MAGN does not reduce the statistical or substantive significance of OLPR; but does return the highly multicolinear PIND to its previous statistical significance (see above).

the notion that OLPR occupies an intermediate position between high accountability FPTP systems and low accountability CLPR systems. In contrast, OLPR outperforms both FPTP and CLPR in restraining political corruption, while the latter two systems are statistically indistinguishable.

## 7 Concluding Discussion

This paper's model suggests that OLPR will, under almost all circumstances, generate greater legislative accountability and lower levels of political corruption than its CLPR and FPTP counterparts, while the latter two systems' relative performance depends on an electorate's partisanship and the size of electoral districts, as well as the particular candidate nomination procedures used in CLPR systems (see ftn 13 and ftn 17). In OLPR systems, the constituency service allocations of individual legislators have positive externalities for fellow incumbents from both parties. In turn, the set of incumbents frequently finds themselves in Mutually-Assured Re-election Nash Equilibrium, in which effort allocations by members of one party perfectly balance those from legislators of competing parties. It is these equilibrium dynamics which allow OLPR to generate greater levels of legislative accountability and lower levels of political corruption CLPR and FPTP systems. The empirical results presented in Sect. 6 provide suggestive support for this basic comparative static prediction.

In addition to calling into question a growing consensus as to the consequences of formal electoral institutions, this paper speaks to the frequent conflation, among both policy and academic circles, of political particularism and corruption. Indeed, critics of clientelism and other forms of targeted public policy often use the two terms interchangeably. As well, it is often suggested that particularism is a precondition for legislative corruption, i.e., that legislators' opportunities for personal material enrichment are particularly strong when public policy is highly targeted. However, a growing body of recent research offers a more nuanced normative and empirical appraisal of particularistic forms of accountability. Keefer and Vlaicu (2008) argue that the targeted public policies often improve aggregate social welfare when politicians cannot credibly commit to the provision of public goods. Fernandez and Pierskalla (2009) find that countries with high levels of political particularism in fact *outperform* their counterparts on select dimensions of economic and human development (e.g. infant mortality and literacy). More generally, an ambitious project by Herbert Kitschelt and collaborators (Freeze et al. 2008) aimed at gathering data on alternative forms of political accountability in a sample of 90 contemporary democracies takes as a starting point the distinction between political particularism and corruption and/or other pernicious governance practices. The current paper shares with this research the undercurrent that at times particularistic accountability may serve as a "second-best" policy alternative when the exogenous environment is not conducive to more normatively palatable forms governance and accountability.

## Appendix 1: Regional Vote Shares and Open Sets

### *Regional Vote Shares*

I develop a model of voter choice in which voters are influenced by *both* their regional incumbent's first-stage effort allocation and their own "partisan" biases for or against parties $A$ and $B$, biases determined by considerations *independent* of their regional incumbent's effort allocation.[26] In the paper, Sects. 2 through 5 treat these partisan attitudes as exogenous and ask the following question: *given* some distribution of partisan attitudes in the electorate, what are the equilibrium levels of $f_{j,d}^P$ and $c_{j,d}^P$ provided by parties' current legislative incumbents? Section 7 addresses situations in which partisan attitudes emerge endogenously.

I follow Persson and Tabellini (2000) in modeling partisanship with a single parameter capturing voter $i$ in region $j$'s *partisan attitude*. Label this partisan attitude $\sigma_{i,j}$, and let higher values of $\sigma_{i,j}$ correspond to more favorable attitudes for party $A$, and lower values to more favorable attitudes for party $B$. Let partisan preferences in region $j$ vary according to a *uniform distribution* over the support set $[\underline{\sigma}_j, \bar{\sigma}_j]$[27] where $\bar{\sigma}_j$ represents the attitude of the voter in region $j$ who is most inclined to choose party $A$, and $\underline{\sigma}_j$ that of the voter in region $j$ most inclined to party $B$. Without loss of generality assume that $-1 < \underline{\sigma}_j < 0 < \bar{\sigma}_j < 1$ and that $\bar{\sigma}_j - \underline{\sigma}_j = 1$, where the latter implies that both the "height" and the "width" of the uniform distribution in region $j$ are equal to 1. For example, a region whose partisan attitudes are distributed over the support set $[-\frac{1}{10}, \frac{9}{10}]$ is heavily "biased" towards party $A$; a region with partisan support set $[-\frac{9}{10}, \frac{1}{10}]$ is heavily "biased" towards party $B$; and a region with partisan support set $[-\frac{1}{2}, \frac{1}{2}]$ is "neutral" (Fig. 6).

I will model voter choice as the decision to "accept" or "reject" *the party of their regional incumbent*, i.e., the party of the legislator who may provide their region particularistic goods and services.[28] Given this approach, I must specify distinct utility functions for voters depending on the party affiliation of the relevant regional incumbent:

$$u_{i,j}^P(f_{j,d}^P) = \begin{cases} f_{j,d}^A + \sigma_{i,j} & \text{if} \quad P = A \\ f_{j,d}^B - \sigma_{i,j} & \text{if} \quad P = B \end{cases}. \tag{6}$$

---

[26]Most obvious among such considerations are voter preferences for parties' respective national-level policy platforms. Also potentially relevant are voters' symbolic and affective "identification" with one party or another, grounded for example in family history, the party's ideological/ historical legacy, etc.

[27]Though the model is robust to alternative distributional assumptions, the straight-forward calculus of uniform distributions greatly simplifies the analysis.

[28]One might object that in PR systems voter choice between competing parties often has little to do with legislative particularism; as will become evident below, in many systems this very pattern emerges in equilibrium, i.e. voter choice is grounded completely in their partisan attitudes $\sigma_{i,j}$.
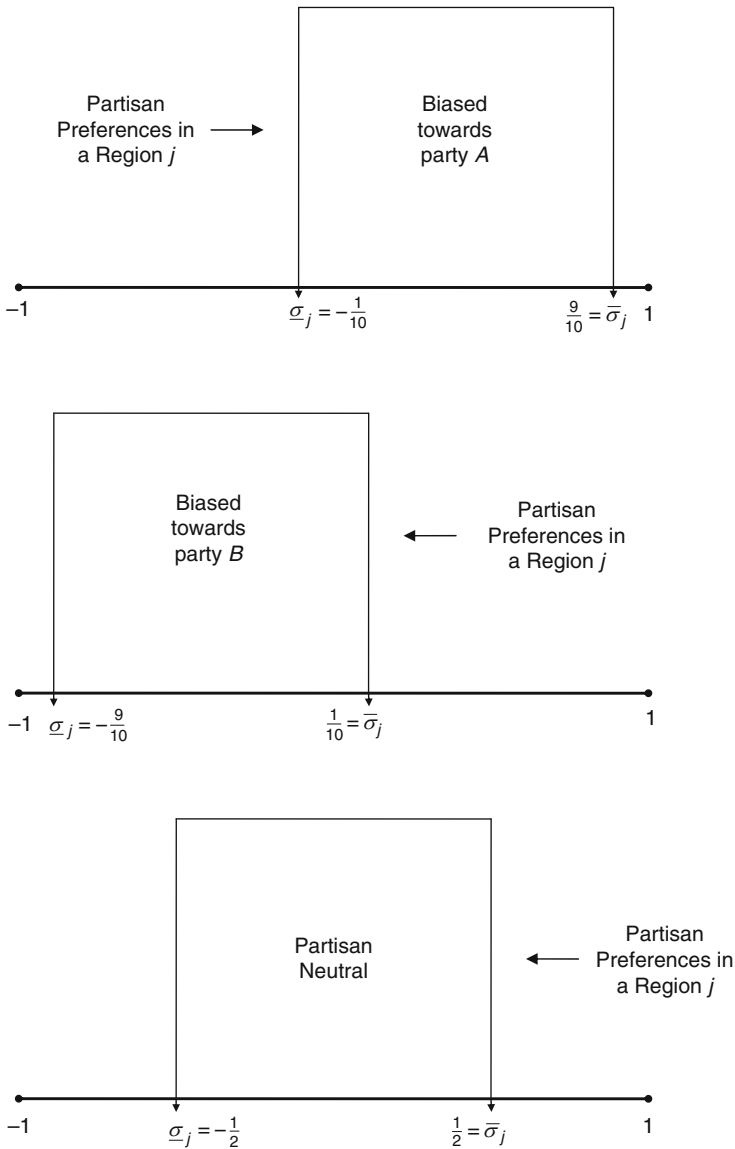
**Fig. 6** Regional distributions of partisan attitudes

Voter utility thus changes linearly with both the level of goods/services targeted to his or her region and with his or her partisan attitudes. Recall that, by construction, higher values of $\sigma_{i,j}$ correspond to more favorable attitudes for party $A$. For this reason $\sigma_{i,j}$ enters (6) in an additive manner for regions where the regional incumbent is affiliated with party $A$, such that the utility is higher among voters with higher values of $\sigma_{i,j}$. Since by construction lower values of $\sigma_{i,j}$ correspond to more favorable

attitudes for $B$, in regions whose legislative incumbent is from party $B$ partisan attitudes enter as a subtracted rather than an additive term: utility for $B$ is higher among voters with lower values of $\sigma_{i,j}$; and voters for whom $\sigma_{i,j} < 0$ will have "positive" utility for $B$.

Voters must thus ask themselves whether or not the combined satisfaction derived from an incumbent's first-stage constituency service effort $f_{j,d}^P$ and their own partisan attitude $\sigma_{i,j}$ is sufficient to vote for this incumbent's party in the game's second-stage election. The notion of a *reservation utility* provides a useful mechanism for modeling the process by which voters make this assessment. Define the reservation utility $\eta$ as the satisfaction level at which voters feel sufficiently pleased with the party of their region's incumbent legislator to choose that party in the game's election. As such, in the game's electoral stage, voters in regions whose incumbent is from party $A$ ($B$) will vote for this party if $u_{i,j}^A \geq \eta$ ($u_{i,j}^B \geq \eta$). Without loss of generality I will normalize the game's reservation utility to $\eta = 0$.

Define $\ell_{j,d}^P$ as the percentage of *party loyalists* in region $j$: the percentage of regional voters whose utility for the party of their regional incumbent surpasses the reservation level $\eta = 0$ even if $f_{j,d}^P = 0$. Given our distributional assumptions, it is straightforward to see that, in regions whose incumbent is affiliated with party $A$, the loyalist percentage is simply equal to $\ell_{j,d}^A = \bar{\sigma}_{j,d}^A$; and in those whose incumbent is affiliated with party $B$ it is equal to $\ell_{j,d}^B = -\underline{\sigma}_{j,d}^B$ (trivial proof omitted).

For a region $j$ whose incumbent is from party $P$, given some effort allocation $f_{j,d}^P$ in the game's first stage we can derive $V_{j,d}^P(f_{j,d}^P)$, party $P$'s regional *vote percentage* in the game's subsequent electoral stage (i.e., the portion of $j$'s voters for whom $u_{i,j}^P \geq 0$). I now prove the following Lemma:

**Lemma 1.** *for a region $j$ whose incumbent is from party P, given some effort allocation $f_{j,d}^P$ in the game's first stage $V_{j,d}^P(f_{j,d}^P)$ can be expressed as follows:*

$$V_{j,d}^P(f_{j,d}^P) = \begin{cases} (f_{j,d}^P + \ell_{j,d}^P) & \text{if} \quad f_{j,d}^P < 1 - \ell_{j,d}^P \\ 1 & \text{if} \quad f_{j,d}^P \geq 1 - \ell_{j,d}^P \end{cases}$$

**Proof of Lemma 1:**
Given some allocation $f_{j,d}^P$, define $\sigma_{s,j}(f_{j,d}^P)$ as the partisan attitude of region $j$'s *swing voter*, i.e., the voter whose utility *exactly* reaches the reservation level $\eta = 0$:

$$u_{s,j}^P(f_{j,d}^P) = 0 = \begin{cases} f_{j,d}^A + \sigma_{s,j}(f_{j,d}^A) & \text{if} \quad P = A \\ f_{j,d}^B - \sigma_{s,j}(f_{j,d}^B) & \text{if} \quad P = B \end{cases}. \tag{7}$$

Given some allocation $f_{j,d}^A$ in regions whose incumbent is from party $A$, voters with values of $\sigma_{i,j}$ *greater* than $\sigma_{s,j}(f_{j,d}^A)$ will vote for party $A$, and those with values of $\sigma_{i,j}$ *less* than $\sigma_{s,j}(f_{j,d}^A)$ will vote for party $B$. In contrast, given some allocation $f_{j,d}^B$ in regions whose incumbent is from party $B$, voters with values of $\sigma_{i,j}$ *less* than $\sigma_{s,j}(f_{j,d}^B)$ will vote for party $B$, and those with values of $\sigma_{i,j}$ *greater* than $\sigma_{s,j}(f_{j,d}^A)$ will vote for party $A$. Since partisan preferences in region $j$ are uniformly distributed, the

percentage of region $j$'s voters that choose the regional incumbent's party can be expressed as follows:

$$V_{j,d}^P(f_{j,d}^P) = \begin{cases} \dfrac{\bar{\sigma}_j - \sigma_{s,j}(f_{j,d}^A)}{\bar{\sigma}_j - \underline{\sigma}_j} & \text{if} \quad P = A \\[3mm] \dfrac{\sigma_{s,j}(f_{j,d}^B) - \underline{\sigma}_j}{\bar{\sigma}_j - \underline{\sigma}_j} & \text{if} \quad P = B \end{cases}. \tag{8}$$

Figures 7a, b display these vote shares visually for regions whose incumbents are from $A$ and $B$ respectively.

Since $\bar{\sigma}_j - \underline{\sigma}_j = 1$ by construction, substituting into (8) using (7) yields the expressions $V_{j,d}^A(f_{j,d}^A) = f_{j,d}^A + \bar{\sigma}_j$ and $V_{j,d}^B(f_{j,d}^B) = f_{j,d}^B - \underline{\sigma}_j$. These expressions can both be represented as $V_{j,d}^P(f_{j,d}^P) = f_{j,d}^P + \ell_{j,d}^P$. Finally, since $V_{j,d}^P(f_{j,d}^P)$ cannot be greater than 1, any value of $f_{j,d}^P$ such that (8) is greater than 1 implies a vote share of 1 (i.e., devoting more effort than $f_{j,d}^P = 1 - \ell_{j,d}^P$ is unnecessary to secure 100% voter support in region $j$), thus establishing Lemma 1. ∎
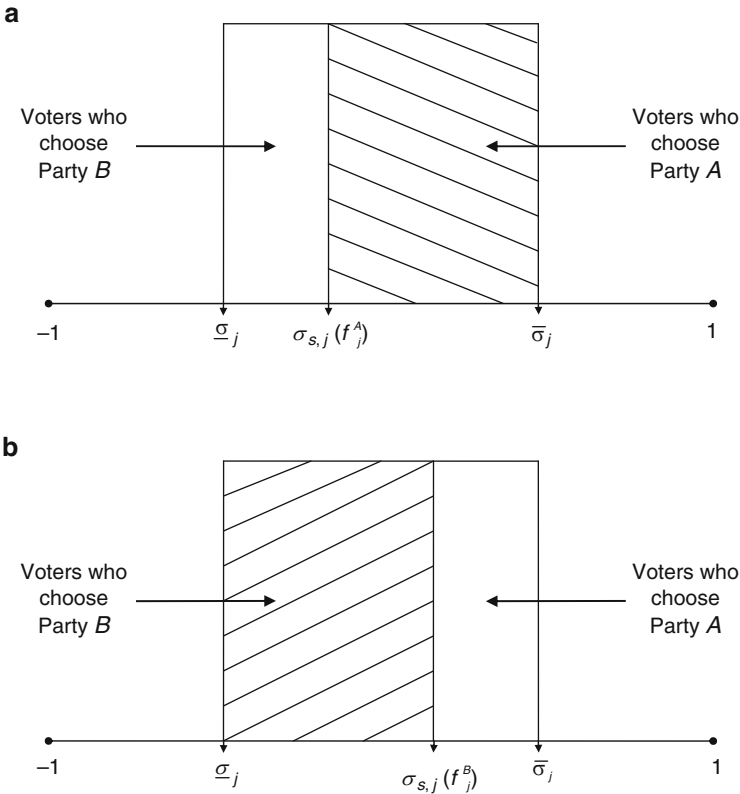


Fig. 7 Regional vote shares of incumbents' parties

## The Open-Set Problem

The term $\varepsilon \to 0$ appears in my definition of $\hat{f}_j^P$ on page 9 of the text, where $\hat{f}_j^P$ is defined as the effort level the incumbent from region $j$ must choose to secure the electoral support of just over half of her region's voters. However, in the strictest sense $\hat{f}_j^P$ does not in fact exist, since $\varepsilon$ can always be made infinitesimally closer to 0, i.e., we have what game theorists generally label an *open-set problem*. As is often noted, this problem is purely technical, and can be eliminated by assuming that incumbents' action spaces are composed of measurable but minute effort increments.

This technicality aside, in the following Appendices I will continue to assume that incumbents' action spaces are continuous so as to avoid added numerical complexity, a common approach in game theoretic analyses. All results are generalizable to situations with non-continuous action spaces composed of measurable but minute effort increments.

## Appendix 2: CLPR Nash Equilibria

I now derive NE results of the game whose sequential structure is outlined in page 12 of the text, in which *incumbents are assumed to be placed higher on electoral lists than their parties' non-incumbents* (**Assumption 3**). Recall from the text that, when Assumption 3 is employed, at least one of the two parties has all of its district-level incumbents re-elected in the game's electoral stage. More specifically, two basic electoral outcome types can emerge when Assumption 3 is employed:

1. *Type 1 outcomes*: in one party *all* incumbent candidates and *at least one* non-incumbent candidate are re-elected; while in the opposing party *not all* incumbents and *zero* non-incumbents are re-elected.[29]
2. *Type 2 outcomes*: all incumbent candidates from *both* parties are re-elected. This occurs if parties $A$ and $B$ win $\mathbf{x_d^A} = \bar{A}_d$ and $\mathbf{x_d^B} = \bar{B}_d$ seats respectively by the quota-remainder rule.

I first prove Theorem 1 from the text, which applies when the electoral outcome associated with the strategy vector $\mathbf{F_d^o}$ is a Type 1 outcome, i.e., in districts where one party (denoted $P_d^+$) has all of its incumbents re-elected if $\mathbf{F_d^o}$ is played, but the other (denoted $P_d^-$) does not; I then extend the result to situations in which the outcome associated with the strategy vector $\mathbf{F_d^o}$ is a Type 2 outcome. Before proceeding consider the following Lemma, which is of use in both Appendices B and C.

---

[29]For example, consider a district $d$ in which parties $A$ and $B$ hold $\bar{A}_d = 7$ and $\bar{B}_d = 3$ current seats. If an election is held in which $A$ and $B$ win $\mathbf{x_d^A} = 4$ and $\mathbf{x_d^B} = 6$ seats, then (given Assumption 3) all 4 of $B$'s incumbents and 3 of its non-incumbents secure re-election, while only 4 of $A$'s incumbents and none of its non-incumbents are re-elected.

**Lemma 2.** *Any strategy vector* $\mathbf{F_d}$ *in which at least one incumbent sets* $f_{j,d}^P>0$ *but does not gain reelection is not a NE; and no incumbent will ever deviate from* $f_{j,d}^P = 0$ *if this deviation does not result in re-election.*

*Proof:* if $f_{j,d}^P>0$ but the incumbent in question does not secure reelection, then she will always prefer deviating and choosing $f_{j,d}^P = 0$, since $U_{j,d}^P(f_{j,d}^P) = (1 - f_{j,d}^P)$ is less than $U_{j,d}^P(0) = 1$. For the same reason, deviating from $f_{j,d}^P = 0$ without winning re-election is strictly-dominated by keeping one's choice at $f_{j,d}^P = 0$.                    ∎

## Theorem 1: Proof of Existence

I begin the Existence proof by identifying the marginal candidate's incentives to deviate from the strategy vector $\mathbf{F_d^o}$. Define $\ell_{m,d}^{P-}$ as the number of party loyalists in the marginal candidate's region. The marginal candidate needs her party's aggregate district vote share to reach the following level in order to gain re-election (at this level her party's electoral remainder just outpaces that of the opposing party $P_d^+$, thus securing $P_d^-$ a total of $\mathbf{x_d^{P-}} = (S_d + 1)$ legislative seats):

$$\hat{\mathbf{v}}_{\mathbf{d}}^{\mathbf{P}-}(S_d + 1) = (S_d/M_d) + (1/2M_d) + \varepsilon \quad (\varepsilon \to 0). \tag{9}$$

### 1. Case 1: Categorical Non-Deviation of the Marginal Candidate

Define $P_d^-$'s aggregate district vote share when the full-shirking vector $\mathbf{F_d^o}$ is played as $\mathbf{v_d^{P-}}(\mathbf{F_d^o})$. If $\{\mathbf{v_d^{P-}}(\mathbf{F_d^o}) + [(1 - \ell_{m,d}^{P-})/M_d]\}< \hat{\mathbf{v}}_{\mathbf{d}}^{\mathbf{P}-}(S_d + 1)$, then even if the marginal candidate deviated from $\mathbf{F_d^o}$ so as to secure 100% voter support in her district, $P_d^-$ would still receive only $\mathbf{x_d^{P-}} = S_d$ seats, and the marginal candidate would not be re-elected. Put otherwise, there are not enough undecided voters in the marginal incumbent's region to secure $P_d^-$ an additional seat. In this situation, by Lemma 2 the marginal incumbent has no incentive to defect from $\mathbf{F_d^o}$.

### 2. Case 2: Potential Deviation by the Marginal Candidate

If $\{\mathbf{v_d^{P-}}(\mathbf{F_d^o}) + [(1 - \ell_{m,d}^{P-})/M_d]\}> \hat{\mathbf{v}}_{\mathbf{d}}^{\mathbf{P}-}(S_d + 1)$, then the marginal candidate may be able to deviate from $\mathbf{F_d^o}$ so as to secure her own re-election. Recall that $\hat{f}_{m,d}^{P-}$ represents the *critical* level of constituency service the marginal candidate must exert in order to push $P_d^-$'s vote total up to $\hat{\mathbf{v}}_{\mathbf{d}}^{\mathbf{P}-}(S_d + 1)$, i.e., the effort necessary to move $P_d^-$'s seat total from $\mathbf{x_d^{P-}} = S_d$ to $\mathbf{x_d^{P-}} = (S_d + 1)$.[30]

In this case, the marginal incumbent will have the incentive to deviate from $\mathbf{F_d^o}$ and choose $\hat{f}_{m,d}^{P-}$ as long as $\hat{f}_{m,d}^{P-}<\beta_{m,d}^{P-}, E^{P-}$, i.e., as long as both: (a) the payoff from

---

[30]This critical value does not, technically, exist due to the open-sent problem discussed in Appendix A. As discussed above, we could address this trivially by making incumbents' action sets non-continuous.

deviating to $\hat{f}^{P-}_{m,d}$ and gaining re-election, $U^{P-}_{m,d}(\hat{f}^{P-}_{m,d}) = [(1 - \hat{f}^{P-}_{m,d}) + \beta^{P-}_{m,d}]$, is higher than the payoff accrued from devoting all effort to the pursuit of personal wealth $U^{P-}_{m,d}(0) = 1$; and (b) the marginal candidate has sufficient effort capacity to secure the necessary votes. On the other hand, if *either* $\hat{f}^{P-}_{m,d} > \beta^{P-}_{m,d}$ or $\hat{f}^{P-}_{m,d} > E^{P-}$ she will not have the incentive to deviate from $\mathbf{F^o_d}$. Finally, if $\hat{f}^{P-}_{m,d} = \beta^{P-}_{m,d}$ but $\hat{f}^{P-}_{m,d} \geq E^{P-}$ then the marginal candidate will be indifferent between deviating to $\hat{f}^{P-}_{m,d}$ and remaining at $f^{P-}_{m,d} = 0$.

### 3. *Case 3: Potential Deviation by the Marginal Candidate (contd)*

If $\{\mathbf{v^P_d}(\mathbf{F^o_d}) + [(1 - \ell^{P-}_{m,d})/M_d]\} = \hat{\mathbf{v}}^{P-}_{\mathbf{d}}(S_d + 1)$, then the marginal candidate may be able to deviate from $\mathbf{F^o_d}$ so as to move her party into a "remainder tie" with the opposing party: both parties will have identical remainders of 5% if the marginal candidate secures 100% electoral support in her own region. For this particular case, we will thus redefine $\hat{f}^{P-}_{m,d} = (1 - \ell^{P-}_{m,d})$ as the critical level of constituency effort necessary to secure the marginal candidate re-election with $\pi^{P-}_{m,d} = \frac{1}{2}$.

In this case, the marginal incumbent will have the incentive to deviate from $\mathbf{F^o_d}$ and choose $\hat{f}^{P-}_{m,d} = (1 - \ell^{P-}_{m,d})$ as long as $(1 - \ell^{P-}_{m,d}) < \frac{1}{2} \cdot \beta^{P-}_{m,d}$, $E^{P-}$, i.e., as long as both: (a) the payoff accrued from deviating to $\hat{f}^{P-}_{m,d} = (1 - \ell^{P-}_{m,d})$ and gaining re-election with probability $\pi^{P-}_{m,d} = \frac{1}{2}$ is higher than the payoff accrued from devoting all effort to the pursuit of personal wealth $U^{P-}_{m,d}(0) = 1$; and (b) the marginal candidate has sufficient effort capacity to secure the necessary votes. On the other hand, if *either* $(1 - \ell^{P-}_{m,d}) > \frac{1}{2} \cdot \beta^{P-}_{m,d}$ or $(1 - \ell^{P-}_{m,d}) > E^{P-}$ she will not have the incentive to deviate from $\mathbf{F^o_d}$. Finally, if $(1 - \ell^{P-}_{m,d}) = \frac{1}{2} \cdot \beta^{P-}_{m,d}$ and $(1 - \ell^{P-}_{m,d}) \leq E^{P-}$ then the marginal candidate will be indifferent between deviating and remaining at $f^{P-}_{m,d} = 0$.

The preceding analysis tells us that, when all other incumbent legislators choose $f^P_{j,d} = 0$, the marginal candidate will choose either $f^{P-}_{m,d} = 0$ or $f^{P-}_{m,d} = \hat{f}^{P-}_{m,d}$, depending on the game's exogenous circumstances. Based on this analysis, I now show that no incumbent has the incentive to deviate from the vector $\mathbf{F^*_d}$ defined in Theorem 1, in which $f^{P-*}_{m,d} \in \{0, \hat{f}^{P-}_{j,d}\}$.

None of the incumbents from $P^+_d$ would defect from $\mathbf{F^*_d}$, as they secure re-election without devoting any effort to constituency service.[31] Similarly, none of the top $S_d$ incumbents from $P^-_d$ has any incentive to defect from $\mathbf{F^o_d}$, as they also secure re-election without devoting any effort to constituency service.

The following expression captures the aggregate vote share $P^-_d$ needs to gain $\mathbf{x^{P-}_d} = (S_d + 2)$ legislative seats (at this level, her party's electoral remainder just

---

[31]Even if the marginal candidate chooses $\hat{f}^P_{m,d}$ and secures her party $P^-_d$ one seat more than it gains at the full-shirking vector $\mathbf{F^o_d}$, this seat will be taken from one of $P^+_d$'s *non-incumbent* candidates, since incumbents are placed above non-incumbents on their electoral list by **Assumption 3** (and since at least one non-incumbent from $P^+_d$ is elected when $\mathbf{F^o_d}$ is played).

outpaces that of the opposing party $P_d^+$, thus securing $P_d^-$ a total of $\mathbf{x_d^{P-}} = (S_d + 2)$ legislative seats):

$$\hat{\mathbf{v}}_{\mathbf{d}}^{\mathbf{P-}}(S_d + 2) = (S_d/M_d) + (3/2M_d) + \varepsilon \quad (\varepsilon \to 0). \tag{10}$$

Define $\ell_{S+2,d}^{P-}$ as the number of party loyalists in the region whose incumbent occupies position $(S_d + 2)$ on $P_d^-$'s electoral list. As well, define $\mathbf{v_d^{P-}}(\mathbf{F_d^*})$ as the aggregate district vote share received by $P_d^-$ when the NE strategy vector from Theorem 1 is played. In turn, regardless of the marginal candidate's choice $f_{m,d}^{P-*} \in \{0, \hat{f}_{m,d}^{P-}\}$, it is straight-forward to show that (algebra omitted) that $\{\mathbf{v_d^{P-}}(\mathbf{F_d^*}) + [(1 - \ell_{S+2,d}^{P-})/M_d]\} < \hat{\mathbf{v}}_{\mathbf{d}}^{\mathbf{P-}}(S_d + 2)$. In words, even if the candidate at list position $(S_d + 2)$ deviated from $\mathbf{F_d^*}$ in Theorem 1 so as to secure 100% voter support in her region, $P_d^-$ would still receive no more than $\mathbf{x_d^{P-}} = (S_d + 1)$ legislative seats, and the incumbent candidate at list position $(S_d + 2)$ would not be re-elected. As such, by Lemma 2 the candidate at list position $(S_d + 2)$ has no incentive to deviate from $\mathbf{F_d^*}$ in Theorem 1.[32]

This in turn implies that candidates from $P_d^-$ at list positions $(S_d + 3)$, $(S_d + 4)$, and so on cannot secure their own re-election, even if they receive 100% voter support in their respective regions. As such, by Lemma 2 no candidate from $P_d^-$ below the marginal list position ever has the incentive to deviate from $\mathbf{F_d^*}$.

We have now established that no non-marginal incumbent has the incentive to deviate from a vector $\mathbf{F_d^*}$ in which the marginal candidate chooses from $f_{m,d}^{P-*} \in \{0, \hat{f}_{m,d}^{P-}\}$ while other incumbents choose $f_{j,d}^{P*} = 0$. The analysis in cases $i$, $ii$, and $iii$ above demonstrates that, as long as all other incumbents choose $f_{j,d}^P = 0$, the marginal candidate's best response will come be $f_{m,d}^{P-*} \in \{0, \hat{f}_{m,d}^{P-}\}$. This establishes Existence: at the vector $\mathbf{F_d^*}$ all incumbents' are playing mutual-best responses.  ∎

---

[32]Why is this the case? The marginal candidate never has the incentive to push $P_d^-$'s vote share higher than **(B1)**, since this is all that is required for her re-election. As such, in order for the candidate at list position $(S_d + 2)$ to secure her party an additional seat at $\mathbf{F_d^*}$, she would need to secure her party the equivalent of an entire additional electoral quota. Given our distributional assumptions from the above model of voter choice, the marginal candidate's region will have some non-zero number of party loyalists (this construction is highly plausible: one can hardly imagine a region in which an incumbent's party would receive a vote share of 0, regardless of her behavior during the previous legislative term). Since we know that $\ell_{S+2,d}^{P-} > 0$, i.e. there is some non-zero number of party loyalists in the candidate at list position $(S_d + 2)$'s region, we also know that there will not be enough undecided voters in this candidate's region to secure an entire electoral quota. As such, by Lemma 2 this candidate will never deviate from $\mathbf{F_d^*}$.

### Theorem 1: Proof of Uniqueness

Consider a strategy vector $\mathbf{F_d}$ in which some number $K>1$ incumbents from $P_d^-$ choose $f_{j,d}^{P-}>0$. At any such vector either: (a) all $K$ gain re-election; or (b) at least one of the $K$ does not gain re-election. If (b), then at least one incumbent has the incentive to defect by Lemma 2. If (a), then only the incumbent from among these $K$ with the *lowest list position* might not have the incentive to defect. In contrast, all those incumbents from among the $K$ with higher list positions would be able to decrease $f_{j,d}^{P-}$ without losing re-election by *free-riding* on the regional vote share of the incumbent from among these $K$ with the lowest list position.

For example, consider a situation in which the candidates with list positions $(S_d + 1)$ and $(S_d + 2)$ both choose $f_{j,d}^{P-}>0$ and gain e-election, such that $\mathbf{x_d^{P-}} = (S_d + 2)$. In this case, the candidate with list position $(S_d + 1)$ could reduce her regional effort to $f_{m,d}^{P-} = 0$ without losing re-election, since she only needs $P_d^-$ to win $\mathbf{x_d^{P-}} = (S_d + 1)$ seats in order to be re-elected. The same is true anytime some number $K>1$ incumbents from $P_d^-$ choose $f_{j,d}^{P-}>0$ and all $K$ gain re-election. As such, no strategy vector in which $K>1$ incumbents from $P_d^-$ choose $f_{j,d}^{P-}>0$ can be a NE.

It is straight-forward to show (algebra omitted) that, if $K \geq 2$ incumbents from $P_d^-$ choose $f_{j,d}^{P-}>0$, then all incumbents from $P_d^+$ will be re-elected even if they choose $f_{j,d}^{P+} = 0$. In turn, since we've established that no strategy vector $\mathbf{F_d}$ in which $K>1$ incumbents from $P_d^-$ choose $f_{j,d}^{P-}>0$ can be a NE, it follows that no $\mathbf{F_d}$ at which *any* incumbents from $P_d^+$ choose $f_{j,d}^{P+} = 0$ will be a NE.

As such, in equilibrium at most one incumbent, an incumbent from $P_d^-$, ever chooses $f_{j,d}^{P-}>0$. This will never be an incumbent with a list position higher than $(S_d + 1)$ since they receive safe seats when all incumbents from $P_d^+$ choose $f_{j,d}^{P+} = 0$. As well, this will never be an incumbent with a list position lower than $(S_d + 1)$ since they would be choosing $f_{j,d}^{P-}>0$ without gaining re-election (see the above proof of Existence), which is ruled out by Lemma 2. As a result, $\mathbf{F_d^*}$ in Theorem 1 is the game's unique NE.                                    ∎

### NE When All Incumbents Are Re-elected at the Full-Shirking Vector

Theorem 1 applies to districts in which party $P_d^-$'s incumbents are not all re-elected when the full-shirking vector $\mathbf{F_d^o}$ is played (Type 2 electoral outcomes). However, as noted at the outset, given **Assumption 3**, it is also possible that incumbents from both parties in a particular district are all re-elected when $\mathbf{F_d^o}$ is played (Type 1 electoral outcomes). In this case, it is straightforward to show that the full-shirking vector itself is the CLPR game's unique district-level NE (i.e. $\mathbf{F_d^*} = \mathbf{F_d^o}$). The proof of Existence is trivial: since all incumbents secure re-election without devoting any

effort to constituency service, choosing $f_{j,d}^P > 0$ would represent a needless diversion of effort away from the pursuit of personal material gain. The proof of Uniqueness (available upon request) is largely identical to the Uniqueness proof presented in the Section "Theorem 1: Proof of Uniqueness" above, demonstrating first that no number $K > 1$ of either party's incumbents ever devote positive effort to constituency service; and in turn that, since at most 1 incumbent from either party ever devotes effort to constituency service, in equilibrium no incumbents devote any effort to constituency service.                                                                ∎

## Appendix 3: Nash Equilibria Under OLPR

This Appendix derives NE properties of the OLPR game laid out on page 22 of the text. In order to keep the Appendix to a reasonable length, the following material will not be included:

1. Proofs of *NE Uniqueness*. The Appendix outlines Uniqueness proofs' methodology, but does not present them in full. All Uniqueness proofs are available upon request.
2. Proofs of NE results in districts where $\bar{A}_d = \bar{B}_d$, i.e., in which both parties have the same number of current incumbents. The below results all pertain to districts in which one party has more current incumbents than the other. Strategically equivalent results for districts where $\bar{A}_d = \bar{B}_d$ are also available upon request.

As noted in the text, this Appendix assumes that $\ell_d$ is identical in all of district $d$'s regions (**Assumption 7**), and that $\beta_d, E^P \geq 1$ for all of district $d$'s incumbent legislators (**Assumptions 8** and **9**). Kselman (2010) generalizes the game to situations in which $\ell_{j,d}^P$ varies across regions, in which $\beta_{j,d}^P < 1$ for some or all incumbents, and in which $E^P < 1$ for one or both parties.

As well, throughout this Appendix I will assume that, if an incumbent candidate and a non-incumbent candidate receive identical candidate vote shares and only one of the two can win a seat, the seat will go to the incumbent candidate (**Assumption 6** from the text). This Assumption is purely expository, and serves to mitigate the open-set problem that arises when actor's have continuous action sets. The game can be generalized to a situation in which incumbent legislators action sets are made up of infinitesimal but finite effort increments, in which case **Assumption 6** would be unnecessary. Candidate vote ties between incumbent candidates will be decided randomly and without bias (**Assumption 5** from the text).

### OLPR in Districts with High Levels of Party Loyalty

For a district $d$ in which one party has more current incumbents than the other, define $\bar{P}_d^{MA}$ ($\bar{P}_d^{MI}$) as the number of seats currently held by the district's majority (minority) party (i.e., $\bar{P}_d^{MA} > \bar{P}_d^{MI}$).

**Proposition 2a:** *For districts in which $\bar{P}_d^{MA} > \bar{P}_d^{MI}$, if $\ell_d > \left\{ 1 - \frac{1}{2 \cdot (2\bar{P}_d^{MA} - M_d)} \right\}$ then the full-shirking vector $\mathbf{F_d^o} = \mathbf{F_d^*}$ is the OLPR game's unique district-level NE.*

*Proof of Existence*: When $\mathbf{F_d^o}$ is played the majority party receives aggregate district vote share:

$$\mathbf{v}_\mathbf{d}^\mathbf{MA}(\mathbf{F_d^o}) = \left\{ \frac{(\bar{P}_d^{MA} \cdot \ell_d) + [\bar{P}_d^{MI} \cdot (1 - \ell_d)]}{M_d} \right\}. \tag{11}$$

By the quota and remainder rule, the majority party needs the following aggregate district vote share to win $\mathbf{x}_\mathbf{d}^\mathbf{MA} = \bar{P}_d^{MA}$ seats:[33]

$$(\bar{P}_d^{MA}/M_d) - \left( \frac{1}{2} \cdot M_d \right) + \varepsilon \ (\varepsilon \to 0). \tag{12}$$

By setting **(11)** and **(12)** equal to one another and solving for $\ell_d$ we obtain the expression in Proposition 1. As long as $\ell_d$ surpasses this level, at $\mathbf{F_d^o}$ the majority and minority parties win $\mathbf{x}_\mathbf{d}^\mathbf{MA} = \bar{P}_d^{MA}$ and $\mathbf{x}_\mathbf{d}^\mathbf{MI} = \bar{P}_d^{MI}$ seats respectively via the quota remainder rule (algebra omitted), and these seats are allocated to the parties' incumbent candidates rather than their challengers, since at this level of party loyalty incumbents have higher candidate vote percentages than challenger candidates (i.e., $V_{j,d}^P(0) > \{1 - V_{j,d}^P(0)\}$). As such, no incumbent wishes to deviate (Existence). ∎

The proof of Uniqueness demonstrates first that no vector $\mathbf{F_d}$ at which either party wins more seats than it currently holds is a NE; and then that no vector $\mathbf{F_d}$ other than $\mathbf{F_d^*}$ from Proposition 2a at which both parties win back their current number of seats is a NE. ∎

**Proposition 2b:** *For districts in which $\bar{P}_d^{MA} > \bar{P}_d^{MI}$, if $\ell_d = \left\{ 1 - \frac{1}{2 \cdot (2\bar{P}_d^{MA} - M_d)} \right\}$ then the in the OLPR game's unique district-level NE one majority party incumbent chooses $f_{j,d}^{MA\,*} = \varepsilon \ (\varepsilon \to 0)$ and all the remaining incumbents choose $f_{j,d}^{P\,*} = 0$.*

*Proof of Existence*: If incumbents behave as stipulated in Proposition 2b, the majority and minority parties win $\mathbf{x}_\mathbf{d}^\mathbf{MA} = \bar{P}_d^{MA}$ and $\mathbf{x}_\mathbf{d}^\mathbf{MI} = \bar{P}_d^{MI}$ seats via the quota remainder rule (algebra omitted), and these seats are allocated to the parties' incumbent candidates rather than their challengers, since at this level of party loyalty incumbents have higher candidate vote percentages than challenger candidates (i.e., $V_{j,d}^P(0) > \{1 - V_{j,d}^P(0)\}$). Trivially, no incumbent who chooses $f_{j,d}^{P\,*} = 0$

---

[33]The above discussion of infinitesimal effort increments and the open-set problem is germane to all of this Appendix's derivations.

wishes to deviate, as they gain re-election without devoting any effort to constituency service.

If the majority party incumbent who chooses $f_{j,d}^{MA *} = \varepsilon$ $(\varepsilon \to 0)$ were to drop her constituency effort level to $f_{j,d}^{MA} = 0$, then the minority and majority parties would have identical remainder levels after their full set of quotas were subtracted from their aggregate district vote shares $\mathbf{v_d^P(F_d^o)}$, implying that the majority party would win $\mathbf{x_d^{MA}} = \bar{P}_d^{MA}$ seats with probability $\frac{1}{2}$ and $\mathbf{x_d^{MA}} = (\bar{P}_d^{MA} - 1)$ seats with probability $\frac{1}{2}$. If the latter were to occur, each majority party incumbent would have identical candidate vote shares, and would thus gain re-election with probability $\pi_{j,d}^{MA} = [(\bar{P}_d^{MA} - 1)/\bar{P}_d^{MA}]$. Thus, the majority incumbent who chooses $f_{j,d}^{MA *} = \varepsilon$ $(\varepsilon \to 0)$ would rather exert infinitesimal effort increment and receive a seat with certainty than jeopardize her chances at re-election (Existence).    ∎

The proof of Uniqueness demonstrates first that no vector $\mathbf{F_d}$ at which either party wins more seats than it currently holds is a NE; and then that no vector $\mathbf{F_d}$ other than $\mathbf{F_d^*}$ from Proposition 2b at which both parties win back their current number of seats is a NE.    ∎

## OLPR in Districts with Intermediate Levels of Party Loyalty

For districts in which $\bar{P}_d^{MA} > \bar{P}_d^{MI}$, I now identify the OLPR game's NE when district-level party loyalty is in the following range:

$$1 - \left\{ \frac{\bar{P}_d^{MA} - \frac{1}{2}}{M_d} \right\} \geq \ell_d < 1 - \left\{ \frac{1}{2 \cdot (2\bar{P}_d^{MA} - M_d)} \right\}. \tag{13}$$

Using the quota remainder rule it is straightforward to show that, when $\ell_d$ is in this range and $\mathbf{F_d^o}$ is played, the minority party will win some number $\mathbf{x_d^{MI}} > \bar{P}_d^{MI}$ seats and the majority party will win some number $\mathbf{x_d^{MA}} < \bar{P}_d^{MA}$ seats (algebra omitted). As such, by definition not all majority party incumbents will be re-elected when $\mathbf{F_d^o}$ is played. By subtracting (11) from (12) we obtain:

$$\hat{\mathbf{v}}_\mathbf{d}^\mathbf{MA}(\mathbf{F_d^o}) = \left\{ \frac{[(2\bar{P}_d^{MA} - M_d) \cdot (1 - \ell_d)] - \frac{1}{2}}{M_d} \right\} + \varepsilon \quad (\varepsilon \to 0). \tag{14}$$

Assuming all minority party incumbents continue to choose $f_{j,d}^{MI} = 0$, expression (14) represents the additional vote share needed by the majority party to secure itself $\mathbf{x_d^{MA}} = \bar{P}_d^{MA}$ via the quota-remainder rule. Define $\hat{f}_d^{MA}$ as the constituency effort level that majority party incumbents would choose if they each were to devote *identical* levels of effort to constituency service, such that their combined efforts were *just sufficient* to win them $\mathbf{x_d^{MA}} = \bar{P}_d^{MA}$ seats. Formally, $\hat{f}_d^{MA}$ is thus the constituency effort level at which $(\bar{P}_d^{MA} \cdot \hat{f}_d^{MA})/M_d = \hat{\mathbf{v}}_\mathbf{d}^\mathbf{MA}(\mathbf{F_d^o})$. Solving this for $\hat{f}_d^{MA}$ yields:

$$\hat{f}_d^{MA} = \left\{ \frac{[(2\bar{P}_d^{MA} - M_d) \cdot (1 - \ell_d)] - \frac{1}{2}}{\bar{P}_d^{MA}} \right\}. \tag{15}$$

Thus, if all majority party incumbents choose $\hat{f}_d^{MA}$ they *divide evenly* the costs, in terms of constituency effort, of just barely gaining $\mathbf{x_d^{MA}} = \bar{P}_d^{MA}$ seats.

---

**Proposition 3:** *If $\ell_d$ is in the range (13), then in the OLPR game's unique district-level NE one majority party incumbent chooses $f_{j,d}^{MA*} = \hat{f}_d^{MA} + \varepsilon$ ($\varepsilon \to 0$), all remaining majority party incumbents choose $f_{j,d}^{MA*} = \hat{f}_d^{MA}$, and all minority party incumbents choose $f_{j,d}^{MI*} = 0$.*

---

*Proof of Existence*: When incumbents behave as stipulated in Proposition 3, the majority and minority parties win $\mathbf{x_d^{MA}} = \bar{P}_d^{MA}$ and $\mathbf{x_d^{MI}} = \bar{P}_d^{MI}$ seats respectively via the quota remainder rule, and these seats are allocated to the parties' incumbent candidates rather than their non-incumbents, since incumbents have higher candidate vote percentages. Trivially, no minority party incumbent has the incentive to deviate, since they gain re-election without devoting any effort to constituency service. Furthermore, no majority party who chooses incumbent $f_{j,d}^{MA*} = \hat{f}_d^{MA}$ has the incentive to deviate[34]:

1. devoting $f_{j,d}^{MA} > \hat{f}_d^{MA}$ effort to constituency service is unnecessary for re-election;
2. when incumbents behave as stipulated in Proposition 3, the majority party just barely secures enough district-level votes to win $\mathbf{x_d^{MA}} = \bar{P}_d^{MA}$ legislative seats. As such, by dropping their constituency effort $f_{j,d}^{MA} < \hat{f}_d^{MA}$, a majority party incumbent would *either* drop their party's district-level vote share into a remainder tie with that of the minority party, or drop their seat share to $\mathbf{x_d^{MA}} = (\bar{P}_d^{MA} - 1)$. Furthermore, it would drop their own candidate vote share below that of the incumbents who choose $f_{j,d}^{MA*} = \hat{f}_d^{MA}$, implying that the deviating incumbent would no longer gain re-election with certainty. Since $\beta_{j,d}^P \geq 1$, there is no level $f_{j,d}^{MA} < \hat{f}_d^{MA}$ at which the increased utility from personal enrichment outweighs this opportunity cost.

Finally the majority party incumbent who chooses $f_{j,d}^{MA*} = \hat{f}_d^{MA} + \varepsilon$ has no incentive to deviate:

1. devoting $f_{j,d}^{MA} > \hat{f}_d^{MA}$ effort to constituency service is unnecessary for re-election;
2. when incumbents behave as stipulated in Proposition 3, the majority party just barely secures enough district-level votes to win $\mathbf{x_d^{MA}} = \bar{P}_d^{MA}$ legislative seats.

---

[34]Once again, the fact that no majority party incumbent who chooses $f_{j,d}^{MA*} = \hat{f}_d^{MA}$ requires the assumption that $\varepsilon$ is effectively 0, i.e. that no majority party incumbent who chooses $f_{j,d}^{MA*} = \hat{f}_d^{MA}$ could drop her constituency effort by a "lower" increment than $\varepsilon$. As with other "open-set" issues, it is straight-forward (but less theoretically parsimonious...) to eliminate this problem by assuming that incumbents' action spaces are not continuous, but in fact composed of discreet but infinitesimal effort increments.

As such, by dropping their constituency effort $f_{j,d}^{MA} < \hat{f}_d^{MA}$, a majority party incumbent would *either* drop their party's district-level vote share into a remainder tie with that of the minority party, or drop their seat share to $\mathbf{x_d^{MA}} = (\bar{P}_d^{MA} - 1)$, implying that the deviating incumbent would no longer gain re-election with certainty. Since $\beta_{j,d}^P \geq 1$, there is no level $f_{j,d}^{MA} < \hat{f}_d^{MA}$ at which the increased utility from personal enrichment outweighs this opportunity cost (Existence). ∎

The proof of Uniqueness demonstrates first that no vector $\mathbf{F_d}$ at which either party wins more seats than it currently holds is a NE; and then that no vector $\mathbf{F_d}$ other than $\mathbf{F_d^*}$ from Proposition 3 at which both parties win back their current number of seats is a NE. ∎

## OLPR in Districts with Low Levels of Party Loyalty

For districts in which $\bar{P}_d^{MA} > \bar{P}_d^{MI}$, I now identify the OLPR game's NE when district-level party loyalty is in the following range:

$$\ell_d < 1 - \left\{ \frac{\bar{P}_d^{MA} - \frac{1}{2}}{M_d} \right\}. \tag{16}$$

Up to this point, all NE to the OLPR game have been *Unique*. I now demonstrate that, when $\ell_d$ is in the range (**16**), the district-level OLPR game will have more than one possible NE. Nonetheless, these NE are confined to a narrow range of incumbents' action spaces, i.e., the model continues to generate precise and useful predictions as to the level of effort incumbent legislators will devote to constituency service in OLPR systems.

---

**Proposition 4: Mutually-Assured-Reelection in OLPR Districts**
*For districts in which $\bar{P}_d^{MA} > \bar{P}_d^{MI}$, if $\ell_d$ is in the range (**16**) then any district-level strategy vector $\mathbf{F_d}$ which satisfies the following two criteria must be a NE to the OLPR game, and any NE to the OLPR game must satisfy the following two criteria (i.e. the criteria are both Necessary and Sufficient for the Existence of NE):*

1. *all majority party incumbents choose an identical level of constituency effort $f_{j,d}^{MA *}$, and this level of constituency effort is in the range $(\underline{\ell}_d - \ell_d) < f_{j,d}^{MA *} \leq (\bar{\ell}_d - \ell_d)$;*
2. *all minority party incumbents choose an identical level of constituency effort $f_{j,d}^{MI *}$, and this level of constituency effort is equal to $f_{j,d}^{MI *} = [1 - (2\ell_d + f_{j,d}^{MA *})]$.*

---

*Proof of Sufficiency*: When incumbents behave as stipulated in Proposition 4, the majority and minority parties win $\mathbf{x_d^{MA}} = \bar{P}_d^{MA}$ and $\mathbf{x_d^{MI}} = \bar{P}_d^{MI}$ seats respectively via the quota remainder rule. Furthermore, the criteria $f_{j,d}^{MI *} = [1 - (\ell_d + f_{j,d}^{MA *})]$

implies that both majority party incumbents and minority party incumbents receive an *identical* number of candidate votes as their parties' respective non-incumbent candidates (algebra omitted). As a result, all incumbent candidates are re-elected.

To prove that criteria (a) and (b) are Sufficient for the existence of NE, I establish that no incumbent has the incentive to deviate from any vector $\mathbf{F_d}$ at which these criteria are satisfied:

1. devoting $f_{j,d}^P > f_{j,d}^{P}{}^*$ effort to constituency service is unnecessary for re-election;
2. at any strategy vector $\mathbf{F_d}$ which satisfies these criteria, were any incumbent to drop their constituency effort to a level lower than $f_{j,d}^P < f_{j,d}^{P}{}^*$, they would drop their candidate vote share to just below that of their party's non-incumbent challengers, implying that the deviating incumbent would no longer gain re-election. Since $\beta_{j,d}^P \geq 1$ then, there is no level of effort $f_{j,d}^P < f_{j,d}^{P}{}^*$ at which the increased utility from personal enrichment outweighs this opportunity cost (Existence). ∎

To prove that criteria (a) and (b) are Necessary conditions for the existence of NE, I must establish that any strategy vector $\mathbf{F_d}$ which does satisfy these criteria is not a NE. As with the above proofs of Uniqueness, this derivation is omitted for reasons of space but available upon request. The proof of Necessity first demonstrates that no vector $\mathbf{F_d}$ at which either party wins more seats than it currently holds is a NE; and then that no vector $\mathbf{F_d}$ at which the parties win back their current number of seats, but which does not satisfy Proposition 4's criteria, is a NE. ∎

# Appendix 4: Simulation Analyses

This Appendix begins by presenting the exogenous restrictions used in generating the simulation results plotted in Figs. 2 and 3 (Sect. 1), then moves to a step-by-step elaboration of the simulation process itself (Sect. 2), and finally demonstrates that the qualitative hypotheses uncovered in Figs. 2 and 3 are generalizable to any similar simulation analysis (Sect. 3).

## *Exogenous Restrictions*

All of these simulations employ the following assumptions: the country's Legislature has $N = 200$ seats; levels of party loyalty are uniform across all of a country's $N = 200$ seats regions (define this uniform level of loyalty as $\ell \in [0, 1]$); and both $\beta_{j,d}^P \geq 1$ for all incumbents and $E^P \geq 1$ for both parties. The size of the Legislature and the uniformity of party loyalty are purely technical and have no bearing on the following simulations' generality.

As demonstrated in Kselman (2010), once incumbents' re-election utilities and parties' effort capacities fall below a certain level, the district-level OLPR game's NE properties *may* (or may not...) change depending on a district's incumbency status quo. For most values of $\beta_{j,d}^P$ and $E^P$ the OLPR game still generates stable NE outcomes, and higher aggregate levels of constituency effort than their FPTP and CLPR counterparts. On the other hand, for unusually low values of $\beta_{j,d}^P$, the OLPR game may under certain circumstances *have no NE*.

That said, this absence of NE does not in fact violate the basic comparative static hypotheses presented here: at these extremely low values of $\beta_{j,d}^P$, neither FPTP nor CLPR systems generate *any* constituency service (i.e., at these extremely low re-election utilities $T^*(FPTP) = T^*(CLPR) = 0$). As such, at these very low values of $\beta_{j,d}^P$, the fact that OLPR competition does not generate a stable outcome in fact makes it *more* constituency oriented then either FPTP or CLPR systems, which generate stable outcomes characterized by the categorical absence of constituency service. As such, like the assumption that $N = 200$ and that $\ell \in [0, 1]$ is uniform across regions, the assumptions that $\beta_{j,d}^P \geq 1$ and $E^P \geq 1$ do not affect the following analyses' generality.

## *Deriving Total Aggregate Constituency Effort*

Figures 2 and 3 present the statistic $T^*(I)$ for all three institutions $I \in \{FPTP, CLPR, OLPR\}$ at all possible values of $\ell \in [0, 1]$. In FPTP systems, the $N = 200$ incumbents represent single-member districts and the calculation of $T^*(FPTP)$ is straightforward. For any value of $\ell > \frac{1}{2}$ all incumbent legislators will choose $f_{1,d}^{P*} = 0$, since by the definition of plurality rule they can do so and still win back their legislative seat. In turn, if $\ell > \frac{1}{2}$ then $T^*(FPTP) = 0$. On the other hand, recalling the analysis of FPTP elections from the text, if $\ell < \frac{1}{2}$ then individual incumbents will have to devote $\hat{f}_{1,d}^P = \frac{1}{2} - \ell + \varepsilon$ ($\varepsilon \to 0$) in order to gain re-election. Since by construction $\beta_{j,d}^P \geq 1$, incumbents will always choose to exert the effort necessary for re-election. As such, for values of $\ell < \frac{1}{2}$ we know that $T^*(FPTP) = 200 \cdot \hat{f}_{1,d}^P \cong 100 - 200 \cdot \ell$. These facts yield the plots of $T^*(FPTP)$ in Figs. 2 and 3.

In PR systems, Fig. 2 begins with the case in which a country's 200 seats are divided into $D = 20$ distinct regions, each with a magnitude of $M_d = 10$, and characterized by the following arbitrarily chosen district-by-district incumbency breakdown:

This district-by-district breakdown implies a slim 103-to-97 legislative majority for party A. The first thing to note about the calculation of $T^*(CLPR)$ and $T^*(OLPR)$ is that, in both CLPR and OLPR systems, district-level NE outcomes are qualitatively unaffected by the identity of the district-level majority party. For example, the NE outcomes in districts where party A has $\bar{A}_d = 6$ seats and party B has $\bar{B}_d = 4$ seats parallel those of districts in which party A has $\bar{A}_d = 4$ seats and

party $B$ has $\bar{B}_d = 6$ seats; the NE outcomes in districts where party $A$ has $\bar{A}_d = 7$ seats and party $B$ has $\bar{B}_d = 3$ seats parallel those of districts in which party $A$ has $\bar{A}_d = 3$ seats and party $B$ has $\bar{B}_d = 7$ seats; and so on.

To demonstrate this point, begin with CLPR systems and consider a district in which party $A$ has $\bar{A}_d = 6$ seats and party $B$ has $\bar{B}_d = 4$ seats, and in which regional loyalty is $\ell = \frac{1}{5}$. In this case, in equilibrium the marginal incumbent is the incumbent candidate from party $A$ at list position 5: when the full-shirking vector is played, $A$ receives a district-level vote share of $\mathbf{v_d^A}(\mathbf{F_d^o}) = [(\frac{1}{5} \cdot 6 + \frac{4}{5} \cdot 4)/10] = .44$, which in turn implies that at the full-shirking vector party $A$ wins $\mathbf{x_d^A}(\mathbf{F_d^o}) = 4$ seats. In order to secure his or her re-election, the marginal candidate must thus devote $\hat{f}_{m,d}^A = \frac{1}{10} + \varepsilon \; (\varepsilon \to 0)$, so as to push party $A$'s remainder just above that received by party $B$. Since $\beta_{j,d}^P \geq 1$ by construction, in the unique district-level NE the marginal candidate will choose $f_{j,d}^{A\,*} = \hat{f}_{j,d}^A$ and all other incumbents will choose $f_{j,d}^{P\,*} = 0$.

Now move to a situation in which party $A$ has $\bar{A}_d = 4$ seats and party $B$ has $\bar{B}_d = 6$ seats, and in which regional loyalty is $\ell = \frac{1}{5}$. In this situation the marginal candidate will be the candidate from party $B$ at list position 5, and this candidate will face the same choice just described when the marginal candidate was from party $A$ (i.e., $\hat{f}_{m,d}^B = \frac{1}{10} + \varepsilon$, where $\varepsilon \to 0$). In the unique district level NE the marginal candidate will choose $f_{j,d}^{B\,*} = \hat{f}_{j,d}^B$ and all other incumbents will choose $f_{j,d}^{P\,*} = 0$.

As such, regardless of district level majority party's identity, the total NE amount of constituency effort generated across all incumbents in a CLPR district with a 6-to-4 incumbency status quo when $\ell = \frac{1}{5}$ will be equal to $\frac{1}{10} + \varepsilon$: in equilibrium $f_{m,d}^{P\,*} = \frac{1}{10} + \varepsilon$ is the marginal incumbent's choice and all other incumbents choose $f_{j,d}^{P\,*} = 0$. For any particular level of $\ell \in [0,1]$, define $T_{6/4}(CLPR, \ell)$ as the total NE constituency effort generated in a CLPR district with a 6-to-4 incumbency status quo, such that for example $T_{6/4}(CLPR, \frac{1}{5}) = \frac{1}{10} + \varepsilon$. Similarly, for any level of $\ell \in (0,1)$, define $T_{7/3}(CLPR, \ell)$ as the total NE constituency effort generated in a district with a 7-to-3 incumbency status quo; $T_{8/2}(CLPR, \ell)$ as the total NE constituency effort generated in a district with a 8-to-2 status quo; and so on.

The district-by-district breakdown in Fig. 8 implies a total of 5 districts with an 8-to-2 incumbency status quo, 6 districts with a 7-to-3 incumbency status quo, 6 districts with a 6-to-4 incumbency status quo, and 3 districts with a 5-to-5 incumbency status quo. In turn, for any particular value of $\ell \in [0,1]$ the following expression generates the statistic $T^*(CLPR)$:

$$T^*(CLPR) = 5 \cdot [T_{8/2}(CLPR, \ell)] + 6 \cdot [T_{7/3}(CLPR, \ell) + 6 \cdot [T_{6/4}(CLPR, \ell)]$$
$$+ \; 3 \cdot [T_{5/5}(CLPR, \ell)] \tag{17}$$

Figure 2 presents the statistic $T^*(CLPR)$ defined in (17) for all possible values of $\ell \in (0,1)$. The specific calculations of district-level effort under all posited incumbency situations, and for any particular value of $\ell \in [0,1]$, are available upon request.

Now move to the calculation of $T^*(OLPR)$. As in the CLPR case, the identity of the district-level majority party has no qualitative consequence for the model's

> ➢ 3 districts in which party $A$ has $\bar{A}_d = 8$ seats and party $B$ has $\bar{B}_d = 2$ seats.
> ➢ 3 districts in which party $A$ has $\bar{A}_d = 7$ seats and party $B$ has $\bar{B}_d = 3$ seats.
> ➢ 3 districts in which party $A$ has $\bar{A}_d = 6$ seats and party $B$ has $\bar{B}_d = 4$ seats.
> ➢ 3 districts in which party $A$ has $\bar{A}_d = 5$ seats and party $B$ has $\bar{B}_d = 5$ seats.
> ➢ 3 districts in which party $A$ has $\bar{A}_d = 4$ seats and party $B$ has $\bar{B}_d = 6$ seats.
> ➢ 3 districts in which party $A$ has $\bar{A}_d = 3$ seats and party $B$ has $\bar{B}_d = 7$ seats.
> ➢ 2 districts in which party $A$ has $\bar{A}_d = 2$ seats and party $B$ has $\bar{B}_d = 8$ seats.

**Fig. 8** Proportional simulation 1

comparative static. For example, referring back to Proposition 4 in Appendix 3, consider a district in which party $A$ has $\bar{A}_d = 6$ seats and party $B$ has $\bar{B}_d = 4$ seats, and in which regional loyalty is $\ell = \frac{1}{5}$. In this case, given that by construction $\beta_{j,d}^P \geq 1$ and $E^P \geq 1$, in any NE all of party $A$'s incumbents choose $f_{j,d}^{A*} \in [\frac{7}{20}, \frac{9}{20}]$ and all of party $B$'s incumbents choose $f_{j,d}^{B*} = [1 - (\frac{2}{5} + f_{j,d}^{A*})]$. Similarly, in a district where party $A$ has $\bar{A}_d = 4$ seats and party $B$ has $\bar{B}_d = 6$ seats, and in which regional loyalty is $\ell = \frac{1}{5}$, in any NE all of $B$'s incumbents choose $f_{j,d}^{B*} \in [\frac{7}{20}, \frac{9}{20}]$ and all of party $A$'s incumbents choose $f_{j,d}^{A*} = [1 - (\frac{2}{5} + f_{j,d}^{B*})]$.

Recall from Proposition 4 that, when $\ell \in [0, 1]$ is sufficiently small, the OLPR game generates NE within a specific range of incumbents' action sets, rather than unique NE. Without loss of generality, in such situations I will employ the NE outcome which represents the *mean level* of constituency service for the relevant value of $\ell \in [0, 1]$ in computing $T^*(OLPR)$. For example, and continuing with the above example, for a district with a 6-to-4 incumbency status quo and regional loyalty level of $\ell = \frac{1}{5}$, in calculating $T^*(OLPR)$ I will adopt the NE outcome in which incumbents from the majority party choose $f_{j,d}^{MA*} = \frac{2}{5}$ and incumbents from the minority party choose $f_{j,d}^{MI*} = \frac{1}{5}$.

For any particular level of $\ell \in [0, 1]$, define $T_{6/4}(OLPR, \ell)$ as the total constituency effort generated in the "mean" NE to the OLPR game for a district with a 6-to-4 incumbency status quo, such that for example $T_{6/4}(OLPR, \frac{1}{5}) = [(f_{j,d}^{MA*} \cdot 6) + (f_{j,d}^{MI*} \cdot 4)] = [\frac{2}{5} \cdot 6 + \frac{1}{5} \cdot 4] = 3.2$. Similarly, for any level of $\ell \in [0, 1]$, define $T_{7/3}(OLPR, \ell)$ as the total constituency effort generated in the "mean" NE to the OLPR game for a district with a 7-to-3 incumbency status quo; $T_{8/2}(OLPR, \ell)$ as the total constituency effort generated in the "mean" NE to the OLPR game for a district with a 8-to-2 incumbency status quo; and so on. In turn, given the district-by-district incumbency breakdown presented in Table 2, the following formula expresses $T^*(OLPR)$ for any level of $\ell \in [0, 1]$:

$$T^*(OLPR) = 5 \cdot [T_{8/2}(OLPR, \ell)] + 6 \cdot [T_{7/3}(OLPR, \ell] + 6 \cdot [T_{6/4}(OLPR, \ell)] + 3 \cdot [T_{5/5}(OLPR, \ell)] \tag{18}$$

**Table 2** Proportional simulation 2

- 10 districts in which party $A$ has $\bar{A}_d = 3$ seats and party $B$ has $\bar{B}_d = 0$ seats
- 10 districts in which party $A$ has $\bar{A}_d = 2$ seats and party $B$ has $\bar{B}_d = 1$ seats
- 10 districts in which party $A$ has $\bar{A}_d = 1$ seats and party $B$ has $\bar{B}_d = 2$ seats
- 10 districts in which party $A$ has $\bar{A}_d = 0$ seats and party $B$ has $\bar{B}_d = 3$ seats
- 15 districts in which party $A$ has $\bar{A}_d = 2$ seats and party $B$ has $\bar{B}_d = 0$ seats
- 15 districts in which party $A$ has $\bar{A}_d = 1$ seats and party $B$ has $\bar{B}_d = 1$ seats
- 10 districts in which party $A$ has $\bar{A}_d = 0$ seats and party $B$ has $\bar{B}_d = 2$ seats

Figure 2 presents the statistic $T^*(OLPR)$ defined in (18) for all possible values of $\ell \in [0, 1]$. The specific calculations of district-level effort under all posited incumbency situations, and for any particular value of $\ell \in [0, 1]$, are available upon request.

Figure 3 undertakes an identical process to that described with respect to Fig. 2, except that PR systems are now divided into $D = 80$ distinct regions, 40 of which have a magnitude of $M_d = 3$ and 40 of which have a magnitude of $M_d = 2$, characterized by the following arbitrarily chosen district-by-district incumbency breakdown:

This district-by-district breakdown implies a slim 105-to-95 legislative majority for party $A$. The FPTP simulation is identical to that presented in Fig. 2. As for the PR simulation, district-level NE outcomes are once again qualitatively unaffected by the identity of the district-level majority party. As such, to derive the plots captured in Fig. 3 we simply re-stipulate expressions (17) and (18) for this altered district-by-district incumbency breakdown. Figure 3 presents threes plots for all possible values of $\ell \in [0, 1]$. The specific calculations of district-level effort under all incumbency situations, and for any particular value of $\ell \in [0, 1]$, are available upon request.

## The Generality of Institutional Hypotheses

The plots presented in Figs. 2 and 3 paint a consistent and telling picture: at unusually high levels of partisanship all three systems generate little to no constituency service; and as partisanship begins to drop, OLPR quickly begins to outperform its counterparts in generating constituency service, whereas the relative performance of CLPR and FPTP systems depends on both a country's partisanship and its district magnitude.

These comparative static hypotheses do not emerge as a result of the specific, arbitrarily chosen simulations described above, i.e., these results emerge regardless of the size of the Legislature and the district-by-district incumbency status quo. To demonstrate this, note first that that, without aggregating across the entire Legislature, in any particular district-to-district comparison from the above simulations OLPR outperforms its counterparts. For example, given a situation in which

PR districts are of magnitude $M_d = 10$, were we to compare $T_{6/4}(OLPR, \ell)$ with $T_{6/4}(CLPR, \ell)$, and then compare both of these to the constituency effort generated in 10 individual FPTP districts of size $M_d = 1$, we would obtain identical comparative statics to those uncovered in Figs. 2 and 3: at all but unusually high levels of partisanship OLPR outperforms its counterparts in generating constituency service, while the relationship between CLPR and FPTP systems depends on the level of partisanship.

It is straight-forward to show that these district-level dynamics are generalizable: regardless of a district's magnitude and the district-level incumbency breakdown, at all but unusually high levels of partisanship, OLPR generates greater district-level constituency service than its counterparts, while the relative performance of CLPR and FPTP systems depends on both a country's partisanship and its district magnitude. Put otherwise, the aggregate plots pictured in Figs. 2 and 3 simply reproduce comparative static relationships which exist in all possible district-by-district comparisons. As such, we could have simulated any district-by-district incumbency status quo and obtained the same results. For reasons of space I omit the full formal proof of this statement; it is available upon request.

That said, while the choice of simulation environment does not affect the model's qualitative hypotheses, it does affect their quantitative size. The district-by-district analysis demonstrates that OLPR has a particularly strong impact on constituency service incentives in districts tilted heavily towards one party or another, i.e., where the majority party's current seats far outweigh those of the minority party. For example, in terms of the simulation plotted in Fig. 2, the effect of OLPR on constituency service incentives would be slightly stronger if all PR districts were characterized by an 8-to-2 majority to minority party seat ratio, but would be slightly weaker if all PR districts were characterized by an 6-to-4 majority to minority seat ratio. That said, the basic comparative static hypothesis, that OLPR outperforms the other two systems at all but the highest partisanship levels, obtains in all exogenous environments.

## Appendix 5: Data and Measurement

Figure 9 contains all countries values on the variables FPTP, CLPR, OLPR, and HYBRID. Countries are placed in four distinct columns depending on their *predominant system* (see text page 8). For countries with mixed systems, their values on the distinct institutional variables are labeled in parentheses. These measures were coded using a variety of different sources for the sake of cross-checking, including but not limited to: Golder (2004); Seddon et al. (2002); the data Appendix in Cox (1997); and the Inter-Parliamentary Union's online database, which can be found at http://www.ipu.org/english/home.htm.

In keeping with the dependent variable's time point, countries are coded according to the electoral system present during the years 1994–1997. Four countries

| **FPTP** | **CLPR** | **OLPR** | **HYBRID** |
|---|---|---|---|
| Bahamas | Argentina | Brazil | Australia |
| Bangladesh | Austria | Chile | Cyprus |
| Barbados | Belgium | Denmark | Malta |
| Belize | Bolivia | Czech Republic | Mauritius |
| Botswana | Bulgaria | Estonia | Taiwan (HYBRID=.58; CLPR=.42) |
| Canada | Colombia | Finland | Thailand |
| Fiji | Costa Rica | Greece (OLPR=.94; FPTP=.006) | |
| France | Dominican Republic | Latvia | |
| Gambia | Ecuador | Luxemburg | |
| Ghana | El Salvador | Poland (OLPR=.85; FPTP=.15) | |
| India | Germany | Slovak Republic | |
| Jamaica | Guatemala | Sri Lanka | |
| Japan (FPTP=.6; CLPR=.4) | Honduras | Switzerland (OLPR=.975; FPTP=.025) | |
| Malawi | Hungary (CLPR=.54; FPTP=.46) | | |
| Malaysia | Iceland | | |
| Mexico (FPTP=.6; CLPR=.4 ) | Israel | | |
| Nepal | Italy | | |
| Pakistan | Namibia | | |
| Papua New Guinea | Netherlands | | |
| Philippines | New Zealand | | |
| Singapore | Nicaragua | | |
| South Korea | Norway | | |
| St. Vincent | Paraguay | | |
| Trinidad | Peru | | |
| USA | Portugal | | |
| Uganda | Romania | | |
| UK | Senegal (CLPR = .583; FPTP=.417) | | |
| Ukraine | South Africa | | |
| Zambia | Spain | | |
| Zimbabwe | Sweden | | |
| | Turkey | | |
| | Uruguay | | |
| | Venezuela | | |

**Fig. 9** Electoral formula and ballot structure

undertook major institutional reforms in 1993: New Zealand went from an FPTP system to a mixed FPTP-PR system in which the upper-tier serves as a *corrective tier* for any disproportionality introduced in the FPTP tier (see discussion of corrective tiers immediately following); Italy went from an OLPR system to mixed FPTP-CLPR system, also with a corrective PR tier; Venezuela went from a pure CLPR system to a mixed FPTP-CLPR system with a corrective upper-tier; and Japan went from using the Single Non-Transferable-Vote in multi-member districts to a mixed FPTP-CLPR system in which the two tiers are independent (i.e. the PR-tier is not corrective). I have re-run all of the paper's empirical analyses on a sample in which these three cases are coded as intermediate, i.e., their values

are weighted equally by the system in place before 1993 and that in place after 1993. The paper's empirical results are completely unaffected. Bolivia and the Philippines both experienced institutional change in 1996, but these changes did not become effective for electoral competition until after 1997.

A number of countries which used ostensibly mixed systems are here coded as pure system types: Germany, New Zealand, Italy, Venezuela, and South Korea. This is due to the fact that a parties' seat allocation in one-tier is not *independent* from their performance in the alternative tier (all cases can be recoded as mixed without changing the paper's empirical results). The first four cases use a corrective, national-level PR tier to correct for any disproportionality in vote shares which arise in the lower FPTP tier. Political parties thus have every incentive to engage in vote-seeking *as if* the system were purely proportional, since in the end seats will be allocated on a purely proportional basis. Similarly, the small upper-tier in South Korean elections serves to amplify the seat majority of whichever party wins a plurality of FPTP seats, such that parties' real emphasis will be on the lower tier (i.e., South Korea is coded as pure FPTP).

# References

Ames B (1995) Electoral strategy under open-list proportional representation. Am J Polit Sci 39:406–433

Austen-Smith D (2000) Redistributing income under proportional representation. J Polit Econ 108:1235–1269

Cox G (1987) Electoral equilibria under alternative voting institutions. Am J Polit Sci 31:82–108

Cox G (1990) Centripetal and centrifugal incentives in electoral institutions. Am J Polit Sci 34:903–935

Cox G (1994) Strategic voting equilibria under the single Non-transferable-vote. Am Polit Sci Rev 88:608–621

Cox G (1997) Making votes count. Cambridge University Press, Cambridge

Duverger M (1959) Political parties, their organization and activity in the modern state. Wiley, London

Fenno RF (1978) Home style: house member in their districts. Little Brown, Boston

Fernandez M, Pierskalla J (2009) Partisan strategies and political economic performance. Do modes of democratic accountability affect economic fortunes? Paper presented at the 2009 annual meeting of the American Political Science Association, Toronto, 3–6 September 2009

Fraenkel J, Grofman B (2004) A neo-downsian model of the alternative vote as a mechanism for mitigating ethnic conflict in plural societies. Public Choice 121:487–506

Freeze, K, Kitschelt H, Wang Y (2008) Measuring democratic accountability: an initial report on an emerging data set. Paper prepared for the workshop on data sets in comparative politics, Duke University Dept of Political Science, 24–25 April 2009

Gingerich D (2009) Ballot structure, political corruption, and the performance of proportional representation. J Theor Polit 21:509–541

Golder M (2004) Democratic electoral systems around the world. Data set and codebook available at http://homepages.nyu.edu/~mrg217/elections.html

Horowitz D (1985) Ethnic groups in conflict. University of California Press, Berkeley

Ingall RE, Crisp BF (2001) Determinants of home-style: the many incentives for going home in Colombia. Legislative Stud Q 26:487–512

Iversen T, Soskice D (2006) Electoral institutions and the politics of coalitions: Why some democracies redistribute more than others. Am Polit Sci Rev 100:165–182

Keefer P, Vlaicu R (2008) Credibility, clientelism and democracy. J Law Econ Organ 24:371–406

Kselman D (2009) Electoral institutions, party organizations, and political instability. Doctoral Thesis, Duke University Department of Political Science

Kselman D (2010) Electoral institutions, legislative accountability, and political corruption. Juan March Institute Working Paper no. 249

Kunicova A, Rose-Ackerman S (2005) Electoral rules and constitutional structures as constraints on corruption. Br J Polit Sci 35:573–606

Lijphart A (1977) Democracy in plural societies. Yale University Press, New Haven

Lijphart A (1999) Patterns of democracy: government forms and performance in 36 countries. Yale University Press, New Haven

Long Jusko K (2009) The electoral foundations of poverty relief in contemporary democratic societies. Working Paper, Stanford University Dept of Political Science

Martin S (2010) Electoral rewards for personal vote cultivation under PR-STV. West Eur Polit 33 (2):369–380

Milesi-Ferretti GM, Perotti R, Rostagno M (2002) Electoral rules and the composition of public spending. Q J Econ 117:609–657

Myerson R (1993a) Incentives to cultivate favored minorities under alternative electoral systems. Am Polit Sci Rev 87:856–869

Myerson R (1993b) Effectiveness of electoral systems for reducing corruption. Game Econ Behav 5:118–132

Palfrey TR (1989) A mathematical proof of Duverger's law. In: Ordeshook P (ed) Models of strategic choice in politics. University of Michigan Press, Ann Arbor

Persson T, Tabellini G (2000) Political economics: explaining economic policy. MIT, Cambridge

Persson T, Tabellini G (2003) The economic effects of constitutions. MIT, Cambridge

Persson T, Tabellini G, Trebbi F (2003) Electoral rules and corruption. J Eur Econ Assoc 1:958–989

Reilly B (2001) Democracy in divided societies: electoral engineering for conflict management. Cambridge University Press, Cambridge

Riker WH (1982) The Two-party system and Duverger's law: an essay on the history of political science. Am Polit Sci Rev 76:753–766

Seddon J, Gaviria A, Panniza U, Stein E (2002) Political particularism around the world. Inter-American Development Bank Working Paper # 463

Szwarcberg M (2009) The political effects of campaign rallies: clientelism and electoral competition in Argentina. Working Paper, Yale University Program on Democracy

Treisman D (2007) What have we learned about the causes of corruption from ten years of cross-national research. Annu Rev Polit Sci 10:211–244

# A Model of Party Entry in Parliamentary Systems with Proportional Representation

**Daniel M. Kselman and Joshua A. Tucker**

## 1 Spatial Models of Party Entry

Spatial models of electoral competition have long-been the workhorse in political science. Black (1958) and Downs (1957) first extended the notion of 'median-voter convergence' to democratic elections with office-seeking candidates, though the basic theoretical mechanics date to Hotelling's model of geographic dispersion in economic markets (1929). Since this seminal research, developments in spatial modeling have, to a large measure, examined the extent to which this convergent tendency does (or does not. . .) emerge in more complex theoretical environments. An exhaustive review of this literature is well-beyond our current scope. Instead, we focus on one particular set of extensions to the traditional Downsian framework: those in which the set of candidates is not exogenously stipulated, i.e., in which some set of potential candidates must choose whether or not to *enter* the party system altogether.

Spatial models with a party entry decision largely fall into one of two classes. The first of these preserves the Downsian assumptions that candidates are essentially office-seeking (though note the forthcoming distinction between vote- and rank-maximization), and can announce policy positions anywhere in the policy

D.M. Kselman (✉)
Juan March Institute, CEACS, Calle Castelló 77, 28006 Madrid, Spain
e-mail: dkselman@march.es

J.A. Tucker
Department of Politics, New York University, 19 West 4th Street, 2nd Fl., New York, NY 10012, USA
e-mail: joshua.tucker@nyu.edu

space. Palfrey (1984) examines competition between two status quo organizations and a single potential entrant, demonstrating that the threat of entry by a third party incentivizes status quo-parties to choose policy positions well-removed from the median voter's ideal point. The phenomenon has been labeled 'entry-deterring dispersion'.[1] For plurality-rule electoral systems with $K$ seats,[2] Greenberg and Shepsle (1987) seek to uncover equilibrium in which $K$ candidates choose positions that maximize their rank-order while also deterring entry by additional candidates. Although the case exists in which $K + 1$ may at times generate entry-deterring dispersion in the aforementioned sense, Shepsle and Greenberg's assumption of rank-maximizing rather than vote-maximizing candidates allows for the more reasonable result that third-party candidates who have no chance of winning will choose not to enter the electoral fray. Osborne (2000) examines a dynamic three-candidate model of entry and spatial positioning, demonstrating that convergent two-party equilibria can be sustained under certain assumptions as to the inter-temporal sequencing of candidates' decisions.

A distinct class of models features what are now known as "citizen-candidates" (Osborne and Slivinsky 1996; Besley and Coate 1997; Eguia 2007). These models do away with the distinction between voters and candidates, and instead endogenize the decision of citizens *to become candidates*. While candidates in the above models were primarily vote- or office-seeking, those which emerge in citizen-candidate models have clear preferences over *policy outcomes*. Also distinct from aforementioned work, those citizens that decide to become electoral candidates must pay a fixed cost for competing in the election. Finally, citizen candidates cannot credibly commit to implementing any policy other than their own most-preferred policy, and thus cannot choose where to position themselves in ideological space (i.e., they are constrained to adopt their 'ideal point' as a campaign platform). This framework in hand, citizen-candidate models seek to identify the number of candidates which emerge in equilibrium, along with their relative dispersion in policy space.

The model developed in this paper employs assumptions and mechanisms from each of these two classes of spatial analyses. Like the first class of models more loyal to the Downsian tradition, the primary strategic actors in our model are political leaders rather than citizens. Conceptually speaking, we thus take the choice by citizens to become political leaders as theoretically prior to our analysis. However, like citizen-candidate models, we allow party leaders to receive utility from both policy outcomes and office, and we require new entrants to pay a cost for choosing to compete in electoral campaigns. Also in keeping with the citizen-candidate framework, we constrain candidates to adopt their own ideal point as a campaign platform. We distinguish ourselves from both streams of research by

---

[1]Lee (2007) extends models entry as a probabilistic process, demonstrating that status quo parties' distance from the median voter's ideal point increases in the probability of third party entry.

[2]When $K = 1$ this implies a traditional first-past-the-post system. When $K > 1$, this implies the single-non-transferable vote system.

introducing into our model a third stage where parties form coalition governments and implement policies.

The framework thus developed allows us to answer important questions concerning the entry of new party organizations in Parliamentary democracies: Is entry more or less-likely when status quo organizations are widely dispersed around the median voter's ideal point? How will the incentives for entry vary with candidates' relative emphasis on 'policy-seeking' in contrast to 'office-seeking'? How will the location of a potential entrant's ideal point (which by assumption will be her campaign platform) affect her incentives to join the election? In the current paper, we present the model primitives (Sect. 2), and then present a series of examples to demonstrate the model's Nash Equilibrium mechanics, which allow us to provide a set of preliminary answers to the aforementioned questions (Sect. 3). We then conclude with a discussion of future research which will extend the modeling framework to a wider range of strategic and institutional environments (Sect. 4).

## 2 Actors and Utility Functions

The model's strategic actors are *party leaders*, who are further subdivided into those who control pre-existing political parties and a *single* leader who must decide whether or not to create a new political party.[3] More specifically, we use the marker $P \in \{1, 2, 3\}$ to denote any one of three party leaders, and arbitrarily define $P = 1$ as the single party leader not already in control of a pre-existing organization. As we will soon see, unlike models in which party leaders' strategic flexibility leads to instability when politics is multi-dimensional, the following results are robust to multi-dimensional competition. That said, for simplicity we model political competition as occurring in a unidimensional space; denote this space $x \in [0, 1]$, and label a party leader's *ideal point* (i.e., most-preferred policy position) in this policy space as $x_P$. Importantly, we assume that leaders' ideal points are *common knowledge*, i.e., that they are known both to voters and to competing candidates. In turn, and in contrast to the traditional Downsian model in which party leaders are free to adopt any desired policy stance in political campaigns, we assume leaders are restricted to their own ideal point in announcing campaign platforms. The game will proceed in three stages: (a) in a first stage, leader 1 must decide whether or not to 'Enter' the party system by creating a new party organization; (b) in a second stage an election is held and voters must choose between the available partisan options; (c) in the third stage an Executive is formed, which may be either a single-party Cabinet or a coalition Cabinet.

---

[3]The setup of a game played between pre-existing organizations and a *single* potential entrant is identical to that in Palfrey's original model (1984) along with most formal research since (e.g., Lee 2007; Hug 2001).

Regarding this final stage, we adopt a simple and standard model of government formation in which the governing party or coalition must be supported by no less than one-half of all incumbent legislators.[4] As such, if any one party emerges from the second stage election with a parliamentary majority, then only it will be called upon to act in the game's third stage, and this act will involve the formation of a single-party government. On the other hand, if no party emerges from the game's electoral stage with a parliamentary majority, then a coalition formation process ensues. Volumes of theoretical and empirical research exist on the topic of coalition formation in Parliamentary democracies (e.g., Baron and Ferejohn 1989; Stromm 1990; Laver and Shepsle 1996; Martin and Stevenson 2001; Carroll and Cox 2007; Bassi 2008). A full review of this literature is well beyond the current paper's scope, as is an original contribution to our theoretical understanding of coalition formation processes. Rather, the goal here is to specify a simple but justifiable model of coalition formation, which in turn allows me address the paper's stated questions regarding the likelihood of party entry in different strategic environments.

As such, we take as a baseline Baron and Ferejohn's seminal model (1989) in which the party designated as 'formateur' (i.e., the party allocated the power to 'propose' coalitions to its competitors) chooses a coalition which maximizes its utility, given the constraint that it be supported by no less than a Parliamentary majority. In the authors' original analysis the formateur was assumed to be the party with a plurality of Parliamentary seats. However, more recent research (Martin and Stevenson 2001) has indentified a variety of formal and informal mechanisms by which parties are allocated the right to make coalition proposals. So as to operationalize the fact that plurality parties are, in fact, often granted formateur status, while at the same time avoiding an overly deterministic specification, we will assume that each party has a probability of being assigned formateur status, and that this probability is an increasing function of its legislative seat share (see below).

We can now specify a party leader's strategy, i.e., a complete plan of action for all situations in which a leader might be called upon to act, which we will denote $s_P$. Similarly, denote a strategy profile, i.e., a set of strategies for all party leaders, as $\mathbf{s} = \{s_1, s_2, ..., s_p\}$. In the game's first stage, leader 1 must choose between 'Entering' the party system and 'not Entering' the party system; label these choices as $E$ and $\sim E$ respectively. If 1 chooses $\sim E$ in the game's first stage, then by definition she will under no circumstances be called upon to act in the game's government formation stage, i.e., anytime 1 chooses $\sim E$ in the game's first stage her strategy will simply be $s_1 = \{\sim E\}$. Define $G$ as an Executive Government. For example, a single-party government formed by party 1 is denoted $G = 1$, a coalition government formed between party 1 and party 2 is denoted $G = 12$, a coalition between parties 1 and 3 is denoted $G = 13$, and a coalition between parties 2 and 3 is denoted $G = 23$. In turn, if 1 chooses to Enter in the game's first stage and there is some non-zero probability that she will be named formateur in the game's third stage, her strategy must include a coalition proposal $G_1(E)$ for the eventuality in

---

[4]We thus do not consider the possibility of minority governments (see Stromm 1990).

which she is in fact assigned formateur status, such that $s_1 = \{E, G_1(E)\}$. For example, the strategy in which 1 Enters and, if named formateur, offers to form a coalition with 2, is denoted $s_1 = \{E, 12\}$.

Given the assumption that party leaders are constrained to adopt their own ideal point as a campaign platform, the only strategic move held by leaders in control of pre-existing organizations arrives in the game's third stage. Define $G_{P\geq2}(E)$ and $G_{P\geq2}(\sim E)$ as party leader $P$'s Government proposals given that 1 chooses $E$ and $\sim E$ respectively in the game's first stage. If there is a positive probability that $P$ will be named formateur regardless of 1's choice, then her strategy is denoted $s_{P\geq2} = \{G_{P\geq2}(\sim E), G_{P\geq2}(E)\}$. In contrast, if in the game's second stage a competing organization secures a single-party majority in Parliament, then leader $P$ will not be called upon to act *at all* in the third stage.[5]

Voters, like party leaders, are defined by their ideal point in the space $x \in [0, 1]$, and naturally the further away a policy is from this ideal point the more it is disliked by the voter in question. In particular, define a voter $i$'s utility for policy $x \in [0, 1]$ with the loss function $u_i(x) = -(x_i - x)^2$, and assume without loss of generality that voter ideal points are distributed *uniformly* over $x \in [0, 1]$ (i.e., $x \sim \text{unif}[0, 1]$). This model assumes that voters are non-strategic: in the game's electoral stage they will simply choose the party whose campaign platform is closest to their ideal point. Given the assumption that a party $P$'s campaign platform is constrained to be $x_P$, we can then express a party's electoral vote share as:

$$v_P = \sum_x x_i : \forall P, P', \quad (x_i - x_P)^2 < (x_i - x_{P'})^2. \tag{1}$$

The translation of electoral votes into legislative seats is rarely a perfectly proportional process, and the extent to which this translation provides extra-proportional seat bonuses to large or small organizations depends on the electoral rule in question. For the sake of simplicity we assume a perfectly proportional translation of votes into seats, such that parties' parliamentary seat shares are also represented by the function $v_P$. In turn, if $v_P \geq \frac{1}{2}$ for some party then that party will form a single-party government. If no party has a vote share of $v_P \geq \frac{1}{2}$, assume that a party's probability of being assigned formateur status is identical to its vote share $v_P$.[6] In turn, let $\pi_G(E)$ and $\pi_G(\sim E)$ represent the probability that a particular government $G$ emerges given the choices 'Enter' and 'not Enter' respectively. Importantly, these

---

[5]If $P$ only has a positive probability of being assigned formateur status given the choice $E$ in the game's first stage, her strategy is denoted $s_{P\geq2} = \{\phi, G_{P\geq2}(E)\}$, where the empty set marker $\phi$ tells us that $P$ has no move in the game's third stage if 1 chooses $\sim E$. Finally, if $P$ has zero probability of being assigned formateur status regardless of 1's first stage choice her strategy is denoted $s_{P\geq2} = \{\phi, \phi\}$

[6]As with the assumption that Parliamentary seat shares perfectly reflect electoral vote shares, this assumption can be relaxed such that larger parties have a disproportionately higher chance of being assigned formateur status than smaller parties; or that smaller parties' chances of being named formateur are disproportionate to their vote shares.

probabilities are not fixed exogenously, but rather emerge as a function of the game's Nash Equilibrium strategies.

Once in office, a government must implement a policy position $x^*(G) \in [0, 1]$. If $G$ is a single-party government, it chooses its own ideal point as the Executive policy $x^*(G) = x_G$. As with research on formation of coalition governments, research on how governing coalitions distribute power internally has a long history. Gamson (1961) famously argued that the internal distribution of Cabinet portfolios tended to be proportional to participating parties' relative parliamentary strength. Although frequently criticized for failing to capture common insights from the literature on strategic bargaining (such as the importance of a first-mover advantage, the importance of being the 'pivotal' legislative party, etc.), recent research has substantially rehabilitated Gamson's basic argument (Carroll and Cox 2007; Bassi 2008). We follow this research in letting a coalition's policy $x^*(G)$ be a function of participating parties' relative vote shares. Define the 'weight' of a party $P'$ in any governing coalition $G$ as follows:[7]

$$\omega_{P'}(G) = \frac{v_{P'}}{\sum\limits_{P \in G} v_P}. \tag{2}$$

In turn, for a two-party coalition consisting of parties $P$ and $P'$ where $x_P > x_{P'}$ we define the Executive policy outcome as follows:

$$x^*(PP') = \omega_P(\cdot) \cdot x_P + \omega_{P'}(\cdot) \cdot x_{P'}. \tag{3}$$

Since by construction coalition weights are proportional to vote shares, it is natural here to incorporate as an assumption in the current model a result which emerges as a Subgame Perfect Nash Equilibrium (SPNE) in Baron and Ferejohn's original piece, namely that *no party ever refuses a coalition proposal*.[8]

As in citizen candidate models, we assume that candidates prefer policy outcomes closer to their ideal points.[9] In addition to their preferences over policy outcomes, political leaders also gain utility from participating in 'office', defined as any incumbency payoff not related to the ultimate policy a government implements (often labeled as 'ego-rents' in the citizen-candidate literature). Define $\beta$ as the benefit a party leader accrues from having his or her party at the head of

---

[7]By definition $\omega_P(\cdot) = 1$ for single-party Cabinets and $\omega_P(\cdot) = 0$ for parties not in government.

[8]This, of course, is a simplification. However, it is justifiable as both an empirical and theoretical regularity. Again, the current aim is not to develop a novel understanding of coalition formation, which might identify certain circumstances in the assumption could be violated, but rather to present a simple but justifiable model of government formation which permits investigation of the paper's stated theoretical questions.

[9]Although as will be noted shortly, the relative preference for policy outcomes as opposed to ego-rents from serving in the government is a parameter in the model, and as such can be set to 0 for those interested in obtaining predictions for situations in which candidates do not care about policy; we provide examples in Sect. 3.

a *single-party government*. For coalition governments, we will weight a party $P$'s office utility by her weight in the coalition, such that she receives utility $[\omega_P(G) \cdot \beta]$. A final consideration enters only into the utility calculation of political leaders not yet in control of their own organization, who must choose whether or not to enter the political fray in stage 1: namely, the fixed financial, legal, and organizational costs associated with creating new party organizations. Denote these fixed costs as $c$. Given any strategy profile $\mathbf{s} = \{s_1, s_2, ..., s_p\}$ we can now specify political leaders' expected utility over all possible outcomes:

$$U_P(\mathbf{s}) = \begin{cases} -\sum_G \pi_G(\sim E) \cdot \left( \phi \cdot [x_1 - x^*(G)]^2 \right) & \text{if } P = 1 \text{ and} \sim E \\ \sum_G \pi_G(E) \cdot \left( \omega_1(G) \cdot \beta - \phi \cdot [x_1 - x^*(G)]^2 \right) - c & \text{if } P = 1 \text{ and } E \\ \sum_G \pi_G(\cdot) \cdot \left( \omega_P(G) \cdot \beta - \phi \cdot [x_P - x^*(G)]^2 \right) & \text{if } P \geq 2 \end{cases} \quad . \quad (4)$$

Like voters, party leaders' policy preferences are thus characterized by a quadratic loss function. This loss-function will be weighted by the parameter $\phi$, which we define as the 'importance' that candidates attribute to policy-seeking motivations. In turn, the relative importance candidates attribute to policy-seeking to office-seeking can be written as $[\phi/\beta]$. The payoff associated with the choice $\sim E$ for party leader 1 (i.e., the potential entrant) is determined solely by the 'expected' government policy (i.e., weighted by each government's probability of emerging given strategy vector $\mathbf{s}$) to be implemented in the game's third stage: by not entering the leader in question forgoes both the fixed cost and the potential 'office' benefits associated with the choice $E$. On the other hand, the choice to enter implies paying the fixed cost $c$, but also potentially gaining access to the spoils of Executive incumbency and potentially influencing the game's ultimate policy outcome. Finally, regardless of whether or not 1 chooses to compete, the expected utility of candidates $P \geq 2$ will be a function of the utilities they receive given various government outcomes, each weighted by the probability that government emerges given strategy vector $\mathbf{s}$.

## 3    Subgame Perfect Nash Equilibrium

Given the game's sequential structure, the solution concept employed is Subgame Perfect Nash Equilibrium (SPNE). Define a SPNE strategy profile $\mathbf{s}^* = \{s_1^*, s_2^*, ..., s_p^*\}$ as a profile in which all coalition proposals $G_P(\cdot)$ represent Nash Equilibrium in the game's third stage, and in which the potential entrant 1 has no incentive to deviate from the relevant entry decision. As already noted, the framework developed in Sect. 2 allows us to answer a number of questions as to entry incentives in Parliamentary democracies: Is entry more or less-likely when status quo organizations are widely dispersed around the median voter's ideal

point? How will the incentives for entry vary with candidates' relative emphasis on 'policy-seeking' in contrast to 'office-seeking'? How will the location of a potential entrant's ideal point affect her incentives to join the election? In order to answer these questions fully, we would have to solve the model for all possible exogenous situations, which in turn implies solving game's third subgame (in which a coalition government is formed) in all possible strategic situations. For reasons that will soon become clear, this turns out to be quite an involved process, and exhaustive analysis of the game for all exogenous cases must await future research. However, by solving for Nash Equilibrium outcomes in a series of examples, we demonstrate here the viability of the analytic framework and develop some preliminary insights regarding the above-stated questions.

Note from the outset that parties 2 and 3 may be either 'bunched' on one or the other side of the political spectrum, i.e., both the left or to the right of the median voter's ideal point at $x = 0.5$, or may fall on opposite sides of the median voter's ideal point. Here we focus on the more likely scenario in which two status quo organizations fall on opposite sides of the median voter's ideal point (all results can be extended to the 'bunched' scenario). Define party 2 (3) as an organization to the left (right) of the median voter. Leader 1's ideal point may be anywhere in the range $x_1 \in [0, 1]$. We first present a general result which stipulates a set of conditions under which entry never occurs. Begin with the case in which $(\frac{1}{2} - x_2) < (x_3 - \frac{1}{2})$, i.e., in which the left party is located closer to the median voter's ideal point than the right party.

---

**\*Proposition 1**. *If $P \in \{1, 2, 3\}$, $(\frac{1}{2} - x_2) < (x_3 - \frac{1}{2})$, and $x_1 > (1 - x_2)$, then the game's unique SPNE is $s_1{}^* = \{\sim E\}$, $s_2{}^* = \{2, 2\}$, and $s_3{}^* = \{\phi, \phi\}$.*

---

**\*Proof**. If $x_1 > (1 - x_2)$ then political leader 1 cannot affect the election's outcome by Entering, i.e., party 2 will win the election with certainty even if she chooses $E$. As such 1 has no incentive to defect from the strategy profile defined in Proposition 1, since this would imply paying the entry cost $c$ without either changing the policy $x^*(G)$ or gaining any access to the spoils of incumbency (SPNE Existence); and any strategy vector at which 1 chooses $E$ is not a SPNE, since she would have the incentive to alter her choice from $E$ to $\sim E$ so as to avoid uselessly paying the costs of entry (SPNE Uniqueness).                                                                  ∎

A qualitatively identical result emerges when $(x_3 - \frac{1}{2}) < (\frac{1}{2} - x_2)$, i.e., when the pre-existing right party is located closer to the median voter than the pre-existing left party, the only difference being that it obtains when 1's ideal point $x_1$ is sufficiently 'left-leaning'.

Naturally, things become more complicated when these conditions don't obtain, i.e., when the decision $E$ by candidate 1 would actually induce a coalition bargaining process.[10] We will continue with the assumption $(\frac{1}{2} - x_2) < (x_3 - \frac{1}{2})$ and

---

[10]In this paper we will not investigate situations in which candidates 2 and 3 are equidistant from the median voter's ideal point such that, absent the choice to enter by candidate 1, election's outcome would be a tie.

consider two examples: (a) one in which $x_2 = 0.2$ and $x_3 = 0.9$, and (b) a second in which $x_2 = 0.5$ and $x_3 = 0.6$. Note right away that, by Proposition 1, the range of policy space over which entry will be a viable option is significantly reduced in the second as compared to the first scenario: in the first scenario entry may be a viable consideration for any $x_1 \in [0, 0.8]$, since at any point in this range the choice to enter will in fact induce a coalition bargaining process. On the other hand, in the second scenario entry will be a viable consideration only if $x_1 \in [0, 0.5]$, since only in this range the choice to enter will in fact induce a coalition bargaining process. As a result, we see from the outset that the existence of at least one party with centrist inclinations significantly reduces the set of candidates who might choose to form new party organizations.

If named formateur, a party's optimal coalition proposal $G_P(\cdot)$ will depend on the competing parties' relative vote shares and their relative proximity to $P$'s ideal point. *Ceteris Paribus*, formateur parties prefer choosing coalition partners close to their own ideal point, since by (3) above this reduces the distance between $x_P$ and $x^*(G)$. Also *ceteris paribus*, formateur parties prefer choosing coalition partners with lower votes shares, since this increases their 'weight' $\omega_P(G)$ in an eventual coalition, and thus both their policy and office utility. Given these two facts it is straightforward to see that, if one of the two remaining parties is *both* closer to the formateur $P$ in policy space *and* has a smaller vote share then the remaining party, then by definition the coalition which maximizes $P$'s utility is with this party. On the other hand, if one of the competing parties is 'closer' to $P$ while the other remaining party has a smaller vote share, identifying $P$'s optimal coalition proposal requires an explicit utility comparison.

Define a *coalition profile* $\{ G_1(E), G_2(E), G_3(E) \}$ as a set of optimal coalition proposals by all three candidates in the game's third stage given the decision by candidate 1 to enter in the game's first stage. For example, the profile in which 1's optimal coalition partner is 3, 3's optimal coalition partner is 1, and 2's optimal coalition partner is 1 would be denoted $\{13, 12, 13\}$. Naturally, we would expect these optimal coalition profiles to vary according to 1's ideal point, and to the relative emphasis that parties place on 'office-seeking' as opposed to 'policy-seeking'. For the scenario in which $x_2 = 0.2$ and $x_3 = 0.9$, Table 1 describes how coalition profiles vary with the position of $x_1$ for the cases in which $\beta = 0$ and $\phi = 1$ and in which $\beta = 1$ and $\phi = 0$, i.e., in which candidates are purely policy-seeking and purely office-seeking respectively.

The straightforward numerical proofs that these coalition proposals are indeed optimal are available upon request. Specific description of a few cells helps to internalize the coalition mechanisms at work. For example consider the case in which $x_1 = 0.7$, $\beta = 0$, and $\phi = 1$. Here the coalition profile will be $\{13, 23, 13\}$: if named formateur, candidates 1 and 3 would choose one another as coalition partners while candidate 2 would choose 3 as a coalition partner. Note that 2 chooses 3 despite the fact that 1's policy position is slightly closer to 2's ideal point: since $v_3$ is significantly smaller than $v_1$ candidate 2 is able to secure a more favorable policy outcome from $x^*(23)$ than she is from $x^*(12)$. On the other hand when $x_1 = 0.4$, $\beta = 0$, and $\phi = 1$, all three candidates will prefer to form a coalition

**Table 1** Coalition profiles when $x_2 = 0.2$ and $x_3 = 0.9$

|  | $v_P$ if 1 enters | Coalition profile if $\beta = 0$ and $\phi = 1$ | Coalition Profile if $\beta = 1$ and $\phi = 0$ |
|---|---|---|---|
| $x_1 = 1$ | $v_1 = 0.05$, $v_3 = 0.40$, $v_2 = 0.55$ | 1 chooses $\sim E$ (by Proposition 1) | 1 chooses $\sim E$ (by Proposition 1) |
| $x_1 = 0.9$ | $v_1 = 0.225$, $v_3 = 0.225$, $v_2 = 0.55$ | 1 chooses $\sim E$ (by Proposition 1) | 1 chooses $\sim E$ (by Proposition 1) |
| $x_1 = 0.8$ | $v_1 = 0.35$, $v_3 = 0.15$, $v_2 = 0.50$ | $\{13, 23, 13\}$ | $\{13, 23, 13\}$ |
| $x_1 = 0.7$ | $v_1 = 0.35$, $v_3 = 0.20$, $v_2 = 0.45$ | $\{13, 23, 13\}$ | $\{13, 23, 13\}$ |
| $x_1 = 0.6$ | $v_1 = 0.35$, $v_3 = 0.25$, $v_2 = 0.40$ | $\{13, 23, 13\}$ | $\{13, 23, 13\}$ |
| $x_1 = 0.5$ | $v_1 = 0.35$, $v_3 = 0.30$, $v_2 = 0.35$ | $\{12, 12, 13\}$ | $\{13, 23, 13\}$ |
| $x_1 = 0.4$ | $v_1 = 0.35$, $v_3 = 0.35$, $v_2 = 0.30$ | $\{12, 12, 13\}$ | $\{12, 12, 23\}$ |
| $x_1 = 0.3$ | $v_1 = 0.35$, $v_3 = 0.40$, $v_2 = 0.25$ | $\{12, 12, 23\}$ | $\{12, 12, 23\}$ |
| $x_1 = 0.2$ | $v_1 = 0.275$, $v_3 = 0.45$, $v_2 = 0.275$ | $\{12, 12, 23\}$ | $\{12, 12, 23\}$ |
| $x_1 = 0.1$ | $v_1 = 0.15$, $v_3 = 0.45$, $v_2 = 0.40$ | $\{12, 12, 13\}$ | $\{12, 12, 13\}$ |
| $x_1 = 0$ | $v_1 = 0.10$, $v_3 = 0.45$, $v_2 = 0.45$ | $\{12, 12, 13\}$ | $\{13, 12, 13\}$ |

with the candidate closest to their own ideal point in policy space (the coalition profile is $\{12, 12, 13\}$). Once we move to the case in which $\beta = 0$ and $\phi = 1$, i.e., in which candidates are purely office-seeking, ideological proximity no longer plays a role in candidates' coalition proposals: since all they care about are the spoils of office, they always choose to coalition with the smaller of their two competitors, regardless of that competitor's ideological position.

Define $\hat{x}_1(E)$ as the expected policy utility candidate 1 receives if she enters. For example, for the case in which $x_1 = 0.7$, $\beta = 0$, and $\phi = 1$ we know that the coalition profile will be $\{13, 23, 13\}$. Furthermore, we know that candidates 1, 2, and 3 will be named formateur with probabilities $v_1 = 0.35$, $v_2 = 0.45$, and $v_3 = 0.2$ respectively. In turn, this tells us that $\pi_{13}(E) = 0.55$ and $\pi_{23}(E) = 0.45$. In the event that $G = 13$, candidate 1 receives policy utility $-\left(x^*(13) - x_1\right)^2$; in the event that $G = 23$, candidate 1 receives policy utility $-\left(x^*(23) - x_1\right)^2$. As such, we can write the expected policy utility 1 receives if she decides to enter as follows: $\hat{x}_1(E) = -0.55 \cdot \left(x^*(13) - x_1\right)^2 - 0.45 \cdot \left(x^*(23) - x_1\right)^2$. A similar computation can be made for any values of $x_1$, $\beta$, and $\phi$.

When $\phi = 0$ the value $\hat{x}_1(E)$ will have no impact on candidate 1's decision to enter the race or not, since policy outcomes are completely discounted in her utility function. On the other hand, for $\phi > 0$ the potential entrant must consider the impact of her entry decision on the game's expected policy. By construction, if she chooses $\sim E$ candidate 2 wins the election with certainty and implements her ideal point, which implies that 1's policy utility from not entering can be written as $\hat{x}_1(\sim E) = -(x_2 - x_1)^2$. We can thus compare the relative policy utility associated with the decisions $E$ and $\sim E$ using the metric $\hat{x}_1(E) - \hat{x}_1(\sim E)$. If this metric is greater than zero, candidate 1 receives a *policy benefit* from entering, i.e., she prefers the expected policy outcome which follows the decision $E$ to having $x_2$ implemented with certainty. If this metric is less than zero, candidate 1 pays a *policy cost* for entering, i.e., she prefers having $x_2$ implemented with certainty to the
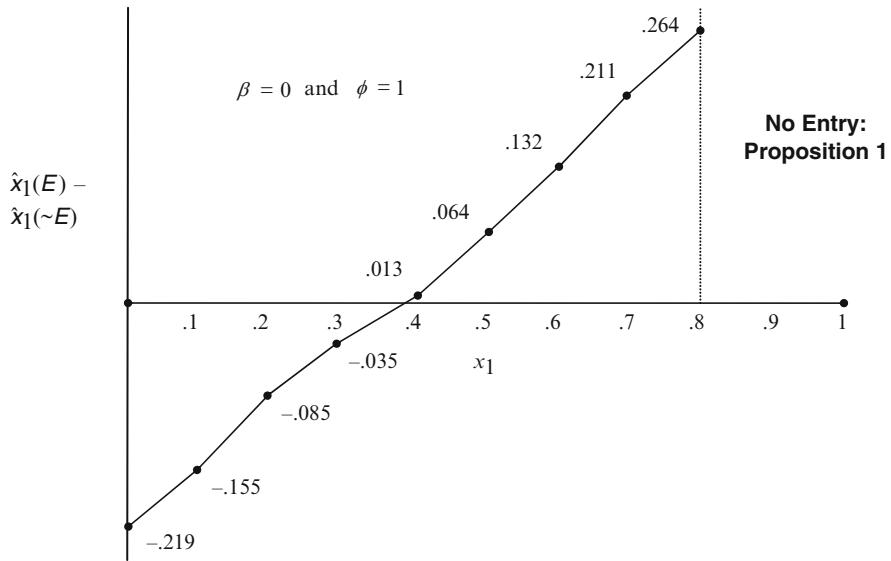
**Fig. 1** Policy consequences of entry ($x_2 = 0.2$ and $x_3 = 0.9$)

expected policy outcome following the decision $E$. For the case in which $\beta = 0$ and $\phi = 1$, Fig. 1 plots $\hat{x}_1(E) - \hat{x}_1(\sim E)$ for various values of $x_1$.

As is in clear from the Figure, for values $x_1 < 0.8$ (at which entry leads to coalition bargaining) there is a monotonic relationship between the potential entrant's ideal point and the policy consequences of entry: policy consequences of entry become less and less appetizing as leader 1's ideal point $x_1$ moves further to the left, and thus becomes more distant from $x_3$. The intuition is as follows: as 1's ideal point $x_1$ becomes more and more distant from $x_3$, two things happen. First, party leader 1's distaste for $x_3$ increases; and secondly $v_3$ increases, thus increasing the likelihood that candidate 3 will be named formateur following the decision $E$. Eventually, a point is reached at which the single-party government associated with the choice $\sim E$ is actually preferred to the to the expected policy outcome of coalition bargaining after the choice $E$, given the prospect of empowering as formateur a large party of competing ideological persuasion.

A similar expected utility comparison can be conducted for the case in which $\beta = 1$ and $\phi = 0$. Here, by definition the utility associated with the decision $\sim E$ for candidate 1 is 0: she doesn't care about policy outcomes, has no chance of securing the spoils office, and pays no entry costs. On other hand, as long as $x_1 \in [0, 0.8]$, by entering candidate 1 guarantees herself at least a probability of executive participation. Consider the case in which $x_1 = 0.7$, $\beta = 1$, and $\phi = 0$. In this case, the operative coalition profile is once again $\{13, 23, 13\}$. As such, we know that $\pi_{13}(E) = 0.55$ and $\pi_{23}(E) = 0.45$. In turn, since $\omega_1(13) = 7/11$ and $\omega_1(23) = 0$, candidate 1 receives an expected office utility of $0.55 \cdot \left(\frac{7}{11}\right) + 0.45 \cdot (0) = 0.35$.
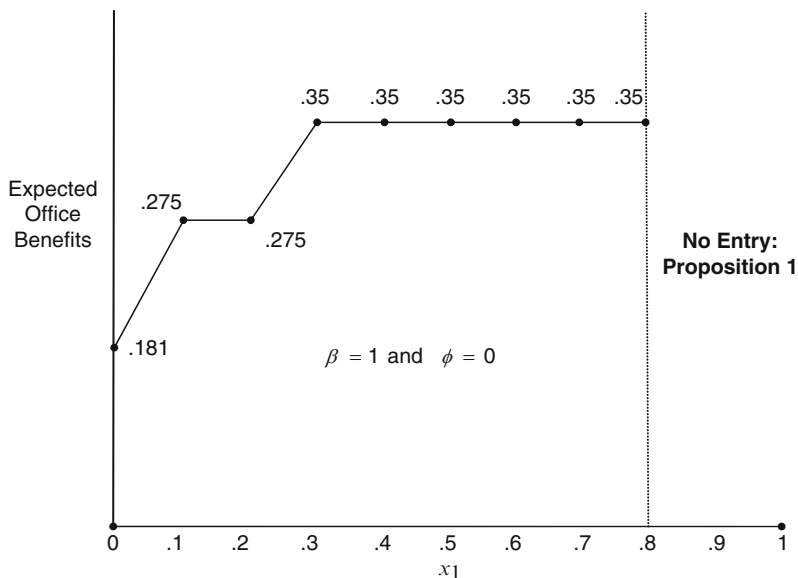
**Fig. 2** Office benefits of entry ($x_2 = 0.2$ and $x_3 = 0.9$)

Figure 2 plots the expected office utility associated with entry by candidates at various positions of $x_1$.

Although the trend is not as strong, we notice again that incentives to enter are weakly decreasing as $x_1$ becomes more leftist. The intuition here is that for any policy $x_1 \leq x_2$ the candidate 1's vote share $v_1$ will drop precipitously, reducing her weight and subsequent office payoff in any coalition.

Based on the information in Figs. 1 and 2 we can specify the game's SPNE outcomes when $x_2 = 0.2$ and $x_3 = 0.9$ for the polar cases in which either $\beta = 0$ or $\phi = 0$. Begin with the case in which $\beta = 0$ and $\phi = 1$. By Proposition 1 we know that candidate 1 never enters if $x_1 > 0.8$, such that in the game's the unique SPNE $s_1^* = \{ \sim E \}$, $s_2^* = \{ 2, 2 \}$, and $s_3^* = \{ \phi, \phi \}$. Similarly, if $x_1 < 0.374$, then candidate 1 actually prefers having the policy $x_2$ implemented with certainty to receiving the expected policy utility associated with the decision to enter. Since $\beta = 0$, i.e., we are dealing with a world in which only policy outcomes matter, candidate 1 will have no incentive to enter in this policy range, and the game's unique SPNE involve non-entry. On the other hand, when $0.374 < x_1 < 0.8$, the policy consequences of entry are positive, and the comparison between the policy benefits and the costs of entering $c$ becomes relevant. For $x_1$ in this range, if $c < \hat{x}_1(E) - \hat{x}_1(\sim E)$ then in the game's unique SPNE candidate 1 chooses to enter. For example, when $\beta = 0$, $\phi = 1$, $x_1 = 0.7$ and $c < 0.211$, the game's unique SPNE will be $s_1^* = \{ E, 13 \}$, $s_2^* = \{ 2, 23 \}$, and $s_3^* = \{ \phi, 13 \}$. Similarly, when $\beta = 0$, $\phi = 1, x_1 = 0.5$, and $c < 0.064$, then the game's unique SPNE will be SPNE $s_1^* = \{ E, 12\}$, $s_2^* = \{ 2, 12 \}$, and $s_3^* = \{ \phi, 13 \}$. In either the case, if the

relevant restriction on $c$ is not met, then the game's unique SPNE includes $s_1{}^* = \{\sim E\}$. Naturally, the restriction becomes harder to satisfy as the potential entrant's ideal point $x_1$ becomes more leftist: as the policy benefits of entry decrease, so do the incentives to pay $c$ for the right to compete.

Move now to the case in which $x_2 = 0.2$ and $x_3 = .9$, but where $\beta = 1$ and $\phi = 0$. Here, it continues to be the case that candidate 1 never enters if $x_1 > 0.8$, such that in the game's the unique SPNE $s_1{}^* = \{\sim E\}$, $s_2{}^* = \{2, 2\}$, and $s_3{}^* = \{\phi, \phi\}$. However, unlike the purely policy-seeking case in which the possibility of entry was confined to a set of minimally centrist positions, here entry may occur at any point $x_1 \in [0, 0.8]$, and furthermore the entry criterion on the costs of entry $c$ is easier to satisfy than it was in the policy-seeking case. For example $x_1 = 0.7$ and $c < 0.35$, the game's unique SPNE will be $s_1{}^* = \{E, 13\}$, $s_2{}^* = \{2, 23\}$, and $s_3{}^* = \{\phi, 13\}$. Similarly, when $x_1 = 0.1$and $c < 0.275$, the game's unique SPNE will be $s_1{}^* = \{E, 12\}$, $s_2{}^* = \{2, 12\}$, and $s_3{}^* = \{\phi, 13\}$. In either the case, if the relevant restriction on $c$ is not met the game's unique SPNE will include $s_1{}^* = \{\sim E\}$. Once again, the restriction becomes 'weakly' harder to satisfy as the potential entrant's ideal point $x_1$ becomes more leftist.

Generally speaking, for the case in which $x_2 = 0.2$ and $x_3 = 0.9$ we can thus say that entry becomes less likely the more leftist is the potential entrant's ideal point, and the more this entrant weights office as compared to policy in her utility calculations. The first point can be restated in slightly more precise sense: if the 'would-be' plurality winner of a two-party contest between candidates 2 and 3 is on the ideological left (right), entry becomes less likely the more 'left-leaning' ('right-leaning') is the potential entrant's ideal point. We now repeat the entire exercise for the case in which status quo parties are significantly more moderate: $x_2 = 0.5$ and $x_3 = 0.6$. For reasons of redundancy we do not present the equivalent of Table 1 again, and instead move straight to the SPNE analysis.[11] For $\beta = 0$ and $\phi = 1$, Fig. 3 plots the metric $\hat{x}_1(E) - \hat{x}_1(\sim E)$ for various values of $x_1$.

For values of $x_1 > 0.5$ the metric $\hat{x}_1(E) - \hat{x}_1(\sim E)$ is simply 0, since entry has no effect on the government outcome (single-party government by candidate 2). For all values of $x_1 < 0.5$, the choice to enter carries with it very mild policy costs. These costs are small compared those uncovered above, since in a party system with two centrist status quo organizations the policy outcome will be fairly centrist regardless. Nonetheless, it is the case that entry in the stipulated conditions *never carries with it a policy benefit*. As such, regardless of the position of $x_1$, the unique SPNE to the game in which $x_2 = 0.5$, $x_3 = 0.6$, $\beta = 0$, and $\phi = 1$will *never* involve entry.

Move now to the case in which $x_2 = 0.5$ and $x_3 = 0.6$, but where $\beta = 1$ and $\phi = 0$. Paralleling Fig. 2 above, Fig. 4 plots the expected office utility associated with entry by candidates at various positions of $x_1$.

Here, for values $x_1 > 0.5$ entry has no effect on the government outcome, such that in the game's the unique SPNE $s_1{}^* = \{\sim E\}$, $s_2{}^* = \{2, 2\}$, and

---

[11]The vote shares and coalition numerics which inform the following figures are available on request.
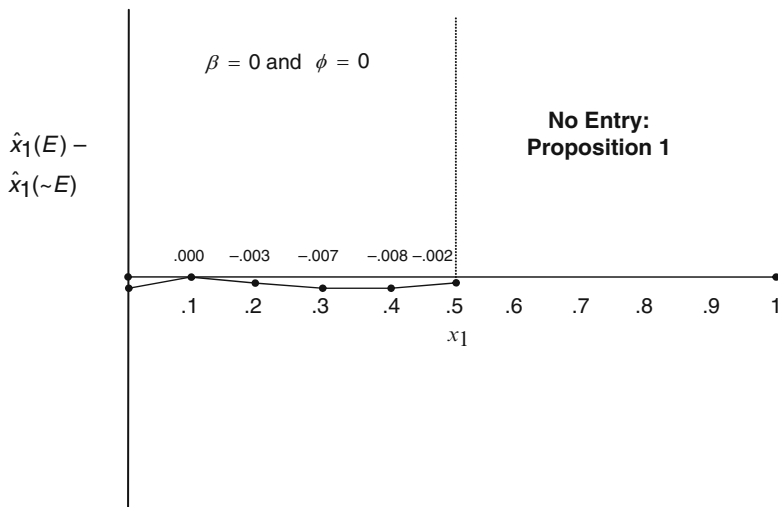
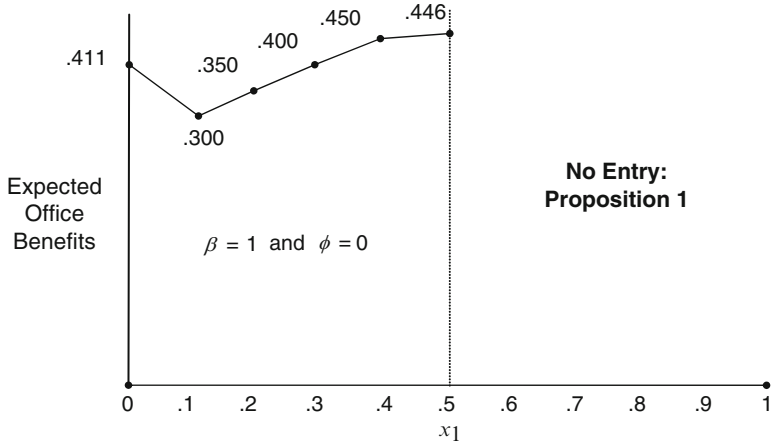**Fig. 3** Policy consequences of entry ($x_2 = 0.5$ and $x_3 = 0.6$)



**Fig. 4** Office benefits of entry ($x_2 = 0.5$ and $x_3 = 0.6$)

$s_3{}^* = \{\phi, \phi\}$. However, unlike the purely policy-seeking case in which entry never occurred, entry can occur at any value of $x_1 < 0.5$; furthermore, the entry criterion on the costs of entry $c$ is easier to satisfy than it was for the office-seeking case in which $x_2 = 0.2$ and $x_3 = 0.9$. For example $x_1 = 0.3$ and $c < 0.4$, the game's unique SPNE will be $s_1{}^* = \{E, 12\}$, $s_2{}^* = \{2, 12\}$, and $s_3{}^* = \{\phi, 23\}$. Similarly, when $x_1 = 0.1$ and $c < 0.3$, the game's unique SPNE will be $s_1{}^* = \{E, 12\}$, $s_2{}^* = \{2, 12\}$, and $s_3{}^* = \{\phi, 23\}$. In either the case, if the relevant restriction on $c$ is not met the game's unique SPNE implies that $s_1{}^* = \{\sim E\}$.

These examples allow us to provide some preliminary answers to the questions posed at the paper's outset: Is entry more or less-likely when status quo organizations are widely dispersed around the median voter's ideal point? How will the incentives for entry vary with candidates' relative emphasis on 'policy-seeking' in contrast to 'office-seeking'? How will the location of a potential entrant's ideal point affect her incentives to join the election? In fact, what we uncover here are predictions in which party system elements interact with candidates' utility parameters in determining the incentives for entry. For example, in a purely policy-seeking world, entry *never* occurred when the status quo parties were both moderate, because it never led to expected policy improvements. On the other hand, when both status quo parties were fairly extreme, there were a fixed range of centrist positions at which entry by candidate 1 implied policy benefits. As such, in a purely policy-seeking world we might conjecture that entry will be more likely in dispersed than in convergent party systems, a prediction which directly contradicts past results from the 'entry-deterring-dispersion' literature.

Things are less clear when candidates were purely office-seeking. On the one hand, the range of ideal points $x_1$ over which entry was feasible is higher when candidates are more dispersed (by Proposition 1). However, within the confined space in which entry is possible when both candidates were moderate ($x_1 < 0.5$), the condition on the cost parameters $c$ necessary for entry to be optimal was easier to satisfy than it was when status quo parties were more extreme. Thus, we conjecture that moving from a convergent to a dispersed status quo party system in a purely office-seeking world increases the likelihood of entry for potential entrants with ideal points in the range $0.5 < x_1 < 0.8$; conversely it decreases the likelihood of entry when $x_1 < 0.5$. Furthermore, note that the tendency for entry to decrease in likelihood as the position $x_1$ becomes more leftist does not emerge in the case with two moderate status quo parties. Finally, and fairly trivially, the overall likelihood of entry increases as candidates weight office more heavily than policy.

## 4   Conclusion

This paper develops a game theoretic model of party entry which builds off of elements of the entry-deterring dispersal and citizen-candidate literatures, but also departs from these literatures by embedding our model in a richer institutional framework by studying party entry in Parliamentary regimes with Proportional Representation that include a stage of government formation. Via a series of examples, it provides some preliminary insights into the strategic and parametric conditions which make entry more or less likely. Naturally, future research will generalize this analysis, solving the game in all possible strategic and parametric conditions so as to provide more exhaustive answers to the paper's stated questions. In addition, we look forward to extending the model to alternative institutional environments. Note that while our express interest here was developing a model applicable to parliamentary regimes with proportional representation, the very same set of

assumptions as to candidate preferences and action sets can be investigated in a pure plurality rule context. Furthermore, by altering our assumptions as to the relationship between vote shares and seat shares, as well as the relationship between seat shares and formateur probabilities, we can begin to investigate a multiplicity of institutionally distinct parliamentary/proportional regimes. Finally, we have already undertaken some initial work in the direction of extending the model to situations with more than two status quo parties, uncovering a fairly strong tendency for entry to become more likely as the effective number of status quo parties increases. The material presented in the current paper will form the foundation of these subsequent analyses.

# References

Baron DP, Ferejohn JA (1989) Bargaining in legislatures. Am Polit Sci Rev 83:1181–1206

Bassi A (2008) A model of endogenous government formation. Manuscript, New York University Department of Political Science

Besley T, Coate S (1997) An economic model of representative democracy. Q J Econ 112:85–114

Black D (1958) The theory of committees and elections. Cambridge University Press, Cambridge

Carroll R, Cox GW (2007) The logic of Gamson's law: pre-election coalitions and portfolio allocations. Am J Polit Sci 51:300–313

Downs A (1957) An economic theory of democracy. Harper and Row, New York

Eguia J (2007) Citizen candidates under uncertainty. Soc Choice Welfare 29:317–331

Gamson WA (1961) A theory of coalition formation. Am Soc Rev 26:373–382

Greenberg J, Shepsle K (1987) The effect of electoral rewards in multiparty competition with entry. Am Polit Sci Rev 81:525–538

Hotelling H (1929) Stability in competition. Econ J 39:41–57

Hug S (2001) Altering party systems. University of Michigan Press, Ann Arbor

Laver M, Shepsle KA (1996) Making and breaking governments: cabinets and legislatures in parliamentary democracies. Cambridge University Press, Cambridge

Lee D (2007) Anticipating entry: major party positioning and third party threat. Manuscript, Duke University Department of Political Science

Martin LW, Stevenson RT (2001) Government formation in parliamentary democracies. Am J Polit Sci 45:33–50

Osborne MJ (2000) Entry-deterring policy differentiation by electoral candidates. Math Soc Sci 40:41–62

Osborne MJ, Slivinsky A (1996) A model of political competition with citizen candidates. Q J Econ 111:65–96

Palfrey TR (1984) Spatial equilibrium with entry. Rev Econ Stud 51:139–56

Strom K (1990) A Behavioral Theory of Competitive Political Parties. Am J Polit Sci 34 (2):565–598

# Moving in Time: Legislative Party Switching as Time-Contingent Choice

**Carol Mershon and Olga Shvetsova**

What would induce a sitting legislator to leave the party under whose banner she has won election, and to join another party? The premise here is that a politician's strategic calculus on party affiliation involves not only what goods she stands to gain or lose, but also what times are best or worst to get the goods. In investigating when incumbents switch party during the legislative term, we shed new light on why they switch.

Distinguished traditions of research on party competition for votes and government have been grounded in the assumptions that parties operate as unitary actors whose legislative memberships are fixed from one national legislative election to the next. In recent years, a growing literature has relaxed these long unquestioned assumptions in order to explain the phenomenon of party switching. Scholars now concur that multiple factors motivate individual representatives to jump from one legislative party into another. The available theoretical and empirical work points in particular to the benefits of office, policy influence, and electoral advantage as incentives to abandoning the party ship (e.g., Desposato 2006; Desposato and Scheiner 2008; Heller and Mershon 2005, 2008, 2009c; Laver and Benoit 2003; Mershon and Shvetsova 2008, 2009a, b; Reed and Scheiner 2003; Schofield 2009). This literature shows that, trivial exceptions aside, individuals' moves from one party to another effect system-level changes in the balance of power among legislative parties and in the array of preferences within parties (cf. Heller and Mershon 2009a, b). We thus take it as established that members of parliament (MPs) can move between parties while in office, and that, if they do so, they move in pursuit of advantage.

C. Mershon (✉)
Political Science Program, National Science Foundation, Arlington, VA, USA
and
Department of Politics, University of Virginia, Charlottesville, VA 22904, USA
e-mail: camershon@gmail.com

O. Shvetsova
Department of Political Science, Binghamton University, P.O. Box 6000, Binghamton, NY 13902, USA
e-mail: shvetsova@gmail.com

We emphasize as well that political parties are central to democratic governance. What a party does and seeks to do depend in turn to no small degree on the people who represent the party in elective office. Incumbent legislators are responsible for translating party promises into enacted policy and for overseeing the implementation of policy once approved. Hence understanding what drives legislators' decisions on party membership holds one key to comprehending how democracy works.

The chapter is organized as follows. The first section presents our theoretical model, which demonstrates that a legislative incumbent engages in a strategic calculus to time a shift in party allegiance in order to minimize losses and maximize gains. The second part of the chapter offers empirical illustrations of our logic. The concluding section discusses the implications of our argument.

# 1   The Strategic Timing of Changes in Party Affiliation

Our model of parliamentary parties as endogenous coalitions of incumbents is premised on the notion that the incumbent's choice of party should maximize her utility in any given period (e.g., month). This assumption accords with scholarly wisdom, of course. The novelty here is that we use this assumption as the basis for analyzing the strategic timing of party affiliation. As we show, at some moments, the MP expects that voter scrutiny is keenest, so that a switch of allegiance will likely cost her electoral support; the incumbent then stays with her original party. At other moments, the legislator calculates that a response to beneficial opportunities knocking at the door will likely invite little damage, and so she steps out and moves into another party.

## 1.1   The Utility Function of Incumbent i, the Potential Switcher

Any decision to move among parties derives from the cost–benefit calculations made by individual legislators. A politician chooses to transfer to a different party whenever the benefits from the move outweigh the costs. We can express the utility function of incumbent MP $i$ as incorporating the three types of rewards that the literature commonly identifies as politicians' goals: policy, office (and its perks), and election (e.g., Desposato 2006; Heller and Mershon 2009c; Reed and Scheiner 2003; Schofield 2009; cf. Strøm 1990 on parties). That is,

$$u_i(\tau, \lambda, \theta) = \tau \sum_{k=1}^{n} p_k^i + f(\lambda, \theta), \qquad (1)$$

where $\tau$ is the parameter attached to the incumbent's prospects for re-election, expressed as the summation across all voters of the probabilities that an individual voter in the electorate will vote for incumbent $i$ in the next election. The parameter

or weight $\lambda$ indicates the importance to an MP of policy gains, while $\theta$ is the weight attached to the benefits of office, which include retention of the legislative seat, perks, and positions within the legislature (such as committee posts) and, in parliamentary systems, within the executive. We assume that each parameter exceeds 0 and can vary to 1: $0 < \tau, \lambda, \theta \leq 1$.

We now develop the model by addressing specifics on its three main moving parts, all time-contingent. We start with re-election and so treat voter preferences.

## 1.2 Voters' Calculus: Agency Risks and Rewards for Partisan Constancy

Voters balance two tasks when evaluating their prospective representatives on the eve of an election: they assess the policy position that a particular candidate promises; and they judge the policy risk associated with that candidate. The second task arises from the possibility that the policy the representative will actually pursue once in office might diverge from her policy promises to voters at election time.[1] The policy variance associated with the candidate reflects both the limitations on the voter's information about the candidate's true policy position and the risk of the candidate's deviation from the platform on which she was elected. The partisan history of the candidate influences both components of policy variance.

Thus we specify the utility function of voter $k$ as reflecting the importance of policy and variance associated with candidate $i$. It can be expressed as

$$u_k^i = u_k(x_i) - \alpha\sigma_i^2(I_i, L_i), \qquad (2)$$

where $x_i$ is candidate $i$'s policy position and the policy variance of candidate $i$ is $\sigma_i^2(I_i, L_i)$. In turn, $I_i$ is an indicator function taking the value of 1 if a candidate belongs to a political party and 0 otherwise. We assume that the fact of partisanship gives a partisan candidate a lower perceived policy variance than the perceived policy variance of an independent competitor. $L_i$ is a loyalty variable indicating the dependability of the candidate's current party membership in the voter's eyes; it is a personal attribute of an incumbent that depends on his partisan behavior during the term, that is, on his history of moving in and out of parties. Loyalty (the degree to which partisanship remains constant) also serves to reduce variance and hence to increase the incumbent's electoral appeal; changing parties instead reduces $L_i$, and so increases $\sigma_i^2(I_i, L_i)$). Denote the utility to a voter from unquestionably loyal incumbents (those who have never switched in a given term) as $L_o(\alpha) > 0$. The utility loss felt by a voter at the moment immediately after an

---

[1]Voters might well also consider other attributes of candidates that do not pertain to policy, such as personal qualifications for office and demographic traits. We assume that these variables enter the utility function of a voter directly, and do not change with a politician's change of party.

incumbent's switch in month $j$ of the term due to the erosion of her perception of the incumbent's loyalty and thus to an increase in the variance associated with this incumbent comes from two sources: the MP has betrayed his status-quo or electoral party, which demonstrates the opportunistic nature of the incumbent's affiliation in the first place; and the novelty of the MP's new party affiliation means that it has unproven strength.

Denote this utility loss due to the switch as $\Delta L(\alpha) < 0$. To unpack the impact of the two factors on the voter's utility loss, we make two reasonable assumptions. First, the reputation loss for the MP who abandons his status-quo party will be greatest if the switch occurs as close as possible to the election initiating the legislative term – in the first month of the life of the legislature. Second, we assume that the reputational damage diminishes with the distance from the initiating election by a factor of $0 < \phi < 1$ with each passing month of the term, so that a switch that takes place in month $j$ will signify betrayal of the status-quo party of the incumbent with the factor of $\phi^j$. Intuitively, $\phi^j$ shows the degree to which the voters view the evidence of the switcher's disloyalty to the original party label as disappointing the voters' trust, given that the switch occurs $j$ months from the moment that voters elected the legislature.

Since a history of past switching increases the negative term in expression (2), it reduces the utility to voters from supporting candidate $i$ if her policy position, $x_i$, is fixed.[2] This holds regardless of the value of $\alpha$, the salience that the voter assigns to the candidate's policy variance (given that $\alpha > 0$, by assumption). Note that the more voters focus on candidates rather than parties when making their vote choice, the higher $\alpha$ is; the value of $\alpha$ thus depends on electoral laws. In other words, the consistency of an MP's partisan attachment is especially valuable where it matters the most – where elections are candidate-centered and where rules require voters to evaluate the promise of individual candidates (e.g., Carey and Shugart 1995).

So far we have discussed the voter's appraisal of an incumbent's switch as hinging on the move's timing relative to the election starting a given legislative term. Now consider how the voter also assesses a switch according to its temporal proximity to the election closing the term. Denote as $T^i = \{t_1^i, ..t_m^i\}$ the history of interparty moves by incumbent $i$, measured by the temporal distance between the move and the month of the election, $m$, that ends the given term. Define a generic component of $T^i$ $t_j^i = 1$ if a switch took place in month $j$ of the term and 0 otherwise. The loyalty parameter with which incumbent $i$ approaches the next election is thus

---

[2]The discussion here suggests that the loyalty variable has two dimensions: (a) stability of party affiliation, the focus of this paper; and (b) propensity to vote with the party of record. That is, if $x_i$ is stable, then changing parties implies either that the (old and new) parties occupy the same observed positions in voting terms (always vote together) or that the switcher must necessarily regularly buck the party line in voting in at least one of the parties in which her membership defines $L_i$. Investigation of (b) lies outside the scope of this paper.

functionally linked to the history of that incumbent's inter-party moves during the term as follows:

$$L_i(t_1^i, ..t_m^i) = L_o + \sum_{j-1}^{m} t_j^i(\delta^{m-j}\Delta l + \phi^j\Delta L). \tag{3}$$

In this expression, terms $\phi^1, ...\phi^m \in (0, 1)$ are, again, the weights on the loss of loyalty arising from switching, respectively, in months 1 to $m$ of the term. An MP who jumps party in the same month as the election inaugurating the term, $t_1$, will suffer the maximal reputation loss.

Such loss diminishes, month by month, as time elapses after the initiating election. Terms $\delta^{m-1}, ...\delta^{m-m} \in (0, 1)$ are the weights on the loyalty loss parameter that indicate the length of time the switcher has been enrolled in her new party and thus the extent to which she has been able to compile a history of attachment with it so as to compensate for the disloyalty to her old party. Maximal loyalty loss due to the unproven loyalty to the current party is inflicted if the switch takes place in month $m$ of the term, and it is weighed by $\delta^{m-j} < 1$ if the switch took place in month $j$.

The voter's perception of a candidate's loyalty thus declines with the number of interparty moves and with the moves' temporal proximity to either the election closing a given term or the election opening the term. The voter's overall utility of voting for the candidate also declines with an increase in the number of moves and the temporal proximity of the move(s) to the two elections bracketing a term.

We now express the probability of voter $k$ voting for a candidate $i$, $p_k^i(\bullet)$, as a function of $k$'s utility from that candidate, $u_k^i(x_i, I_i, L_i)$, which in turn depends on the candidate's policy, partisan status, and past history of consistent partisanship. Assuming that the higher the utility the more likely the voter is to vote for the candidate, we observe that $p_k^i(x_i, I_i, T^i)$ declines in $\sum_{j-1}^{m} t_j^i(\delta^{m-j}\Delta l + \phi^j\Delta L)$. That is, the more times the candidate switched in the past, and the closer the move(s) occurred to either of the elections bracketing the parliamentary term (either the election initiating or that ending a term), the less likely, all else equal, the voter is to vote for that candidate. What does this result imply for the calculations and behavior of strategic incumbents?

## 1.3  Constraints on the Timing of Switches Due to Incumbents' Re-election Calculus

An incumbent who contemplates a move from one party to another compares two states, one actual and the other not (yet) observed. In our framework, she compares the net benefits available to her from membership in her current party $q$ and the estimated net benefits to be had from belonging to an alternative party $r$. Only when party $r$ promises greater benefits, will the incumbent carry out the transfer

considered. *If* the MP's re-election prospects were unaffected by a switch, she would move whenever $f(\lambda_r, \theta_r) > f(\lambda_q, \theta_q)$.

Yet that "if" does not apply, and that inequality is woefully incomplete. Once more, voters are less likely to vote for switchers than for loyalists, with the probability of voting depending on the timing and number of moves. To see the impact of a move on the probability of an incumbent's re-election, the first step is to incorporate the probability of receiving voter $k$'s vote into the utility function of incumbent legislator $i$. Expression (1) thus becomes:

$$u_i(\lambda, \theta, \tau) = f(\lambda, \theta) + \tau \sum_{k=1}^{n} p_k^i(x_i, I_i, T^i). \tag{4}$$

Assuming that all voters assign equal value to partisanship and loyalty, and that those factors additively contribute to the probability of voter $k$ voting for incumbent $i$, we can rewrite the above expression as

$$u_i(\cdot, \tau, T^i) = f(\lambda, \theta) + \tau \sum_{k=1}^{n} g_k^i(x_i, x_k) + \tau \alpha \sigma^2(I_i, T^i). \tag{5}$$

The incumbent politician's utility in expression (5) is a function of the parliamentary (policy and office) benefits she receives, $f(\lambda, \theta)$, and of her expected vote in the next election, $\tau \sum_{k=1}^{n} g_k^i(x_i, x_k) + \tau \alpha \sigma^2(I_i, T^i)$. Her expected vote, in turn, reflects voters' reaction to her history of party loyalty – that is, steadfast (vs. inconstant) party affiliation – in addition to her policy proximity to her voters. We use $g_k^i(x_i, x_k)$ to designate the policy-related element of the probability that voter $k$ would vote for candidate $i$. For each individual voter, this is a function of the distance between the policy locations of the voter and the candidate, $|x_i, x_k|$, where the policy location of the voter is her ideal point, and the policy location of the candidate is his perceived policy platform. Expression $\alpha \sigma^2(I_i, T^i)$ is the component in a voter's probability of voting for candidate $i$ that depends on the candidate's policy variance.[3] For convenience, we assume that all voters, regardless of their policy ideals, respond to candidates' variance in a similar way, which allows us to take $\alpha \sigma^2(I_i, T^i)$ from under the summation sign and treat it as an additive component of the utility function of an incumbent politician. This is useful, since this component is the only part of the politician's utility (negatively) affected by her history of switching during the term.

Note that $\alpha \sigma^2(I_i, T^i)$ and thus also $u_i(\cdot, \tau, T^i)$ are declining in $\sum_{j-1}^{m} t_j^i(\delta^{m-j} \Delta l + \phi^j \Delta L)$. The more recently and frequently an incumbent has changed parties near the initiating and the upcoming elections, the greater the utility loss she will suffer,

---

[3]Our model does not incorporate variation across party systems in relationships between candidates' policy stances and party platforms.

other things equal. For incumbent legislator $i$ in party $q$, who has remained loyal to the original electoral label:

$$u_i(\lambda q, \theta q, \tau q) = f(\lambda q, \theta q) + \tau \sum_{k=1}^{n} p_k^i(x_i, I_i, L_0). \tag{6}$$

For the mobile incumbent now in party $r$, who switched just once during the given legislative term, in month $j$, this expression becomes

$$u_i(\lambda r, \theta r, \tau r) = f(\lambda r, \theta r) + \tau \sum_{k=1}^{n} p_k^i(x_i, I_i, (L_o + \delta^{m-j}\Delta l + \phi^j \Delta L)). \tag{7}$$

Observe that expression (7) depicts the single move made by the recruit to party $r$ as occurring before $m$, the month of the election for the subsequent legislature, yet after $t_1$, the month of the term coinciding with election initiating the legislature. The logic here indicates that the minimum in electoral damage conditional on executing the switch would correspond to the maximum of $L_i(j) = L_o + \delta^{m-j}\Delta l + \phi^j \Delta L$, or, since $\Delta l, \Delta L < 0$, to the minimum of $\delta^{m-j}\Delta l + \phi^j \Delta L$. Differentiating by $j$, we obtain

$$\partial[\delta^{m-j}\Delta l + \phi^j \Delta L]/\partial j = \phi^j \log \phi \Delta L - \delta^{m-j} \log \delta \Delta l. \tag{8}$$

This function attains the value of zero somewhere in the middle of the parliamentary term, i.e., the interval $[1; m]$. To illustrate, setting $\delta = \phi, \Delta l = \Delta L$, it becomes zero at $\phi^j \log \phi \Delta L = \phi^{m-j} \log \phi \Delta L$, or at $j = m/2$.

This argument thus leads to the conclusion that legislative incumbents' utility loss from switching parties is minimized and benefits are maximized, all else equal, if moves are timed near the middle of a legislative term.

## 2 Empirical Illustrations: Midterm Mobility Versus Stability Near Election Time

We now turn to empirical examination of the argument we advance. We do not concern ourselves with addressing the absence of change in legislative party membership. Given ingrained scholarly assumptions, that job has been amply performed to date. Nor do we seek to replicate the findings of the new literature on legislative party switching, which, as noted, has shown that legislative incumbents do exit one party and enter another in the quest for office, policy, or electoral advantage. What the literature lacks, with few exceptions, is empirical investigation of incumbents whose strategy extends to the timing of moves among parties (Mershon and Shvetsova 2009a,b).

Our empirical analysis focuses on a small number of country-terms, which gives us access to multiple types of evidence on each term and allows for relatively

confident inferences about posited strategic behavior. The central criterion for our choice of terms is the design of electoral laws, which in our argument, again, determine the salience of candidates' policy variance in the eyes of voters and thus voter preferences for incumbents' party loyalty. We approach single-member district (SMD) rules as the key institutional parameter that isolates voter attention to individual candidates rather than a party slate. Our case selection guarantees variation in electoral laws. Beyond that, the particular choice of terms for our plausibility probes hinges on the knowledge we have of the within-term dynamics of legislative party memberships, given prior and in-progress work (large-N analyses in Mershon and Shvetsova 2009a; case studies in Mershon and Shvetsova 2008, 2009b; Schofield 2009). The reigning scholarly presumption is still of stasis in legislative parties, despite the contributions of the new research on switching. We know where and when stasis did not hold, and the point is to discover if incumbents treated the decision to change party as a time-contingent choice.

## 2.1 SMD Rules: United States and United Kingdom

We begin by investigating the US 104th Congress, which served from 1995 to 1997. During this term, the US House of Representatives witnessed a total of five switches, the most of any term from 1953 to 2002 (Nokken and Poole 2004). The first Member of Congress (MC) to defect did so 5 months after the election, in April 1995. Two representatives next crossed the aisle in, respectively, June and August 1995. The last two switched at roughly midterm, in November and December 1995.[4]

More broadly, of the total of 19 House switchers from 1950 to 2000 counted by Nokken (2009), only one moved after mid-March of a term's second year. The lone late switcher faced an autumn primary (United Press International 1984) and saw that his district had, like him, "become more Republican" (Gillespie in Hickey 2003). He won re-election as "part of the Reagan landslide" (his post-switch campaign manager, Karl Rove, quoted in Gillespie 2005).[5]

---

[4]Data sources are as follows: Ballot Access News (2001a,b), Butler and Butler (2006, 112–114), Camera dei Deputati (2010), Congressional Biographical Directory 2008, INDEM Foundation (2000).

[5]It might be objected that an alternative explanation for the near-absence of late switchers is grounded in MCs' need to compete in primary elections; that is, MCs view the primary, not the general election, as the decisive contest on the horizon that subjects them to voters' judgments at what is, electorally, the de facto end of a term. We note that, if MCs were to try to weaken voter support for challengers in primaries or avoid a contested primary altogether, this would push the optimal timing of a switch to roughly July of the first year (since House primaries run from March through September of the second year, Federal Voting Assistance Program 2010); indeed, primary filing deadlines might bring the optimal date somewhat earlier. Given the timing of actual switches, this alternative and our argument constitute observationally equivalent explanations for the US case. We emphasize, however, that the US institution of primary elections is unusual in comparative perspective, so that ours is the more general explanation of the two.

**Table 1** Mean number of monthly moves per 100 MPs, by 6-month span in the legislative term, United Kingdom 1979–1983, and Italy 1968–1972

| Legislature | Months in legislative term | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 0–6 | 7–12 | 13–18 | 19–24 | 25–30 | 31–36 | 37–42 | >42 |
| UK 79–83 | 0.02 | 0 | 0 | 0.37 | 0.21 | 0.21 | 0.05 | 0.03 |
| Italy 68–72 | 0.23 | 0 | 2.41 | 0.11 | 0.05 | 0.05 | 0.03 | 0.03 |

*Sources*: See footnote 4

We now turn to the UK House of Commons from 1979 to 1983. This legislative term is well known as the first in which Margaret Thatcher held the post of prime minister. It is recognized too for the split in the Labour Party, when in March 1981 four prominent leaders founded the Social Democratic Party (SDP). Less often appreciated is the timing of the decisions of those Labour MPs who followed the so-called "Gang of Four."

To track those choices, Table 1 reports the mean number of monthly moves per 100 MPs for every 6-month span in the British term. We standardize the number of monthly moves since the lower houses we study vary substantially in size. As the table indicates, switchers traveled to the SDP in several waves, which, despite some spread, were clustered near the midterm. Most recruits to the SDP jumped in March 1981, soon after the party was launched, and in October–December 1981.

The table also conveys the decisions of those legislators who did not enter the SDP. The first switcher of the term journeyed from one minor party to another. The last migrant, an erstwhile Labourite, left the SDP for a minor party. Consider Conservatives as well. In a January 1981 television interview, just before the SDP's founding, a "wet" Conservative backbencher estimated that 20 of his colleagues might join a center party (Crewe and King 1995, 114). In autumn 1981, SDP leaders held discussions about prospective membership with several Conservative MPs and Peers. Yet after a turning point in winter 1981–1982, such interest in the SDP ebbed away (Crewe and King 1995, 114–116). The sole Conservative MP who entered the SDP made the leap in the heady days of March 1981.[6]

## 2.2 PR Rules: Italy

Having examined two legislatures elected under SMD rules, we look at two consecutive legislative terms in Italy while proportional representation (PR) electoral laws applied. The first term, serving from May 1968 to May 1972, saw much movement, as Table 1 also displays. In July 1968, the standing orders of the Chamber of Deputies forced MPs from small parties into the Mixed Group.[7] The

[6]Potential recruits who stayed with Labour numbered about 40 (Crewe and King 1995, 479).

[7]The Chamber's standing orders stipulate that individual parties can form their own groups at the start of the term only if they have a minimum of 20 MPs (with few authorized exceptions, Art. 14.2). Otherwise, they enter the Mixed Group (Art. 14.5).

peak in switching corresponds to the July 1969 re-establishment of the Socialist and Social Democratic parliamentary groups, when the short-lived merger of the two parties, consummated in November 1966, fell apart (e.g., Di Scala 1988, esp. 161–165). Once the Socialists and Social Democrats returned to their respective longstanding party homes, several legislators made solo switches, with most such moves clustered near the middle of the term.

Few Italian MPs changed affiliation over the course of the Chamber's May 1972–June 1976 term. In May 1972, four MPs from tiny parties declared membership in the Mixed Group, as required by parliamentary rules. At the term's start, too, four independents elected on the lists of the Communist Party (PCI) chose the Mixed. This behavior aligned with party strategy, for the PCI had deliberately incorporated non-party candidates (often intellectuals) on its party lists in an effort to broaden its electoral appeal, with the understanding that if elected they were free to make their way to the Mixed (cf. Hellman 1977). Otherwise, a mere four deputies executed solo switches, all from April 1974 to March 1975.

## 2.3 Hybrid Systems: Italy and Russia

Last, we take up the 1996–2001 Italian Chamber and the 1993–1995 Russian Duma. For these terms, the lower houses in the two very different countries were elected under remarkably similar rules, combining PR, thresholds for PR, and plurality in SMDs.[8] Granted, only in Italy were the tiers linked: in Italy, not Russia, SMD wins were compensated in the PR tier; and Italy from 1993 to 2005 required that every SMD candidate be associated with at least one party list on the PR ballot, whereas independents could compete in Russian in SMDs.

Figure 1 plots two measures of switching – the number of moves per 100 MPs and the log of one plus the raw count of switches – for every month of the Italian term. Figure 2 does the same for the shorter Russian term. The log is informative because it reduces the pull of outliers while retaining information on all variation; it thus brings variation at lower levels into clear view. We highlight dominant trends in the data with LOWESS smoothing, with tension set at 0.50, so that half of the observations appear above the smoother, and half below.[9]

A basic item to flag in Fig. 1, on Italy, is that MP moves oscillate around a fairly stable central tendency until about a third of the way through the term, when switching begins to rise. In the months featuring the greatest switching, legislators

---

[8]These include sub-national governments with substantial powers (though Italy is not federal), and non-concurrence between parliamentary elections and other important elections.

[9]The smoothers take into account the presence of observations with high values, but for Russia those observations are not plotted so as to maintain the scale and depiction of overall patterns in MP behavior.
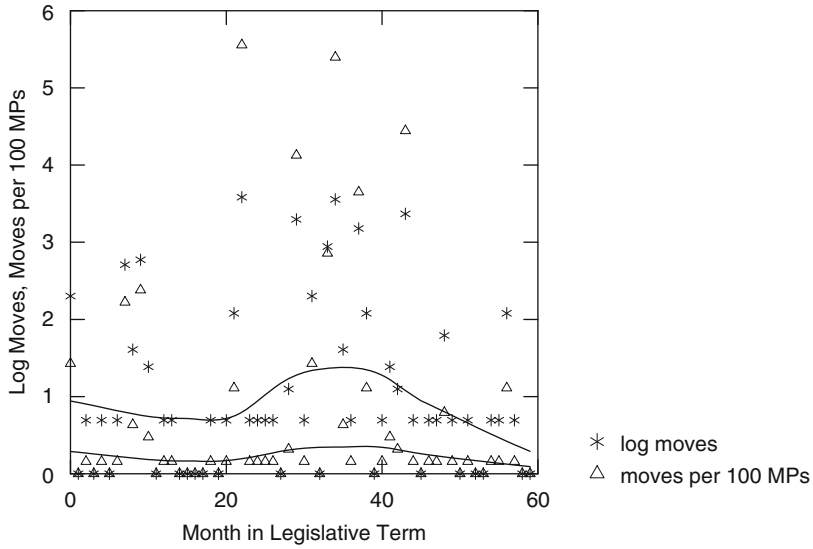
**Fig. 1** Log of raw number of MP switches and number of moves per 100 MPs, by month in the legislative term, Italy 1996–2001 (LOWESS smoothing used, tension = 0.50). *Note*: Here as in Fig. 2, each point represents a monthly observation, either the log of one plus the raw count of switches (denoted by *asterisk*) or the number of moves per 100 MPs (denoted by *triangle*)
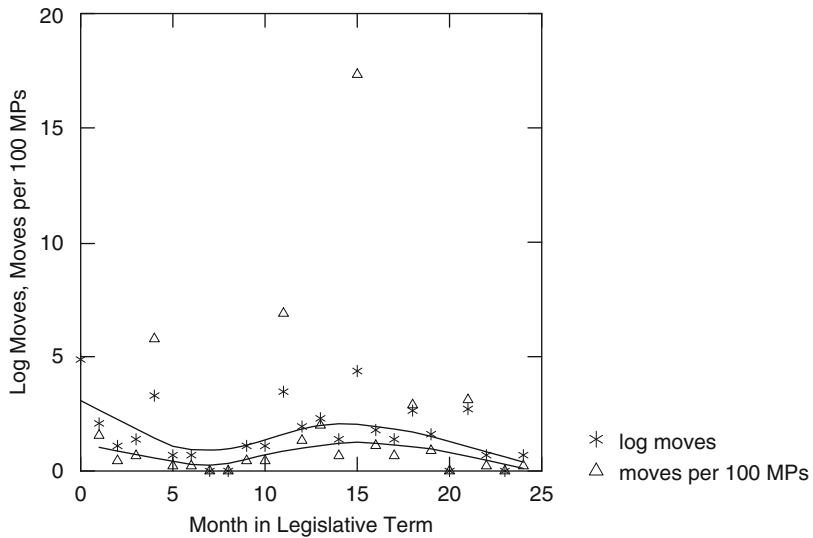


**Fig. 2** Log of raw number of MP switches and number of moves per 100 MPs, by month in the legislative term, Russia 1993–1995 (LOWESS smoothing used, tension = 0.50). *Note*: Russian outlier in week 0 (switches per 100 MPs = 27.78) excluded. The presence of observations with high values is reflected in the smoother, but high-value observations are omitted in order to maintain the vertical scale and depict adequately overall patterns in MP behavior

acted collectively to rupture extant parliamentary party groups and create new ones (Mershon and Shvetsova 2009b). The smoother evinces a peak about halfway through the term. Comparing the first third of the Italian term with the steeper downward slope of the smoother in the last third, we find support for the notion that legislators retain party loyalties when election time is near.

As Fig. 2 shows, in Russia the frequency of switching is extremely high when the legislature opens and then declines for the first third of the term. This initial spike in mobility does not upset our argument that association with party labels induces stability right after an election, since it captures not the behavior of partisan MPs but instead independents' early rush to join parties (cf. Mershon and Shvetsova 2008). In the last third of the Russian term, MP moves wane. Attending specifically to partisan MPs, then, the Russian evidence also comports with the reasoning about constraints on legislative party switching when elections have just occurred or loom on the horizon.

Support is equally clear for the claim of a midterm maximum in switching. We find obvious favorable evidence in Italy. Even with the extraordinary behavior of independents at the outset of the Russian term, we observe the midterm effect in the Duma. Despite some differences, the overall records in both terms reveal a midterm peak in MP mobility.

In sum, the evidence from all legislatures studied here establishes the presence of switching maxima near the midterm and discloses conversely the near-absence of switching in proximity to elections. Save for one noteworthy group, the few early or late moves we observe are isolated in safe electoral districts (US), confined to minor parties (UK), forced by parliamentary rules, or permitted by deals on party lists (Italy). The spectacular class of exceptions does not weaken but rather strengthens the inference that legislators calculate the electoral costs of party disloyalty as tied to the timing of switching: Russian MPs who won as independents frenetically shopped and hopped party early in the term.

## 3   Conclusion

The institutional design of most democracies leaves legislators free to change party at virtually any moment during their term of office. The conventional wisdom on stability in legislative party membership is rooted in the reality that, compared to the near-limitless possibilities, most incumbents, in most places, remain loyal to their status quo parties throughout the life of a legislature. Yet the recent literature on legislative party switching demonstrates that, under some conditions, a legislator reaches for the benefits than an alternative party affiliation offers, and so switches. The achievements of this literature can be enriched, we show, when analysts consider the conditions of time.

We argue that just when sitting legislators switch is a matter not of chance but instead of strategic choice. Representatives who join a new party – and those who refrain from doing so – themselves take time into account. The incumbent's

calculus on party affiliation includes expectations of punishment from voters, and the closer a move to election time, the more severe the penalties it will bring. Hence, we contend, MPs should stick with party labels near election time and should switch near the middle of a legislative term. MPs should move in time to contain electoral costs. The plausibility probes conducted here yield support for our reasoning.

Our logic and findings, moreover, suggest that the basic democratic practice of holding regular elections imparts overall stability to legislative party systems. The electoral connection (Mayhew 1974) binds representatives to the party labels on which they have won their seat. In this way, the treatment of politicians' decision-making on party membership as strategic, time-contingent choice provides new insights on the workings of political parties and of democratic institutions writ large.

# References

Ballot Access News (2001a) Candidates elected to the house, 1902–2000, who weren't nominees of the democratic or republican parties. 17, 3 (June 1). http://www.ballotaccess.org/2011/0601.html. Accessed July 2008

Ballot Access News (2001b) Minor party candidates elected to the house, 1902–2000, who were nominated by their own minor party, plus major party. 17, 3 (June 1). http://www.ballotaccess.org/2011/0601.html. Accessed July 2008

Butler D, Butler G (2006) British political facts since 1979. Palgrave Macmillan, New York

Camera dei Deputati (2010) Legislature precedenti. http://legislature.camera.it/. Accessed July 2010

Carey JM, Shugart MS (1995) Incentives to cultivate a personal vote: a rank ordering of electoral formulas. Elect Stud 14(4):417–439

Crewe I, King A (1995) SDP: the birth, life, and death of the social democratic party. Oxford University Press, Oxford

Desposato S (2006) Parties for rent? Careerism, ideology, and party switching in Brazil's chamber of deputies. Am J Polit Sci 50(1):62–80

Desposato S, Scheiner E (2008) Governmental centralization and party affiliation: legislator strategies in Brazil and Japan. Am Polit Sci Rev 102 (4 Nov):509–524

Di Scala SM (1988) Renewing Italian socialism: from Nenni to Craxi. Oxford University Press, New York

Federal Voting Assistance Program (2010) Primary election calendar. Retrieved from: http://www.fvap.gov/vao/calendar.html. Accessed Dec 2010

Gillespie E (2005) Frontline: Karl Rove: the architect. PBS (online). Retrieved from http://www.pbs.org/wgbh/pages/frontline/shows/architect/interviews/gillespie.html

Heller WB, Mershon C (2005) Party switching in the Italian chamber of deputies, 1996–2001. J Polit 67(2):536–559

Heller WB, Mershon C (2008) Dealing in discipline: party switching and legislative voting in the Italian chamber of deputies, 1988–2000. Am J Polit Sci 52:910–925

Heller WB, Mershon C (2009a) Integrating theoretical and empirical models of party switching. In: Heller WB, Mershon C (eds) Political parties and legislative party switching. Palgrave Macmillan, New York

Heller WB, Mershon C (2009b) Legislator preferences, party desires: the impact of party switching on legislative party positions. In: Heller WB, Mershon C (eds) Political parties and legislative party switching. Palgrave Macmillan, New York

Heller WB, Mershon C (eds) (2009c) Political parties and legislative party switching. Palgrave Macmillan, New York

Hellman SM (1977) The longest campaign: communist party strategy and the election of 1976. In: Penniman HR (ed) Italy at the polls, 1976. American Enterprise Institute, Washington, DC

Hickey JG (2003) Gillespie focuses on growth of GOP. Insight on the News (online). Retrieved from http://www.findarticles.com/p/articles/mi_m1571/is_2003_Oct27/ai_109128676?pi=cal

INDEM Foundation (2000) INDEM-statistics 2000 project. Bolshoi Zlatoustinskii Pereulok, Moscow

Laver M, Benoit K (2003) The evolution of party systems between elections. Am J Polit Sci 47(2):215–233

Mayhew DR (1974) Congress: the electoral connection. Yale University Press: New Haven

Mershon C, Shvetsova O (2008) Parliamentary cycles and party switching in legislatures. Comp Polit Stud 41(1):99–127

Mershon C, Shvetsova O (2009a) Change and stability in parties and party systems between elections. Invited paper presented at the workshop on the political economy of Democratic Institutions, Hoover Institution, May, Stanford, CA

Mershon C, Shvetsova O (2009b) Timing matters: incentives for party switching and stages of parliamentary cycles. In: Heller WB, Mershon C (eds) Political parties and legislative party switching. Palgrave Macmillan, New York

Nokken TP (2009) Party switching and the procedural party agenda in the US house of representatives. In: Heller WB, Mershon C (eds) Political parties and legislative party switching. Palgrave Macmillan, New York

Nokken TP, Poole KT (2004) Congressional party defection in American history. Legis Stud Quart 29(4):545–568

Reed SR, Scheiner E (2003) Electoral incentives and policy preferences: mixed motives behind party defections in Japan. Br J Polit Sci 33(3):469–490

Schofield N (2009) Switching equilibria. In: Heller WB, Mershon C (eds) Political parties and legislative party switching. Palgrave Macmillan, New York

Strøm K (1990) A behavioral theory of competitive political parties. Am J Polit Sci 34(2):565–598

United Press International (1984) Campaign notes: county official wins Florida runoff election. New York Times (Online). Retrieved from: http://query.nytimes.com/gst/fullpage.html?res=9C0CE0DB123BF930A35753C1A962948260&n=Top%2fReference%2fTimes%20Topics%2fOrganizations%2fR%2fRepublican%20Party

United States Congressional Biographical Directory (2008) http://bioguide.congress.gov/biosearch/biosearch.asp. Accessed July 2008

# On the Distribution of Particularistic Goods

**Jon X. Eguia and Antonio Nicolò**

The provision of local public goods financed by general taxation can lead to funding inefficient local projects, commonly denominated "pork". These projects are particularistic goods that greatly benefit a small community, while the costs of providing them are diluted among a much larger set of taxpayers. Because the beneficiaries of the projects do not absorb the full cost of funding the project, these beneficiaries, and the legislators who represent them, avidly advocate for projects that are overtly costly and altogether inefficient.[1]

In the simplest electoral model of pork barrel politics, two candidates compete for office in three districts under a majoritarian rule. Candidates can promise to finance a local project in any subset of districts. Each project generates benefits concentrated in one district but imposes a cost that all districts have to bear and that overcomes the benefits of the projects. Voters vote for the candidate's proposal which maximizes their expected utility. We solve the simultaneous game in which each candidate makes a proposal to the voters. While sequential games in which proposals are sequentially amended and voted on capture the bargaining nature of a legislative procedure, a game in which candidates simultaneously make a proposal to voters better describes an electoral contest.

Our focus is on the effect of electoral competition over pork spending. Do political campaigns lead candidates to engage in pandering to local districts to gain electoral support? Which is the promised amount of inefficient local public goods that we may expect to observe? We answer to these questions, finding the unique Nash equilibrium of the election game in which the strategy of each

---

J.X. Eguia (✉)
Department of Politics, New York University, 19 West 4th Street, 2nd Floor, New York, NY 10012, USA
e-mail: eguia@nyu.edu

A. Nicolò
Department of Economics "M. Fanno", Università degli Studi di Padova, Via del Santo 33, 35123 Padova, Italy

---

candidate is to choose a (possibly empty) subset of districts whose projects the candidate commits to implement if she is elected.

We first show that if projects are extremely inefficient, it is not possible to cobble together a majority of districts that would support executing all the projects in their own districts. If projects are so inefficient that any such majority would rather not implement its own projects, then it is not possible to win an election on the basis of pork promises, and in equilibrium both candidates commit to implement the social optimum of funding no inefficient project.

However, if projects are inefficient, but not too inefficient, we show that pork spending occurs in equilibrium. Both parties propose to implement inefficient projects. The equilibrium is in mixed strategies, and in expectation, candidates commit to execute about 43% of the projects. Ex-post up to two out of three projects are implemented.

Pork barrel politics has long been a subject of interest in political science and public choice (see the classic works by Riker 1962; Buchanan and Tullock 1962; Riker and Odershook 1973; Weingast 1979, and Ferejohn 1974). Riker (1962) and Buchanan and Tullock (1962) suggest that a minimal winning coalition forms in a legislature, so that majorities of the minimum size adopt policies that benefit themselves to the expense of the minorities. Weingast (1979), Shepsle and Weingast (1981) and Weingast et al. (1981) argue that instead of accepting this simple majority outcome, legislators prefer to embrace a universalistic norm such that pork is distributed to every district without excluding any minority. The intuition is that while funding a minimal winning majority of projects is more efficient than the universalistic rule, if legislators are risk averse and they cannot know in advance whether or not they would belong to the minimal winning coalition, then they prefer to accept the uniform distribution of inefficient spending in every district, rather than risk being left out. Ferejohn et al. (1987) and Baron (1991) modify the extensive form of the original simultaneous game looking at particular legislative process where proposals are put to a vote sequentially. They find that provision of local inefficient public goods is concentrated in some districts.

Instead of analyzing legislative bargaining, our theory addresses the relation between distributive policies and elections. In closely related work, Roberson (2008) analyzes the electoral and policy outcomes under different configurations of targetability of distributive policies. They main difference is that Roberson (2008) assumes that the provision of local public goods increases aggregate welfare, and inefficiencies occur when the possibility of transfers across jurisdictions lead to the underprovision of efficient local public goods. We are interested in the case of overprovision of inefficient local public goods, and therefore our analysis complements his results.

In other literature about the effect of political competition on redistributive policies, Lindbeck and Weibul (1987) present a probabilistic voting model of targeted redistribution in which agents face some global uncertainty in the form of a shock to the utility that each candidate generates to every voter, and some additional idiosyncratic uncertainty in the form of another shock to the utility that each candidate generates to each specific voter. Targeted redistribution occurs by

implementing transfers from the rich to the poor, or from groups that are more sensitive to transfers. Extending Lindbeck and Weibul's (1987) theory, Dixit and Londregan (1996) assume that there exist a continuum of voters within each group, and they allow transfers to incur a net loss (some of the wealth transferred is lost in transit), further generalizing the model by letting each party be more efficient at transferring resources to some groups. Voters have complete information of all the terms that are payoff relevant to them.

Lizzeri and Persico (2001) analyze the effect of different voting rules on redistribution in a model in which candidates have the ability to target individual transfers to each of a continuum of voters. McKelvey and Riezman (1992) and Muthoo and Shepsle (2010) explain how districts represented by more senior legislators obtain a greater share of resources. Mitchell and Moro (2006) and Echenique and Eguia (2007) explain redistribute policies that sustain employment in inefficient economic sectors as second best policies given the government's incomplete information or inability to provide individualized transfers. Drazen and Ilzetzki (2010) notes that the provision of inefficient goods can serve a signaling purpose that increases social welfare.[2]

In Sect. 1 we present the model; in Sect. 2 we state our main results. We conclude in Sect. 3 discussing extensions to the theory.

## 1 The Model

Let $A$ and $B$ be candidates in an election. Let $a$, $b$ and $c$ be representative voters in each of three districts. Let $S = \{0, 1\}^3$ be the strategy set of candidate $J$, for $J = A, B$.

A strategy consists of choosing which voters to favor, by implementing a project that generates a particularistic good to the voter, and which voters to ignore.

Projects are financed with taxes, paid equally by all voters. We normalize the cost of projects so that the per voter cost of each project is one unit. We assume that the benefit that her particularistic good generates to a voter is $\beta \in (1, 3)$. Therefore projects are socially inefficient, even though each of them is beneficial to one agent. Social welfare is maximized uniquely if no project is carried out. The degree of inefficiency of the projects depends on the parameter $\beta$. In the following analysis we distinguish the case when the benefit of each project is low, $\beta \in (0, 2)$ and each voter prefers to have no projects at all approved, than to have a majority of projects, including her own, approved, to the case when the benefit of the projects are high (even if projects are still inefficient), $\beta \in [2, 3)$ and each voter prefers to have two projects approved, including her own projects, to no projects approved.

Voters can take one of three actions: Vote for $A$, vote for $B$, or abstain. There is no cost of voting. The utility of citizens depends solely on the benefit of her own

---

[2]See Persson and Tabellini (2000), Chap. 7 for a broader survey and discussion on the topic of redistributive politics, and, more specifically, on the provision of local public goods.

particularistic good, and the taxes, it is additive in these two terms, and linearly decreasing in taxes.

The timing is as follows. First, each candidate $J$ simultaneously choose a policy $p^J \in \{0, 1\}^3$. The chosen policies are common knowledge. Second, each voter $i$ chooses how to vote ($A$, $B$, or abstain). Third, the candidate who obtains most votes wins, and her announced policy is implemented. In case of a tie, the winner is randomly chosen.

Candidates seek to maximize their probability of winning, and, lexicographically, they break ties between two strategies by choosing the strategy that maximizes their expected margin of victory. All agents are fully rational.

As in other voting games, there exist uninteresting equilibria in which all voters vote for the same candidate, so no deviation by voter or candidate is profitable. We ignore these equilibria looking only at equilibria in which every player plays a strategy that is not weakly undominated.

We also assume that voters abstain when they are indifferent between the two candidates.

## 2   Results

The voters' problem is very simple: Voters count the number of projects proposed by each candidate, which is equal to the disutility they experience from taxation, and then add the benefit of their own project, if this one is implemented. Then they compare the values for the two candidates, and vote for the candidate whose policy generates the highest utility. We consider first the case when the projects generates low benefits.

When the benefit of the project is low, $\beta \in [0, 2)$, each minimal winning coalition, which is formed by two districts, prefers having no project implemented that both their projects implemented. Hence, if a candidate promises to implement some projects, the other candidate can win the election by promising to not implement any project. Anticipating this voting behavior, if the benefit of each project is low, candidates propose the efficient policy of not implementing any project, fixing taxes at zero, and not distributing any particularistic goods.

**Proposition 1.** *Assume* $\beta \in [0, 2)$. *There is a unique Nash equilibrium that survives the iterative elimination of weakly dominated strategies. This equilibrium is in pure strategies and symmetric: Both candidates propose to implement zero projects. The outcome maximizes aggregate social welfare.*

*Proof.* For each voter, it is weakly dominated not to vote for the candidate whose policy proposal gives the voter the greatest payoff. Let $s^J \in S$ be a pure strategy by candidate $J$. Let $\{J, -J\} \equiv \{A, B\}$, that is, given candidate $J$, candidate $-J$ is the other candidate. Let the eight feasible strategies by agent $J$ be labeled as follows: $s_1 = (0, 0, 0)$; $s_2 = (1, 0, 0)$; $s_3 = (0, 1, 0)$; $s_4 = (0, 0, 1)$; $s_5 = (1, 1, 0)$; $s_6 = (1, 0, 1)$; $s_7 = (0, 1, 1)$; $s_8 = (1, 1, 1)$.

Consider first the case with $\beta < 1$, every voter strictly prefers $s_1$ to any other strategy. Under the assumption that voters do not use weakly dominated strategies, it follows that for any $k \neq 1$, strategy $s_k$ is weakly dominated by $s_1$ for each candidate $J$. So $(s^A, s^B) = (s_1, s_1)$ is the unique equilibrium obtained by iterative elimination of weakly dominated strategies.

Consider now the case with $\beta \in (1, 2)$. Recalling that we assume that voters do not use weakly dominated strategies and that they abstain when indifferent between two candidates, the following matrix indicates the number of votes for $A$ minus the number of votes for $B$, given that $A$ uses the row strategies and $B$ uses column strategies.

|       | $s_1$ | $s_2$ | $s_3$ | $s_4$ | $s_5$ | $s_6$ | $s_7$ | $s_8$ |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $s_1$ | 0     | 1     | 1     | 1     | 3     | 3     | 3     | 3     |
| $s_2$ | $-1$  | 0     | 0     | 0     | 1     | 1     | $-1$  | 3     |
| $s_3$ | $-1$  | 0     | 0     | 0     | 1     | $-1$  | 1     | 3     |
| $s_4$ | $-1$  | 0     | 0     | 0     | $-1$  | 1     | 1     | 3     |
| $s_5$ | $-3$  | $-1$  | $-1$  | 1     | 0     | 0     | 0     | 1     |
| $s_6$ | $-3$  | $-1$  | 1     | $-1$  | 0     | 0     | 0     | 1     |
| $s_7$ | $-3$  | 1     | $-1$  | $-1$  | 0     | 0     | 0     | 1     |
| $s_8$ | $-3$  | $-3$  | $-3$  | $-1$  | $-1$  | $-1$  | $-1$  | 0     |

It follows from the matrix that for any $k \neq 1$, strategy $s_k$ is iteratively weakly dominated by $s_1$ for each candidate $J$. Thus, $(s^A, s^B) = (s_1, s_1)$ is the unique equilibrium obtained by iterative elimination of weakly dominated strategies. ∎

Therefore, extremely inefficient projects are not implemented in equilibrium. However, consider projects that are inefficient, but not too inefficient. In particular, let $\beta \in (2, 3)$, so that the aggregate benefit of each project is at least two thirds of its aggregate cost. These projects generate a loss of aggregate welfare if they are implemented. In fact, in equilibrium they are implemented. Laffond et al. (1993) show that in a more abstract electoral game where candidates compete in a tournament that corresponds to the majority preferences of society, there exists a unique equilibrium, possibly in mixed strategies. In our application, we find that the equilibrium is indeed in mixed strategies if $\beta \in (2, 3)$.

**Proposition 2.** *Assume $\beta \in (2, 3)$. There is a unique Nash equilibrium that survives the iterative elimination of weakly dominated strategies. This equilibrium is in mixed strategies and symmetric. The outcome does not maximize aggregate welfare: in expectation, 42.9% of projects are implemented. Each candidate proposes no projects with probability 1/7; one (randomly chosen) project with probability 3/7 and two (randomly chosen) projects with probability 3/7.*

*Proof.* Let $\{J, -J\} \equiv \{A, B\}$, that is, given candidate $J$, candidate $-J$ is the other candidate. Let $\sigma^J$ be a mixed strategy by agent $J$. Let $\sigma_k^J$ be the weight assigned to pure strategy $s_k$ in mixed strategy $\sigma^J$.

Under the assumptions that $\beta \in [2, 3)$, that voters do not use weakly dominated strategies and that they abstain when indifferent between two candidates, the

following matrix indicates the number of votes for $A$ minus the number of votes for $B$, given that $A$ uses the row strategies and $B$ uses column strategies.

|       | $s_1$ | $s_2$ | $s_3$ | $s_4$ | $s_5$ | $s_6$ | $s_7$ | $s_8$ |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $s_1$ | 0     | 1     | 1     | 1     | $-1$  | $-1$  | $-1$  | 3     |
| $s_2$ | $-1$  | 0     | 0     | 0     | 1     | 1     | $-1$  | $-1$  |
| $s_3$ | $-1$  | 0     | 0     | 0     | 1     | $-1$  | 1     | $-1$  |
| $s_4$ | $-1$  | 0     | 0     | 0     | $-1$  | 1     | 1     | $-1$  |
| $s_5$ | 1     | $-1$  | $-1$  | 1     | 0     | 0     | 0     | 1     |
| $s_6$ | 1     | $-1$  | 1     | $-1$  | 0     | 0     | 0     | 1     |
| $s_7$ | 1     | 1     | $-1$  | $-1$  | 0     | 0     | 0     | 1     |
| $s_8$ | $-3$  | 1     | 1     | 1     | $-1$  | $-1$  | $-1$  | 0     |

It follows that $s_8$ is weakly dominated by $s_1$, once we assume that voters do not use weakly dominated strategies, and abstain when indifferent between two candidates. Once we eliminate $s_8$, the margin of victory is the same for any candidate strategy profile such that one candidate wins the election. Thus, once we eliminate $s_8$, we can dismiss the lexicographic preferences for a large margin of victory, and assume each candidate maximizes the probability of victory alone, in which case the reduced matrix without the last column and row indicates the payoff to candidate $A$.

By assumption, for any $s \in S$, if $s^A = s^B = s$, then all three voters abstain, and the candidates tie the election. Thus, for any strategy $s^{-J}$, candidate $J$ can always at least tie the election. Thus, in equilibrium candidates must win with probability 50%, otherwise, the candidate who won with a lower probability would deviate to mimic the other candidate and obtain a probability of victory of 50%. Since in equilibrium $A$ wins with probability 50% given the equilibrium $\sigma^B$, it must be that the probability that $A$ wins if $A$ plays strategy $s_k$ and $B$ plays $\sigma^B$ is no more than 50% for any $k \in \{1, \ldots, 7\}$. This observation, together with the matrix, generates the following seven inequalities:

(1) $\sigma_2^B + \sigma_3^B + \sigma_4^B \leq \sigma_5^B + \sigma_6^B + \sigma_7^B$
(2) $\sigma_5^B + \sigma_6^B \leq \sigma_1^B + \sigma_7^B$
(3) $\sigma_5^B + \sigma_7^B \leq \sigma_1^B + \sigma_6^B$
(4) $\sigma_6^B + \sigma_7^B \leq \sigma_1^B + \sigma_5^B$
(5) $\sigma_1^B + \sigma_4^B \leq \sigma_2^B + \sigma_3^B$
(6) $\sigma_1^B + \sigma_3^B \leq \sigma_2^B + \sigma_4^B$
(7) $\sigma_1^B + \sigma_2^B \leq \sigma_3^B + \sigma_4^B$

From (2) + (3) we obtain $\sigma_5^B \leq \sigma_1^B$. From (2) + (4) we obtain $\sigma_6^B \leq \sigma_1^B$. From (3) + (4) we obtain $\sigma_7^B \leq \sigma_1^B$. From (5) + (6) we obtain $\sigma_1^B \leq \sigma_2^B$. From (5) + (7) we obtain $\sigma_1^B \leq \sigma_3^B$. From (6) + (7) we obtain $\sigma_1^B \leq \sigma_4^B$. Given inequality (1), these inequalities can only hold if they all hold with equality. Given that $\sum_{k=1}^{7} \sigma_k^B = 1$, it follows $\sigma_k^B = 1/7$ for any $k \in \{1, \ldots, 7\}$. A symmetric argument implies $\sigma_k^A = 1/7$ for any $k \in \{1, \ldots, 7\}$. Given $\sigma_k^J = 1/7$ for any $k \in \{1, \ldots, 7\}$ and

any $J \in \{A,\ B\}$, and given that each candidate wins with equal probability for any $s_k \in \{s_1,\ \ldots,\ s_7\}$, the probability that the project in district 1 is implemented is $\sigma_2 + \sigma_5 + \sigma_7 = 3/7$, the probability is also 3/7 for the other two districts, and the expected number of implemented projects is 9/7. ∎

The above proposition shows that electoral pressures lead to implementing a variable but in expectation large number of socially inefficient projects when they are not too inefficient. Candidates never propose to provide pork spending for everybody. Rather, the equilibrium outcome leads to implementing zero, one or two projects.

## 3 Conclusion

We have shown that candidates commit to distribute pork among voters even though voters know that the provision of local public goods is inefficient. While distributing pork is a weakly dominated strategy when pork is very inefficient, if the local benefits of pork is higher, even though it remains inefficient, electoral competition leads candidates to pander for votes by committing to distribute pork to a subset of districts. Full information prevents the distribution of pork to every district: Pork for everybody is dominated by not offering pork to anybody.

However, this partially optimistic conclusion depends on the assumption of full information about candidates' policies. In practice, voters may have only vague ideas about candidates' (or parties) proposals about the provision of local particularistic goods. Voters who do not benefit from a local public good in another district, may not be unduly interested or informed about the provision of such goods. Moreover, if candidates try not to spread the information about what they promise in other districts, it becomes costly for voters to become fully informed. Further incentives to provide inefficient local public goods may also arise if each candidate has an exogenously given preference to favor one particular district, which is particularly likely to occur if districts lines correspond to ethnic divisions (Nye et al. 2010).

Our future research agenda is to explore these extensions. We start by analyzing elections with an imperfect informed electorate in a companion working paper (Eguia and Nicolò 2011).

## References

Baron DP (1991) Majoritarian incentives, pork barrel programs and procedural control. Am J Polit Sci 35:57–90

Buchanan J, Tullock G (1962) The calculus of consent. University of Michigan Press, Ann Arbor, MI

Dixit A, Londregan J (1996) The determinants of success of special interests in redistributive politics. J Polit 58:1132–1155

Drazen A, Ilzetzki E (2010) Kosher pork. In: Working paper

Echenique F, Eguia JX (2007) Cohesion, insurance and redistribution. Quart J Polit Sci 2:287–305

Ferejohn J (1974) Pork barrel politics. Stanford University Press, Stanford, CA

Ferejohn J, Fiorina M, McKelvey RD (1987) Sophisticated voting and agenda independence in the distributive politics setting. Am J Polit Sci 31:169–193

Laffond G, Laslier J-F, Le Breton M (1993) The bipartisan set of a tournament game. Games Econ Behav 5:182–201

Lindbeck A, Weibul JW (1987) Balanced-budget redistribution as the outcome of political competition. Public Choice 52:273–297

Lizzeri A, Persico N (2001) The provision of public goods under alternative electoral incentives. Am Econ Rev 91(1):225–239

McKelvey RD, Riezman R (1992) Seniority in legislatures. Am Polit Sci Rev 86:951–965

Mitchell MF, Moro A (2006) Persistent distortionary policies with asymmetric information. Am Econ Rev 96:387–393

Muthoo A, Shepsle KA (2010) Seniority and incumbency in legislatures. Manuscript, Harvard University

Eguia JX and Nicolò A (2011) Information and elections. In: Working paper, New York University

Nye JVC, Rainer I, Stratmann T (2010) Do black mayors improve black employment outcomes? In: SSRN working paper #1585283

Persson T, Tabellini G (2000) Political economics. MIT Press, Cambridge, MA

Riker WH (1962) The theory of political coalitions. Yale University Press, New Haven, CT

Riker WH, Odershook P (1973) An introduction to positive political theory. Prentice-Hall, Englewood Cliffs, NJ

Roberson B (2008) Pork-barrel politics, targetable policies, and fiscal federalism. J Eur Econ Assoc 6:819–844

Shepsle KA, Weingast BR (1981) Political preferences for the pork barrel: a generalization. Am J Polit Sci 25:96–111

Weingast BR (1979) A rational choice perspective on congressional norms. Am J Polit Sci 23:245–262

Weingast BR, Shepsle KA, Johnsen C (1981) The political economy of benefits and costs: a neoclassical approach. J Polit Econ 89:642–664

# Vote Revelation: Empirical Content
of Scoring Rules

**Andrei Gomberg**

## 1 Introduction

Consider an observer trying to make sense of the goings on in a secretive committee, such as the old Soviet Politburo. Such an observer would not have any direct evidence about preferences of individual committee members, nor would he be likely to observe the rules the committee uses to make its decisions. Nevertheless, our Kremlinologist does have some information to work with. For one, he may have a reasonably good idea of the options the committee members are facing. He would also be able to observe the committee decision: perhaps, it would come out in the Pravda. Finally, the committee membership is public knowledge (he could determine it by observing the figures standing on the observation deck of Lenin's Mausoleum during the Revolution Day parade). What sort of deductions would it be possible to make about the unobservable preferences and preference aggregation rules within the committee from this information? In fact, not much could be said from a single observation of the committee decision. However, it turns out that, if a number of observations of decisions taken by a committee with variable membership is available, one can use the available data to test certain hypotheses about the committee functioning.

The approach I use in this paper is closely related to the ideas of revealed preference and rationalizability, that have long been standard foundations of economic analysis. Ever since Houthakker (1950) it has been known that a simple consistency condition on choices (the Strong Axiom of Revealed Preference, SARP) is a necessary and sufficient condition for being able to explain individual choices with rational preference maximization. Over the years a sizeable literature on restrictions on choices implied by various individual and group decision-making procedures developed. Thus, for instance, in the context of social choice rules, Blair et al. (1976) characterized such restrictions as would derive from maximizing

A. Gomberg
Centro de Investigación Económica, Instituto Tecnológico Autónomo de México, Av. Camino a Santa Teresa 930, DF 10700 México, México
e-mail: gomberg@itam.mx

preferences that are merely acyclic, rather than transitive. This, of course, may be interpreted as characterizing choices made by committees of rational members with some of those members exercising veto power. However, though well-established, the tradition of revealed preference approach to group decisions has not been much developed recently. In particular, I am aware of no studies establishing "signatures" imposed on collective decisions by most commonly used voting rules. It is precisely this that I attempt to do in this paper.

In fact, when in recent years concepts of choice and revealed preference have received substantial renewed attention in economics, it was in the context of individual decision-making. This attention has been derived from the new focus on "boundedly rational" decision-making procedures different from the usual rational preference maximization. In this context one might mention Manzini and Mariotti (2007) work on "sequential rationalizability" or Masatlioglu and Ok (2005) study of choice with status-quo bias, both of which attempt to establish restrictions imposed on choices by distinct decision-making procedures. Other recent studies, such as Caplin and Dean (2011), attempt to explore the restrictions that various "boundedly rational" procedures would impose on records that are somewhat more detailed than the usual choice data, though still plausibly observable.

In the situation described at the beginning of this introduction, the group decision data is, in fact, richer than usual: in addition to the record of choices from a given set of alternatives, we have the committee membership at each decision point to consider. Thus, if we want to test a given theory of how the committee works, we have more information to base our testing on. Even on incomplete data (i.e., when not all possible observations might be there), we may observe enough to do this.

In this study I concentrate on a particular class of theories about the internal committee workings. I will generally assume that each committee consists of rational members who decide using some scoring rule (a class of rules, which includes simple "first past the post" plurality, approval voting or the Borda Count), and will try to formulate the natural restrictions (so far incomplete) on my observations implied by these rules. Even when the particular scoring rule is unknown, such restrictions turn out to be nontrivial.

The scoring rules are those in which individuals are asked to provide each alternative with a numeric score (reflecting their preferences), the individual scores are added up and the alternative with the highest aggregate score is chosen. These rules have long been characterized by social choice theorists (see Smith 1973; Young 1975; Myerson 1995).

This work is also related to the study of empirical content of sincere (vs. strategic) voting by Degan and Merlo (2009). In fact, if the formal decision rule is known, this work may be reinterpreted precisely as the test of voter sincerity: if I know how the votes are counted, violations of the conditions established here could only be interpreted as indications that the scores do not directly reflect rational individual preference. Thus, to the extent one maintains the assumption that voters are rational, sincere voting would be falsified in this case. Likewise, this paper is related to Kalandrakis (2010) work on rationalizing individual voting

decisions. This paper crucially differs from both Degan and Merlo (2009) and Kalandrakis (2010), however, in that I do not assume observability of individual votes (nor do I impose anything in addition to rationality on individual preferences). Rather, individual votes are "revealed" here from the observations of the group choices. In fact, by establishing a number of "SARP-like" conditions, I hope to characterize the conditions under which revealed scores are consistent (so far this characterization is incomplete).

## 2 Basic Set-Up

Consider a finite set $N = \{1, 2, \ldots n\}$ of agents and a finite set $X = \{x_1, x_2 \ldots x_m\}$ of alternatives. A set of alternatives to be considered by a *committee* $S \in 2^N \backslash \{\emptyset\}$ is $B \in 2^X \backslash \{\emptyset\}$; following the standard terminology of individual choice theory, I shall call $B$ the *budget set*. If a committee $S$ is offered a choice from the budget set $B$ the committee choice is recorded as $\emptyset \neq C(B, S) \subset B$. The *committee choice structure* is defined as a pair $(\varepsilon, C(.,.))$ where $\varepsilon \subset 2^X \backslash \{\emptyset\} \times 2^N \backslash \{\emptyset\}$ is the record of which budget sets where considered by which committees and $C : \varepsilon \to X$, such that $C(B, S) \subset B$ is the non-empty-valued choice correspondence, recording committee choices.

In order to explain observed committee choice structures I shall, in general, assume that each agent $i \in N$ has rational (complete and transitive) preferences $\succeq_i$ defined over $X$. The committee choice structure provides a record of observed committee choices, which may be used by an observer to deduce the preference profiles and the preference aggregation rules the committee uses. In this paper II concentrate on a particular class of such rules: the scoring rules, a class that includes such distinct procedures as the plurality vote (in which the winner is an alternative that is chosen by the largest number of voters), the Borda Count (in which alternatives get assigned the most points for being someone's top choice, a point less for being a second choice, etc., the scores get summed up over all the voters and the alternative with the largest score wins), or the Approval Voting (in which an individual is allowed to mark alternatives as acceptable or unacceptable, and the alternative which has been marked as acceptable by the largest number of voters gets chosen). Overall, I shall assume that agents are non-strategic, in that they ignore who else is in the committee (as noted above, the conditions I am deriving here might, if the formal rule is observable, be viewed as empirical implications of sincere voting itself). However, I shall allow the votes to depend on the budget sets under consideration (as would be the case in a sincere Borda Count). Thus, if the set of alternatives $B$, a vote of agent $i \in S$ is a function $v_i^B : B \to \mathbb{R}$.

Given a vote from each of its members a committee $S$ chooses an alternative that gets the highest score

$$C^{scoring}(B, S) = \arg\max_{x \in B} \sum_{i \in S} v_i^B(x)$$

where $\sum_{i \in S} v_i^B(x)$ is called the *score* received by an alternative $x \in B$ in voting by committee $S$. Such a choice structure is said to be generated by the scoring rule.

Following Myerson (1995), I shall allow agents to submit votes that are distinct from reporting their preference orderings. In fact, for the purposes of defining a scoring rule one does not need to assume that the votes themselves derive from rational preferences. However, the scoring rules require agents to report a ranking of alternatives in $B$ by means of their votes $v_i^B \in \mathbb{R}^k$. Though in general such a ranking may not necessarily represent a rational preference (and thus, for instance, could be inconsistent over the different budget sets $B$), I shall concentrate on voting that, indeed, can be viewed as a sincere representation of individual preferences. Formally, given a rational preference profile $\succsim = (\succsim_1, \succsim_2, \ldots, \succsim_n)$ I shall say that a committee vote $v_i^B$ is (*weakly*) *consistent with preferences* if $x \succsim_i y$ implies $v_i^B(x) \geq v_i^B(y)$.

If a committee choice structure is such that for any $(B, S) \in \varepsilon$

$$C(B, S) = C^{scoring}(B, S)$$

where the votes are consistent with preferences for some rational preference profile $\succsim$, I shall say that $\succsim$ rationalizes $(\varepsilon, C(.,.))$ via a scoring rule.

It should be noted, that unless the choice structure is extended by allowing observing variations in committee membership, scoring rules would, at first glance, appear particularly unpromising from the standpoint of this research: it would seem that nearly every possible committee decision could be explained by some sort of scoring applied to an unobserved preference profile of a fixed committee. Thus, if one defines, in the spirit of Salant and Rubinstein (2008) work on the choice with frames, the choice correspondence as

$$C_c(B) = \{x : x \in C(B, S) \text{ for some committee } S\}$$

little, if anything appears to be imposed on $C_c(.)$ (some restrictions may be derived from the relative cardinalities of $B$ and $N$, if the latter is observed, but that appears to be it). However, it turns out that more can be said if committee membership and its variations are observed.

## 3 Revealed Scoring

Supposing that committees are making their decisions using scoring rules implies that each committee produces a ranking, represented by the score in question. Of course, if committee members may change their votes arbitrarily, based on either committee membership or the set of alternatives involved, not much could be done here. For this reason, at least for now, I shall assume that voters are restricted to be

weakly consistent with sincere voting. With this assumption, I shall try to "reveal" as much as possible about individual votes.

I shall start by defining the direct preference revelation

- *Direct revelation.* For each $(B, S) \in \varepsilon$ a pair of nested binary relations $P^*_{B,S} \subset R^*_{B,S}$ on $B$ is defined by:

  (i)  Let $x \in C(B, S)$ then $xR^*_{B,S}y$ for any $y \in B$
  (ii) Let $x \in C(B, S)$ and $y \notin C(B, S)$ for some $y \in B$ then $xP^*_{B,S}y$

This constitutes a record of direct preference revelation: if an alternative is chosen, it implies it received at least as high a score as any other feasible alternative and a strictly higher score than any feasible alternative not chosen.

The reinforcement axiom of Smith (1973) and Young (1975) provides us with a way of extending these revealed scoring relations, often even when a particular pair $(B, S)$ is not in $\varepsilon$. This axiom states that, if each of the two disjoint committees makes the same choice, the union of those two committees has to follow it. It is easy to see that every scoring rule would satisfy it: thus, for instance, if $C(\{a, o\}, \{1, 2\}) = C(\{a, o\}, \{3, 4\}) = \{a\}$ we may not have $C(\{a, o\}, \{1, 2, 3, 4\}) = \{o\}$. Furthermore, individual preference revelation is sometimes possible in this framework as well: if one ever observes an individual choosing an alternative when alone, this reveals his/her preference that would be unchanged even when the budget set changes. This motivates the following extension of the score revelation

- *Reinforcement.* The binary relations $P_{B,S} \subset R_{B,S}$ on $B$ are defined by:

  (i)   $xP^*_{B,S}y$ implies $xP_{B,S}y$, $xR^*_{B,S}y$ implies $xR_{B,S}y$
  (ii)  For any $B \in 2^X\backslash\{\varnothing\}$ and any $S, T \in 2^N\backslash\{\varnothing\}$ such that $S \cap T = \varnothing$, $xR_{B,S}y$ and $xR_{B,T}y$ imply that $xR_{B,S\cup T}y$
  (iii) For any $B \in 2^X\backslash\{\varnothing\}$ and any $S, T \in 2^N\backslash\{\varnothing\}$ such that $S \cap T = \varnothing$, $xP_{B,S}y$ and $xR_{B,T}y$ imply that $xP_{B,S\cup T}y$
  (iv)  For any $B \in 2^X\backslash\{\varnothing\}$ and any $S, T \in 2^N\backslash\{\varnothing\}$ such that $S \subset T(T\backslash S \neq \varnothing)$, $xP_{B,S}y$ and $yR_{B,T}x$ imply that $yP_{B,T\backslash S}x$
  (v)   For any $B \in 2^X\backslash\{\varnothing\}$ and any $S, T \in 2^N\backslash\{\varnothing\}$ such that $S \subset T(T\backslash S \neq \varnothing)$, $xR_{B,S}y$ and $yP_{B,T}x$ imply that $yP_{B,T\backslash S}x$
  (vi)  For any $B \in 2^X\backslash\{\varnothing\}$ and any $i \in N, xP_{B,\{i\}}y$ implies $xR_{D,\{i\}}y$ for all $D \in 2^X\backslash\{\varnothing\}$

The statements $xP_{B,S}y$ (respectively, $xR_{B,S}y$) may be understood as "x is revealed (directly or indirectly) to have obtained a higher (respectively, at least as high) score than y in a vote by a committee S over the budget set B". Of course, no matter how obtained, scoring revelation cannot be selfcontradicting. Thus, for instance, if $C(\{a, o\}, \{1, 2\}) = C(\{a, o\}, \{3, 4\}) = \{a\}$ one may not have $C(\{a, o\}, \{1, 2, 3, 4\}) = \{o\}$. In fact, since the binary relations $R_{B,S}$ and $P_{B,S}$ refer to the number of votes, the relation should be transitive (if more people vote for x than for y and more people vote for y than for z more people should be voting for x than for z).

This motivates the following simple axiom:

**Axiom 1 [Committee Axiom of Revealed Preference (CARP)][1].** *For any* $B \in 2^X \setminus \{\emptyset\}$, *any* $S \in 2^N \setminus \{\emptyset\}$ *and any* $x_1, x_2, \ldots x_n \in B$, $x_1 R_{B,S} x_2, x_2 R_{B,S} x_3 \ldots x_{n-1} R_{B,S} x_n$ *implies* $\neg (x_n P_{B,S} x_1)$.

*Example 1.* Consider the budget set $B = \{a, b, c\}$ and the four disjoint committees $S_1$, $S_2$, $S_3$ and $T$. Let $C(B, S_1) = a, C(B, S_2) = b, C(B, S_3) = c, C(B, S_1 \cup T) = b, C(B, S_2 \cup T) = c, C(B, S_3 \cup T) = a$. It is not hard to see that this implies that $b P_{B,T} c P_{B,T} a P_{B,T} b$ which, of course, contradicts Axiom 1: committee $T$ should be giving alternative $b$ a higher score than alternative $c$, alternative $c$ a higher score than alternative $a$, and alternative $a$ the higher score than alternative $i$, which is impossible.

Of course, as noted above, one may be able to make inferences about individual preferences, for instance, from direct or indirect observations of singleton coalitions, by defining an individual revealed preference relation $P_i$ as follows:

- *Individual preference revelation*. If $x P_{B,\{i\}} y$ for some $B \in 2^X \setminus \{\emptyset\}$ define $x P_i y$.

As I have assumed individual rationality it is clear that the standard Strong Axiom of Revealed preference must hold for individual preference as well.

**Axiom 2 [Strong Axiom of Revealed Preference (SARP)].** *For any* $i \in N$ *and any* $x_1, x_2, \ldots x_n \in B$, $x_1, P_i x_2, x_2 P_i x_3 \ldots x_{n-1} P_i x_n$ *implies* $\neg (x_n P_i x_1)$

It is clear that both CARP and SARP would have to hold if a committee of rational individuals is deciding by sincere votes using a scoring rule, since otherwise we'd have to accept either cycles in individual preferences or in group scores (as in the example above). Hence, the main result of this paper follows immediately from the construction.

**Theorem 1.** *A committee choice structure* $(\varepsilon, C(.,.))$ *may be generated by a scoring rule strictly consistent with rational preferences only if the implied* $P_i$ *satisfies SARP for each $i$ and the implied* $R_{B,S}$ *satisfies CARP for each* $(B, S)$.

It should be stressed that this result provides only a necessary and not a sufficient condition for rationalizability with scoring. In fact, counterexamples to the converse are not hard to generate, as possibilities for indirect score revelation are by no means exhausted with application of reinforcement.

*Example 2.* Consider a budget set $B = \{a, b\}$ and four committees $S_1, S_2, T_1$ and $T_2$ such that $S_i \cap T_j = \emptyset$. Suppose $\{a\} = C(B, T_1) = C(B, T_2) = C(B, S_1 \cup T_1) = C(B, S_2 \cup T_2)$, while $\{b\} = C(B, S_1 \cup T_2) = C(B, S_2 \cup T_1)$. Of course, if these decisions were arrived to by scoring, this would imply (by reinforcement) that both $S_i$ would have to be choosing $b$ as well. The votes of $S_1$ are sufficient to overturn the preference of $T_2$, but not of $T_1$ for a, so we can conclude that the advantage in votes

---

[1] I am grateful to Professor Schofield for the naming suggestion for this axiom.

that $T_1$ gives to $a$ is strictly bigger than that given by $T_2$. However, the votes of $S_2$ overturn the choice of $T_1$, but not the choice of $T_2$, so the vote advantage of $a$ in $T_2$ is strictly bigger than that in $T_1$. Clearly, this is impossible, unless the scores assigned are not independent of committee membership.

The establishment of the exact conditions for rationalizability with scoring is, thus, for the moment, an open question.

# 4 Conclusions and Further Research

So far it has been possible to establish a set of properties of committee choice structures that are necessary consequences of sincere scoring-based committee decisions It remains to see if this could be strengthened to a concise sufficient condition for rationalizability with scoring. An interesting further extension of the model would be to consider the consequences of particular scoring rules, such as plurality, approval or the Borda Count.

# References

Blair D, Bordes G, Kelly J, Suzumura K (1976) Impossibility theorems without collective rationality. J Econ Theory 13:361–379
Caplin A, Dean M (2011) Search, Choice and Revealed Preference, Theor Econ 6:19–48
Degan A, Merlo A (2009) Do voters vote ideologically. J Econ Theory 144:1868–1894
Houthakker HS (1950) Revealed preference and the utility function. Economica 17:159–174
Kalandrakis T (2010) Rationalizable voting. Theor Econ 5:93–125
Manzini P, Mariotti M (2007) Sequentially rationalizable choice. Am Econ Rev 97:1824–1839
Masatlioglu Y, Ok E (2005) Rational choice with status quo bias. J Econ Theory 121:1–29
Myerson R (1995) Axiomatic derivation of the scoring rules without the ordering assumption. Soc Choice Welf 12:59–74
Salant Y, Rubinstein A (2008) Choice with frames. Rev Econ Stud 75:1287–1296
Smith JH (1973) Aggregation of preferences with variable electorate. Econometrica 41:1027–1041
Young P (1975) Social choice scoring functions. SIAM J Appl Math 28:824–838

# Biographies

## Editors' Biographies

**Norman Schofield** is Director of the Center in Political Economy, the William R. Taussig Professor of Political Economy, and Professor in the Departments of Economics and Political Science at Washington University in Saint Louis. He has Ph.D.s in both government and economics from Essex University, a LittD from Liverpool University and a doctorat d'etat from the University of Caen. He is currently working on topics in the theory of social choice, political economy, and democracy. He has written several books, including *Multiparty Government* (with Michael Laver, in 1990), *Social Choice and Democracy* (Springer, 1985), and edited or co-edited seven volumes. His book, *Mathematical Methods in Economics and Social Choice*, was published by Springer in 2003. In 2006 he published *Architects of Political Change* as well as *Multiparty Democracy* (with Itai Sened) both with Cambridge University Press. His book, *The Spatial Model of Politics* was published by Routledge in 2008, followed by *The Political Economy of Democracy and Tyranny* in 2009 with Oldenbourg (Munich). He is presently completing a book, *Chaos or Leadership* (co-authored with Maria Gallego). He has been the recipient of a number of NSF awards, most recently one on electoral politics and regime change. In 2003 he was the Fulbright distinguished professor at Humboldt University and, in 2007 and 2008, visiting fellow at ICER in Turin. In Spring 2008 he was the visiting Leitner Professor at Yale University and was a National Fellow at the Hoover Institution in summers 2009 and 2010. He was recipient of the Riker prize for contributions to political theory in 2002, and was elected to the American Academy of Arts and Sciences in 2005. E-mail: schofield.norman@gmail.com.

    **Gonzalo Caballero** is an Associate Professor of Economics at the University of Vigo, Spain. He received his Ph.D. and a B.Sc. in Economics with Honors from the University of Vigo, and he received a M.Phil in Contemporaneous Political Studies from the University of Santiago de Compostela. He has been a Visiting Scholar at the Center for New Institutional Social Sciences, Washington University in St. Louis, the Center for Advanced Study in the Social Sciences in Madrid, the University of California in Santa Barbara and the University of California in Berkeley. He taught a course on "Economics of Institutions" at Charles University

in Prague, at Comenius University in Bratislava and at the University Roma Tre. He has published many papers on institutional analysis in various journals, including *Economic Modelling*, *Journal of Institutional Economics*, *Ocean & Coastal Management*, *Marine Policy*, *Environmetrics*, *El Trimestre Económico*, *Brazilian Journal of Political Economy*, *Política y Gobierno*, *Revista Española de Investigaciones Sociológicas*, *Revista de Estudios Políticos* and *Revista Española de Ciencia Política*. He has written several books and co-edited a volume on the Great Recession. E-mail: gcaballerom@yahoo.com.

## Biographies of Contributors

**Carles Boix** writes and teaches on comparative political economy and comparative politics in Princeton University. He is the author of *Political Parties, Growth and Equality* (Cambridge University Press, 1998), *L'obertura catalana* (Idees, 2002) and *Democracy and Redistribution* (Cambridge University Press, 2003) and the co-editor of the *Oxford handbook of Comparative Politics* (Oxford University Press, 2007). He has received the William Riker award for the best book on political economy twice, the Mattei Dogan award for best book published in the field of comparative research and the Heinz Eulau award for best article published in the *American Political Science Review*. His current research projects include: the formation of the state and its distributional consequences; and the political conditions that led to the emergence of various party systems and electoral institutions in advanced democracies in the nineteenth and twentieth centuries. E-mail: cboix@princeton.edu.

**Christopher Claassen** is a doctoral student in political science at Washington University in St. Louis. His research interests are comparative politics and political psychology. He has co-authored a paper with Schofield in the *International Journal of Mathematics and Mathematical Systems*. E-mail: claassen@wustl.edu.

**Bruno Codenotti** is the Research Director of the Istituto di Informatica e Telematica, Consiglio Nazionale delle Ricerche, Pisa, Italy. His research interests include computation of market equilibria, computational game theory, WEB algorithmics, algorithms in linear algebra and distributed computing. E-mail: bruno.codenotti@iit.cnr.it.

**Alfred Darnell** received his Ph.D. in sociology from the University of Chicago in 1990. He has done research in a variety of areas concerning ethnic identity formation and the impact of cultural influences, specifically religion, on educational attainment. He is currently working with Sunita Parikh on understanding transformations and continuities in apex courts in India. E-mail: adarnell@wustl.edu.

**Desiree A. Desierto** is an Assistant Professor at the University of the Philippines, where she teaches econometrics and mathematical economics at the School of Economics and conducts research on various topics on institutional economics, political economy, law and economics, and development. Some of her current projects include studies on the efficiency of prohibition vs. taxation, the evolution of institutions, and the problem of child soldiers in armed conflict. Prior to her

current appointment, Desierto was a Robert Solow Postdoctoral Fellow at the Cournot Centre for Economic Studies in France and a visiting researcher at the Faculty of Economics and Politics of the University of Cambridge. Desiree obtained her Ph.D. in Economics from the University of Nottingham in 2006, an M.Sc. in Economics for Development from the University of Oxford in 2002, an MEc in Economics from Macquarie University in 2000, and a BS in Business Economics from the University of the Philippines in 1997. E-mail: dadesierto@econ.upd.edu. ph.

**Jon X. Eguia** is an assistant professor in the Department of Politics and affiliated to the Department of Economics at New York University. He obtained his Ph.D. from the California Institute of Technology in 2007. His fields of specialization are political economy, formal political theory, social choice and public economics. His work on party formation, voting and elections, and the spatial model of policy preferences has been published in leading political science journals such as the *American Journal of Political Science* and the *Quarterly Journal of Political Science* and in economic theory journals such as *Social Choice and Welfare*, *Games and Economic Behavior*, *Economic Theory* and the *Journal of Mathematical Economics* among others. E-mail: eguia@nyu.edu.

**Lawrence Ezrow** is a Senior Lecturer of European Politics in the Department of Government at the University of Essex. He has written on political representation, elections, political parties, party strategies, and political institutions, and has published articles in the *American Journal of Political Science*, *British Journal of Political Science*, *Comparative Political Studies*, and the *Journal of Politics*. In 2010 he published a research monograph, *Linking Citizens and Parties*, with Oxford University Press. E-mail: ezrow@essex.ac.uk.

**Maria Gallego** has been concerned with leadership accountability in political economy situations and how the strategic interactions between political and economic agents determine policy outcomes. Using results from her empirical study of leadership transitions (*Canadian Journal of Economics*, 1996, and *Economics and Politics*, 1998) she then developed with Carol Pitchik an economic theory of leadership transition (*Journal of Public Economics*, 2004). With Cressman she extended Nash's bargaining model to allow agents to have single peaked rather than increasing preferences over the policy space (*Mathematical Social Science*, 2009). With Scoones she applied this theory to intergovernmental negotiations (*Social Choice and Welfare*, 2011). More recently, Gallego joined the research project with Schofield and has participated in studying individual elections in various countries including Turkey, Canada and Great Britain. E-mail: mgallego@wlu.ca.

**Andrei Gomberg** (born in Moscow 1974) is an Associate Professor of Economics at ITAM (Mexico City). He obtained his Ph.D. at New York University in 2000. His theoretical and experimental research is in the area of political economy and has been published, among others, in the *Journal of Economic Theory*, *Economic Theory*, the *International Journal of Game Theory* and *Social Choice and Welfare*. E-mail: gomberg@itam.mx.

**Yadira González de Lara's** main research projects inquire on the institutional foundations of markets and states in pre-modern and modern Europe using

comparative and historical institutional analysis. Her works have been published in the *American Economic Review*, the *Journal of Economic History* and the *European Review of Economic History*, among others. She earned her Ph.D. degree at the European University Institute and has taught and undertaken research at Stanford University, Ente Einaudi (Bank of Italy), University of Siena, and University of Alicante. She is currently assistant professor at the University of Valencia and researcher within the PROMETEO project at the University CEU-Cardenal Herrera. E-mail: yadira.gonzalez@uch.ceu.es.

**Avner Greif** is professor of economics at Stanford University and is the Bowman Family Professor in the Humanities and Sciences. Grief received his B.A. in Economics and History of the Jewish People and M.A. in History of the Jewish People from Tel Aviv University, Israel. He received his Ph.D. in Economics at Northwestern University, 1989. His works deal with economic history and role of institutions in economic development, including analysis of trade in medieval Europe and Levant. He specializes in the study of the social institutions that support economic development, and their history, incorporating game theory into his approach. He published *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. E-mail: avner@stanford.edu.

**Jee Seon Jeon** is a graduate student in political science at Washington University in St Louis. She received her B.A. in 2005 and M.A. in 2007 in political science from Seoul National University, Korea. Her research focuses on spatial model of elections, individual voting behavior and electoral institutions. E-mail: jsjeon@wustl.edu.

**Chris Kingston** is associate professor of economics at Amherst College. A native of Ireland, he holds a B.A. in Mathematics from Trinity College Dublin, and received his Ph.D. in economics from Stanford University in 2001. His research focuses on institutional economics, game theory, and economic history, with a particular focus on theories of institutional change, and on the development of the institutions governing marine insurance transactions during the eighteenth century. He lives in Amherst, Massachusetts. E-mail: cgkingston@amherst.edu.

**Daniel Kselman** is a Research Fellow at the Center for Advanced Studies in the Social Sciences, Juan March Institute (Madrid). His research employs both game theoretic and econometric modeling to study the impact if political institutions and political party organizations on a variety of contemporary political phenomena, including but not limited political corruption, legislative instability, and democratic accountability. He is also a student of modern Turkish politics, and has written on Turkish political institutions and party organizations. E-mail: dkselman@march.es.

**Marina Muskhelishvili**, Head of the Centre for Social Studies, Tbilisi, Georgia, took her M.A. in mathematics at Moscow State University in 1979 and a Ph.D. in mathematics in 1984 at Tbilisi State University, Georgia. Since 1995 she has worked in political science, studying various aspects of the institutional transformation of Georgia. From September 2009 to June 2010 she held a Fulbright Fellowship at Washington University in Saint Louis. She has published extensively on democratization and democratic theory. E-mail: marinmus@yahoo.com.

**Carol Mershon** is Associate Professor, Department of Politics, University of Virginia, and also, for 2009–2011, Political Science Program Director, National Science Foundation. A specialist in comparative politics, Mershon conducts research on intraparty competition, the dynamics of party systems, the politics of multiparty government, and legislative politics. Her articles have appeared in the *American Political Science Review*, the *American Journal of Political Science*, *Comparative Politics*, *Comparative Political Studies*, *Electoral Studies*, and the *Journal of Politics*, among others. Mershon is the author of *The Costs of Coalition* (Stanford University Press, 2002) and the co-editor of *Political Parties and Legislative Party Switching* (Palgrave Macmillan, 2009). E-mail: cmershon@nsf.gov.

**Antonio Nicolò** is an associate professor in the Department of Economics at the University of Padova. He obtained his Ph.D. from the Universitat Autonoma de Barcelona in 2000. His fields of specialization are mechanism design, social choice, political economy, contracts and incentives. His work on provision of public goods, problems of fair division and contract theory, have been published in economic journals such as *Journal of Economic Theory*, *Games and Economic Behavior*, *Economic Theory* and *European Economic Review* among others. E-mail: antonio.nicolo@unipd.it.

**John V.C. Nye** holds the Frédéric Bastiat Chair in Political Economy and is Professor of Economics at George Mason University. He is a specialist in economic history and the New Institutional Economics. He served on the faculty of Washington University in St. Louis for two decades, has been a National Fellow at the Hoover Institution at Stanford University and serves on the International Advisory Council of the Higher School of Economics, Moscow. He has done research on a variety of topics from firm size in France and the rise of the British fiscal state to Soviet collusion in championship chess, demography and superstition in Asia, and problems of reform in developing nations. He is the author of *War*, *Wine*, *and Taxes: The Political Economy of Anglo-French Trade 1689–1900* (Princeton, 2007) and was co-editor with John Drobak of *Frontiers in the New Institutional Economics* (Academic Press, 1997). E-mail: jvcnye@gmail.com.

**Ugur Ozdemir** is a graduate student in political science at Washington University in Saint Louis. He received his M.Sc. in Economics in 2005 from Istanbul Bilgi University and his B.S. in Industrial Engineering in 2003 from Bogazici University. He has published papers with Schofield in the *International Journal of Mathematics and Mathematical Systems*, *Social Choice and Welfare* and the *Czech Economic Review*. E-mail: uozdemir@wustl.edu.

**Sunita Parikh** is Associate Professor of Political Science at Washington University in St. Louis. She has published research on ethnicity in politics, political violence, and political and fiscal federalism. She is currently conducting research with Alfred Darnell on the emergence of judicial institutions in colonial and independent India. E-mail: saparikh@wustl.edu.

**Giovanni Resta** is a senior researcher in Computational Mathematics section of the Institute of Informatics and Telematics of the Italian CNR. He specializes in wireless and ad hoc networks, computational complexity, parallel and distributed algorithms, and computational methods in number theory. E-mail: g.resta@iit.cnr.it.

**Olga Shvetsova** is an associate professor of political science and economics at Binghamton University, USA. She studies political mechanisms as they are shaped by institutional choices. Her long-term research focus is on incentives to political elites and how those can lead to amelioration of societal conflict. She publishes in the fields of political economy of federalism and determinants of party system stability and change in democracies. She co-authored *Designing Federalism: A Theory of Self-Sustainable Federal Institutions* (with Mikhail Filippov and Peter C. Ordeshook, Cambridge University Press, 2004) and is now working on *Parties and Party System Change in Legislatures Worldwide* (with Carol Mershon). E-mail: shvetsova@gmail.com.

**Margit Tavits** took her BA and MA (*cum laude*) in 1999 and 2000 from the University of Tartu, Estonia, and her Ph.D. in Political Science from the University of Pittsburgh in 2004. She was an assistant professor in the Department of Political Science at the University of Missouri, and then a post-doctoral fellow at Nuffield College, Oxford. She joined the Department of Political Science at Washington University in Saint Louis in 2008, where she is now Associate Professor. She has published numerous articles in for example *Political Research Quarterly* (2009), the *American Political Science Review* (2008), the *American Journal of Political Science*, the *British Journal of Political Science* (2008), and *Electoral Studies* (2006). Her book, *Presidents and Prime Ministers*, was published by Oxford University Press in 2009. E-mail: tavits@wustl.edu.

**Fernando Toboso** is Senior Lecturer at the Department of Applied Economics, University of Valencia, Spain. His research interests are in methodological and epistemological issues in economic analysis as well as the organization of policy tasks and public services provision, with an emphasis on the new Spanish federalism. He has contributed to several edited volumes on these topics and has published in various academic journals including the *Cambridge Journal of Economics*, *Regional and Federal Studies*, *European Journal of Law and Economics*, *Journal of Interdisciplinary Economics*, *Hacienda Pública Española*, *El Trimestre Económico*, and *Presupuesto y Gasto Público*. He has been a visiting scholar at the Public Choice Center, George Mason University, Leeds University, and the University of Bologna, Italy. He has also served as invited editor and author for a special issue published in the *Chilean Revista de Análisis Económico*, and has recently co-edited a book on economics and institutions. E-mail: Fernando.Toboso@uv.es.

**Joshua A. Tucker** is an Associate Professor in the Politics Department with an Affiliated Appointment in the Department of Russian and Slavic Studies at New York University (NYU). Professor Tucker specializes in comparative politics with an emphasis on mass political behavior in East-Central Europe and the former Soviet Union, including elections and voting, the development of partisan attachment, and public opinion formation. He is the author of *Regional Economic Voting: Russia*, *Poland*, *Hungary*, *Slovakia*, *and the Czech Republic*, *1990–99* (Cambridge University Press, 2006). His work has appeared in numerous academic journals, including the *American Journal of Political Science*, *The Annual Review of Political Science*, *Electoral Studies* and *Post-Soviet Affairs*, and his opinions have been published in *The New Republic* and the *International Herald Tribune*. In 2006,

he was awarded the Emerging Scholar Award for the top scholar in the field of Elections, Public Opinion, and Voting Behavior within 10 years of the doctorate. E-mail: joshua.tucker@nyu.edu.

**Ignacio Urquizu-Sancho** completed his M.A. at the Juan March Institute and, in 2008, and received his Ph.D. in Sociology from Complutense University of Madrid. He is now a member of the Juan March Institute in Madrid and is Lecturer in Sociology at the Complutense University of Madrid.

He has been Visiting Fellow at Harvard University (EEUU), University of Essex (U.K.) and the European University Institute in Fiesole in Italy. His recent publications discuss Spanish electoral behavior and coalition governments. His works has been published in various journals, including *Electoral Studies*, *South European Society & Politics* and *Revista Española de Ciencia Política*.

His fields of interest include comparative politics, electoral behavior and the theory of democracy. E-mail: ignacio.urquizu@gmail.com.