Lecture Notes in Economics and Mathematical Systems 652

Sjoukje Osinga
Gert Jan Hofstede
Tim Verwaart *Editors*

Emergent Results of Artificial Economics

Springer

For further volumes:
http://www.springer.com/series/300

Sjoukje Osinga • Gert Jan Hofstede • Tim Verwaart
Editors

Emergent Results of Artificial Economics

Springer

Editors
Drs. Sjoukje Osinga
Assoc. Prof. Gert Jan Hofstede
Department INF/bode 143
Wageningen University
Hollandseweg 1
6706KN Wageningen
The Netherlands
sjoukje.osinga@wur.nl
gertjan.hofstede@wur.nl

Dr. Tim Verwaart
Agricultural Economics Institute
Wageningen UR
Alexanderveld 5
2585DB Den Haag
Netherlands
tim.verwaart@wur.nl

ISSN 0075-8442
ISBN 978-3-642-21107-2 e-ISBN 978-3-642-21108-9
DOI 10.1007/978-3-642-21108-9
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011932220

Cover design: eStudio Calamar S.L.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Now in its 7th year, the conference series in Artificial Economics is alive and kicking. A number of research topics are generating sustained interest. Others are emerging. We received 33 submissions, and based on almost 100 reviews (three per paper, on average), we now find 17 excellent papers in this volume that we called *Emergent Results of Artificial Economics*.

Artificial Economics can play a pioneering role in tackling precisely those issues that challenge the wider field of economics in general. In recent years, the international economy has repeatedly been shaken by crises that involve both the private and the public financial system on a worldwide scale. These crises take systems outside of their boundaries. In finding ways out, economists have to leave their comfort zones in which rational economic behaviour can be assumed. One can hear them speak of responsibility, of checks, of public control over the financial sector, of re-allocating risks. These are issues that impose challenges on economists, because they require both economic theory on markets and institutions, and deep understanding of human social behaviour. The practical issues to solve are at the nexus where these various areas meet. In order to address today's challenges, economists need to deal with the bigger picture.

Artificial economics is well poised to take on this challenge. Of necessity, artificial economists must integrate various bodies of knowledge into their models. The step to add another set of concepts, such as human social imitation behaviour, is not such a large one. Small improvements in modelling this kind of aspect can dramatically improve predictions of emergent system behaviour. This is the first reason why we included the word 'emergent' into the title of this volume. Through the use of models that integrate levels of aggregation, our field can build on existing knowledge and innovate at the same time. On the other hand, such advances do not come for free. Artificial Economics involves creatively developing new concepts for integrating into models. Then, extensive experimentation and painstaking validation of those models is required. When more elements are modelled together, the validation of the resulting models will be more intrinsically difficult. To really advance the field will require a lot of hard work. So far, some of that work is brought together in this volume. Hence the second reason why we liked to call it 'emergent results'.

The articles are grouped according to the following themes:

Humans in the system These contributions explicitly seek to include the modelling of human behaviours that might not be primarily economical in nature. This ranges from individual-based to social relationship-based. Jablonska and Kauranne simulate human emotions in econometric models of an electricity spot market. Thiriot, Lewkovicz, Caillou and Kant investigate the role of social networks on the labour market. Osinga, Kramer, Hofstede and Beulens model the pork cycle from farmer perspective and the potential of information management measures in the sector.

Financial markets Financial markets, a model for the most perfect of markets, are one of the better-researched areas in artificial economics. This area was also much affected by recent financial crises. Teglio, Raberto and Cincotti have captured this in an analysis of interactions among economic actors and their effects under varying leverage ratios of banks and at varying time horizons. Mallett extends existing work by including loan repayments into a model of a fractional reserve banking system. Hauser and LiCalzi study the evolution of trading strategies in unbalanced double auctions.

Organization design Artificial economics can also investigate the design of institutions such as departments, organizations or supply chains. Leitner and Wall are concerned with departmentalization and decision interdependencies in organizations, in relation to the desirability of stability or of discontinuity. McCain, Hamilton and Linnehan investigate whether one could have too many people in emergency departments. Valeeva and Verwaart model the adoption of food safety practices.

Macroeconomics This session assembles some issues that are fundamental to economics. Bersini and van Zeebroek address the 'hot topic' of inequalities in resource allocation. They do this in a model of a free market that can either be competitive or random. Desmarchelier, Djellal and Gallouj investigate the vicious historic cycle of growth by waste generation that has operated in the USA in the last half-century. Guerrero and Axtell assess the relative merits of neoclassical and computational models for exploring firm and labour dynamics.

Market dynamics Here we have another well-studied application of artificial economics. Garca-Daz and van Witteloostuijn investigate the starting conditions of a simulation of market structure. It turns out that these are of paramount importance for these markets' evolution. Fano and Pellizzari investigate the role of time constraints on optimality of order strategy in a continuous double auction. Provenzano presents a model of a wholesale electricity market and focuses on bilateral contracting.

Games The topic of games ranges from abstract game theory to actual simulation games involving human players. This session consists of an instance of both flavours. Cotla investigates the role of spatial structure on the performance of a network of prisoner's dilemma agents. Meijer, Raghothama, King and Palavalli present a gaming simulation of the Indian mango chain that allows to mix agents and people in one game run.

Wageningen,
May 2011

Tim Verwaart
Sjoukje Osinga
Gert Jan Hofstede

Acknowledgements

We would like to thank all the members of the Scientific Committee who generously donated their time as reviewers and who helped in various stages of the production of this volume:

- **Frédéric Amblard**, Université Toulouse 1, France
- **Luis Antunes**, Universidad de Lisboa, Portugal
- **Bruno Beaufils**, Université Lille 1, France
- **Giulio Bottazzi**, Scuola Superiore Sant' Anna, Italy
- **Olivier Brandouy**, Université Lille 1, France
- **Charlotte Bruun**, Aarhus University, Denmark
- **Andrea Consiglio**, Università degli Studi di Palermo, Italy
- **Giorgio Fagiolo**, Scuola Superiore Sant' Anna, Italy
- **Jose Manuel Galán**, Universidad de Burgos, Spain
- **Florian Hauser**, Universität Innsbruck, Austria
- **Cesàreo Hernandez**, Universidad de Valladolid, Spain
- **Gert Jan Hofstede**, Wageningen University, The Netherlands
- **Juergen Huber**, Universität Innsbruck, Austria
- **Wander Jager**, University of Groningen, The Netherlands
- **Marco Janssen**, Arizona State University, USA
- **Michael Kirchler**, Universität Innsbruck, Austria
- **Marco LiCalzi**, Università CaFoscari Venezia, Italy
- **Thomas Lux**, Christian-Albrechts-Universität Kiel, Germany
- **Rosario Mantegna**, Università degli Studi di Palermo, Italy
- **Philippe Mathieu**, Université Lille 1, France
- **Javier Pavares**, Universidad de Valladolid, Spain
- **Juan Pavón**, Universidad Computense de Madrid, Spain
- **Paolo Pellizzari**, Università CaFoscari Venezia, Italy
- **Denis Phan**, GEMAS, France
- **Paolo Pin**, Università degli Studi di Siena, Italy
- **Marta Posada**, Universidad de Valladolid, Spain
- **Juliette Rouchier**, GREQAM, France

- **Annalisa Russino**, Università degli Studi di Palermo, Italy
- **Klaus Schredelseker**, Universität Innsbruck, Austria
- **Elpida Tzafestas**, University of Athens, Greece
- **Tim Verwaart**, LEI, Wageningen UR, The Netherlands
- **Eugene Westerhof**, LEI, Wageningen UR, The Netherlands
- **Murat Yildizoglu**, Université Paul Cézanne, GREQAM, France

Contents

List of Contributors

Robert L. Axtell
Department of Computational Social Science, George Mason University, USA
e-mail: rax222@gmu.edu

Hugues Bersini
IRIDIA-CODE, Université Libre de Bruxelles, CP 194/6,50 av. Franklin Roosevelt, 1050 Bruxelles, Belgium
e-mail: bersini@ulb.ac.be

Adrie J.M. Beulens
Wageningen Univerity, P.O. Box 8130, NL-6700 EW Wageningen, The Netherlands
e-mail: Adrie.Beulens@wur.nl

P. Caillou
LRI, Universite Paris Sud, France
e-mail: caillou@lri.fr

Silvano Cincotti
DIBE-CINEF, Università di Genova, Via Opera Pia 11a, 16145 Genova, Italy
e-mail: cincotti@dibe.unige.it

Chenna Reddy Cotla
Department of Computational Social Science, George Mason University, USA
e-mail: ccotla@gmu.edu

Benoît Desmarchelier
Lille 1 University, 59655 Villeneuve d'Ascq Cedex, France
e-mail: benoit.desmarchelier@ed.univ-lille1.fr

Faridah Djellal
Lille 1 University, 59655 Villeneuve d'Ascq Cedex, France
e-mail: faridah.djellal@univ-lille1.fr

Shira Fano
Dept. of Economics, Bocconi University Milan, Italy
e-mail: shira.fano@gmail.com

Faïz Gallouj
Lille 1 University, 59655 Villeneuve d'Ascq Cedex, France
e-mail: Faiz.Gallouj@univ-lille1.fr

César García-Díaz
Management Department (ACED), Faculty of Applied Economics, University of Antwerp, Prinsstraat 13, BE-2000 Antwerp (Belgium),
e-mail: cesar.garcia-diaz@ua.ac.be

Omar A. Guerrero
Department of Computational Social Science, George Mason University
e-mail: oguerrer@gmu.edu

Richard Hamilton
Department of Emergency Medicine, Drexel University College of Medicine, 245 N. 15th Street, Mailstop 1011, Phila. PA 19102, USA
e-mail: richard.hamilton@drexelmed.edu

Florian Hauser
Dept. Banking and Finance, Universität Innsbruck,
e-mail: florian.hauser@uibk.ac.at

Gert Jan Hofstede
Wageningen Univerity, P.O. Box 8130, NL-6700 EW Wageningen, The Netherlands
e-mail: GertJan.Hofstede@wur.nl

Matylda Jabłońska
Lappeenranta University of Technology, Department of Mathematics and Physics, P.O.Box 20, 53850 Lappeenranta, Finland,
e-mail: matylda.jablonska@lut.fi

J.-D. Kant
LIP6, Universite Paris VI, France
e-mail: kant@lip6.fr

Tuomo Kauranne
Lappeenranta University of Technology, Department of Mathematics and Physics, P.O.Box 20, 53850 Lappeenranta, Finland,
e-mail: tuomo.kauranne@lut.fi

Robin King
CSTEP, Bangalore
e-mail: robin@cstep.in

Mark R. Kramer
Wageningen Univerity, P.O. Box 8130, NL-6700 EW Wageningen, The Netherlands
e-mail: Mark.Kramer@wur.nl

Stephan Leitner
Alpen-Adria Universitaet Klagenfurt, Department of Controlling and Strategic Management, Universitaetsstrasse 65-67, 9020 Klagenfurt, Austria
e-mail: stephan.leitner@uni-klu.ac.at

Zach Lewkovicz
LIP6, Universite Paris VI, France
e-mail: zach.lewkovicz@lip6.fr

Marco LiCalzi
Dept. Management, Università Ca' Foscari Venezia,
e-mail: licalzi@unive.it

Frank Linnehan
Lebow College of Business, Drexel University, Philadelphia, PA, 19104, USA
e-mail: linnehf@drexel.edu

Jacky Mallett
University of Reykjavik, Institute of Intelligent Machines, Klapparstigur 18, 101 Reykjavik, Iceland
e-mail: warlock@alum.mit.edu

Roger A. McCain
Lebow College of Business, Drexel University, Philadelphia, PA, 19104, USA
e-mail: mccainra@drexel.edu

Sebastiaan. A. Meijer
TUDelft, Netherlands
e-mail: Sebastiaan.Meijer@tudelft.nl

Sjoukje A. Osinga
Wageningen Univerity, P.O. Box 8130, NL-6700 EW Wageningen, The Netherlands
e-mail: Sjoukje.Osinga@wur.nl

Bharath Palavalli
CSTEP, Bangalore
e-mail: bharath@cstep.in

Paolo Pellizzari
Dept. of Economics, University Ca' Foscari of Venice
e-mail: paolop@unive.it

Davide Provenzano
Dipartimento di Scienze Statistiche e Matematiche "Silvio Vianelli", University of Palermo, Italy
e-mail: provenzano@unipa.it

Marco Raberto
DIBE-CINEF, Università di Genova, Via Opera Pia 11a, 16145 Genova, Italy
e-mail: marco.raberto@dibe.unige.it

Jayanth Raghothama
CSTEP, Bangalore
e-mail: jayanth@cstep.in

Andrea Teglio
Departament d'Economia, Universitat Jaume I, Av. Sos Baynat, Castellón de la Plana, Spain
e-mail: teglio@eco.uji.es

Samuel Thiriot
TAO Project, INRIA Saclay, France
e-mail: samuel.thiriot@res-ear.ch

Friederike Wall
Alpen-Adria Universitaet Klagenfurt, Department of Controlling and Strategic Management, Universitaetsstrasse 65-67, 9020 Klagenfurt, Austria
e-mail: friederike.wall@uni-klu.ac.at

Natalia I. Valeeva
LEI Wageningen UR, Postbus 29703, 2502 LS den Haag, The Netherlands
e-mail: natasha.valeeva@wurl.nl

Tim Verwaart
LEI Wageningen UR, Postbus 29703, 2502 LS den Haag, The Netherlands
e-mail: tim.verwaart@wur.nl

Arjen van Witteloostuijn
Management Department (ACED), Faculty of Applied Economics, University of Antwerp, Prinsstraat 13, BE-2000 Antwerp (Belgium),
e-mail: arjen.vanwitteloostuijn@ua.ac.be

Nicolas van Zeebroeck
ECARES and IRIDIA-CODE, Université Libre de Bruxelles, CP 114/04, 50 av. Franklin Roosevelt, 1050 Bruxelles, Belgium
e-mail: nivzeebr@ulb.ac.be

Part I
Humans in the System

Multi-Agent Stochastic Simulation for the Electricity Spot Market Price

Matylda Jabłońska and Tuomo Kauranne

Abstract The Great Recession of 2008-2009 has dented public confidence in econometrics quite significantly, as few econometric models were able to predict it. Since then, many economists have turned to looking at the psychology of markets in more detail. While some see these events as a sign that economics is an art, rather than a science, multi-agent modelling represents a compromise between these two worlds. In this article, we try to reintroduce stochastic processes to the heart of econometrics, but now equipped with the capability of simulating human emotions. This is done by representing several of Keynes' Animal Spirits with terms in ensemble methods for stochastic differerential equations. These terms are derived from similarities between fluid dynamics and collective market behavior. As our test market, we use the price series of the Nordic electricity spot market Nordpool.

1 Introduction

After the Great Recession of 2008-2009, many mathematical and econometric models used in economy have received a lot of criticism, since they were not able to predict the emergence of the asset bubble in the U. S. housing market. As a result of this, econometricians have increasingly turned towards seeking explanations to what happened in the psychological element in market traders' actions. This has repeatedly brought up the idea of emotions that influence human economic behavior. These emotions are also known as *animal spirits* and were originally introduced by John Maynard Keynes in his 1936 book [15].

Not only global economic upheavals display such behavior. A different type of extreme event can be observed in deregulated electricity spot markets which are known to be one of the most volatile financial markets. This distinctive phenomenon

Matylda Jabłońska and Tuomo Kauranne
Lappeenranta University of Technology, Department of Mathematics and Physics, P.O.Box 20, 53850 Lappeenranta, Finland e-mail: matylda.jablonska@lut.fi, e-mail: tuomo.kauranne@lut.fi

S. Osinga et al. (eds.), *Emergent Results of Artificial Economics*, Lecture Notes in Economics and Mathematical Systems 652, DOI 10.1007/978-3-642-21108-9_1,

is the appearance of price spikes, i.e. sudden price changes to values up to dozens of times higher within only an hour, and again falling back to the previous level within a couple of hours or days. After each spike, market specialists are able to find a reason that caused it in hindsight. But few of those reasons are reliable predictors for future spikes [4, 25]. Nor has any econometric model shown any skill in forecasting those sudden price changes, to these authors' knowledge.

In the current study, we investigate the possible origin of price spikes in animal spirits that rule the behavior of all traders. Our modelling attempt of electricity spot markets is not directly based on any notion of intelligent – albeit emotional – agents. Instead, we have sought to equip classical Ornstein-Uhlenbeck type mean-reverting econometric models with new non-linear terms that emulate the impact of a distinct animal spirit each. Although the motivation for these models is not grounded in the psychology of intelligent agents, they represent very similar collective behavior to that of models based on intelligent agents. So does the real price history of electricity on the NordPool Spot market, the largest electricity spot market in the world.

This modelling approach is not limited to only spot markets for electricity, where prominent spikes can partly be understood as a consequence of the non-storability of electricity. After suitable normalizing transformations, identical models simulate accurately the behavior of other commodity markets. We have conducted such a study also on the oil spot market for Brent crude. In both cases, the free parameters in the terms of the models are calibrated with a Bayesian Maximum Likelihood principle from the real time series. After this step, the resulting simulated price series reproduces the distribution of the real price series almost exactly up to the sixth statistical moment.

The article is organized as follows. Sect. 2 presents theoretical aspects used in our proposal. Sect. 3 describes briefly the data set, as well as the two proposed models and their results. Sect. 4 concludes.

2 Theoretical Framework

2.1 Electricity Spot Market Price

As we already mentioned, electricity spot markets and their prices are very distinctive due to electricity non-storability. Therefore, defining appropriate models that would relevantly forecast spot prices is very challenging in both short and long term. Some researchers took trials in analyzing specific auction theories to investigate the electricity market power. The two most common ones are supply function equilibrium and multi-unit independent private value [27]. Some claim that information available to power traders is asymmetric on both auction sides which violates the goal of lowering volatility and marginal price [1]. An idea of modelling several competitive traders in an electricity market as a coupled system of mathematical programs with equilibrium constraints was proposed in [9], however, without explicit numerical results.

Depending on type of electricity market, there have been different attempts based in classical time series theory [28], methodologies utilizing price periodicities [37, 21], basic and more elaborate stochastic mean reversion models [10], wait-jump structures [13] or regime switching approaches [35]. What is certain is that a lot of electricity price local trends and part of the volatility can be explained by historical information on factors known to be driving the prices [31]. Nevertheless, even though we know how important specific patterns are in the electricity prices [20], the most challenging part of modeling their dynamics lies in the non-explained part which we claim to be highly influence by traders' psychology. Therefore, as presented later in Sect. 3.1, we calibrate our simulation based on a time series from which we remove all known deterministic factors.

2.2 Animal Spirits in Financial Markets

The term *animal spirits* appeared in literature already in 1936, introduced by Keynes [15]. However, as we can read from [22], many specialists did not want to accept importance of psychology as one of major economy drivers. An attempt to formalize Keynes' forces can be found in a work related to catastrophe theory [8]. Some researchers focused on the idea of risk-aversion [17]. Others have found an inverse relation between consumer and business confidence and national unemployment rate [23].

The issue gained notoriety when Akerlof and Shiller published their book discussing the trading psychology [2]. They claim that real financial dynamics is strongly based in irrational, emotional and often intuitive decisions by human agents. Even if agent-based modelling builds upon individual psychology, also government decisions still have human factors behind and, therefore, economies fall globally [16].

Other authors underline the importance of trust and confidence [33] or rational expectations [18] as crucial forces pulling markets towards or away from economic crisis. Multi-agent models have become a popular method to address human emotions, and they have been applied to macroeconomy [7], where agents adaptively learn from their mistakes. Other models specifically cater for transaction taxes, greed and risk aversion in [6]. A number of studies apply agent-based models to electricity markets [32]. However, they have focused on the technical aspect of meeting demand and supply. Such models often work well only for regular price evolution. We argue that price spikes originate in human psychology. In this study, we present an ensemble model that accounts for some of the animal spirits in the spot markets.

2.3 Capasso-Morale-Type Population Dynamics

We know, that people, as other animal species, have animal spirits. These, in financial market mean mostly fear and greed, influenced then by our collective trading biases: herding, overconfidence and short-term thinking. The type of model we have adopted are based on the Capasso-Morale system of stochastic differential equations (1), used so far for modelling animal population dynamics (see [24]). Its basic equation has the form

$$dX_N^k(t) = [\gamma_1 \nabla U(X_N^k(t)) + \gamma_2(\nabla(G - V_N) * X_N)(X_N^k(t))]dt + \sigma dW^k(t), \qquad (1)$$

for $k = 1, \ldots, N$. The equation describes physical herding of animal populations.

In our case, the population is a group of traders in the spot market, and a measure of their spatial distance is the price. Traders do observe one another and thus create a mean price path, which could be also understood as the global (in *macroscale*) population formation. However, there is also a limit for overcrowding (in *microscale*) which in power trading could be interpreted as physical impossibility of two market participants to buy the same asset of electricity. Each individual price path simulated from the model proposed in this paper represents a single trader, and the simulation of the ensemble would provide coupling between the participants (in *mesoscale*). The movement of each particle is driven by an external information coming from the environment, expressed via suitable potentials.

2.3.1 Momentum in Financial Markets

Since the 1980s researchers have been repeatedly noticing that, on average, stocks performing well keep doing so over some further time. It is called the momentum effect. Similar behaviour can be observed in any commodity markets. In many funds the managers are rewarded for good performance and for beating the market. Thus they must be holding the most popular and rapidly appreciating stocks. When they perform well, clients invest even more money, which again goes into the same investments and boosts shares that have already performed well even further. Simply put, investors are buying stocks just because their price has risen. This is the essence of the momentum effect.

A physical analogy to the momentum phenomenon can be found in fluid dynamics. The Burgers' equation (2) is a one-dimensional form of the Navier-Stokes equations without the pressure term and volume forces. It is widely used in various areas of applied mathematics, such as modeling of fluid dynamics and traffic flow [5, 11].

$$u_t + \theta u u_x + \alpha u_{xx} = f(x,t) \qquad (2)$$

To build an analogy between markets and fluids, the price represents some one dimensional measurement of fluid, such as pressure, along a periodic domain. This characterization is not far from stock market reality. Worldwide trading takes place in a periodical domain of the earth. With one exchange closing in one time zone, another one is opening for another trading day on the next continent. The information circulates in a periodic fashion around the world.

In the Burgers' equation (2) u stands for the price, $f(x,t)$ describes the fundamentals (often of a periodic character), αu_{xx} is the diffusion term that is related to the fact that the spot market tends to reach an equilibrium price, u_x is the spread between any given day's average and most common bids, i.e. the mean and the mode of the bid distribution, uu_x is the momentum term that expresses traders' tendency to move towards the most common price. This effect is magnified at higher prices.

The momentum effect should occur when a sufficiently big subgroup of the whole population has significantly different behavior (external information) that deviates from the total population mean. This has been noticed in studies related to animal and human spatial dynamics, when a large group of people is asked to keep moving randomly around a big hall. When a sub-group of five per cent or more are silently told to move towards a given target, the whole population will follow. Again, in terms of prices this could be understood as considerable departure of the mode of the population price distribution from its mean price. Then the rest of the individuals may follow that trend and this effect unexpectedly amplifies that deviation to a scale of a prominent price spike.

3 Multi-Agent Simulations for Electricity Spot Market

3.1 Data

The data used for this study is the daily system price from the Nord Pool electricity spot market covering a horizon of over 10 years from January 1999 to February 2009. However, for model and simulation calibration we do not take the original prices, but its detrended and deseasonalized version. We motivate it by the fact that electricity spot prices are known to be highly periodical seasonally [37], as well as weekly [21]. Some of these effects may be removed by regression models, based on high correlation of electricity prices with specific background variables, such as load [34], production type [3, 12], temperatures [29], and other different climatic factors [19].

Therefore, we first remove specific types of periodicity with use of classical time series additive decomposition, and then build a regression model employing available explanatory variables, that is temperatures and water reservoir levels. Moreover, the regression is not run globally but in a moving fashion with half a year horizon (182 days). This is motivated by the fact that except for obvious periodicities, there are also other cycles (like economic) driving electricity prices. Also, influence of specific factors on the prices changes over years.

The fit and resulting residual series, also claimed to be a pure trading price series reflecting more clearly the electricity spot market dynamics, are presented in Fig. 1.

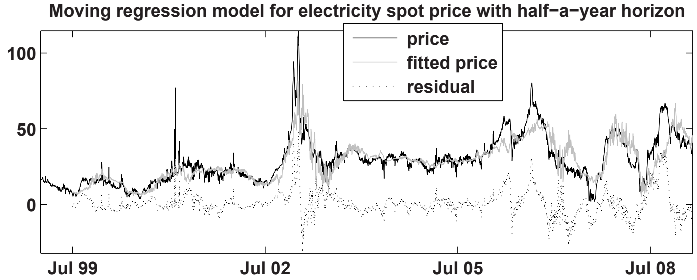

Fig. 1 Moving regression fit for Nord Pool system price with half-a-year horizon window.

3.2 *Mean-Reverting Jump Diffusion Ensemble Simulation*

In this study we propose to represent individual spot price traders as an ensemble. Price realizations of all of them are described with a system of stochastic differential equations (*Lagrangian representation*). As it was mentioned in [24], this approach makes sense for small or medium-sized populations. To reflect reality, we set the ensemble size to 300, because currently the Nord Pool market has approximately 330 participants.

In particular, each of those differential equations has form (3)

$$dX_t^k = \gamma_t[(X_t^* - X_t^k) + (f(k, \mathbf{X}_t) - X_t^k)]dt + \sigma_t dW_t^k + {}^+J_t^k dN_t + {}^-J_t^k dN_t, \quad (3)$$

for $k = 1, \ldots, N$, where X_t^k is the price of trader k at time t, X_t^* is the global price reversion level at time t, γ_t is the mean reversion rate at time t, $\mathbf{X}_t$ is the vector of all traders' prices at time t, $f(k, \mathbf{X}_t)$ is a function describing local interaction of trader k with his neighbors (small range of individuals from vector $\mathbf{X}_t$), W_t^k is the Wiener process value for trader k at time t, σ_t is the standard deviation for Wiener increment at time t, ${}^+J_t^k$ is the positive jump for trader k at time t, ${}^-J_t^k$ is the negative jump for trader k at time t, N_t is the count process for jumps at time t.

The model parameters of the mean reverting part are estimated with use of Maximum Likelihood (MLE) approach. The log-likelihood function for Ornstein-Uhlenbeck process can be found from [26]. The probabilities of jumps are generated from Poisson distribution based on probability of spike occurrence from specific price levels. The jump sizes are estimated from empirical distribution of the original prices.

In this model we follow the global mean reversion level X_t^* and rate γ_t in a moving fashion with half a year historical horizon (182 days). This feature represents *short-term thinking*, that is one of the main trading biases characterizing market participants. The local interaction $f(k, \mathbf{X}_t)$ is based on following the mean value of neighbors within price range equal to 10% of the total price range, and it stands for the *herding* bias. The jump processes ${}^{+}J$ and ${}^{-}J$ are dependent on current price level at each time t, as we know that electricity spot price is more likely to spike from higher levels than from lower [14]. Therefore, spikes generated by the jumps are reflecting *panic* reaction of traders in the uncertain environment, on both positive and negative side. We could claim that they originate from human *fear and greed* emotions.

In Fig. 2 we can see the original price and example simulated trajectory (for one out of 300 traders) together with their respective histograms. We can see that the simulation nicely follows the original data, both in the long term and in the appearance of spikes.

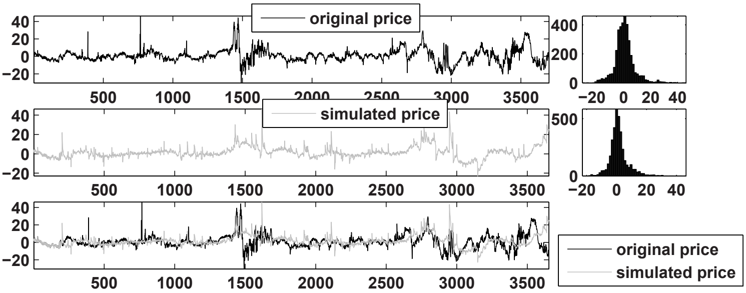

Fig. 2 Ensemble simulation: global reversion to moving mean level with moving rate, and local to neighbors' mean.

The original and simulated histograms are similar. However, we want to quantify the difference as well. Therefore, Table 1 collects comparison of the basic statistics for original pure trading prices and the mean ensemble values. These are mean, standard deviation and five consecutive central moments. We observe that especially skewness and kurtosis have values very close to one another. The 5th moment is still comparable. Only 6th moment starts to be higher for the ensemble than the original data by the factor of nearly 1.5 and the 7th moment by the factor of 2.

To complement the whole analysis we employ one more comparison measure, i.e. the probability of the series to exceed specific levels. We slice both the original pure price series and each simulated ensemble series into ten intervals each, from their respective minimum to maximum values. Then we calculate the percentage of observations falling above each slice threshold. These probabilities are illustrated in Fig. 3, together with the respective mean value and standard deviation limits of the original price.

Table 1 Original and ensemble statistics: global reversion to moving mean level with moving rate, and local to neighbors' mean.

	Original	Ensamble
Means	0.72	1.69
St dev	7.47	6.14
Skewness	0.92	0.94
Kurtosis	6.97	6.91
5th moment	18.81	22.09
6th moment	104.91	150.43
7th moment	423.45	832.39

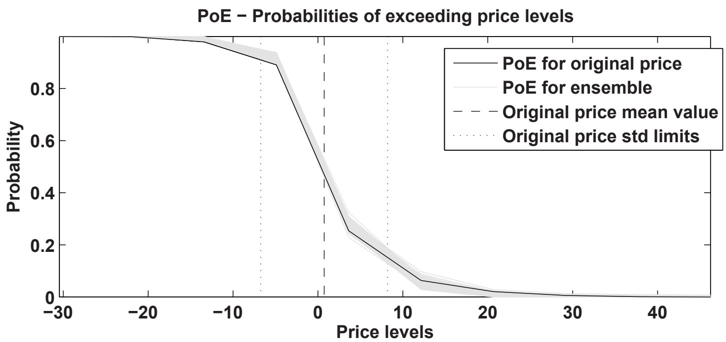

Fig. 3 Original pure price and ensemble probabilities of exceeding specific levels.

Clearly, the real data's probabilities fall within the envelope of the whole ensemble. That confirms statistical accuracy and robustness of our approach.

3.3 Ensemble Simulation with Burgers'-Type Interaction

The ensemble model proposed in Sect. 3.2 reproduces the real price dynamics very well. However, it has a weakness in that the jump components are just superimposed on the base mean-reverting process. They should rather be based directly on price dynamics. We propose to eliminate the jump processes from the model (3) and replace the mean-based local interactions $f(k,\mathbf{X}_t)$ with a Burgers'-type momentum component $h(k,\mathbf{X}_t)$. Thus the model takes the form

$$dX_t^k = [\gamma_t(X_t^* - X_t^k) + \theta_t(h(k,\mathbf{X}_t) - X_t^k)]dt + \sigma_t dW_t^k, \tag{4}$$

where $h(k,\mathbf{X}_t) = \mathbf{E}(\mathbf{X}_t)\cdot[\mathbf{E}(\mathbf{X}_t) - \mathbf{M}(\mathbf{X}_t)]$ and $\mathbf{M}(X)$ stands for the mode of a random variable X. Also, θ_t represents the strength of that local interaction at time t.

The model estimation is also done by MLE. Following the solution of a mean reverting process we get that the process has mean value and variance in form of (5) and (6).

$$\mathbf{E}(\mathbf{X}_t) = [\gamma X^* + \theta \mathbf{E}(\mathbf{E}(P_t)(\mathbf{E}(P_t) - \mathbf{M}(P_t)))] \cdot \frac{1}{\gamma+\theta} \cdot (1 - e^{-(\gamma+\theta)t}) \tag{5}$$

$$\mathbf{Var}(\mathbf{X}_t) = \frac{\sigma^2}{2(\gamma+\theta)} \cdot (1 - e^{-2(\gamma+\theta)t}) \tag{6}$$

When assuming for simplicity that the process is normally distributed, the population mean $\mathbf{E}(P_t)$ is expected to equal the population mode $\mathbf{M}(P_t)$, and thus the log-likelihood function takes form (7).

$$\mathscr{L}(\mathbf{X}, X^*, \gamma, \theta, \sigma) = n \ln\left(\frac{1}{\sqrt{2\pi \mathbf{Var}(\mathbf{X}_t)}}\right) - \sum \frac{(X_i - \mathbf{E}(\mathbf{X}_t))^2}{\mathbf{Var}(\mathbf{X}_t)} \tag{7}$$

The simulation results for this model can be seen in Fig. 4. The general price level follows the original data. Moreover, the simulation spikes, even though the model does not have any jump component. Thanks to interactions of the individuals, the price spikes are based on the pure price dynamics. One can see that spikes in the simulation are not as frequent as in original data. We can blame here the normality assumption for MLE. Future work could consider some numerical methods for MLE of parametric distributions like g-and-h.

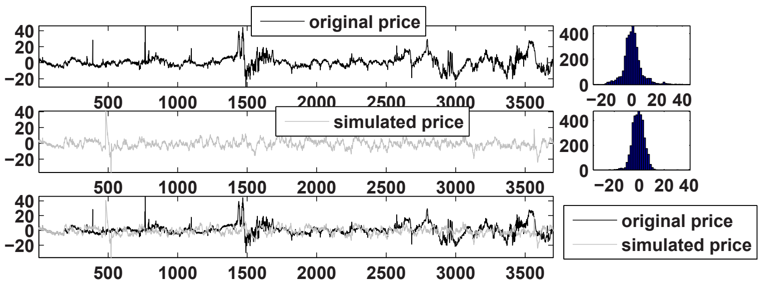

Fig. 4 Ensemble simulation: global reversion to moving mean level with moving rate, and Burgers'-type local interaction.

Table 2 collects the statistics for the original price and the ensemble. The values of central moments are not as close to the original ones as the first model was able to produce, but the distribution of the simulated price is leptokurtic on a level similar to that of the true price.

When we measure the probability of the price to exceed a specific price level, see Fig. 5, it is visible that the ensemble envelope does not cover the original data in the range of positive extreme values well, but the results are still promising. The model

Table 2 Original and ensemble statistics: global reversion to moving mean level with moving rate, and Burgers'-type local interaction.

	Original	Ensamble
Means	0.72	-0.43
St dev	7.47	5.01
Skewness	0.92	0.35
Kurtosis	6.97	7.68
5th moment	18.81	24.40
6th moment	104.91	242.69
7th moment	423.45	380.57

can probably improved by changing the MLE function or by enriching the model with suitable potential functions representing the economic situation.

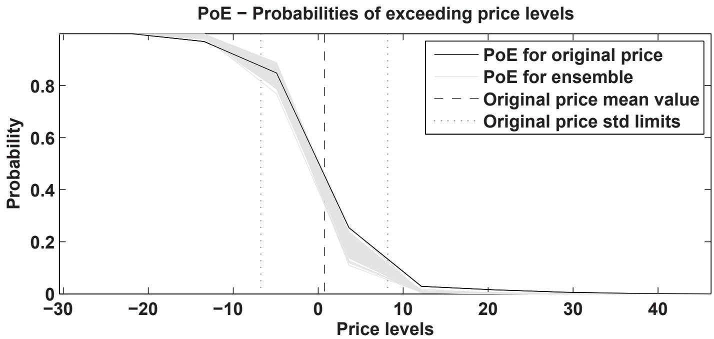

Fig. 5 Original pure price and ensemble probabilities of exceeding specific levels.

4 Conclusions

In this work we have presented two possible multi-agent models that simulate bids of electricity spot market participants. The study was carried out on data originally coming from Nord Pool market. However, we have removed any known deterministic factors from the available price time series, so that we are left only with data reflecting best the true spot market dynamics.

The multi-agent models that we proposed in this paper were based on a Capasso-Morale-type population dynamics Lagrangian approach, where movement of each individual is described by a separate stochastic differential equation. However, these agents keep interacting with each other at each time instant, on both local and global basis. Our model caters for the most common trading biases, i.e. short-term thinking and herding. Also, we have included terms that represent panic that originates from market uncertainty. Finally, the second model has eliminated the need for a separate jump component in the simulation. Instead, it uses an interaction term that has been borrowed from fluid dynamics that represents market momentum.

Simulation results presented in our paper prove that our approaches reproduces well many statistical features of the real spot price time series. This was measured by comparing distribution histograms of the original and simulated series, through statistical central moments up to the 6th order, as well as by the probability of the prices to exceed a specific level. All these showed remarkable resemblance. Also, the second simulation was able to reproduce price spikes based only on price dynamics and ensemble behavior.

As to suggestions for future work, we hope to improve the multi-agent model still by employing more elaborate functions for the local interaction of traders, as well as by the inclusion of potentials that would represent market information available to the traders. Also, the MLE assumptions of model (4) should be revised.

References

1. Abbink, K., Brandts, J., and McDaniel, T.: Asymmetric Demand Information in Uniform and Discriminatory Call Auctions: An Experimental Analysis Motivated by Electricity Markets. Journal of Regulatory Economics. **23**(2), 125 – 144 (Kluwer Academic Publishers, 2003)
2. G.A. Akerlof, R.J. Shiller, *Animal Spirits. How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism* (Princeton University Press, 2009)
3. Batlle, C. and Barquin, J.: A strategic production costing model for electricity market price analysis. IEEE Trans. Power Systems. **20**(1), 67 – 74 (2005)
4. Baya, H., Buchasia, C., Rykfors, M., Saint-Aubain, P.A., Vecchio, I., Wangwe, I., and Yedeg, E.L. (2009) Influence of physical factors on electricity spot market price. In: ECMI Student Mathematical Modelling Week, Wrocław, Poland, 23-30 Aug
5. Cole, J.D., On a quasilinear parabolic equation occurring in aerodynamics. Quart. Appl. Math. **9**(3), 225–236 (1951)
6. Demary, M.: Transaction taxes, greed and risk aversion in an agent-based financial market model. J. Econ. Interact. Coord. (2010) doi: 10.1007/s11403-010-0071-9
7. Grauwe, P. de: Animal spirits and monetary policy. Econ. Theory. (2010) doi: 10.1007/s00199-010-0543-0
8. Harris, L.: Catastrophe theory, utility theory and animal spirit expectations. Australian Economic Papers. Dec 268 – 282 (1979)
9. Henrion, R. and Römisch, W.: On M-stationary points for a stochastic equilibrium problem under equilibrium constraints in electricity spot market modeling. Applications of Mathematics. **52**(6), 473 – 494 (2007)
10. Higgs, H. and Worthington, A.: Stochastic price modeling of high volatility, mean-reverting, spike-prone commodities: The Australian wholesale spot electricity market. Energy Economics. **30**, 3172 – 3185 (2008)
11. Hopf, E.: The partial differential equation ut + uux = uxx, Comm. Pure and Appl. Math. **3**, 201–230 (1950)
12. Hreinsson, E.B. (2006) Supply Adequacy Issues in Renewable Energy and Hydro-Based Power System. In: International Conference on Power System Technology. (pp. 1 – 6)
13. Jabłońska, M., Mayrhofer, A., and Gleeson, J.: Stochastic simulation of the Uplift process for the Irish Electricity Market. Mathematics-in-Industry Case Studies. **2** 86 – 110 (2010)
14. Jabłońska, M., Nampala, H., and Kauranne, T.: Multiple mean reversion jump diffusion model for Nordic electricity spot prices. Journal of Risk. To be published (2011)

15. Keynes, J.M., *The General Theory of Employment, Interest and Money* (London: Macmillan, 1936)
16. Kling, A.: Animal Spirits. The Independent Review. **14**(1) 135 – 140 (2009)
17. Kupiec, P.H. and Sharpe, S.A.: Animal Spirits, Margin Requirements, and Stock Price Volatility. The Journal of Finance. **46**(2) (1991)
18. Kurz, M.: Symposium: on the role of market belief in economic dynamics, an introduction. Econ. Theory. (2010) doi: 10.1007/s00199-010-0560-z
19. Laitinen, K., Hovila, J., Mannila, T., and Korpinen, L. (2000) The influences of climatic factors on electricity prices in liberalized market in Finland. In: International Conference on Electric Utility Deregulation and Restructuring and Power Technologies. (pp. 544 – 548)
20. Lucia, J.J. and Schwartz, E.S.: Electricity Prices and Power Derivatives: Evidence from the Nordic Power Exchange. Review of Derivatives Research. **5**, 5 – 50 (Kluwer Academic Publishers, 2002)
21. Mandal, P., Senjyu, T., Urasaki, N., Funabashi, T., and Srivastava, A.K.: A Novel Approach to Forecast Electricity Price for PJM Using Neural Network and Similar Days Method. IEEE Trans. Power Systems. **22**(4), 2058 – 2065 (2007)
22. Marchionatti, R.: On Keynes' Animal Spirits. Kyklos. **52** 415 – 439 (1999)
23. Middleton, E.: Adaptation level and 'animal spirits'. Journal of Economic Psychology. **17** 479 – 498 (1996)
24. Morale, D., Capasso, V., and Oelschlager, K.: An interacting particle system modelling aggregation behavior: from individuals to populations. Journal of Mathematical Biology. (2004)
25. Murara, J-P. (2010) Regime switching models for electricity time series that capture fat tailed distributions. Master's Thesis, Lappeenranta University of Technology.
26. Øksendal, B. (1995) *Stochastic Differential Equations: An introduction with applications*, 4th ed., Springer Verlag, New York, NY.
27. Parisio, L and Bosco, B.: Market Power and the Power Market: Multi-Unit Bidding and (In)Efficiency in Electricity Auctions. International Tax and Public Finance. **10**, 377 – 401 (Kluwer Academic Publishers, 2003)
28. Ptak, P., Jabłońska, M., Habimana, D., and Kauranne, T. (2008) Reliability of ARMA and GARCH Models of Electricity Spot Market Prices. In: European Symposium on Time Series Prediction, Porvoo, Finland
29. Ruibal, C.M. and Mazumdar, M.: Forecasting the Mean and the Variance of Electricity Prices in Deregulated Markets. IEEE Trans. Power Systems. **23**(1), 25 – 32 (2008)
30. Schwartz, H.: Does Akerlof and Shillers *Animal Spirits* provide a helpful new approach for macroeconomics? The Journal of Socio-Economics. **39** 150 – 154 (2010)
31. Strozzi, F., Gutiérrez, E., Noè, C., Rossi, T., Serati, M., and Zaldívar, J.M.: Measuring volatility in the Nordic spot electricity market using Recurrence Quantification Analysis. Eur. Phys. J. Special Topics **164**, 105 – 115 (EDP Sciences, Springer-Verlag 2008)
32. Tesfatsion, L. (2011) ACE Research Area: Electricity Market Research. http://www2.econ.iastate.edu/tesfatsi/aelect.htm
33. Tonkiss, F.: Trust, Confi dence and Economic Crisis. Intereconomics. (2009) doi: 10.1007/s10272-009-0295-x
34. Vucetic, S., Tomsovic, K., and Obradovic, Z.: Discovering price-load relationships in California's electricity market. IEEE Trans. Power Systems. **16**(2), 280 – 286 (2001)
35. Weron, R.: Heavy-tails and regime-switching in electricity prices. Math. Meth. Oper. Res. **69**, 457 – 473 (2009)
36. Yang, T. and McDonough, J.M. (2002) Solution filtering technique for solving Burgers' equation. In: Proceedings of the Fourth International Conference on Dynamical Systems and Differential Equations (pp. 951 – 959)
37. Zhou, Z. and Chan, W.K.V.: Reducing Electricity Price Forecasting Error Using Seasonality and Higher Order Crossing Information. IEEE Trans. Power Systems. **24**(3), 1126 – 1135 (2009)

Referral Hiring and Labor Markets: a Computational Study

Samuel Thiriot, Zach Lewkovicz, Philippe Caillou and Jean-Daniel Kant

Abstract Many empirical studies emphasize the role of social networks in job search. The social network implicated in this process is known to be characterized by diversified properties, including communities, homophily or ties having various strength levels. Nevertheless, previous models of the labor markets fail to capture the complexity of social networks, as each specific network requires the development of specific algorithms. In this paper, we rather rely on an independent generic network generator for creating detailed networks describing friendships, colleagues, communities and various degrees of connectivity. We build a simple model of the labor market in which individuals find positions solely through their acquaintances, and update their network when being hired. This original experimental setting facilitates the analysis of various characteristics of networks in the labor market, including various sizes, the number of friendship links or the impact of communities. Experiments confirm the ”strength of weak ties” phenomenon. However, the initial characteristics of the network like the existence of communities are shown to be destroyed by the implausible mechanisms integrated into this simplistic model; this suggests that the impact of plausible networks on models' dynamics may only be studied when the mechanisms of this economic model are plausible as well - in other words, ”a model is only as descriptive as its most implausible component”. Addition of probabilities on ties allow to study close friends configuration and weak, or Facebook friends, configurations. Experiment show that, whereas friend are always

S. Thiriot
TAO Project, INRIA Saclay, France, e-mail: samuel.thiriot@res-ear.ch

Z. Lewkovicz
LIP6, Universite Paris VI, France e-mail: zach.lewkovicz@lip6.fr

P. Caillou
LRI, Universite Paris Sud, France e-mail: caillou@lri.fr

J.-D. Kant
LIP6, Universite Paris VI, France e-mail: kant@lip6.fr

S. Osinga et al. (eds.), *Emergent Results of Artificial Economics*, Lecture Notes in Economics and Mathematical Systems 652, DOI 10.1007/978-3-642-21108-9_2,

more useful than colleagues, weak friends are especially useful when the market is stable and with low unemployment and vacancy rate.

1 Introduction

Field studies on job search highlighted several stylized facts. First, Stylized Fact 1 [SF1] : searching and finding a job implies the use of social acquaintances to retrieve information (see [9] for a review). This social structure is commonly represented using the social network metaphor: each individual is represented as a node, and communication links as edges in this network. From an economic perspective, the communication of job opportunities through social relationships may lead to incomplete information in the market, thus possibly leading to a suboptimal market efficiency. Many recent models of the labor market already described such a social network and studied its influence on the market's efficiency (e.g. [10, 4]).

Secondly, [SF2] : all ties (links) are not equally useful for job seekers, nor lead to the same information. Since the famous Granovetter's studies on job search [7, 8], it became common to distinguish *weak* and *strong* ties; strong ties in a social network reflect frequent interactions between individuals, while weak ties typically lead to less frequent and less personal relationships. Until now, only few models (notably [12]) described several link types to study the impact of this stylized fact on the market.

Moreover, strong ties are more local, because they are mainly created and maintained because of common workplaces (co-workers), life-places or other activities (near family, friends); typically, the clusters observed in social networks are mainly made of strong ties. Weak ties are more random in the network; they correspond to old friendships born at school or for family. Granovetter observed that despite of long distance and rare interactions, weak ties are more efficient for finding job opportunities than strong ties: strong ties correspond to less diversified people who may communicate easily but receive the very same information, while weak ties link more different people exposed to different types of information. These observations were replicated on different countries and populations (see [13, p.5] and [9] for detailed reviews).

Thirdly, [SF3] shared by all empirical studies on social networks [5, 16], underlines the complex nature of these social networks: the positions of agents in their social environment is far from being random; at the dyadic level, it appears that people tend to bond together when they have close socio-demographic characteristics or interests (homophily), or more generally that the existence of a social tie depends on the properties of individuals (assortativness). Some recent models included this kind of complex properties such as homophily on employment status [3] or ethnicity [13] (see also [9] for a detailed review of existing models). Nevertheless, because of the difficulty to generate plausible networks[1] including individuals with several

[1] We define "plausible networks" as networks which, given state-of-art metrics and observation tools, complies with our limited knowledge on the real networks.

interdependent attributes[2], these ”more detailed” networks remain very simplified compared to real ones.

The use of social acquaintances to search for jobs often changes with location and demographic characteristics. Living in the same location increases the probability of co-working, as do similar socio-demographic characteristics [2]. Moreover, complex patterns are robustly observed in real networks at the scale of the triad (strong clustering or transitivity rate, intuitively corresponding to the ”friends of my friends are also my friends” effect). The recent stream of statistical analysis of large networks [11] also highlighted network-scale properties of real networks, including the frequent presence of biased distribution of degree of connectivity (most people have few ties, while few trust a big number of relationships).

The numerous studies listed before suggest a strong impact of the structure of social network on the efficiency of job search, including the strong impact of specific properties of networks on this market[3].

Previous studies focus either on the impact of initial networks detailed in various levels on the markets’ dynamics or on the strategic evolution of simple networks. In this paper, we *describe both a rich initial network and its evolution*. Among others, this approach aims to provide possible answers to the question: “*how realistic initial networks should be in labor market models where these networks evolve endogenously ?*”. In order to simplify the generation of “rich” networks, based on properties (e.g. assortativness, homophily or communities) matching those of real networks, we use an independent software generator dedicated to the reconstruction of plausible networks for social simulation. This original experimental setting enables to explore easily the impact of different networks.

In the next section we will describe the two components of our experimental setting: the use of an generic network generator for constructing the initial networks and a simple model of the labour market. In section 3 we show some results of experiments focused on the efficiency of various link types for finding a job, and on the evolution of networks.

---

[2] For instance, let us suppose we wish to generate a multiplex network containing strong and weak ties, with agents being tagged with a ”red” or ”blue” color supposed to reflect their ethnicity, and a ”gender” attribute that influences their friendship links. Should we first generate the network and then set the color attribute of agents in order to reflect the homophily observed from field studies, as done in [13] ? Once this first attribute is set, how could we assign gender attributes in a plausible way ? The addition of each supplementory attribute requires new algorithms, thus making the exploration of the sensitivity of markets to rich networks time-costly and hard to reproduce and communicate. In short, this technical limitation on networks’ generation makes almost impossible the computationnal study of the impact of numerous complex properties which are both observed in real networks and presumed to impact job search.

[3] Note that some authors used networks, not for describing the structure of communications, but rather to analyze patterns of employment by describing them as networks [14].

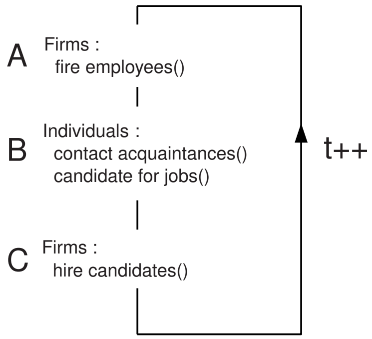

Fig. 1 A whole cycle in the simulation

2 Model and Experimental Settings

2.1 Labor Market Model

We describe a simple multi-agent model of the labor market, in which agents represent both the *firms* (characterized by a number of jobs) and *individuals* (employed or unemployed). This market is designed such that the total number of individuals N_i is proportional to the total number of jobs N_j: $N_i = MarketTension * N_j$.

The social network describes the structure of communication used by the agents to communicate job opportunities. Tacking stock of empirical studies (see before), we describe different kinds of relationships in this multiplex network: friendship links, colleagues (including both current colleagues and former ones) and spouses.

The behaviour of an agent, ran at each step, depends of his employment status (see Fig. 1).

A cycle in the simulation takes place in 3 parts (see Figure 1). First (A), Firms randomly lay off some of their employees with a random probability $FiringProba$. Otherwise the employee is kept in the Firm.

In the second part (B) Individual agents interact. If they are unemployed, they look for a job. In this model, individuals may only find jobs using their social acquaintances (as done previously [1]). The Individual agents may contact their friends, colleagues etc. In order to comply with [SF2], a probability of interaction $p(linktype)$ is associated to each link type in order to communicate positions. Then they candidate to all the vacancies they are aware of.

In the last part of the cycle (C) Firms iterate all their vacancies. If a vacancy has no candidature requests, it stays vacant. When a vacancy has several candidates, the Firm chooses randomly a candidate to be hired under the condition that it will not hire an individual that it has just laid off. As soon as an individual is hired, his new colleagues are added to his set of ”colleague” acquaintances; however, in order to forbid agents to know the entire population, colleagues are them removed randomly from this list of acquaintances, in order to keep the list of colleagues to at most *max_colleagues* (parameter). As a consequence, individuals remember past colleagues from past positions; however, the older the colleague, the higher the probability to break this tie.

2.2 Social Network Generation

We represent as a social network the social relationships that may lead to the communication of job offers. In this network, vertices are agents (either individuals or firms) while edges represent a dyadic relationship between these agents.

Networks used as parameters are built with the YANG[4] network generator [15] which stands as a generic tool dedicated to the generation of plausible networks for social simulation. This flexible tool enables the tuning of the networks' properties, including the definition of dense clusters of colleagues and the random creation of friendships links across the network. Throughout experiments, the size of firms (number of jobs) or the number of friends will be tuned for assessing their impact on the simulation.

The network generator accepts as many *discrete agents' attributes* as desired. Attributes of agents are described in YANG as random variables in a Bayesian network. This formalism enables the description of interdependencies between attributes. Probabilities associated with these variables are defined as follows: agentTypes takes value 'firm' with probability 0.1 and 'individual' with probability 0.9, leading the generator to create one firm per nine individuals. In the same way, gender take 'male' and 'female' values with probability 0.5 for individuals and value 'notRelevant' for firms. At initialization, 10% of the workers are not tied to firms and will have to find a job[5]. In practice, the degree for friendship (attribute *auto_friends_degree*) will be set to 5 or 2, depending to the experiments. In and out degree of connectivity for the matching of firms (*auto_eco_indegree* and *auto_eco_outdegree*) respectively describe the number of links getting out of an individual (1 if employed, 0 else) and going in a firm (9 for all firms in the first experiments).

The last parameters of the generator are the *generation rules*[6], which describe how the links are actually created in the population. YANG accepts two types of generation rules: "attributes rules" refer to generation rules that match two agents depending to their attributes, while "transitivity rules" propose the creation of links at the triadic scale by transitivity. We define the generation rules described in Table 1. The spirit of these rules, which will be applied in this order, is to create wedding links; then, to attribute to each worker a firm; then, to create links between all the colleagues; last, to create friendship links randomly across the population.

[4] http://yang.res-ear.ch

[5] Which will generate initially 10% of unemployment.

[6] Note that attributes rules always implicitly take into account the degree described before as an attribute of the agent. Laso Note that some of these rules are changed in some experiments.

rule name	method	principle
wedding	attributes	create links 'spouses' between males and females for 80% of agents with max degree 1
match	attributes	create links 'worksInFirm' between individuals having 'employed' as salarialStatus and firms
colleagues	transitivity	when an agent A1 'worksInFirm' A2, and another agent A3 'worksInFirm' A2, then create a link 'colleague' A1 and A3
friendsRandom	attributes	create links 'friendship' between individuals in pure random way

Table 1 Generation rules

The YANG network generator uses all of these parameters for generating random networks of size N. It first creates the whole population, each agent being given a combination of the possible attributes values. This population is stored in an SQL database. Then, the generator applies all the generation rules, by retrieving agents that may be tied together by SQL set operations on the population. The software that implements the generator also provides dynamic visualization of the network generation in order to check their plausibility. More details on this algorithm are provided in [15]. The detailed parameters are provided as supplementary material for reproduction purpose.

It is important to note that, as this network generator is random, the generated population may be slightly biased; for instance, the actual proportions of agents and firms may be 85/15 instead of the theoretical 90/10. As a consequence, the number of positions and individuals in sometimes not strictly equal in generated networks. To solve this problem, when networks are loaded, open positions are removed randomly if positions are too numerous, or open positions are added if workers are too numerous.

2.3 Experimental Settings

This model was implemented in Java (1.6) under the platform Repast (http://repast.sourceforge.net). The results we present here were attained by simulations involving 1000 agents: 900 individuals and 100 firms. We measure the aggregated value for all the indicators after 300 steps on 1000 simulations.

Except when stated otherwise, the other parameter of the simulations are chosen randomly with these parameters:

- $max\_colleagues = 5$
- $marketTension = 1$
- $FiringRate \in \{0.01; 0.04; 0.07; 0.1\}$
- $p(spouse) = 1.0$, $p(colleagues) = 1.0$
- $p(friend) = ProbaFriend \in \{0.2; 0.4; 0.6; 0.8; 1\}$

The observed variables are:

- The $UnemployementRate$, measuring the proportion of unemployed individuals divided by the total number of individuals
- The $Diameter$ of the graph
- The proportion of hiring for each link type (the link type which has first informed the hired individual of the existence of the vacant position): $HiringFriend$, $HiringSpouse$ and $HiringColleague$.
- The previous proportion has to be normalize to take into account the links proabilities and the fact that there is a different number of colleagues and friends. We define a $RelativeFriendEfficiency$ variable, measuring the efficiency of information given by friends, compared to information given by colleagues:
  $RelativeFriendEfficiency = \frac{HiringFriend}{HiringColleague} * \frac{Max\_colleague}{NumberOfFriends*ProbaFriend}$

3 Results

3.1 Efficiency of Link Types

In this first set of experiments, we explore which links, in this model of the labor market, allow individuals to find positions after being fired. The first simulations are run with the same number of friends and colleagues (5 each). The probability of interaction for each link type is here systematic (perfect interaction).

The unemployment rate stabilizes around a certain level whilst the agents are being fired and search for positions in other firms through their social acquaintances (see Fig. 3).

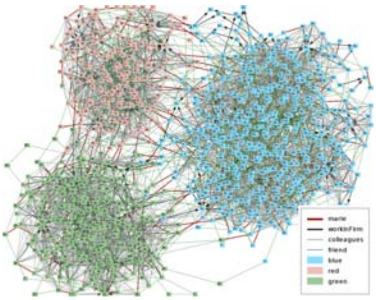

Fig. 2 Example of network used for simulations

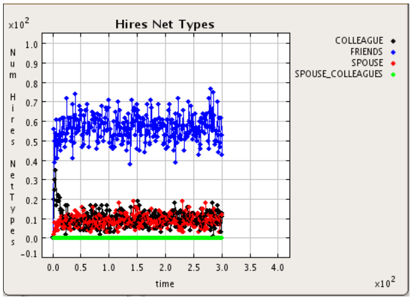

Fig. 3 Efficiency of the various link types during a typical run of the model for 5 friends and 5 colleagues

Simulations reflect the Granovetter's strength of weak ties theory: even if individuals have the very same number of friends and colleagues, they actually find most of their positions (∼55%) through friendships, which is as much as through colleague links (∼34%). This effect is confirmed by other experiments. With 2 friends, the relative effiency of friendship remains similar, each friend is ∼1.5 times more useful than a colleague (see Table 2). This unemployment rate remains also unchanged (slightly higher) when heterogeneity is introduced into the size of firms.

parameters	Unemployment rate	**links efficiency** colleagues	friends	spouse	Friend vs Colleague $RelativeFriendEfficiency$
same size for firms					
5 friends	1.8%	37.1%	54.5%	8.4%	1.48
2 friends	3.2%	54.5%	33.0%	12.3%	1.54
fat-tailed distribution of firms' sizes					
5 friends	2.0%	39.3%	51.5%	9.3%	1.52
2 friends	3.3%	53.8%	33.8%	12.6%	1.58

Table 2 Unemployment rate and the efficiency of link types for various combinations of parameters (with 5 colleagues). By links' efficiency we mean the proportion of hires done through this link type.

3.2 *Perfect Interactions, Weak Sensitivity to Networks' Structure*

Surprisingly however, *the structure of the network did not appear to have a strong impact on the unemployment rate* nor on the relative importance of each link type, despite of variations like the inclusion of three communities having a large number of endogenous links (Fig. 2). More examples with networks like the one depicted in Fig. 4 did not shift the unemployment rate as expected, despite of the strong local limitation of individuals acquaintances and of the big diameter of the graph (24 in average). These results contradict the conclusions from previous studies [3, 13] which underlined the impact of segregation on simulation results (when the network remains static during the simulation).

Fig. 4 Example of unrealistic network used for investigating network change during simulation

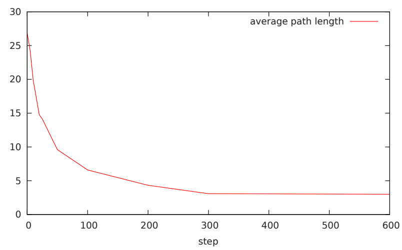

Fig. 5 Evolution of network size during a simulation

An analysis of the social network at different steps reveals a dramatic drop in the network size (Fig. 5), which shows how powerful the evolution of the network is, and explains the low impact of the initial network on the model. Whatever the initial properties of the network are, they will be quickly destroyed by the evolution of the network during the simulation. This may reveal a flaw in the initial model; among others, where we neglected to take into account the different possible natures of links [SF2]. These experiments led us to the idea of weighting the various types of links of interactions with probabilities (strength).

Beyond the case of labor markets, this phenomenon underlines an interesting methodological point for agent-based simulation: *providing networks with many properties - even if plausible or real - is useless if the networks dynamics is not realistic*.

3.3 Networks with Probabilistic Interactions

In order to both limit this destruction of the network and enhance the plausibility of the model, we switch from a perfect to probabilistic interaction by changing the probabilities $p(linktype)$, for instance with $p(frienship) = ProbaFriend$, $p(marie) = 1.0$, $p(colleagues) = 1.0$. Given this parameters' setting, individuals have now a lower probability to discover open positions through each acquaintance, as defined in [SF2]. We can thus compare the situation with few "strong" friends configurations, with whom we communicate frequently, and more "weak" friends configurations, with less frequent interactions.

Table 3 $RelativeFriendEfficiency$ (relative efficiency of friend information compared to colleague information) with 5 colleagues and 5 friends. For example, when there is 10% more workers than positions (tension=1.1), when friends transmit information with probability 0.2, each friend information transmission leads to 1.91 more hiring than a colleague information transmission.

$MarketTension \backslash ProbaFriend$	**0.2**	**0.4**	**0.6**	**0.8**	**1**	**Average**
0.7	2.00	1.76	1.68	1.62	1.60	1.73
0.8	1.96	1.80	1.65	1.63	1.58	1.72
0.9	1.98	1.71	1.63	1.56	1.53	1.68
1	2.06	1.77	1.63	1.53	1.50	1.70
1.1	1.91	1.68	1.56	1.51	1.47	1.62
1.2	1.83	1.61	1.53	1.48	1.44	1.58
1.3	1.75	1.60	1.47	1.41	1.40	1.53
Average	**1.90**	**1.68**	**1.57**	**1.51**	**1.48**	**1.63**

Without surprise, information coming from friends is always more useful than information coming from colleagues (hiring comes more often from friend than from colleagues after normalization by number and probability). More interestingly, the configuration where the friends are the most useful changes with the probability of interaction. With close friends (high probability), friends are the most efficient when the market is saturated (more jobs than employees). Friend efficiency decreases when unemployment rate increases (following market tension) because friends are more likely to be unemployed than colleagues (which were employed when they were added to the acquaintances). With weak friends, a second and stronger effect appears: the speed of information transfer. When there is a high unemployment rate or a high vacancy rate, to wait several steps before transmitting information has a stronger impact. With high unemployment, an other employee is likely to have already taken the job. And when there is a high vacancy rate (low market tension), a colleague is likely to have found an other job before the friend information arrives. The situation when weak friends are the most efficient is when there is the same number of jobs and workers.

This property is confirmed by the analysis of the $RelativeFriendEfficiency$ sensibility toward the $FiringRate$ parameter (the probability for each employee to be fired). Friends are more efficient than colleagues when the market is stable (low firing rate), which is understandable because the colleagues stay in the same company, and their information is thus more redundant. What is more interesting is to see that this effect is stronger for weak friends (+55% relative efficiency from 0.1 to 0.01 firing rate) than for close friends (+20%). Again, the time effect explains this difference: weak friend are more useful when a delay in the information has less impact (when the market is stable). Extreme cases can illustrate this: With $FiringRate = 0.01$ and $MarketTension = 1$, weak friends ($ProbaFriend = 0.2$) are 2.89 times more useful than colleagues, whereas this value is only 1.47 with $MarketTension = 1.3$ and $FiringRate = 0.1$. Close friends efficiency even drops to 1.19 with this configuration (colleagues are almost as useful as friends when the market is saturated and changes very quickly).

Table 4 $RelativeFriendEfficiency$ (relative efficiency of friend information compared to colleague information) with 5 colleagues and 5 friends. For example, with a probability to be fired for each worker of 1%, when friends transmit information with probability 0.2, each friend information transmission leads to 2.49 more hiring than a colleague information transmission.

$FiringRate \backslash ProbaFriend$	**0.2**	**0.4**	**0.6**	**0.8**	**1**	**Average**
0.01	2.49	2.03	1.83	1.73	1.66	1.95
0.04	1.83	1.63	1.54	1.47	1.46	1.58
0.07	1.67	1.55	1.48	1.44	1.42	1.51
0.1	1.61	1.51	1.44	1.41	1.38	1.47
Average	**1.90**	**1.68**	**1.57**	**1.51**	**1.48**	**1.63**

4 Discussion

In this paper, we assess the impact of a ”rich” network on a labor market, whose network is partly evolving when hiring employees. This study is based on an original experimental setting coupling a generic network generator and a model of the labor market. Contrary to our initial expectations, first simulations had a weak sensitivity to the initial network; we explained this fact by a fast and unrealistic evolution of the network that quickly destroyed the initial structure. Further studies with probabilistic interaction, attributed to link types, enhance the plausibility of the model and limits the evolution of the network. Further experiments highlight the impact of various properties of the initial network. For example, experiment show that, whereas friend are always more useful than colleagues, weak friends are especially useful when the market is stable and with low unemployment and vacancy rate. These experiments suggest that when the evolution of the network is not too simplistic, the description of attributes of agents and various link types in the initial network has a strong impact on the dynamics of the labor market.

References

1. Arcaute E, Vassilvitskii S (2009) Social networks and stable matchings in the job market. Internet and Network Economics pp 220–231
2. Bayer P, Ross S, Topa G (2005) Place of work and place of residence: Informal hiring networks and labor market outcomes. Tech. rep., NBER working paper
3. Bramoullé Y, Saint-Paul G (2010) Social networks and labor market transitions. Labour Economics 17(1):188–195
4. Cahuc P, Fontaine F (2009) On the efficiency of job search with social networks. Journal of Public Economic Theory 11(3):411–439
5. Carrington PJ, Scott J, Wasserman S (2005) Models and Methods in Social Network Analysis. Cambridge University Press
6. Gilbert N (ed) (1994) Simulating societies : the computer simulation of social phenomena. University of Surrey, Guilford (GBR)
7. Granovetter M (1973) The Strength of Weak Ties. The American Journal of Sociology 78(6):1360–1380
8. Granovetter M (1983) The strength of weak ties: A network theory revisited. Sociological Theory 1:201–233
9. Ioannides Y, Loury D (2004) Job information networks, neighborhood effects, and inequality. Journal of Economic Literature 42(4):1056–1093
10. Montgomery J (1991) Social networks and labor-market outcomes: Toward an economic analysis. The American economic review 81(5):1408–1418
11. Newman M, Barabasi AL, Watts DJ (2006) The Structure and Dynamics of Networks: (Princeton Studies in Complexity). Princeton University Press, Princeton, NJ, USA
12. Tassier T (2006) Labor market implications of weak ties. Southern Economic Journal 72(3):704–719
13. Tassier T, Menczer F (2008) Social network structure, segregation, and equality in a labor market with referral hiring. Journal of Economic Behavior & Organization 66(3-4):514–528
14. Tesfatsion L (2001) Structure, behavior, and market power in an evolutionary labor market with adaptive search* 1. Journal of Economic Dynamics and Control 25(3-4):419–457
15. Thiriot S, Kant JD (2008) Generate country-scale networks of interaction from scattered statistics. In: The Fifth Conference of the European Social Simulation Association, Brescia, Italy
16. Wasserman S, Faust K (1994) Social network analysis, methods and applications. Cambridge: Cambridge University Press

An Agent-Based Information Management Approach to Smoothen the Pork Cycle in China

Sjoukje A. Osinga, Mark R. Kramer, Gert Jan Hofstede and Adrie J.M. Beulens

Abstract The objective of our research is to study the relationship between (a) the spread of information at farmer level and (b) the emerging behaviour at sector level, applied to the case of the pork cycle in China, using an agent-based model. For this paper, we investigate the effect of farmers' individual supply decisions on the overall supply pattern in the sector. We apply a basic agent-based supply- and demand model populated with pig farmers, where supply is based on price expectations that include a time lag. The farmers decide upon their future supply (at farm level) using the price expectations they are able to make based on the information at their disposal. We compare our agent-based model with the classical cobweb model, which exhibits periodical over- and under-supply. This periodicity is not desirable, as is illustrated by a realistic example from the pork sector in China. The Chinese government tries to smoothen the overall supply and demand pattern by acting as a speculator as soon as price imbalance at total system level exceeds a threshold value, hence intervening at system level. Our agent-based model displays similar behaviour, and we can conclude that smoothening the supply curve by diminishing periodicity is also possible at individual level. An emergent result from the comparison is that mapping of economic supply and demand functions to individual agents' decisions is not straightforward. Our model is a fruitful basis for further research, which will include social interaction, imitation behaviour and a more sophisticated information diffusion process that reflects the rate at which a farmers population adopts information.

Sjoukje A. Osinga, Mark R. Kramer, Gert Jan Hofstede and Adrie J.M. Beulens
Wageningen Univerity, P.O. Box 8130, NL-6700 EW Wageningen, The Netherlands
e-mail: Sjoukje.Osinga@wur.nl

S. Osinga et al. (eds.), *Emergent Results of Artificial Economics*, Lecture Notes in Economics and Mathematical Systems 652, DOI 10.1007/978-3-642-21108-9_3,

1 Introduction

1.1 Information Management Objective

To run their businesses, farmers take decisions for which they need information. Information includes a wide range of product and market information, managerial and technological know-how, knowledge of how to comply with legal requirements and regulations, and more. Information management is defined as the activity of acquiring and applying information to improve business performance. In this research we assume that better use of information management strategies will improve business performance [14].

Information management theory applies to activities within business firm environments, and often assumes state-of-the-art development [10]. Family farms usually work with only basic information management systems, if at all. However, availability of adequate and accurate information and the know-how to apply it is of vital importance to any farmers' business performance [8]. Similar to an innovation diffusion process [16], such information spreads over a population: a percentage of innovative farmers learn and decide to adopt, while others adopt later due to social influence, i.e. by imitating behaviour of others. The speed of diffusion depends among other things on social relationships between farmers: they may exchange information with geographical neighbours, with friends, or with both.

The objective of our research is to study the relationship between (a) the availability of information at farm level and (b) the emerging behaviour at sector level, applied to the case of the pork cycle in China, using an agent-based model [5]. For this paper, we investigate the effect of farmers' individual supply decisions on the overall supply pattern in the sector. We apply a basic agent-based supply- and demand model populated with pig farmers, where supply is based on price expectations that include a time lag. The farmers decide upon their future supply (at farm level) using the price expectations they are able to make based on the information they have at their disposal. We compare our agent-based model with the classical cobweb model, which exhibits periodical over- and under-supply. This periodicity is not desirable, as is illustrated by a realistic example from the pork sector in China. The Chinese government tries to smoothen the overall supply and demand pattern by acting as a speculator as soon as price imbalance at total system level exceeds a threshold value, hence intervening at system level.

It is assumed that the quality of farmers' price expectations depends on the amount of information they have at their disposal: the emergent result of an information diffusion process. Individuals are modelled at the level of stylized facts to yield a mid-range model [5]. Heterogeneous attributes, or personality characteristics such as personality ('openness') and susceptibility to social influence, affect the outcome of decision processes [11]. At individual level, these attributes affect each farmer's decision to adjust supply according to price expectations. At population level, the mean values of these attributes express the cultural characteristics of the population [7]. The model in this paper does not include yet personality and culture heterogeneity.

2 Background Literature

In economics, the term pork cycle, or hog cycle, describes the phenomenon of cyclical fluctuations of supply and prices in livestock markets [6], [2]. A classical model to study this phenomenon is based on the cobweb theorem [3], [9]. When prices are high, producers have an incentive to invest in more pigs, but with delayed effect due to the necessary breeding time. Next, the market becomes saturated, which leads to a decline in prices. As a result of this, production is reduced, which in turn takes time to be noticed, and then leads to supply falling below the demand and results in higher prices. This procedure repeats itself cyclically. The resulting supply-demand graph resembles a cobweb. The cobweb models distinguish several expectation schemes by which farmers predict future prices: naive expectation (assuming that price remains the same as it was), adaptive expectation (including previously made forecasting errors) and rational expectation (including all currently available information).

The cobweb model has been extended in several directions over the past decades. Westerhoff and Wieland describe a model that introduces heterogeneous speculators into the traditional cobweb framework [21]. They argue that most primary producers in commodity markets are mainly concerned with the production process, but that speculators (e.g. in stock markets) actively predict commodity price movements. The speculators may apply technical analysis (extrapolating past price trends) or fundamental analysis (assuming that prices converge towards their fundamental values, i.e. the equilibrium price in a perfect market). Westerhoff and Wieland conclude that speculators may indeed have an impact on price dynamics: in most cases the impact is destabilizing, but in some situations it may stabilize the dynamics.

2.1 Pork Cycle in China

More than half of the world's pork is produced and consumed in China. The pork cycle also occurs in the Chinese market [13]. The seasonal fluctuation is significant: every year prices are high in January, then drop until the summer, then start increasing again for a second peak in October (see Figure 1a). This fluctuation is closely related to meat consumption patterns: when temperatures become higher, more vegetables are available and meat consumption decreases; after the summer, the weather turns cold and traditional festivals stir up the price (Spring Festival in January, Moon Festival in October). Price fluctuations also occur due to unexpected events or contingency shocks, e.g. after a disease outbreak, or when the grain price goes up (see Figure 1b). Grain price is a direct indicator of feed price, which represents half of the costs for pig production in China [18].

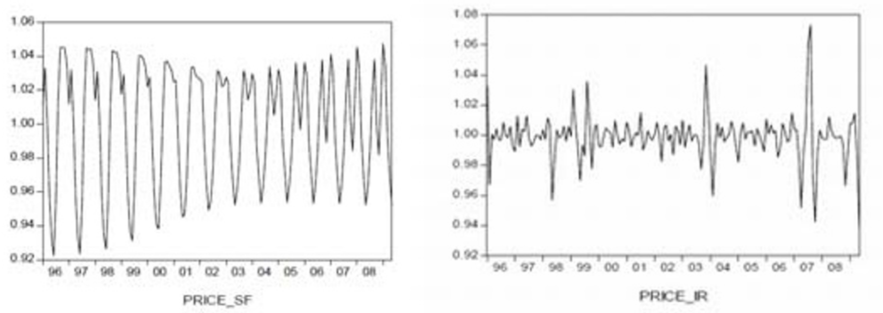

Fig. 1 a: Regular, seasonal fluctuation; b: Irregular fluctuation. *Graphs are based on the monthly pork price (Jan 1996 to May 2009), taken from [13].*

2.2 Interventions from Government

The Chinese government is taking an effort to control these fluctuations. In January 2009, the National Development and Reform Commission and Ministries of Finance, Agriculture, Commerce, Commercial Bureau, and AQSIQ jointly issued a document containing intervention measures [19]: if the ratio of hog to grain prices falls below 6.0 for more than four weeks, the government will consider purchases for reserve stocks to reduce market supply and increase prices, to set them out in the market again in times of supply shortage. Indeed, at least two such government purchases to restrain supply and bump up prices were made in April 2009 [19], and interventions occurred in 2010 as well [12]. Newspapers reported by the end of 2010 that pork supply and demand were essentially in balance in 2010 thanks to the governmental interventions [17]. However, industry experts question the long-term feasibility of these short-term solutions and fear even worse imbalance [1].

2.3 Information Management Based Approach

Instead of using the overall ratio of hog to grain price, which is the key indicator for the government to intervene, intervention can also be divided over the population of farmers. The majority of Chinese pork is produced by small-scale farmers, who run a family business of pig farming [4]. Improving information management, i.e. to make use of available information, may help to increase farmers' awareness and knowledge of how to estimate future prices, hence to anticipate the cyclical nature of demand and supply. In this way, the overall pork demand and supply curve will exhibit the emerging result of all individual farmers' adjusting behaviour, instead of the result of governmental control. Price expectations are calculated based on information, but not each farmer has the same information at his disposal, nor acts upon the calculated outcome in an equal way. This approach is different from Westerhoff and Wieland [21], who include a different agent-type, i.e. the speculator, into their model. In our approach, their are no explicit speculators, but differences in price expectation occur due to heterogeneity among the farmers' accuracy of predicting future prices.

3 Research Questions

The research questions of this paper are:

1. Can the periodic pattern of cyclical supply and demand fluctuations as described in the literature be generated by means of an agent-based model?
2. As an alternative to government intervention, can this periodicity be eliminated when farmers use improved price expectations?
3. What is the impact of heterogeneous information diffusion among farmers on the periodicity, assuming that more information leads to more accurate price expectations?

4 Model

For our agent-based model of the pork cycle in line with the cobweb theorem, we apply a basic linear demand and supply system:

$$Q_d = D_0 - h_D * P \tag{1}$$

$$Q_s = S_0 + h_S * P \tag{2}$$

where Q_d and Q_s, D_0, S_0, h_D and h_S are the quantities, intercepts and slopes of the linear demand and supply functions, respectively. The actual values for intercepts and slopes were taken from a textbook example [15], chapter 2. P represents the price. Assuming market clearing, Q_d equals Q_s, hence there exists an equilibrium price. When a time lag is involved, the cobweb problem arises: when price turns out to be not at equilibrium level, quantities are either too high or too low, resulting in a cyclical demand and supply pattern. At each time step, price and supply are calculated as:

$$P_t = \frac{D_0 - S_t}{h_D} \tag{3}$$

$$S_t = hS * P_{t-5} - S_0 \tag{4}$$

where the value 5 represents a time lag of 5 time steps, i.e. P_{t-5} is the price farmers assume to receive for pigs that still need a maturing time of 5. The expected price for each farmer is based on all information available to him. When there is no information available, farmers adopt naive expectation, i.e. equal to the current observed price. With more information available, the expected price becomes a better estimate.

In the supply- and demand equations, when $h_D < h_S$, the periodic cycle is divergent, which in the movement along the (linear) supply and demand curves resembles a cobweb starting in the centre and spiralling outwards. When $h_D > h_S$, the cycle is convergent, which resembles a cobweb starting from the outside and spiralling inwards. When the slopes are equal ($h_D = h_S$), a stable cycle results. As we wish to investigate the effect on a stable situation, h_S and h_D were set to equal values for all our models.

4.1 Information Management Approach

We assume in this research that better information management leads to better price estimates. In our model, each farmer receives a number of information items during setup. The more different information items available, the better accuracy for the estimated price. No information items available is equivalent to using the current price, maximum information is equivalent to using the (perfect) equilibrium price that is expected to exactly match supply and demand, and anything between minimum and maximum information leads to a proportional accuracy of the equilibrium price.

Accuracy (α) is calculated per farmer i as

$$\alpha_i = \frac{info\text{-}items_i}{\sum_i info\text{-}items} \tag{5}$$

where $info\text{-}items_i$ is the number of information items that farmer i has available, and $\sum_i info\text{-}items$ is the total number of information items available in the system.

The estimated price is calculated as:

$$EP = \alpha * P^{eq} + (1-\alpha) * P_t \tag{6}$$

where α is in interval $(0,1)$; P^{eq} is the equilibrium price, and P_t is the current price.

4.2 Research Models

To answer the research questions we develop three models, each using a different price expectation on which farmers base their supply decisions:

1. the current-price-model, where $EP_{t+5} = P_t$
2. the perfect-price-model, where $EP_{t+5} = P^{eq}$
3. the estimated-price-model, where $EP_{t+5} = \alpha * P^{eq} + (1-\alpha) * P_t$

In fact, the current-price-model is the estimated-price-model with a value of α equal to 0, and the perfect-price-model is the estimated-price-model with a value of α equal to 1.

4.3 Decision to Restock

In our model, farmers without stock decide every time step whether they will restock or not. If they restock, they do so with a fixed capacity, which is the same for all farmers. After restock, pigs need 5 time steps to mature, during which time no restocking decision can be made. In reality, farmers maintain several pig batches of different age, but we assume that this refinement would not fundamentally lead to different outcomes at system level, hence we prefer the simpler model. Also, farmers would have different restock capacities, but we keep this value equal for simplicity reasons.

The decision to actually restock or not depends on the total supply S that will be needed, which each farmer calculates according to equation (2) using his estimated price (which differs per model):

$$Restock = r < \frac{S}{maxS} \tag{7}$$

where r is a discretized random number uniformly drawn from interval $(0,1)$, and $maxS$ is the maximum supply possible, i.e. the supply in case all farmers currently without stock would choose to restock. This implies that farmers know how many other farmers are currently making the decision to restock, but they do not know the actual outcome of the other farmers' decisions.

4.4 Simulation Process

Before simulations start, farmers are created with attributes $stock$, $pig\text{-}age$, $info\text{-}items$ and $estimated\text{-}price$. There are 5 batches of farmers, each having pigs of a certain age. Their $info\text{-}items$ attribute is assigned to them at random. The higher the number, the more informed they are. At each time step:

- farmers sell their pigs if these have reached maturity, which is the case for a different batch of farmers at each time step. When the pigs have been sold, those farmers' stock becomes 0, so they will be the ones who later must decide whether they will restock or not.
- the current price P_t is calculated according to equation (3).

- *maxS* is calculated, i.e. the total capacity of all farmers who currently have no stock.
- in the estimated-price model, farmers calculate their estimated price, based on their number of *info-items*. In the other two models, the previous price or the equilibrium price are used, respectively.
- farmers decide to restock or not according to equation (7), and update their stock accordingly.
- total supply currently present in the system is the sum of all stock of farmers who now have mature pigs.

Because the total supply changes at every time step, the price also changes, which in its turn affects the decision to restock or not. This determines the total supply present at each time step, and defines the pattern in the resulting graph.

4.5 Fourier Transformation

Simulation outputs of model runs show total stock changes per time step, which sometimes appears to be a periodical wave, and sometimes resembles a random walk. In order to determine whether a waveform is still periodical or not, we transform the total stock data from the simulation output into frequency diagrams, using Fast Fourier Transformation. FTT diagrams show the frequencies within the waveform period, from which it is much easier to identify periodicity.

5 Results

The models were implemented in Netlogo and were all run with 500 farmers. Figure 2 shows the results for the current price model. In this model, where farmers use the

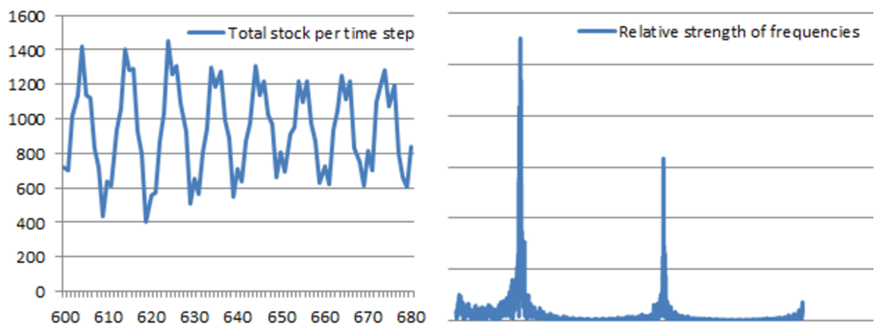

Fig. 2 (a) Simulation output and (b) Fourier frequency diagram for the *current price model*.

current price to estimate future stock, the pattern is clearly periodical, as expected (figure 2a). The Fourier frequency graph (figure 2b) confirms this.

Figure 3 shows the results for the perfect price model. In this model, farmers use

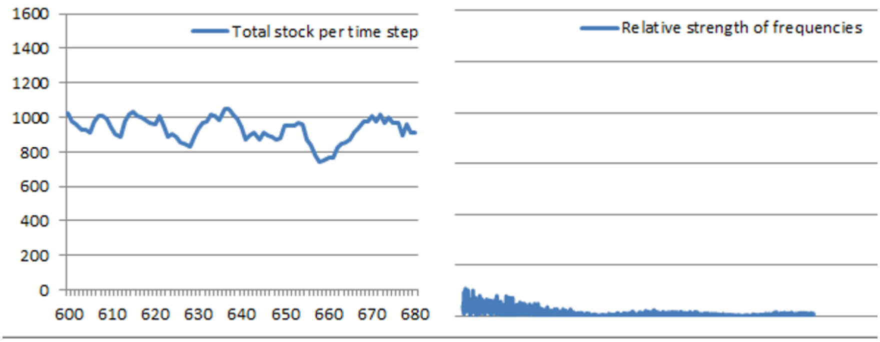

Fig. 3 (a) Simulation output and (b) Fourier frequency diagram for the *perfect price model*.

the equilibrium price to estimate future stock, and only their decision to restock or not was randomized. The result of this seems to be a random walk (figure 3a), as expected. However, in the Fourier frequency graph (figure 3b) we can observe that there is still a small remainder of periodicity left in the model.

Figure 4 displays the results for the estimated price model. In this model, we

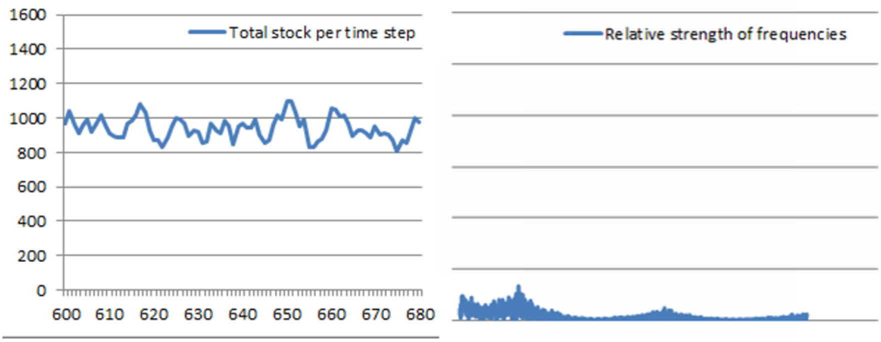

Fig. 4 (a) Simulation output and (b) Fourier frequency diagram for the *estimated price model*, with information parameter α set to 0.5.

experimented with several values for the information parameter α. The more information available, the less periodicity we expected to appear. This was consistent with the outcomes. Figure 4a shows the result with α set to 0.5. We can observe that periodicity has indeed diminished when compared with figure 2a, but that it is clearly more present than in figure 3a. The Fourier frequency graph (figure 4b) confirms that some periodicity is left in the model.

6 Conclusion and Discussion

Coming back to the research questions, we can conclude that the pattern of cyclical supply and demand fluctuations as described in the literature can indeed be generated by means of an agent-based model. When farmers assume naive expectations, there is clearly a periodic pattern that resembles the literature, as figures 2a + 2b show.

However, there are differences as well. The mapping of economic supply and demand functions to individual agents' decisions is not straightforward. The theory works with average total supply quantities at each time step. But on individual agent level, each farmer will not restock an average quantity of pigs. More likely is that he will decide either to (a) fully restock his stables when (expected) prices are worth the effort, or to (b) not restock at all and find income from alternative labour until prices increase. The alternative labour option is realistic in the Chinese situation [20]. The fact that restock is a discrete choice and happens at full capacity, is an emergent property from modelling at individual level.

An implicit result from this finding is that the number of farmers actually involved in restocking may be limited, dependent on the ratio of farmers, and their maximum capacity, with respect to total demand. When this ratio is high, only a few farmers restock, and the others have to find alternative labour. This was also the case in our model experiments. In models that do not include individual farmers, the issue that a majority of farmers may be out of business does not appear. But in the real world, such a scenario would cause huge social problems. It would be worth experimenting in the model with the parameter values for maximum restock capacity and number of farmers with respect to demand, to obtain a model with an higher average of participating farmers. The agent-based model easily allows for such experiments.

Interesting in this respect is the transition point present in our model: with the current settings, there is no periodicity when there are fewer than 185 farmers, because then all farmers always restock at maximum capacity, which is still less than demand. So the cyclic pattern only appears when there is actually an overcapacity of stock. This is another emergent result from the agent-based model.

As for the second research question, we can conclude from figure 3 that periodicity would indeed almost be eliminated if farmers were able to use the perfect price in their estimations. The small remaining periodicity visible in the Fourier frequency graph may be a result of the fact that our model works with batches of farmers who have pigs of the same age. When the pigs reach maturity, batches of farmers make the decision to restock or not, which may cause some periodicity.

The third research question addresses the issue to what extent heterogeneous information diffusion among farmers leads to diminishing periodicity, assuming that more information leads to more accurate price expectations. Figure 4 shows that heterogeneous distribution of information over farmers leads indeed to a decrease in cyclical pattern. The more information, the stronger the effect.

The inspiration for this paper was the government intervention that happens in China in order to smoothen the over- and under-supply pattern. From our research questions, we can conclude that an alternative way to smoothen that pattern could be to improve the price expectations among farmers by well-informing them. Our model is too rudimentary to make a serious claim here, but it offers a good start for further refinement.

An addition to the model would be to include external influence, which happens event-wise instead of in regular cycles. One way to do that is to actually use the grain price in the model, which we currently assume as constant. Grain price changes as an independent variable, and is responsible for 50% of the farmer's pig production costs, which would be an additional factor in farmers' decision whether to restock or not.

Our model currently contains no explicit interaction mechanisms. However, there is implicit interaction, because we assume that farmers know what the maximum possible supply is. It would be interesting to see what actual interaction and information exchange would lead to. We tried this (in a rudimentary way), but we did not reach a point beyond 'what we put in, is what we get out'. We would need to add network structures among the farmers to be able to observe worthwhile interaction effects. It is realistic to assume that farmers let their decision to restock or not influence by what their peers - those in their social network - decide. Heterogeneous personality attributes and cultural population attributes that affect decision making would make the model reflect social behaviour both at individual and at population level. Finally, it would be interesting for future research to include a more sophisticated information diffusion process, reflecting the rate at which a population of farmers adopts information.

References

1. China Economic Review (2010). Pig run: will imbalances in supply and government controls on pork prices contrive to push up inflation?, *China Economic Review*. Available at http://www.chinaeconomicreview.com/
2. Coase RH, Fowler RF (1937). The Pig-Cycle in Great Britain: An Explanation. *Economica* 4: 55-82.
3. Ezekiel MJB (1938). The cobweb theorem. *Quarterly Journal of Economics* 52: 255-280.
4. Fabiosa JF, Hu D, Fang C (2005). A Case Study of Chinas Commercial Pork Value Chain, Midwest Agribusiness Trade Research and Information Center (MATRIC), Iowa State University, Iowa.
5. Gilbert N (2008). Agent-Based Models. *Quantitative Applications in the Social Sciences* 153. Sage Publications, Los Angeles.
6. Hanau A (1928). Die Prognose der Schweinepreise, Institut fr Konjunkturforschung, Berlin.
7. Hofstede G, Hofstede GJ, Minkov M (2010). Cultures and Organizations: Software of the Mind. McGraw-Hill, New York.
8. Isaac ME, Erickson BH, Quashie-Sam SJ, Timmer VR (2007). Transfer of knowledge on agroforestry management practices: the structure of farmer advice networks. *Ecology and Society* 12(2)
9. Kaldor N (1938). The Cobweb Theorem. *Quarterly Journal of Economics* 52(2): pp. 255-280
10. Laudon KC, Laudon JP (2009). Management Information Systems: Managing the digital firm. Pearson Education, Inc., Upper Saddle River, New Jersey

11. McCrae RR, Costa PT Jr (2003). Personality in adulthood: a five-factor theory perspective. The Guildford Press, New York.
12. Newsroom (2010). China - Government intervention pork buying, Meat Trade News Daily. Available at http://www.meattradenewsdaily.co.uk/, Devon, UK.
13. Nie F, Dong L, Bi J (2009). Fluctuation and Cycle of Pork Price in China. International Association of Agricultural Economists Conference Beijing, China.
14. Osinga SA, Roozmand O, Kramer MR, Hofstede GJ (2010). An agent-based model of information management in the Chinese pig sector: top-down versus bottom-up. In: Wubben, EFM et al. (eds) 9th Wageningen International Conference on Chain and Network Management (WICaNeM 2010). Wageningen Academic Publishers, Wageningen, The Netherlands
15. Perloff JM (2009). Microeconomics. Pearson Education, Boston.
16. Rogers EM (2003). Diffusion of innovations, 5th edition. Free Press, New York
17. Swineweb.com (2010). Hog prices in China to hold steady, Swineweb.com. Available at http://swineweb.com/hog-prices-in-china-to-hold-steady/.
18. The Pig Site (2007). The story behind China's rising pork prices, Markets and Economics Featured Articles. Available at http://www.thepigsite.com/articles/7/markets-and-economics/2014/the-story-behind-chinas-rising-pork-prices
19. USDA (2009). China Plans to Purchase Pork for State Reserves. GAIN Report Number: CH9048. USDA FAS, USA.
20. Wang J, Watanabe M (2008). Pork Production in China - A Survey and Analysis of the Industry at a Lewis Turning Point, ASEDP report nr 77: Joint studies on economic development policies in ASEAN and neighboring countries.
21. Westerhoff F, Wieland C (2010). A behavioral cobweb-like commodity market model with heterogeneous speculators. *Economic Modelling* 27(5): 1136-1143

Part II
Financial Markets

Do Capital Requirements Affect Long-Run Output Trends?

Andrea Teglio, Marco Raberto and Silvano Cincotti

Abstract The macroeconomic implications of capital requirement for banks have drawn remarkable attention after the financial crisis started in 2007. In particular, a considerable effort has been devoted by the scientific community and by the central banks in order to understand the effects of different capital requirements on long term growth. This paper aims to contribute to the debate, proposing an analysis based on an agent-based macroeconomic model, i.e., the Eurace model, that takes into account the complex pattern of interactions among different economic agents in a realistic and complete way. The institutional setting considered in the computational experiment consists in varying the allowed leverage ratio for commercial banks, i.e. the ratio between the value of the loan portfolio held by banks, weighted with a measure of the loan riskiness, and the banks net worth or equity, along the lines of capital adequacy ratios set by the Basel II agreement. The outcomes of the analysis show that bank's capital requirement affects both the level of output and the output variability. In particular, limiting bank leverage by means of stricter capital requirements has a negative impact on output in short term that fades if the medium term and disappears in a long run scenario characterized by higher GDP levels.

Andrea Teglio
Departament d'Economia, Universitat Jaume I, Av. Sos Baynat, Castellón de la Plana, Spain
e-mail: teglio@eco.uji.es

Marco Raberto
DIBE-CINEF, Università di Genova, Via Opera Pia 11a, 16145 Genova, Italy
e-mail: marco.raberto@dibe.unige.it

Silvano Cincotti
DIBE-CINEF, Università di Genova, Via Opera Pia 11a, 16145 Genova, Italy
e-mail: cincotti@dibe.unige.it

S. Osinga et al. (eds.), *Emergent Results of Artificial Economics*, Lecture Notes in Economics and Mathematical Systems 652, DOI 10.1007/978-3-642-21108-9_4,

Introduction

The study aims to investigate the impact on the economy of an important regulatory setting characterizing the functioning of credit markets. In particular, the attention is focused on the leverage of the banking sector, i.e., the ratio between the value of the risk-weighted assets held by banks and the banks' net worth or equity. It is worth noting that the banks' capital adequacy ratio, i.e., the inverse of the previously defined leverage ratio, is an important issue addressed by the regulation of the banking system, both at the national and international level, as testified by the well-known Basel II agreement which stipulates the maximum leverage ratio, or equivalently the minimum capital adequacy ratio, that banks are allowed to take.

The debate on economic implications of capital adequacy requirements has been growing in the last decades (Berger et al. [3]; Blum and Hellwig [4]). In particular Hellwig [8] and Adrian and Shin [2] focus their attention on the new scenario after the financial crisis of 2007-2009. Some of the main controversial aspects about the implications of higher equity requirements are clearly resumed by Admati et al. [1]. The main benefit of increased equity capital requirements is claimed to be the weakening of systemic risk, while potential drawbacks are, among others, the reduction of return on equity (ROE) for banks, the restriction of lending, and the increase of funding costs for banks. In particular, the long-term macroeconomic impact of capital requirements has been intensively examined after the last financial crisis by central banks or by international institutions performing policy analysis and financial advising [1].

The agent-based computational approach is the methodological approach adopted in this study. In particular, we investigate the impact of capital requirements for banks by means of the Eurace artificial economy, which is an agent-based model and simulator of the macroeconomy characterized by different types of agents, i.e., private agents, like households, firms, and commercial banks, as well as policy makers, like the the Government and the Central Bank, which interact through a set of decentralized markets, such as the labor, the goods and the credit market, and a centralized financial market where government bonds and firms equity shares are exchanged (Cincotti et al. [5]; Teglio et al. [11]).

In particular, this study presents the main features of the Eurace credit market and analyzes the impact of its regulatory setting on the macroeconomic activity. Results show how the main macroeconomic variables characterizing the Eurace economy, i.e. the real Gross Domestic Product (GDP) and the unemployment rate, are affected by the aggregate amount of loans provided by banks, which in turn depend, through the leverage ratio allowed by the regulatory provisions,on banks net worth (equity).

The paper is organized as follows. Section 1 presents an overall description of the credit market model, considering both the demand and the supply side of the market.

[1] See, among others, the reports on capital requirements implications by the Basel Committee on Banking Supervision (*http://www.bis.org/publ/bcbs173.htm*); by the Institute of International Finance (*http://www.iif.com/press/press+151.php*); or by the Bank of Canada (*http://www.bank-banque-canada.ca/en/publication/strengthening.html*)

Results of the computational experiment within the Eurace economy are presented and discussed in Section 2, while the concluding remarks are drawn in section 3.

1 The Eurace Credit Market Model

The credit market in the Eurace model is a decentralized market where banks are on the supply side and consumption goods producers are on the demand side; demand and supply of credit are matched by means of pairwise interactions among banks and producers. The law of one price (interest rate) does not apply in the Eurace credit market and lending rates depend, among other things, on the creditworthiness of the borrower. Banks are price makers and quantities takers; however, credit may be in short supply depending on the fulfilment of capital adequacy requirements by lending banks. Therefore, borrowers may be rationed in the market.

1.1 Credit Demand

Consumption good producers (CGP) demand credit to finance their production and investment plans as well as their financial commitments.

Production and investment plans are made on a monthly basis, while the elementary time unit of the Eurace simulator is conventionally the trading day. At different days during the month, each CGP makes a production plan, based on the expected future demand and on the stock of inventories. Then the CGP determines the amount of capital investment necessary to fulfill the production plan, considering a Cobb-Douglas technology. The costs associated to the production and investment plans are then computed considering the given cost of inputs, labor and capital (see [6, 7] for details).

At the same time, each CGP computes its monthly income statement and its monthly balance sheet, taking into account the revenues and the expenses of the previous month, and determines its monthly financial commitments, which are given by dividends payments, taxes, interests on debt, as well as the debt instalments set by the monthly schedule of loans repayment.

The CGP production and investment costs summed to the financial commitments bill give the monthly liquidity needs of any CGP. If the available liquidity is not sufficient to cover the above liquidity needs, the CGP tries to collect the residual liquidity needs first by borrowing a new loan in the credit market, then, if rationed in the credit market, by issuing new shares in the stock market. This order of priority, i.e., first internal funding, then new loans, and finally, only as last resort, new equity shares, recalls a well-known behavioral patters of firms, as rationalized by the pecking-order theory [9]. If the CGP is rationed in the credit market, and not all new equity shares have been successfully sold in the stock market, then the collected liquidity is not enough to face the scheduled payments. In that case, the CGP first reduces its discretionary payments, i.e., dividends, then scales down production and investments eventually to zero, and if these actions are still not sufficient to meet its compulsory payments, i.e. taxes, interests and debt instalments, with the available liquidity, the CGP goes bankrupt.

Finally, it is worth noting that two types of producers are considered in the Eurace model, namely, consumption goods producers, as described above, and capital goods producers. Capital goods producers employ labor to produce capital goods for CGPs according to a job production schedule. Indeed, capital goods producers have no financing needs, because they do not undertake capital investments and because they pay the labor input with the revenues of capital goods sales that, given the job production assumption, are always equal to the amount of capital goods produced. Therefore, capital goods producers never enter into the credit market and consumption goods producers are the only players in the credit market demand side.

We stipulate that a CGP, when entering into the credit market, asks the needed loan to more than one bank, so to be able to select the lending bank(s) offering the best loan conditions, i.e., the lowest lending rate(s), and to reduce the likelihood of rationing. However, in order to take into account search costs as well as incomplete information, we assume that the subset of banks to whom a loan is asked by a CGP is smaller than the entire set of banks populating the Eurace economy. The subset of banks is chosen randomly at any time.

In the following, we will refer to CGPs generally as firms.

1.2 Credit Supply

The supply side of the credit market is made by banks. Banks collect money in form of deposits from the private sector, namely households, capital goods and consumption goods producers, and provide loans to the consumption goods producers, henceforth firms, that have insufficient liquidity to meet their financial obligations. The liabilities side of banks balance sheets is made by deposits, debt with the central bank, and equity (net worth), while the assets side is given by the portfolio of loans to firms and liquidity deposited at the central bank.

Liquidity is not a constraint for banks in providing loans. If a bank has insufficient liquidity to grant a requested loan, it borrows the needed amount of money from the central bank. We stipulate that the central bank can provide unlimited liquidity to banks at the given policy rate i^c. However, banks are constrained in providing loans by a Basel II-like capital adequacy rule, which sets the maximum leverage, measured as risk-weighted assets to equity ratio, that they are allowed to take. Therefore, banks' net worth as well as borrowers' credit-worthiness determine the maximum amount of loans that banks are able to grant. In the following, we will describe in details how the limiting mechanism works.

A bank, when receiving a loan request from firm f, evaluates the creditworthiness of the firm by examining its present balance sheet and, in particular, its equity level E^f and total debt D^f. Both equity and debt of firm f are assumed to be known with certainty by the bank and no informational asymmetries are supposed to apply here. Given E^f, D^f and the loan request amount λ^f, the bank computes the probability that firm f will default on the loan as:

$$\pi^f = 1 - e^{-\left(\frac{D^f+\lambda^f}{E^f}\right)}. \tag{1}$$

It is worth noting that the default probability π^f correctly increases with the leverage of the firm.

The modeled Basel II-like regulatory capital requirement sets the maximum amount of money that any bank can lend. Let us identify the previous bank with the index b. Given the capital (equity) of bank b, henceforth E^b, the capital adequacy rule states that the total value of the bank's risk-weighted assets, henceforth A^b, has to be lower (or equal) than α times the level of E^b. We consider α, which can be interpreted also as a leverage ratio, as the key policy parameter of this study. As stated before, bank assets are given by loans and liquidity; therefore, considering a risk weight equal to zero for liquidity, A^b is given by the sum of outstanding loans granted by bank b weighted by their riskiness. We stipulate that the risk weight of each loan is equal to its default probability, computed according to Eq. 1, where the equity and the debt of the firm are the ones related to the time when the loan was granted.

Granting a new loan λ^f would increase the total value of risk-weighted assets of the lending bank by the amount $\pi^f \lambda^f$; therefore, if firm f asks for a loan λ^f to bank b, the bank is allowed to supply a credit amount $\ell^b \leq \lambda^f$ determined as follows:

$$\ell^b = \begin{cases} \lambda^f & \text{if} \quad A^b + \pi^f \lambda^f \leq \alpha E^b, \\ \frac{\alpha E^b - A^b}{\pi^f} & \text{if} \quad A^b + \pi^f \lambda^f > \alpha E^b > A^b, \\ 0 & \text{if} \quad A^b \geq \alpha E^b. \end{cases} \tag{2}$$

Finally, bank b computes the interest rate i^{λ^b} that would be applied to the loan ℓ^b if accepted by the firm. The lending rate i^{ℓ^b} is then determined as follow:

$$i^{\ell^b} = i^c + \gamma^b \pi^f, \tag{3}$$

where i^c is the policy rate set by the central bank and $\gamma^b \pi^f$ is a rate spread depending on the riskiness of the loan through π^f. Parameter γ^b sets the sensitivity of the interest rate spread to the credit worthiness of the firm and is an evolving parameter that basically adjusts in order to reinforce the previous choices that were successful in increasing the bank's profits.

1.3 Matching Demand and Supply of Credit

As stated in the two previous paragraphs, when a firm f enters into the credit market to borrow a needed amount of money λ^f, it contacts a random subset of the whole set of banks operating in the economy. Each contacted bank, indexed by b, then determines the loan conditions, i.e., the maximum amount of money it is allowed to lend $\ell^b \leq \lambda^f$, according to the Basel II-like capital adequacy rule described by Eq. 2, and the interest rate associated to the loan i^{ℓ^b}, which, as stated by Eq. 3, foresees a spread on the central bank policy rate as a reward for the risk. Firm f then first accepts the loan characterized by the best conditions, i.e. by the lowest interest rate, and, if this loan is smaller than the total amount λ^f needed, the firm accepts the loan of the bank offering the second best interest rate, and so on if the money borrowed so far is still lower than the amount λ^f needed, till the loan characterized by the highest interest rate has been accepted. If the firm is still rationed in its demand of loans in the credit market, even taking into account the loan at the worst conditions, then it tries to raised the residual amount of needed money by issuing new shares and selling them in the stock market.

On banks' side, they receive demands by firms sequentially and deal with them in a "first come, first served" basis.

The rationale of banks' capital requirement is to provide a cushion in the case of firm bankruptcy and consequent debt restructuring. In that case, banks face a write-off in their loan portfolio and a consequent reduction of their equity capital. In this respect, the value of α acts like a multiplier of risk, e.g., in the case of $\alpha = 10$, a relative small write-off of the 10 % of the loan portfolio may cause the wipe out of the entire net worth of the bank. Therefore, while on one hand high values for α mean loose credit constraints which may give rise to a debt-fueled economic boom, on the other hand, high α increase significantly the systemic risk of the economy, e.g., even a single bankruptcy may trigger a severe credit crunch which in turn causes further bankruptcies and a consequent collapse of economic activity.

Finally, it is worth noting that opposite forces are also in place to stabilize the system. The bank can choose if to pay or not to pay dividends to shareholders and this choice is crucial for driving the equity dynamics. In particular, if a bank is subject to credit supply restriction due to a low net worth compared to the risk-weighted assets portfolio, then it stops paying dividends so to raise its equity capital and increaser the chance to match in the future the unmet credit demand. Furthermore, loans are due to be paid back by firms in a predetermined and fixed number of constant installments.

The following computational experiment will present the impact of different settings of α on the Eurace economy.

2 The Computational Experiment

Results presented in this section have been obtained using a computational setting of 2000 households, 21 firms (20 consumption goods producers and 1 producer of investments goods), 3 commercial banks, a central bank, a centralized clearing house for the financial market and a government.

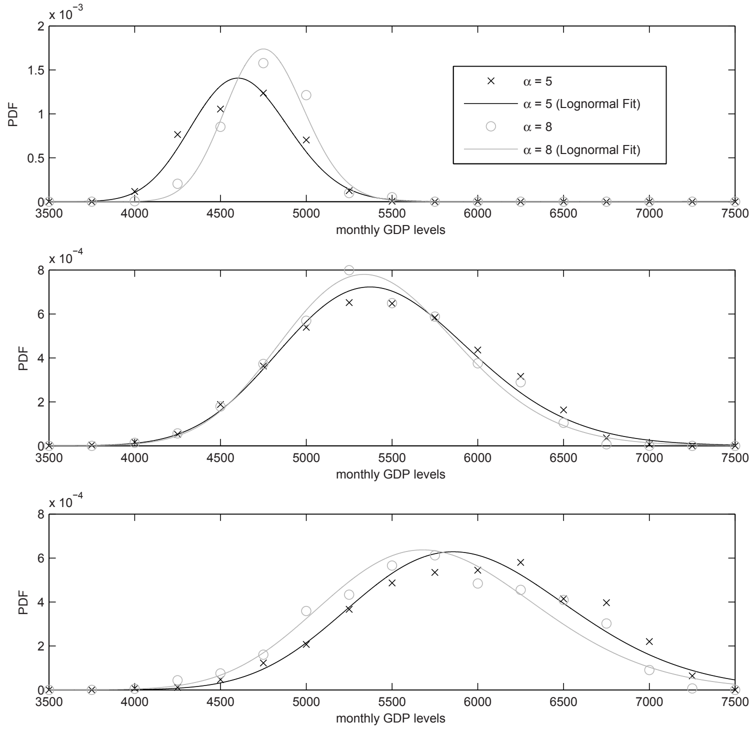

Fig. 1 Empirical PDF of monthly real GDP levels computed over 15 different seeds of the random number generator. Top panel represents GDP levels for the first 60 months. Central panel includes months from 61 to 240, and bottom panel from 241 to 480. Two values of α have been considered, i.e., $\alpha = 5$ (black line) and $\alpha = 8$ (gray line).

For six values of the parameter α, ranging from 4 to 9, a set of simulations of the Eurace model has been performed. Each set is composed by 15 different runs, corresponding to 15 different random seeds, in order to increase the robustness of results. It is worth reminding that $\alpha = 4$ corresponds to the lowest permitted leverage (strict capital requirement) while $\alpha = 9$ to the highest. Each simulation is composed by 9600 time steps, corresponding to 40 years (we consider 20 working days per month).

The model has been calibrated by using realistic empirical values both for the parameters of the agents and for the state variables initialization. Moreover, financial state variables of the various agents have been submitted to a crossed balance sheet consistency test in order to check the overall coherence of the model (see Teglio et al. [10] for details).

α	Real GDP 1 - 5	6 - 20	21 - 40	Unemployment (%) 1 - 5	6 - 20	21 - 40
4	4651 (22)	5519 (44)	6167 (100)	15.0 (0.5)	12.0 (0.4)	12.6 (0.2)
5	4628 (31)	5456 (25)	5959 (88)	16.5 (0.6)	11.4 (0.4)	13.9 (0.4)
6	4682 (12)	5411 (34)	5887 (65)	15.1 (0.2)	11.6 (0.4)	13.8 (0.5)
7	4756 (20)	5484 (37)	5966 (82)	13.6 (0.4)	11.4 (0.2)	13.9 (0.5)
8	4767 (20)	5410 (28)	5781 (103)	13.5 (0.4)	11.6 (0.3)	13.4 (0.6)
9	4722 (16)	5469 (27)	5990 (93)	14.3 (0.3)	11.0 (0.3)	13.2 (0.4)

Table 1 Mean values and standard errors (in brackets) of real GDP and unemployment rate (%) computed in the three time considered time intervals for different values of α. Statistics are calculated over 15 seeds of the random number generator.

Figure 1 shows the empirical probability distribution function of the monthly output levels during the 40 years of simulation for $\alpha = 5$ and $\alpha = 8$. They are grouped according to three different time spans: the first five years (top panel), from year 6 to year 20 (central panel), from year 21 to year 40 (bottom panel). For each α, all the 15 simulations corresponding to different random seeds have been included.

α	Banks' equity 1 - 5	6 - 20	21 - 40	Outstanding credit 1 - 5	6 - 20	21 - 40
4	4106 (52)	4662 (86)	2398 (156)	18271 (194)	20137 (278)	18425 (632)
5	3651 (30)	3823 (57)	1831 (111)	18544 (177)	19628 (123)	17158 (631)
6	3589 (22)	3696 (52)	1679 (111)	20202 (118)	19795 (175)	17116 (519)
7	3286 (21)	3135 (67)	1561 (194)	20976 (183)	20564 (249)	17632 (981)
8	3066 (10)	2968 (45)	1287 (113)	21001 (167)	20311 (214)	16666 (726)
9	2971 (11)	2916 (62)	1194 (94)	20827 (141)	20255 (196)	16994 (631)

Table 2 Mean values and standard errors (in brackets) of banks' aggregate equity and outstanding credit computed in the three time considered time intervals for different values of α. Statistics are calculated over 15 seeds of the random number generator.

Comparing the three panels it emerges that the economy is characterized by a long-run growth both in the case of $\alpha = 5$ and $\alpha = 8$. However, significative differences come out when considering the two curves within the same panel. In the case of higher capital requirement, i.e., $\alpha = 5$, the output level is lower and characterized by more variability in the first 5 years (top panel). This can be explained keeping into account that the risk weighted assets of each of the three banks have been initialized to be five times the initial level of equity, with the implication that for $\alpha = 5$ the constraint on bank leverage is binding and it is not possible for banks in the short run to increase the supply of credit in order to match the demand by firms. Table 2 corroborates this interpretation showing that the outstanding bank credit in the first period is significantly lower for $\alpha = 5$ than for $\alpha = 8$. The lower level of credit supply reduces the opportunities for firms to invest and to increase production, therefore affecting the short-run values of GDP and creating more instability.

The central panel of figure 1 clearly shows that in the medium term (6 - 20 years) the loss of output given by credit rationing is completely recovered. This can be

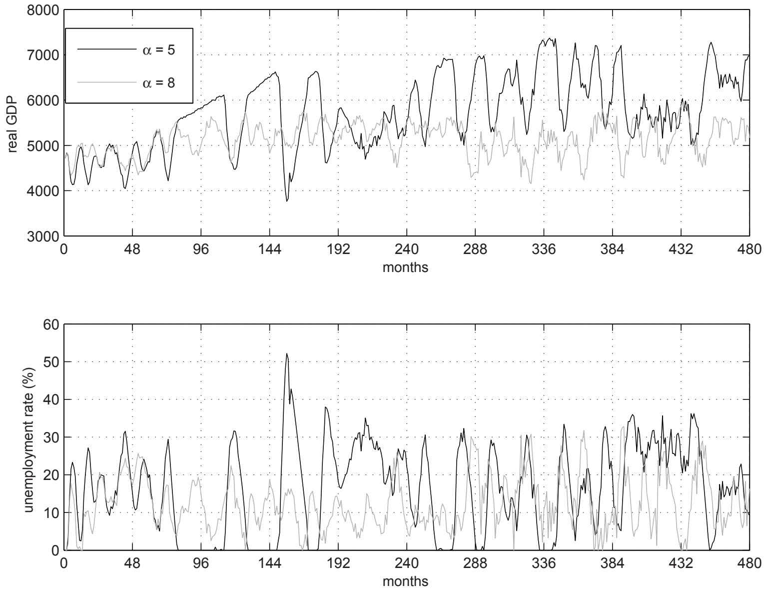

Fig. 2 Real output (top panel) and unemployment rate (bottom panel). Two different values of α have been considered for the same seed, i.e., $\alpha = 5$ (black line) and $\alpha = 8$ (gray line).

interpreted again by looking at table 2, showing that the aggregate equity level of banks is decreasing with the leverage level α in the short run. This is due to the fact that when banks face a credit demand higher than their supply constraints (low α), they stop to payout dividends in order to raise net worth so to meet the future demand of credit.

The bottom panel of figure 1 shows that the short run implications are reversed in the second half of the simulation, where a better economic welfare is observed in the case of higher capital requirement ($\alpha = 5$). Table 1, examined along with table 2, shows that lower capital requirements reduce banks' equity and limit economic expansion in the long run. Banks' equity is reduced when firms go bankruptcy, and debt write-offs are clearly higher for more indebted firms. The higher level of debt accumulated in the short run in the case of high α therefore causes a more severe reduction of banks' equity. Moreover, a low capital requirement reduces credit rationing in the short run, inducing banks to payout their dividends and to keep low equity levels, being more exposed to further equity reductions due to firms' bankruptcies in the long run.

Figures 2 and 3 show the simulation paths of output, unemployment, banks' equity and outstanding loans for the same specific seed of the random number generator. The previous general considerations deduced by examining average outcomes

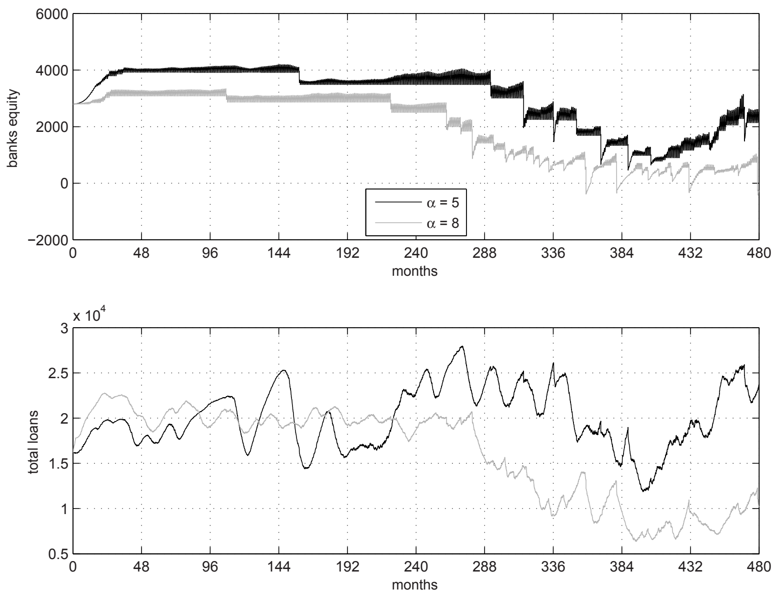

Fig. 3 Banks' aggregate equity level (top panel) and aggregate amount of outstanding loans (bottom panel). Two different values of α have been considered for the same seed, i.e., $\alpha = 5$ (black line) and $\alpha = 8$ (gray line).

are confirmed by the time trajectories of a realization. In particular, a higher capital requirement compels banks to raise equity, leading to a higher capacity for banks to meet credit demand in the long run. The dynamics of total credit is consistent with the equity path, showing contractions of loans caused by insufficient levels of equity on the side of banks. Real output, characterized by long-run growth and irregular business cycles, appears to be correlated with the total amount of loans. From around month 240, the total loans path for $\alpha = 5$ clearly dominates the corresponding path for $\alpha = 8$, and a similar relation can be observed comparing the two output trajectories.

3 Conclusions

The aim of this paper is to approach the highly topical issue of the macroeconomic implications of capital requirement for banks, using the complex economic environment of the agent-based Eurace model. The methodology has been to run a what-if computational experiment varying the level of capital adequacy requirements. In the presented credit market model banks are price makers and quantity takers and the lending rates depend on the creditworthiness of the borrower. Credit demand can be rationed if the equity capital of the bank is too small with respect to bank's weighted assets.

Results of the experiment show that capital requirements have important macroeconomic implications that depend also on the considered time span. In general, while leading to an expansion of output in the short run, regulations allowing for a high leverage of the banking system tend to be depressing in the medium and in the long run.

Our conclusions suggest that an appropriate set of regulations of the banking system could have strong potential benefits for growth and economic stability. Moreover, we can affirm that agent-based models (and in particular the Eurace model), based on a complex pattern of interaction among economic agents, can be considered as an alternative to standard DSGE models for performing policy analysis and economic forecasting.

Acknowledgement

This work was carried out in conjunction with the Eurace project (EU IST FP6 STREP grant: 035086) which is a collaboration lead by S. Cincotti (Universit di Genova), H Dawid (Universitaet Bielefeld), C. Deissenberg (Universit de la Mditerrane), K. Erkan (TUBITAK-UEKAE National Research Institute of Electronics and Cryptology), M. Gallegati (Universit Politecnica delle Marche), M. Holcombe (University of Sheffield), M. Marchesi (Universit di Cagliari), C. Greenough (Science and Technology Facilities Council Rutherford Appleton Laboratory).

References

1. Admati, A. R., DeMarzo, P. M., Hellwig, M. F., Pfleiderer, P., Sep. 2010. Fallacies, irrelevant facts, and myths in the discussion of capital regulation: Why bank equity is not expensive. Working paper series of the max planck institute for research on collective goods, Max Planck Institute for Research on Collective Goods.
2. Adrian, T., Shin, H. S., July 2010. Liquidity and leverage. Journal of Financial Intermediation 19 (3), 418–437.
3. Berger, A. N., Herring, R. J., Szego, G. P., June 1995. The role of capital in financial institutions. Journal of Banking & Finance 19 (3-4), 393–430.
4. Blum, J., Hellwig, M., April 1995. The macroeconomic implications of capital adequacy requirements for banks. European Economic Review 39 (3-4), 739–749.
5. Cincotti, S., Raberto, M., Teglio, A., 2010. Credit money and macroeconomic instability in the agent-based model and simulator Eurace. Economics: The Open-Access, Open-Assessment E-Journal 4 (26), http://dx.doi.org/10.5018/economics-ejournal.ja.2010-26.

6. Dawid, H., Gemkow, S., Harting, P., Neugart, M., 2009. Spatial skill heterogeneity and growth: an agent-based policy ananlysis. Journal of Artificial Societies and Social Simulation 12 (4).
7. Dawid, H., Neugart, M., Wersching, K., Kabus, K., Harting, P., Gemkow, S., april 2007. Eurace Working Paper 7.1: Capital, Consumption Goods and Labor Markets in Eurace. Tech. rep., Bielefeld University.
8. Hellwig, M., Jul. 2010. Capital regulation after the crisis: Business as usual? Working paper series of the max planck institute for research on collective goods, Max Planck Institute for Research on Collective Goods.
9. Myers, S., Majluf, N., June 1984. Corporate financing and investment decisions when firms have information that investors do not have. Journal of Financial Economics 13 (2), 187–221.
10. Teglio, A., Raberto, M., Cincotti, S., 2010a. Balance sheet approach to agent-based computational economics: the EURACE project. Vol. 77 of Advances in Intelligent and Soft Computing. Springer Verlag, pp. 603–610.
11. Teglio, A., Raberto, M., Cincotti, S., 2010b. Endogeneous credit dynamics as source of business cycles in the Eurace model. Vol. 645 of Lecture Notes in Economics and Mathematical Systems. Springer Verlag, pp. 203–214.

Modeling the Textbook Fractional Reserve Banking System.

Jacky Mallett

Abstract Banking systems based on the fractional reserve banking process have been in use for several hundred years. However textbook models of these systems do not include either loan repayments or loan defaults, and predict the evolution over time of a stable, asymptotically converging process governing credit and money supplies to the general economy for which there is no empirical evidence in long run monetary time series. In this paper we describe a computer simulation of a simplified model of a fractional reserve banking system that includes loan repayment. We show that this demonstrates several issues in the accepted model including instabilities arising from flows of money and credit between different banks, a discrepancy with the accepted value of the money multiplier within the system, and the appearance of a cyclic governing function over time.

1 Introduction

Fractional reserve banking has been the basis for western banking systems for over 200 years, evolving from gold based payment and storage facilities developed by medieval goldsmiths [3] The fractional reserve banking process is usually modelled in introductory Economic textbooks as a straightforward recursive function which is created by the redeposit of money within the banking system as banks create loans against deposits. Banks are represented in this model as being permitted to extend loans in proportion to the amount of money they have in their customer's deposits whilst retaining a required reserve. Although the required reserve is usually described as providing the day to day cash required to meet customer demands for payments and transfers, it also serves to prevent unlimited expansion of bank deposits as loans are extended to their customers, and new deposits are created as a

Jacky Mallett
University of Reykjavik, Inst. of Intelligent Machines, Klapparstigur 18, 101 Reykjavik, Iceland
e-mail: warlock@alum.mit.edu

S. Osinga et al. (eds.), *Emergent Results of Artificial Economics*, Lecture Notes in Economics and Mathematical Systems 652, DOI 10.1007/978-3-642-21108-9_5,

consequence. Considerable analytical complexity then arises both from the interaction of monetary flows within the system as the network of loans and their repayments created by the banks redistributes the sums on deposit between lending institutions, and also from the changing regulatory frameworks which have been used to attempt to control the system over historical periods and in different countries. Instability in the banking system is usually attributed to the problems caused by runs on the bank, when for whatever reason customer demands for cash or transfers to other banks cause the bank to exhaust the fraction of total deposits that it keeps on hand. Historically, the position of the Central Bank as a lender of last resort was one solution to these issues, as was the introduction of public deposit insurance.

It seems fair to observe that the fractional reserve banking system itself appears to be both misrepresented and misunderstood both within accepted economic theory and outside of it, and from this perspective alone a computer model may be useful as a pedagogical tool. Discussion of the mechanics of the banking system has a long history, dating back to the 19th century, but the long duration of loan cycle related effects, in conjunction with a sequence of regulatory changes has presented considerable challenges to analysis, as have other issues including debates over the precise status of bank deposits with respect to physical tokens of monetary exchange. Although statistics from the 19th century clearly showed deposit expansion occurring within the banking system [4] (punctuated with periodic and extremely rapid contractions), their precise cause was a matter for debate until the early 20th century. Even today, debates continue about the endogenous nature of money, and whether the continuous monetary expansion that can be observed in long term statistical series of the money supply is a cause of credit expansion or its consequence [7].

The simple textbook model currently provided in modern introductory textbooks [6] appears to have originated from the description provided by John Maynard Keynes [8] for the 1931 Macmillan Report to the British Parliament [5]. This correctly described the loan/money creation re-deposit process but did not include any description of how loan repayments or loan defaults would effect the behaviour of the system.

Incomplete as it is, the Keynsian model is preferable to those of the Austrian school within Economics, which have been unduly influenced by Murray Rothbard's [9] claim that individual banks can lend a multiple of their deposits rather than a fraction, something that can be disproved by reflection on the size of the resulting exponential expansion if this were the case, or by examining the annual reports of individual Banks. [1]

A further complexity are the many changes and variations in the implementation of the banking system itself. Over the course of the 19th and 20th centuries, several distinct, and very differently regulated frameworks can be distinguished: gold reserve banking without a central Bank (USA pre 1914); with a central bank (Europe); reserve based banking without a fixed relationship to gold (first world

[1] Although the Rothbard claim that individual banks can lend ten times their deposits is factually incorrect, a complication to empirical analysis is that some regulatory frameworks, notably the current Basel treaties do allow individual banks to lend slightly more money than they have on deposit.

war); and with one (the inter-war period) and the post second world war Bretton Woods regime, and lately the three Basel treaties. These last have instituted significant changes in moving the system away from reserve based regulation to a regime where regulation is based on individual bank's capital holdings, with no effective central control on the total amount of such holdings within the banking system.

A final issue that has also added to confusion within Economics over the behaviour of the monetary system and which is evident in the literature, is the changing definition of money. Economists in the 19th and early 20th century generally regarded physical notes and coins, gold and in some cases silver as money - but not deposits recorded at Banks and cheques or other financial instruments drawn against them. Irving Fisher [2] writing in 1911 for example explicitly describes bank deposits as not being equivalent to physical money.

This confusion can also be seen currently in some public measures of the money supply, the United State's M2 measure for example includes money market funds, which are a mixture of cash holdings and investments in short term debt instruments. When monetary expansion is discussed in the context of fractional reserve banking, what is strictly meant is the expansion of deposits within the banking system, rather than the general expansion of physical money. However In the increasingly electronic systems in use today, the distinction between physical money and bank deposits is not a particularly useful distinction. It was though significant in the context of gold reserve based banking systems, which maintained ratios of physical money to gold, based on the price of gold, which price would have also been mediated by the amount of money represented as being on deposit within the banking system. Some of the instabilities experienced during gold standard periods for example may be attributable to the expansion of deposits within the banking system varying in ways that were not generally understood at the time. It is these kinds of issue that makes a detailed understanding of systems based on the fractional reserve significant for modern economic analysis.

Economic and agent based models of the economy that include a banking sector exist. However there do not appear to be any models concentrating purely on the behaviour of the banking system in isolation, based on the known rules governing deposit and loan regulation. Even papers which acknowledge the issue in the existing model such as that by Berardi [1], and attempt to construct alternative models do so with systems based on higher level economic constructs, rather than trying to isolate the peculiar behaviour of the banking system itself in response to changes in configuration. Berardi does show however the presence of network effects in the flows of debt repayments within the system, which are certainly one of the several sources of potential variability in the behaviour of the system. Generally though, existing economic models are complex constructions of economic behaviour at levels of analysis above the banking system, making it difficult to extract the purely mechanical evolution over time of the banking system, from second order effects arising within the economy from its effects on agent behaviour.

2 Textbook Description

In the standard textbook model of the banking system, money is deposited with banks which is then lent out to borrowers. In this model, the definition of money is implicitly the sum of all deposits and reserves held by the banks. As loans are made, additional deposits are created as the loan capital is redeposited within the banking system. In order to prevent unlimited monetary expansion, and also to ensure that banks have sufficient funds to meet day to day demands by customers for access to their funds, a reserve is kept based on the quantity of deposits held by each bank. The possibility of confusion of money with debt within the system - for example by allowing reserves to be held in financial instruments that represent loans (e.g. treasury certificates) which occurs in today's banking system, is not included in the model, which implicitly assumed that money and debt are kept completely separate.

The model is typically presented in the form of a series of deposits, loans and reserves made between a set of banks, with a specified reserve requirement, as shown in Table 1.

Table 1 Expansion of bank deposits with 10% Reserve requirement.

Bank	Deposit (Liability)	Loan (Asset)	Reserve
A	1000	900	100
B	900	810	90
C	810	729	81
D	729	656	72
E	656	590	66
F	590	531	59

As loans are created against each new deposit, the resulting monetary expansion is progressively throttled by the reserve requirement. The limit on total monetary expansion by the banking system is presented in conjunction with this model as the theory of the money multiplier (M), which is expressed as $M = 1/r$ where r is the reserve requirement or ratio expressed as a fraction. [2] For example, where the reserve requirement is 10% or $1/10$ the Money multiplier is $M = 1/r = 10$

The formula is derived from the expansion series:

$$x + x(r) + (x(r))r + ... = x \sum_{k=0}^{\infty} (r)^k \tag{1}$$

which converges to

$$x/(1-r) \tag{2}$$

[2] Basel treaty regulation used in modern systems has effectively removed the reserve requirement as a regulatory factor, replacing it with capital regulation. Consequently many modern banking systems no longer have full reserve requirements.

By extension, and assuming a complete separation of money and debt within the model banking system, the model also predicts limits on the total quantity of loan capital that is issued against the money held by the banking system as shown in Figure 1 [10].

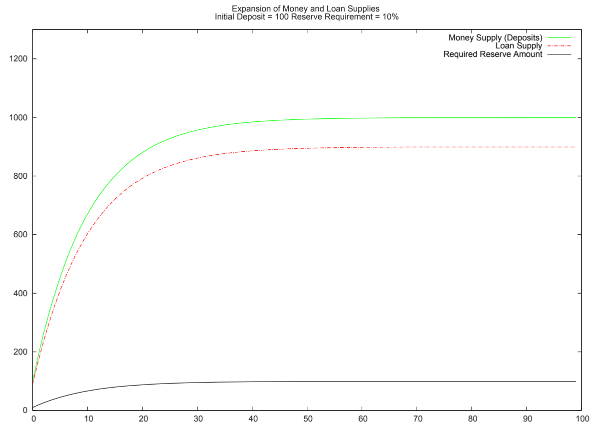

Fig. 1 Theoretical limits of the Money and Loan Supplies

3 A Simple Model of the Banking System

The model presented in this paper is established as a set of banks, each with deposits held by individual deposit holders. Loans are issued to deposit holders loans within the limits of loan availability. In order to provide a flow of money within the system, deposit holders are regarded as bank employees and receive a ”salary” each accounting period that allows them to make their capital and loan repayments for that period, once money within their deposit account is exhausted. For reasons that will be discussed below, loans can only be made to deposit holders at the bank originating the loan.

This is a highly simplified version of the process that would operate in the larger economy where the flow of money would usually be intermediated through several additional transfers within the economy. It does however allow some of the purely mechanical behaviour of monetary expansion and contraction within the banking system as loans are extended against regulatory limits to be isolated and explored.

Both the interest rate and the reserve ratio can be modified for the system, although only a single value is allowed for all loans of the system at any given time point. The number of banks and the number of employee’s per bank can be assigned arbitrarily. A lower limit on the amount being lent is set proportional to the loan duration, to prevent loans becoming too small. In order to prevent problems due to rounding issues, loan capital repayments are restricted to integer quantities, and any necessary adjustments are made in the final loan repayment period. Loan defaults within the system are handled by arbitrarily extending the repayment period so eventually loans are completely repaid. The general issue of how loan defaults are handled is an interesting one, and possibly a critical element in examining the accuracy of the textbook description as a model of the actual banking system, since any failure to repay loans within the system inevitably causes a removal of loan capital and a consequent contraction of the money supply.

Within the flow of deposits and loans at individual, an order of evaluation issue occurs between the repayment of loans, and the creation of new loans. If new loans are made before repayments are credited, then a larger amount of lending can be performed by the bank, than in the reverse case owing to the contraction in deposits caused by loan capital repayment. In this paper it is assumed that loan repayment before the creation of new loans in order to minimize problems of loan default. It is not known how this issue is handled in actual banking systems, or if it is even acknowledged.

Am important restriction in the system is that all loans are made as simple interest loans with uniform capital repayment schedules. The system was implemented with this restriction in order to simplify verification of the results. Given the general sensitivity of the system to capital repayments, it is probable that a compound interest schedule would change its behaviour, and this is something that should be explored further.

4 Results

Possibly the most accurate description that can be made about the fractional reserve banking system is that it appears to be sensitive to almost any condition. Two broad states can be determined, one which is asymptotically stable as presented in the textbooks, and another which shows distinct cyclic properties. Which occurs depends on the balance for any given simulation run between a Bank's ability to lend and the repayment of loan capital. When a bank's ability to make new loans is constrained, either by an absence of depositors to lend to, or by loan default a cyclic pattern emerges, as loan capital repayments cause the money supply to contract, whilst new lending is constrained.

One interesting finding was that the money multiplier for any given configuration of the system was partially dependent on loan duration. This arises from the problem mentioned earlier, that the rate of future lending is a function of previous loan repayment, and so full expansion of the system is dependent on the balance between capital repayments and new lending within each loan accounting period.

4.1 Textbook Description

An immediate issue presents itself in the literal interpretation of the textbook description shown in Table 1, as shown in Table 2, which led to modifications of the model used. Assume that loan repayments begin on the second expansion round, i.e. after Bank A has made a loan to the depositor at Bank B, and that the Banks are making simple interest loans with a duration of 10 months. When the depositor at

Table 2 Expansion of bank deposits with 10% Reserve requirement.

Round	Bank	Deposit (Liability)	Loan (Asset)	Reserve
1	A	1007	810	197
1	B	803	0	0
2	A	1007	810	197
2	B	803	723	80
2	C	722	0	722

Bank B makes the repayment of interest and capital on the loan, 90 in capital and 9 in interest payments are transferred to Bank A. Bank B then makes a loan to a depositor at Bank C, which as previously noted is less than indicated in the textbook example because of the interest and capital repayment just made. The issue that then presents itself occurs as a result of the next round of loan repayments. The repayment made by the depositor at Bank B on the loan from Bank A, is necessarily greater than the repayment received by Bank B from the depositor at Bank C, since the loan amounts are decreasing with each Bank in the cascade. Consequently after these repayments are made, Bank B is no longer in regulatory compliance as its loan amount is more than 90% of its deposits.

This problem arises in the textbook model because it is presented as an artificial cascade of loans in order to illustrate the re-deposit and consequent expansion of money within the banking system as loans are made. All Banks in the cascade, with the exception of the first are in the position of receiving smaller payments on the loans they have originated, than are being made on the loans that were deposited at them. In actual banking systems however, a similar issue could be expected to occur with any set of banks which allowed loans to be made to customers at other banking institutions. For example, if a single bank exists in a geographically separate area, it will be stable as long as all of its loans are made to its own customers. If a new bank opens up with access to its customer base, and either bank makes a loan to a customer at the other institution, then instabilities can be expected to arise at some point in the future purely as a result of unbalanced monetary flows between the two institutions as their customer's repay their loans.

In addition, it is noticeable that there is a tendency for money to become concentrated at the bank that early on originates the largest loan due to the consequent flow of interest payments. This is interesting in that it parallels the historical development of the banking system where typically this bank subsequently became the central bank and came to act as the provider in the last resort of loans to the rest of the banking system.

4.2 Regional Banking Model

In order to avoid issues arising from network flows of debt between different banks, the model was modified to restrict loans so that they are only made to depositors at the same bank which originated the loan. This allowed the simplest form of the money and credit repayment behaviour to be studied, without the additional complication of debt flows.[3]

As this form of the model evolved over time, two clear patterns emerged, as illustrated in Figures 2 and 3. Figure 2 shows a run of the simulation with a loan duration of 12 accounting periods, an interest rate of 10% per annum, and 12 depositors. The initial deposits in the system are 10,000 monetary units, 5,000 of which are held by a single depositor and 5,000 by the bank. There are no constraints on bank lending and the money supply expands to the maximum possible with this duration which varies between 4.2 and 4.8 on successive loan repayment rounds.

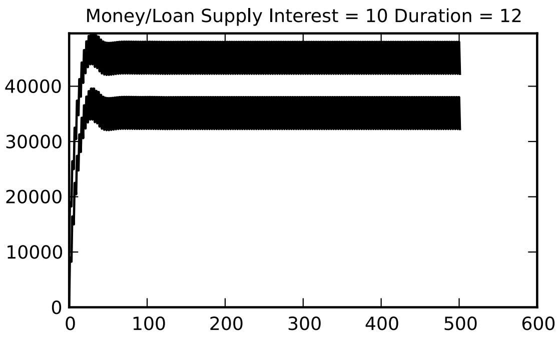

Fig. 2 Monetary Expansion with 12 Depositors

In contrast, Figure 3 shows a run with the same parameters, except that there are only 11 depositors. Consequently since loan duration is 12 accounting periods, the bank is unable to make a loan in the 12th period, as no depositor can borrow. (Depositors may only have one loan at a time.) This triggers a cyclic contraction and expansion in the money supply with the money multiplier for the system varying between 3.3 and 4.9 as loans are repaid. Interestingly, this was not due to any form of loan default occurring, but occurred purely because there were no qualified borrowers for the bank to lend to.

[3] In actual banking systems the short term need for funds to stay within regulation is resolved through the overnight interbank lending system.

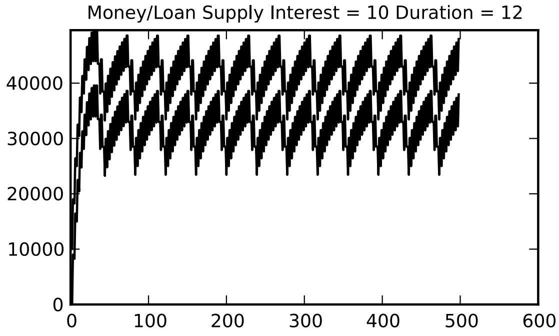

Fig. 3 Monetary Expansion with 11 Depositors

4.3 Evolution of the Money Multiplier

Within the model, capital repayments are treated as a deduction in the loan amount outstanding and a matching deduction in the deposit amount held by the debtor. Interest payments effectively represent a movement of money between accounts, and so have no effect on the money supply in a simple interest model. With longer loan duration periods, the ability of the system to expand to its limits were explored, and this showed that the money multiplier was not only a function of the loan duration, but was also able to exceed the predicted theoretical limit of the standard model of $1/ReserveRatio$.

In the standard model which does not include any form of capital repayment, each new loan is made as the difference between the total amount of deposits, minus the reserve requirement, and the amount currently on loan. When loan repayments are introduced to the system, each accounting period causes the loan capital repayment amount to be deducted from both the total loan supply and the total money supply. However, as the money supply is always greater than the loan supply, the percentage change in the amount on loan is slightly lower for the money supply than it is for the loan supply. Consequently, when the next loan is made, based on the difference between the money supply and the loan supply, it is for a slightly larger amount than would have occurred without loan repayments.

For example, as shown in Table 1 with a reserve requirement of 10%, and an initial deposit into the system of 1000, the second loan made without capital repayments is 810. If before the second loan is made, a capital repayment of 100 is made on the first loan, then the total money and loan supplies will be 1800 and 800 respectively. The next loan amount is then $1800 * 0.9 - 800 = 820$, resulting over time and assuming a sufficiently large loan duration in a larger monetary expansion than occurs without capital repayment.

As shown in Figure 4, which is a run with 500 depositors, and a loan duration of 240 accounting periods, since the reserve requirement also acts as a constraint on lending, once the full monetary expansion of the system has been reached, a cyclic pattern is once again seen in the evolution of the money and loan supplies over time.

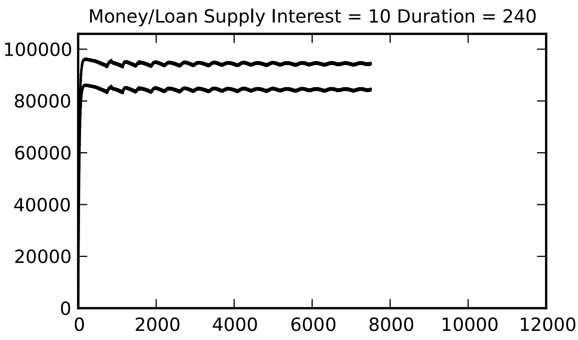

Fig. 4 Behaviour with long duration (20 year) loan periods.

5 Conclusion

Banking systems play a critical role within the economy in terms both of their ability to provide credit against their deposits, and also in their somewhat hidden role in modulating that part of the money supply that is represented as bank deposits. Given the history of instabilities associated with Banking, it is interesting that even a dramatically simplified model of fractional reserve based banking system could show the variability of behaviour shown here, and also significant given the variability both historically, and also between currencies in the regulatory frameworks used by different monetary systems.

The standard textbook description of the banking system currently used in Economics is not an accurate description of the majority of current Basel regulated systems, which no longer rely on reserve regulation.[4] Nevertheless, it is the starting point for any analysis of such systems and is also important for historical comparisons with periods with which it does bear more resemblance, notably the gold standard systems used in late 19th and early 20th century Europe.

The textbook model is notably deficient though, and this is especially evident with respect to both the money multiplier predictions it makes, and also the omission

[4] Basel based systems rely on capital regulation mechanisms, where the amount a Bank can lend is restricted by its regulatory capital holdings. Although some systems, notably those of China and Brazil still attempt to control their monetary expansion with escalating reserve requirements, most Basel countries have either minimized or removed the role of reserve requirements as a constraint on monetary expansion. It is perhaps worth remarking, that as a consequence the monetary base which is defined as the total amount of physical money and reserves in the banking system no longer provides any significant information about its total monetary expansion.

of loan defaults. Loan defaults are a particular problem since they cause a permanent removal of lending capacity from the system, and also trigger a monetary contraction. Neither the textbook model nor the model described here provide explanations for the phenomena of ”endogenous money expansion”, which is the generally observed tendency of banking systems under both gold standard and Basel regulation to experience continuous expansion in the total amount of money represented by their deposits.

The model described in this paper is extremely simple, particularly with respect to interest and loan capital repayments, which are evenly distributed over the duration of the loan, and consequently not realistic of actual banking systems which are almost invariably based on compound interest repayments. It does though allow the influence of different components of the system such as loan duration to be studied in isolation. It is also important, especially with systems such as fractional reserve banking that exhibit sensitivities to multiple conditions, and whose behaviour is clearly both complex and not well understood, not only to be able to isolate different sources of change, but also to be able to rule out or at least minimise one significant confounding source of variation, software bugs. However, this approach does leave open the question of whether normal variability within the general economy of loan periods and interest rates would smooth out the variability shown here or conceivably intensify it.

Given the general sensitivity of the money multiplier to capital repayments, it seems probable that using a compound interest formula to calculate loan and interest repayments would also create changes in the evolution of the system over time, and almost certainly be another source of potential cyclic behaviour, since compound interest formulas cause considerable variation in the amount of loan capital repayments within each accounting period over the course of a loan. Other potential areas for research are the extension of the model to include inter-bank lending with consequent inclusion of network effects in flows between banks, as well as lending by the central bank in its position of lender of the last resort; the introduction of loan sales to non-bank entities (Asset Backed Securities) which effectively de-couples the total bank originated loan supply from the money supply; and the difference in behaviour between theoretical reserve based regulatory mechanisms, gold standard era mechanisms and the Basel based regulatory frameworks currently in use.

References

1. Berardi, M.: Beyond the static money multiplier: In search of a dynamic theory of money. Artificial Markets Modeling, Volume 599, pp 1-16 (2007)
2. Fisher, I.: The Purchasing Power of Money, its Determination and relation to Credit, Interest, and Crises”, The Macmillan Company(1922)
3. Quinn, S.F. : Banking before the Bank: London’s unregulated goldsmith-bankers, 1660-1694, Ph.D Thesis University of Illinois (1994)
4. Higonnet, R.P: Bank Deposits in the United Kingdom 1870-1914, pp 329-367 The Quarterly Journal of Economics Volume 71 No. 3 (August 1957)

5. British parliamentary reports on international finance: The report of the Macmillan committee, pp. 34-35. H.M. Stationary Office, London (1931)
6. Mankiw, N.G.:Principles of Economics,pp 600-604, Dryden Press (1997)
7. Palley, T.I: Endogenous Money: What it is and Why it Matters, pp 152-180 Metroeconomica Volume 53 No. 2, Wiley (May 2002)
8. Stamp, J.C.: The report of the Macmillan Committee, pp 424-435. The Economic Journal Volume 41, London (1931)
9. Rothbard, M.: Fractional Reserve Banking The Freeman, (October 1995)
10. Federal Reserve Bank of Chicago: Modern Money Mechanics: A Workbook on Bank Reserves and Deposit Expansion, Chicago (1992)

Learning to Trade in an Unbalanced Market

Florian Hauser and Marco LiCalzi

Abstract We study the evolution of trading strategies in double auctions as the size of the market gets larger. When the number of buyers and sellers is balanced, Fano et al. [2] show that the choice of the order-clearing rule (simultaneous or asynchronous) steers the emergence of fundamentally different strategic behavior. We extend their work to unbalanced markets, confirming their main result as well as that allocative inefficiency tends to zero. On the other hand, we discover that convergence to the competitive outcome takes place only when the market is large and that the long side of the market is more effective at improving its disadvantaged terms of trade under asynchronous order-clearing.

1 Introduction

Recently, Fano et al. [2] have studied the evolution of trading strategies for a double auction when the number of traders increases. They provide two main results. First, the competitive outcome obtains under different market architectures, provided that the size of the market is sufficiently large. Second, the choice of the order-clearing rule affects trading behavior. Under simultaneous order-clearing, marginal traders learn to act as *price takers* and make offers equal to their valuations or costs. Under asynchronous order-clearing, the intramarginal traders learn to act as *price makers* and make offers equal to the competitive price.

An important feature of their study is the assumption that buyers and sellers populate the market in equal numbers; that is, the market is *balanced*. Moreover, using a simulative approach, agents' learning is modeled by means of a genetic algorithm. Curiously enough, although it is well known that "traders on the long side of a market wind up holding the short end of the stick" [8], the agent-based literature has paid scant attention to the study of unbalanced markets; however, see Gode and Sunder [5] or Anufriev et al. [1] for the special case of one seller, one intramarginal buyer and n extramarginal buyers.

This paper explores what happens in double auctions when we remove the assumption that the market is balanced. A second methodological contribution is that we model agents' learning by means of genetic programming, extending the reach of genetic algorithms. We confirm and extend the main results in Fano et al. [2] when the market is balanced.

Florian Hauser
Dept. Banking and Finance, Universität Innsbruck, e-mail: florian.hauser@uibk.ac.at

Marco LiCalzi
Dept. Management, Università Ca' Foscari Venezia, e-mail: licalzi@unive.it

S. Osinga et al. (eds.), *Emergent Results of Artificial Economics*, Lecture Notes in Economics and Mathematical Systems 652, DOI 10.1007/978-3-642-21108-9_6,

More importantly, we discover two novel effects when the market is unbalanced. Under simultaneous order-clearing, the convergence towards price-taking behavior expands from the marginal traders to the intramarginal traders on the long side, due to the competitive pressure. A second result is about the interaction of strategic behavior and order-clearing rule. When traders are forced to act as price takers, the disadvantageous terms of trade for the long side are worse under asynchronous order-clearing because the price is not uniform across all transactions. However, when we let traders optimize their strategies for the protocol in use, asynchronous order-clearing improves the average transaction price for the long side. In short, by forcing a uniform price over all transaction, simultaneous order-clearing dampens the effects due to strategic behavior.

2 The Model

There are $n = b + s$ traders, where b are buyers and s are sellers. Each trader wishes to maximize expected profits and is in the market to exchange at most one unit of a generic good per round. Each buyer i has a private valuation v_i and each seller j has a private cost c_j. Valuations and costs are drawn from two stochastically independent uniform distributions on $[0,1]$. Profits are $v - p$ for buyers and $p - c$ for sellers, where p is the price at which a transaction occurs.

We speak of intramarginal and extramarginal buyers (sellers) according to their position on the demand (supply) function with respect to the market-clearing price(s). We refine this qualitative distinction into a complete ordering, by defining the *strength* of a buyer with valuation v as the distance from the valuation of the weakest buyer ($v = 0$) and the strength of a seller with cost c as the distance from the valuation of the weakest seller ($c = 1$). Hence, stronger intramarginal traders have valuations (or costs) lying farther away from the market-clearing price(s).

Agents trade using one of two distinct trading protocols, whose main difference is in the nature of the order-clearing rule. Under simultaneous order-clearing, we have a *call market* where the offers made by buyers and sellers are aggregated to form their demand and supply functions and all transactions take place at a unique price (that we choose to be the mid-value of the interval of market-clearing prices).

Under asynchronous order-clearing, we have a *continuous double auction* where agents arrive in random order and sequentially submit binding [1] offers on the selling and buying books. Orders are immediately executed at the outstanding price if they are marketable; otherwise, they are recorded on the books with the usual price-time priority and remain valid unless a cancellation occurs. When a transaction takes place between two traders, their orders are removed from the books and they leave the market. Hence, orders are cleared asynchronously in separate trades, usually at different prices.

[1] See LiCalzi and Pellizzari [6] on the role of an assumption of binding offers.

Agents from a pool of N traders interact repeatedly and anonymously. In each *round*, we randomly draw (without replacement) b buyers and s sellers from the pool of available agents and let them visit the market ($b+s=n \ll N$). Hence, the agents active in the market during a specific round change randomly over time. This increases variety and learning opportunities.

We constrain traders to make offers that are *individually rational*: no buyer can bid more than his valuation and no seller can ask less than her cost. We evolve traders' strategies using genetic programming, henceforth nicknamed GP for brevity. This optimization method is an alternative approach to the genetic algorithm used in Fano et al. [2]. The design of our GP routines follows closely the standard tree-based approach described in Koza [4]. Evolution takes place sequentially over at least 2500 optimization steps.[2] In each step, one trader is randomly selected from the pool and his trading function is optimized by GP. As the complexity of the optimization problem is rather low, we choose conservative parameters for the GP algorithm.

An initial population of 100 trading strategies is generated by the ramped-half-half method with a maximum number of 100 nodes. Input data for a strategy is v or c, as well as a set of constants (0.01, 0.1, 2, 3, 7). To modify the input data, GP can refer to a set of standard arithmetic functions ($+,-,\times,\div,\exp,\ln$). The initial population of strategies is improved by an evolutionary process over 50 generations. To generate a member of a new generation, GP selects successful strategies from the previous generation based on a tournament of size 4 and applies crossover and mutation with probability 0.8 and 0.2, respectively. The fitness function is the average profit of a strategy over 30 rounds. The best strategy from the last generation is selected and the trader stays with it until he is chosen for a further optimization step.

We report results over a fixed experimental design with nine cells, where buyers are on the longer side of the market. Each cell corresponds to a different traders' population, where the number of sellers is equal to $s=1$, $s=5$, and $s=50$, against a number of buyers equal to $b=s$, $b=3s$, and $b=5s$. We usually arrange the nine cells in a 3×3 matrix, where the columns correspond to $s=1$ (left), $s=5$ (center) and $s=50$ (right), and the rows to $b=s$ buyers (top), $b=3s$ (middle), and $b=5s$ (bottom). In each experiment, the pool of agents learns to trade using the GP as described above. At the end of the learning process, for each cell we calculate results over 1000 additional rounds of trading.

3 Convergence to the Competitive Outcome

The *competitive price* obtains under simultaneous order-clearing when all agents truthfully report their costs and their valuations. We denote this by S-TT for brevity.

[2] We go up to 6000 steps for large markets with many agents.

(The rest of our shorthand notation is "A" for the asynchronous order-clearing rule and "GP" for the trading strategies developed by genetic programming.)

Generally speaking, we find that the trading price rapidly stabilizes around the competitive price both for S-GP and A-GP. This convergence is faster under S-GP. To exemplify, we compare in Figure 1 the average trading price (in black) and its standard deviation (in grey) for the case $s = 50$ and $b = 5s$ as a function of the optimization steps. (The statistics are computed over 1000 rounds of trading while freezing the GP process.) It is apparent that, compared to S-GP, it takes more time under A-GP before enough learning takes place to stabilize prices.

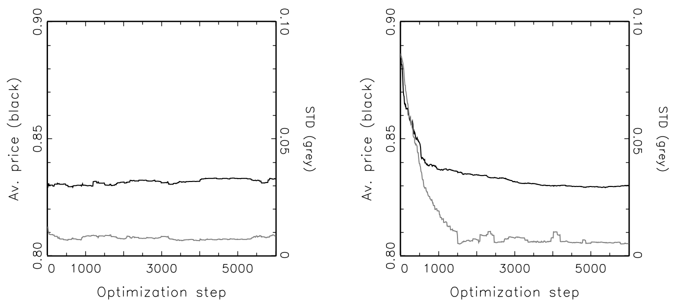

Fig. 1 Average price (black) and standard deviation (grey) under simultaneous (left) and asynchronous (right) order clearing for the case $s = 50$ and $b = 5s$.

More detailed information appears in Table 1, where for each cell we report in the first line the (average) competitive equilibrium price computed using S-TT (left) and the average trading price under S-GP (center). The second line in each cell gives the average trading price under A-TT (left) and A-GP (center). To control for the accumulation of orders in the book, we also give the average closing price on the right of the second line in each cell, using data from the last transaction in each round. Averages are taken over all the transactions from 1000 rounds per at least 10 distinct simulations with different random seeds.

	$s = 1$	$s = 5$	$s = 50$
$b = s$	0.498 0.496 0.495 0.494 0.494	0.501 0.501 0.503 0.500 0.499	0.500 0.499 0.500 0.500 0.500
$b = 3s$	0.601 0.698 0.638 0.630 0.630	0.690 0.723 0.735 0.705 0.710	0.746 0.748 0.807 0.744 0.745
$b = 5s$	0.647 0.774 0.716 0.703 0.703	0.763 0.806 0.818 0.789 0.794	0.829 0.831 0.882 0.829 0.830

Table 1 Each cell exhibits the average transaction price for S-TT (top left), S-GP (top center), A-TT (bottom left), A-GP (bottom center), as well as the average closing price for A-GP (bottom right).

The close alignment of the values in the cells of the first line (where $b = s$ and $s = 1, 5, 50$ as we move rightward) confirms the result in Fano et al. [2]: in a balanced market, the evolution of trading strategies stabilizes prices around the competitive price p^*, particularly when the market grows large. On the other hand, the second and third horizontal lines (with $b = 3s$ and $b = 5s$, respectively) show a mismatch between the price under the baseline case of S-TT (top-left) and the prices under S-GP or A-GP in the central column of the four cells with $s = 1$ or $s = 5$ in the bottom-left corner. This mismatch is supported for any practical level of confidence by a two-sided paired t-test. (By a *practical* level of confidence we mean a p-value lower than 10^{-5}.)

We conclude that there is no convergence of the realized prices to the competitive price p^* under price-taking behavior when the market is unbalanced ($b = 3s, 5s$) and the market size is not large ($s = 1, 5$). When the market is relatively small, the disadvantage borne by the long side turns out to lead to prices higher than p^* under either simultaneous or asynchronous order-clearing. This effect is novel with respect to what is reported by Fano et al. [2] under the assumption of balanced markets.

A careful examination of Table 1 reveals a second novel effect. Comparing the two prices in the central column of each cell for the six unbalanced markets, we find that the average transaction price under S-GP is systematically higher than the average transaction price under A-GP. For five of the six cells, this difference is statistically significant for any practical level of confidence based on a two-sided paired t-test. (The exception is the case $s = 50, b = 3s$ where we find a p-value of 0.004.) In other words, when we take into account that agents try and optimize their trading strategies with respect to the protocol in use, the long side of the market finds the asynchronous order-clearing rule more beneficial.

We come back to this issue in the next sections, but the main intuition is the following. The simultaneous order-clearing forces a uniform price for all traders on the long side, while the asynchronous rule allows different agents to trade at different prices. When both sides of the market optimize their strategies, no one on the long side can earn any advantage under a uniform price. On the other hand, the price dispersion allowed by the asynchronous rule can be put at use to mitigate the disadvantage borne by traders on the long side.

Contrast this with the case where we assume that agents stick to TT trading. Comparing values across the first column in each cell of the six unbalanced markets, we see that simultaneous order-clearing is preferred by the long side. When all agents trade truthfully, asynchronous order-clearing puts the long side at a disadvantage because it increases the competitive pressure each buyer must bear; see Gjerstad [3].

Clearly, the strategic behavior of traders interacts differently with the order-clearing rule when we move from TT to GP. Comparing S-TT and S-GP in the first row of each cell of the unbalanced markets, we see that under simultaneous order-clearing trading based on optimized strategies puts the long side to a disadvantage. On the other hand, comparing A-TT and A-GP in the second row of each cell of the unbalanced markets, we find that asynchronous order-clearing improves the terms of trade for the long side.

Finally, we note that, under both GP and TT, the difference between the average trading price under both the S or the A treatment shrinks as the size of the market grows: expanding the market size swamps the differences in prices because it increases competition among traders on the same side and thus reduces the strategic advantage that either leg of the market may hold.

4 The Evolution of Strategic Behavior

We report the outcome of our simulations about the evolution of strategic behavior in two distinct sections. The first one deals with simultaneous order-clearing. The second one discusses the asynchronous case.

4.1 Simultaneous Order Clearing

Figure 2 reports in black the average trading functions evolved by GP for the nine cells in our experimental design, as well as a sample of the actual trading functions in grey. The horizontal dashed lines denote the average theoretical price (grey), and the average price over all trades (black).

Given our assumption of individual rationality, buyers' functions lie in the lower triangle below the diagonal and sellers' functions in the upper triangle. For visual clarity and computational speed, we assume that an extramarginal agent picks truth-telling over another strategy whenever both options are optimal; e.g., since none of the individually rational bids from a buyer with valuation $v < 0.25$ leads to trading, we let him pick truth-telling. Hence, the initial (final) segment of buyers' (sellers') trading functions is actually *on* the diagonal.

A comparison of the average trading functions across the first line (where $b = s$) confirms the result in Fano et al. [2]: when the size of a balanced market grows, the marginal agents learn to make an offer equal to their valuation/cost and their trading strategies approach price-taking. Additional evidence can be gathered by the actual trading functions depicted in grey. Broadly speaking, their dispersion provides a visual representation of the level of variety in the learning of individual trading functions, and thus on the strength of the evolutionary pressure insisting on them. (Recall that by assumption the dispersion for very weak traders ends up being set to zero.)

As we move from left to right in the first row, the dispersion around the trading functions of the marginal traders shrinks to zero: there is a strong evolutionary pressure on all marginal traders to learn price-taking. The residual dispersion in learning for $s = 1$ and $s = 50$ has different explanations. When $s = 1$, we have a bilateral monopoly with one buyer and one seller where there is very little selection pressure due to lack of competition. When $s = 50$, under simultaneous order-clearing the uniform price p^* is set by the marginal traders; once these have learned to be price-

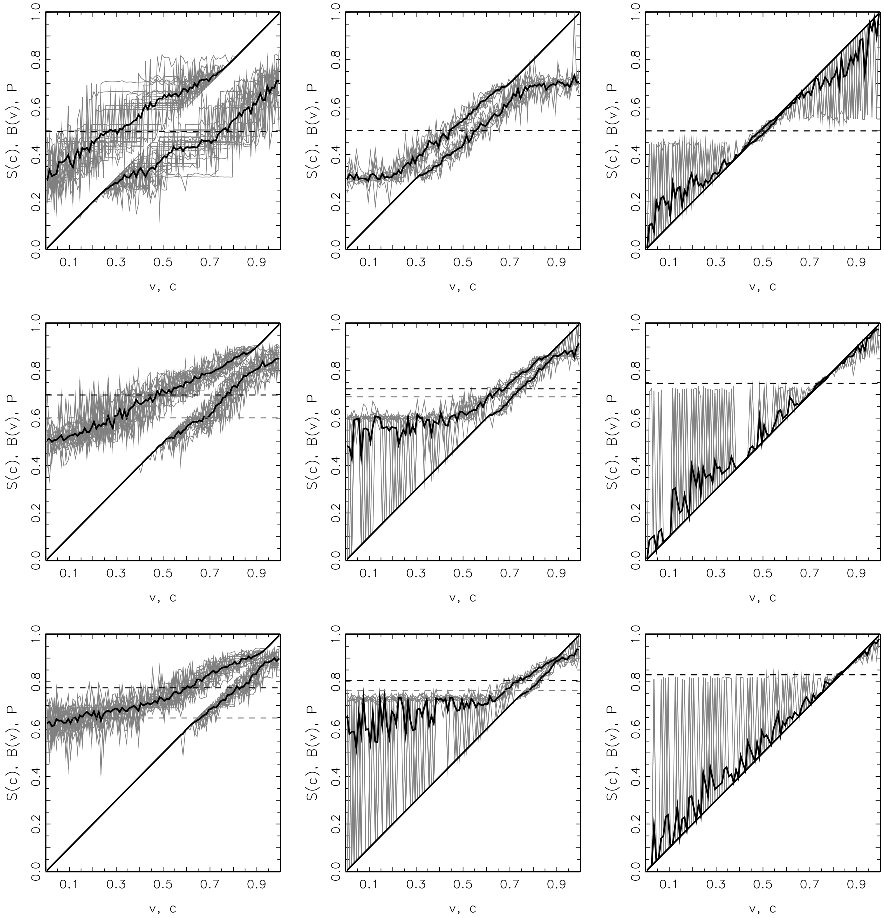

Fig. 2 Average trading functions under simultaneous order-clearing (solid black). The horizontal dashed lines denote the average competitive price (grey), and the average price over all trades (black). Columns are $s = 1$ (left), $s = 5$ (center), $s = 50$ (right); rows are $b = s$ (top), $b = 3s$ (middle), $b = 5s$ (bottom).

takers, a bid above p^* or an ask below p^* is equally profitable for the other traders. Therefore, the main evolutionary pressure experienced by the intramarginal buyers (sellers) is to make offers above (below) p^*; accordingly, we see a lot of dispersion but the intramarginal trading functions never escape such regions. Finally, the only profitable alternative to offering p^* is for buyers (sellers) to bid (ask) slightly above (below) p^* and extract better terms of trade, although this entails a significant risk of no trade. So the actual trading functions oscillate between price-taking and an offer close to p^*.

Consider now the unbalanced markets in the second and third row. Since $b > s$, buyers represent the long side of the market. A comparison of the average trading functions shows that, when the size of the market grows, the trading strategies converge to price-taking both for the marginal sellers and for a large interval of buyers around the marginal ones: that is, price-taking behavior spreads from the marginal traders to most agents on the long side, due to the competitive pressure. Accordingly, the dispersion of intramarginal buyers' trading functions shrinks to zero (except for the strongest ones) while this is not the case for intramarginal sellers. It is apparent that "traders on the long side of a market wind up holding the short end of the stick" [8].

4.2 Asynchronous Order Clearing

Figure 3 reports the average trading functions evolved by GP for the nine cells in our experimental design, using the same format as in Figure 2.

A comparison across the first line (where $b = s$) confirms the result in Fano et al. [2]: when the size of a balanced market grows, the intramarginal agents learn to make a constant offer equal to the competitive price and their trading strategies approach price-making. The dispersion of trading functions makes it clear that the evolutionary pressure applies to all intramarginal traders. The same observation holds when markets are unbalanced.

Combining the results about the effects of the order-clearing rule over the evolution of strategic behavior, we can say that two forces are at work. The first one was pointed out by Fano et al. [2]: the simultaneous rule pushes marginal traders towards price-taking, while the asynchronous rule attracts intramarginal traders towards price-making. The second novel effect is that, when the market is unbalanced and order-clearing is simultaneous, the competitive pressure spreads the price-taking behavior on the long side to all the intramarginal traders.

Different reasons underline these two effects. The first one stems from the asynchronous rule leading to non-uniform prices across trades: this increases the risk of trading using price-taking, and allows many more learning opportunities to run away from it. The second effect is due to the increase in competitive pressure among traders on the long side. As discussed in Sect. 3, a side effect of the asynchronous rule in allowing better learning is to help traders on the long side of small unbalanced markets to achieve better terms of trade than under simultaneous order-clearing. A similar point is made by Pouget [7]), who compares the Walrasian auctioneer trading protocol against the call auction and finds that in the latter one agents may fail to learn the competitive equilibrium if they use reinforcement learning.

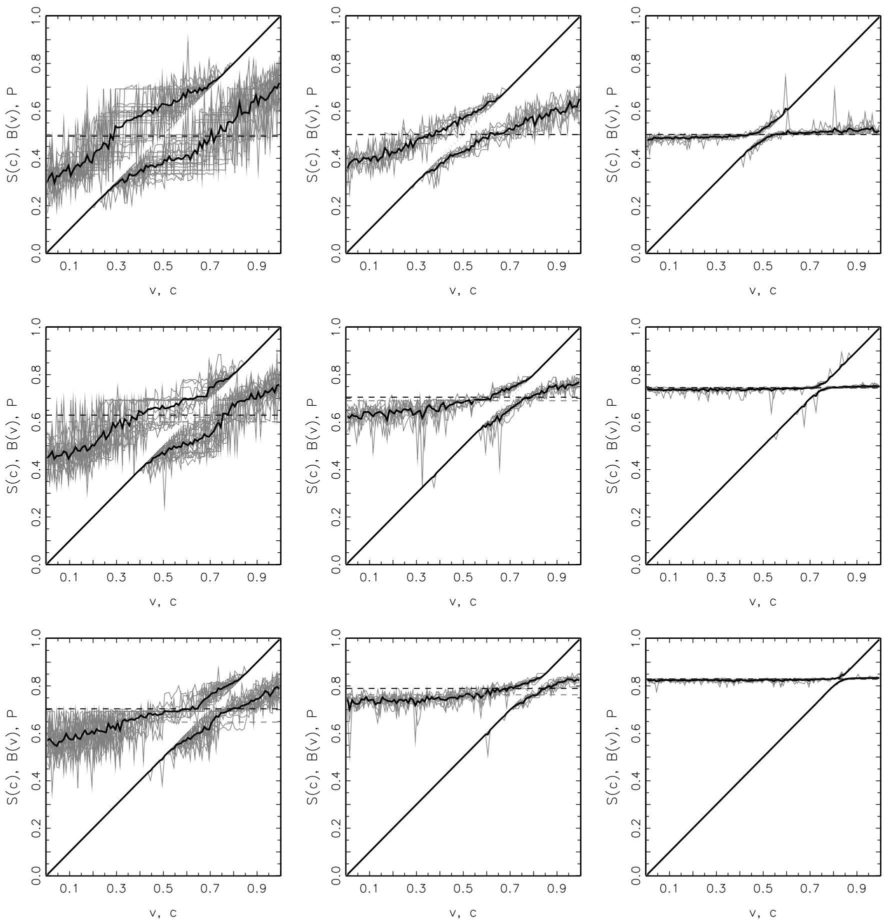

Fig. 3 Average trading functions under asynchronous order-clearing (solid black). The horizontal dashed lines denote the average competitive price (grey), and the average price over all trades (black). Columns are $s = 1$ (left), $s = 5$ (center), $s = 50$ (right); rows are $b = s$ (top), $b = 3s$ (middle), $b = 5s$ (bottom).

5 Allocative Efficiency

Our last batch of results concerns allocative efficiency, that we define as the ratio between the total surplus realized by traders and the theoretical maximum surplus. (This latter one is computed as the surplus realized under S-TT.) We confirm the result from Fano et al. [2] that the allocative inefficiency tends to zero as the size of the market grows regardless of the order-clearing rule, and we find that it holds also for unbalanced markets.

Ceteris paribus, the allocative efficiency is higher under the S treatment than under the A treatment for all cells in our experimental design. Moreover, as a function of the optimization steps, the efficiency remains virtually constant under simultaneous order-clearing while it takes several optimization steps before it levels off under the asynchronous rule. This qualitative difference conforms with our former observation that the effects of a simultaneous order-clearing rule swamp those associated with the evolution of trading strategies, while these latter ones carry much greater impact under the asynchronous rule. To exemplify, Figure 4 depicts the allocative efficiency (in black) and the buyers'/sellers' surplus ratio (in grey, to be discussed below) for the case $s = 50$ and $b = 5s$ as a function of the optimization steps. (The statistics are computed over 1000 rounds of trading after freezing the GP process.)

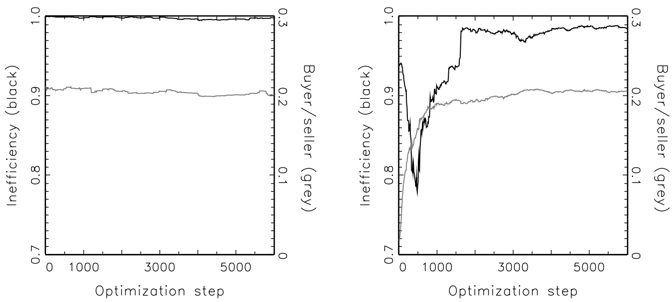

Fig. 4 Efficiency (black) and buyers'/sellers surplus ratio (grey) under simultaneous (left) and asynchronous (right) order clearing for the case $s = 50$ and $b = 5s$.

For a more complete description, we compare the surplus realized under S-GP and under A-GP. Using the surplus realized under S-TT as benchmark, we look at the ratios for the realized surplus in the two GP treatments over the benchmark. Table 2 provides for each cell of our experimental design the average values for such two quantities, computed over at least 10 distinct simulations. We report the ratio of surplus for S-GP over S-TT on the left and for A-GP over S-TT on the right. Apart

	$s = 1$		$s = 5$		$s = 50$	
$b = s$	0.785	0.797	0.971	0.888	0.996	0.968
$b = 3s$	0.863	0.824	0.982	0.923	0.998	0.980
$b = 5s$	0.898	0.854	0.982	0.941	0.998	0.984

Table 2 Average ratios for S-GP vs. S-TT (left) and for A-GP vs. S-TT (right).

from the case where $s = b = 1$ (for which the distinction between simultaneous and asynchronous order-clearing is of little consequence), allocative efficiency is higher under S than under A and grows to 1 as the size of the market increases. This claim is supported for any practical level of confidence by a two-sided paired t-test. The lack of balance among the two sides of the market does not seem to carry discernible effects, except for the increase in allocative efficiency due to the higher number of traders available in the unbalanced markets.

Clearly, allocative efficiency is affected both by the protocol and by the trading strategies adopted by the agents. For instance, since the ratios on the left of each cell of Table 2 are lower than 1, we can conclude that the strategic behavior developed under GP reduces the efficiency that could be achieved if everyone would make truthful offers. Unfortunately, it is well known that the pursuit of individual profits may hurt the common welfare.

For the case of asynchronous order-clearing, we attempt to disentangle the effects of protocol and behavior in Table 3 where we report the (average) ratio of surplus for A-TT over S-TT on the left and for A-GP over A-TT on the right. The ratio

	$s=1$		$s=5$		$s=50$	
$b=s$	1.000	0.797	0.807	1.100	0.702	1.379
$b=3s$	0.844	0.975	0.825	1.119	0.875	1.119
$b=5s$	0.820	1.042	0.864	1.089	0.937	1.050

Table 3 Average ratios for A-TT vs. S-TT (left) and for A-GP vs. A-TT (right).

between A-TT and S-TT measures the loss in allocative efficiency when traders using truth-telling switch from simultaneous to asynchronous order-clearing. Given that this latter rule destroys uniform pricing and adds noise due to the random order of arrival, we expect to see allocative inefficiency.

The second value on the right of each cell gives the (average) ratio between the surpluses realized under A-GP and A-TT. It measures the improvement in allocative efficiency brought over by optimizing trading strategies with respect to the protocol. Since the ratio is greater than 1 in all markets with $s \neq 1$, the pursuit of individual profits is socially beneficial under asynchronous order-clearing because it reduces both missed trades and the number of transactions involving extramarginal traders. Contrasting this with the case of simultaneous order-clearing, we conclude that the social benefits of strategic behavior may depend on protocolary details.

Our last comparison concerns the ratio between the surplus realized by the long side and the surplus realized by the short side. This should be (approximately) 1 in a balanced market and strictly lower than 1 in an unbalanced market. Table 4 reports the (average) ratio of buyers' to sellers' surplus under S-TT on the left, under S-GP at the center, and under A-GP on the right. As expected, the ratio is close to 1 for all balanced markets in the first row. Moreover, it is exactly 1 for S-TT with $s=1$ because the simultaneous order-clearing rule sets the clearing price exactly midway between the best bid and the best ask, so the two agents who complete a transaction must split the surplus equally between themselves.

In all other cases, it is apparent that the buyers on the long side of the market capture a share of the allocative efficiency much lower than the sellers on the short side. Three major effects are worth noting. First, increasing the market size makes the long side worse off by increasing the competitive pressure among buy-

	$s=1$			$s=5$			$s=50$		
$b=s$	1.000	1.012	1.044	1.001	0.994	0.993	1.000	1.004	1.001
$b=3s$	1.000	0.433	0.591	0.542	0.401	0.428	0.347	0.341	0.348
$b=5s$	1.000	0.302	0.422	0.413	0.257	0.274	0.212	0.205	0.209

Table 4 Average ratios for buyers' to sellers' surplus under S-TT (left), S-GP (center), and A-GP (right).

ers. Second, compared to the S-TT benchmark, optimizing trading strategies under simultaneous order-clearing further deteriorates the long side's position because the sellers on the short side learn to exploit their relative advantage. Finally, switching to asynchronous order-clearing under optimized strategies improves the long side's performance because the random order of arrival for traders weakens the bargaining position of sellers. These three effects closely match the observations in Sect. 3, in so far that better terms of trades for the buyers on the long side correspond to capturing a larger share of the realized surplus.

References

1. Anufriev M, Arifovic J, Ledyard J., Panchenko V (2010), Efficiency of continuous double auctions under individual evolutionary learning with full or limited information. Working Paper, CeNDEF, University of Amsterdam
2. Fano S, LiCalzi M, Pellizzari P (2011), Convergence of outcomes and evolution of strategic behavior in double auctions. Journal of Evolutionary Economics, forthcoming
3. Gjerstad S (2007), The competitive market paradox. Journal of Economic Dynamics and Control 31:1753-1780
4. Koza JR (1992), Genetic Programming: On the Programming of Computers by Means of Natural Selection. The MIT Press, Cambridge, Massachusetts
5. Gode DK, Sunder S (1997), What makes markets allocationally efficient? Quarterly Journal of Economics 112:603-630
6. LiCalzi M, Pellizzari P (2008), Zero-intelligence trading without resampling. In: Schredelseker K, Hauser F (eds) Complexity and Artificial Markets, 3–14. Springer, Berlin Heidelberg New York
7. Pouget S (2007), Adaptive traders and the design of financial markets. The Journal of Finance 52:2835-2863
8. Taylor CR (1995), The long side of the market and the short end of the stick: Bargaining power and price formation in buyers', sellers', and balanced markets. Quarterly Journal of Economics 110:837-855

Part III
Organization design

Effectivity of Multi Criteria Decision-Making in Organisations: Results of an Agent-Based Simulation

Stephan Leitner and Friederike Wall

Abstract This paper analyses how different multi criteria decision-making modes affect organisational performance under the regime of alternative organisational design options. We set up a computational model based on the concept of NK-fitness landscapes. Our results indicate that organisational design and especially departmentalisation with respect to decision interdependencies is a fundamental parameter to affect organisational performance. Conventional wisdom indicates that stability positively affects achieved performance. This could not be proved for all cases. We show that in certain organisational setups discontinuity positively affects performance.

1 Introduction, Research Question and Research Method

Recent developments in management evoke an increasing relevance of multi criteria decision-making (MCDM). For example, stakeholder-orientation requires adjusting the interests of multiple stakeholder groups - instead of focusing mainly on firm value for owners as encapsulated in value based management. Furthermore, the claim for sustainability calls for balancing economic, ecological and social objectives. With these developments firms inevitably pursue multiple and potentially conflicting objectives simultaneously [10].

Stephan Leitner
Alpen-Adria Universitaet Klagenfurt, Department of Controlling and Strategic Management, Universitaetsstrasse 65-67, 9020 Klagenfurt, Austria
e-mail: stephan.leitner@uni-klu.ac.at

Friederike Wall
Alpen-Adria Universitaet Klagenfurt, Department of Controlling and Strategic Management, Universitaetsstrasse 65-67, 9020 Klagenfurt, Austria
e-mail: friederike.wall@uni-klu.ac.at

S. Osinga et al. (eds.), *Emergent Results of Artificial Economics*, Lecture Notes in Economics and Mathematical Systems 652, DOI 10.1007/978-3-642-21108-9_7,

A fundamental question of organizational design is how to coordinate interdependent decisions [5, 11]. Due to multiple objectives, interdependencies become more complex and, hence, coordination gets more demanding [16]. Although there is a lot of research on MCDM, it is rarely investigated in an organisational context. In this paper we seek to contribute to the question whether and how organisational performance achieved with different MCDM modes is affected by organisational design options. In particular, we analyse which coordination modes in combination with which MCDM modes are most effective for certain setups of hierarchical organisations. Additionally, our research considers interacting agents and effects of different natures of interdependencies among decisions. We do not seek to optimise search algorithms. Rather we intend to analyse how performance in hierarchical organisations evolves and how MCDM methods and coordination mechanisms, that are widespread in organisational practice, affect achieved performances.

In order to investigate the research question, a research method that allows mapping organisational design elements, different natures of interdependencies among decisions and interacting agents in multi objective setups is needed. With its incorporated multitude of issues, the research question would lead to intractable dimensions of formal modelling [6, 3]. Additionally, controlling the multitude of issues and disentangling effects of variables under research from other effects would be particularly difficult in an empirical research approach [21]. Agent based simulation, on the contrary, appears to be a powerful method to face this complexity [14].

2 Simulation Model

Our simulation model is based on the NK-Model introduced by Kauffman [9, 8]. The representation of organisations mainly corresponds to the model of Siggelkow and Rivkin [18]. In consideration of earlier work on MCDM in organisations (e.g. [4]), the distinctive feature of the model applied in this paper is the representation of multiple objectives in *hierarchical* organisations. In particular, we map MCDM as adaptive walks on multiple performance landscapes where each landscape represents one objective. Consequently, each managerial decision affects performance of multiple objectives. So, a certain decision may increase performance achieved on the one performance landscape, whereas the achieved performance on another landscape may be reduced.

In order to represent MCDM in hierarchical organisations appropriately, the following major components of the computational model have to be specified more precisely: (1) the model of organisations and options for organisational design, (2) the representation of the performance landscapes, and (3) the mapped methods of MCDM.

2.1 Model of Organisations and Options for Organisational Design

Our organisations are conceptualised as systems of interdependent choices [13]. Instead of making decision in a single-dimensional setup, our firms rather have to search along a multidimensional space for optimal configurations of decisons [15]. Especially, our organisations are in scope of making ten (potentially interdependent) decisions per period, where each decision is of binary nature. The dimensionality of the decision problem and the architecture of the performance landscapes are constant along the observation time. In each period $t \in \{1,...,T\}$ organisations make decisions $n^{i,t} \in N$ with $n^{i,t} \in \{0,1\}$, $i \in \{1,...,10\}$ indexing the number of decisions. Due to the fact that decisons are binary solvable, there exist $2^{|N|}$ differing overall configurations. For each period t the chosen strategy is denoted as vector $S^t = \left(n^{i=1,t},...,n^{i=|N|,t}\right)$. The starting configuration $S^{t=0}$ is selected randomly.

Each managerial decision affects performance of all of the multiple objectives. Especially, in every period t and for each decision $n^{i,t}$ one performance contribution $p_z^{i,t}$ to overall performance P_z^t exists, with $z \in Z$ indexing the number of objectives. For the fact that decisions may intercorrelate with each other, the functioning of each performance contribution $p_z^{i,t}$ may additionally to decision $n^{i,t}$ be affected by a number of decisions $n_k^{j,t}$. So, $p_z^{i,t}$ depends on the single decision $n^{i,t}$ and, eventually, on a set of other decisions $n_k^{j,t}$. These epistatic relations are described by parameter K_z^i which stands for the number of decisions $n_k^{j,t}$ that affect $p_z^{i,t}$ additionally to decision $n^{i,t}$. In consideration of interdependencies among decisions, for each period t and each performance contribution $p_z^{i,t}$ the payoff-function f_z^i randomly draws a value from uniform distrubition $U[0,1]$, i.e.

$$p_z^{i,t} = f_z^i\left(n^{i,t}; n_{k=1}^{j,t},...,n_{k=K_z^i}^{j,t}\right) \tag{1}$$

with $i,j \in \{1,...,|N|\}$, $k \in \left\{1,...,K_z^i\right\}$, $0 \leq p_z^{i,t} \leq 1$ and $i \neq j$. Whenever any of the functional dependend decisions changes, the value for the performance contribution $p_z^{i,t}$ is redrawn from the underlying distribution. In our model, each performance contribution $p_z^{i,t}$ has the same weight with respect to the overall performance per objective. So, overall performance P_z^t results as

$$P_z^t = \frac{1}{|N|}\sum_{i=1}^{|N|} p_z^{i,t} \ . \tag{2}$$

Our organisations consist of headquarters h and departments $d \in D$. Considering prior work on MCDM in organisations (e.g. [4]), the mapping of hierarchical structures appears to be a novelty. We map firms where the decentral structure is organised in three departments. Especially, with respect to the ten-dimensional decision problem, one department is in scope of four decisions and two departments are in scope of three decisions. The set of decisions a certain department is in scope of, is denoted as N^{own^d} while the other departments' decisions are denoted as N^{res^d}.

Decentral units aim at maximizing their individual utility functions (cf. section 2.3) via incremental changes. The decision maker's selfishness goes along with economic literature [7], while assuming stepwise improvements is consistent with the literature on organisational learning [2] and prior modelling approaches (e.g. [17]). Due to bounded rationality [19], the number of alternative configurations envisioned by departments is limited. In every period, departments discover two alternative configurations of decision for the own area of responsibility N^{own^d} that differ in one respectively two decisions from the status quo. Along with the status quo, departments evaluate three alternative configurations of decisions. Each department is in charge of providing the headquarters with two proposals for the composition of N^{own^d} for the following period. Hence, departments rank two of the alternative configurations of decisions under evaluation in respect of which strategy is perceived to provide the highest improvement in personal utility. We denote the proposed configurations as vectors $A_r^{own^d,t}$ with $r = \{1,2\}$ standing for the assigned rank.

We map various options for organisational design. Especially, we take into account options in (1) the coordination mode and (2) the incentive scheme as organisational design elements. Furthermore, we consider the structuring of departments with respect to cross-unit interdependencies among decisions (cf. section 2.2) as one further design-option.

One main determinant of organisational design is the choice of coordination mode. This choice specifies how overall configurations of decisions for future periods is build and, hence, might affect performances and speed of performance improvement crucially. Escpecially, our model considers two different coordination modes: (1) a central mode of coordination, and (2) fully decentralised coordination. In case of the *(1) central mode*, headquarters evaluate the ranked alternative configurations of decisions with regard to overall performance (cf. equation 5). Especially, headquarters evaluate concatenations of proposed configurations of departmental decisions $A_r^{own^d,t}$ and residual decisions $D^{res^d,t}$ according to the status quo and choose that proposal that promises the highest overall performance to be the configuration for the following period. In case of the *(2) decentral coordination mode*, departments decide and act autonomously in their area of responsibility N^{own^d}. So, the strategy for period $t+1$ results as concatenation of those proposed partial configurations $A_1^{own^d,t}$ ranked on the first place.

Furthermore, we consider the parameterisation of the reward system as option of organisational design. The departments' compensation is based on a linear incentive scheme where the variable salary components depend on performance out of departmental decisions N^{own^d} and performance out of residual decision N^{res^d} weighted with r^{own^d} and r^{res^d} respectively. Of course, the reward system is reflected in the departments' utility functions and, hence, crucially affects the ranking of alternative configurations for departmental decisions. In the simulation we apply incentive schemes that put more weight on departmental than on residual performance. So, we set $r^{own^d} = 1$ and $r^{res^d} = 0.5$.

2.2 The Representation of the Performance Landscapes

Organisational complexity is critically shaped by interdependencies among decisions [20]. There are several factors influencing the intensity of interdependencies. On the one hand, interdependencies depend on the decision problems organisations face. On the other hand, firms deal with complexity by building departments with respect to cross-unit interdependencies. This managerial task might be crucially complicated in case of multiple objectives.

As outlined in section 2.1 we use parameter K_z^i in order to give a measure for the number of interdependent decisions. The higher the degree of interdependencies is, the more rugged appears the corresponding performance landscape to be. Higher K-values increase the number of status quo traps, i.e. local maxima where no further improvement in the neighbourhood can be observed [4]. Hence, once such a status quo trap is reached, the strategy for the remaining observation period is likely to be steady.

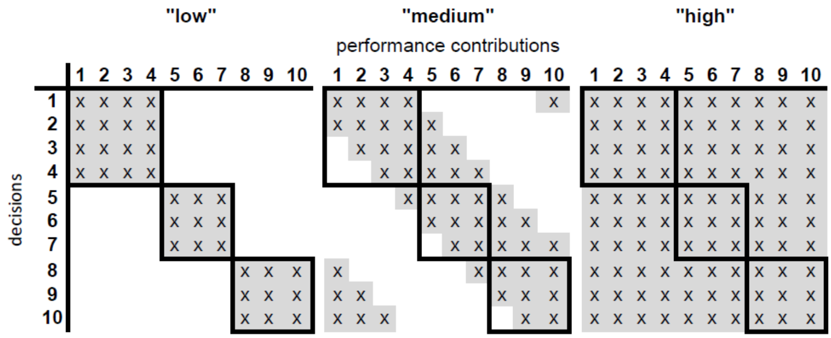

Fig. 1 Interdependence matrices

In order to describe different natures of functional dependencies among decisions we use matrices M (cf. figure 1). The horizontal axis represents performance contributions (cf. equation 1) while the set of decisions is entered on the vertical axis. Indicated by ’x’ along the main diagonale, we map decisions n^i and corresponding performance contributions p_g^i to be interdependent in all cases. Additionally, an ’x’ in cell $m_{i,j}$ with $i, j \in \{1, ..., |N|\}$ and $i \neq j$ indicates that decision n^i affects performance contribution p_g^j.

Our investigation is limited to three exemplary natures of interdependencies among decisions. With reference to modular setups as investigated by Rivkin and Siggelkow [17], in case of nature of interdependencies *(1) low* there is no cross-unit interdependence but decisions within a department are fully interdependent (cf. figure 1, panel ’low’). So, for each department K-values are constant along departmental decisions, i.e. $K_z^i = \left|N^{own^d}\right| - 1$. Nature of interdependencies *(2) medium* refers to small world networks and, hence, is characterised by a high level of clustering [22]. Especially, each department's decisions are partly independent from other departmental decisions but they also partly interact with other departments' decisions (cf. figure 1, panel 'medium'). K-values are constantly 4 along all decisions. We furthermore map a setup with *(3) full interdependence* where all decisions and all performance contributions are reciprocally dependend. So, $K_z^i = 9$ for all decisions (cf. figure 1, panel 'high').

2.3 Methods of Multi Criteria Decision Making

This paper analyses effectivity of different MCDM modes. Especially, we map three MCDM-methods: (1) equal weighting, (2) long-run schism and (3) short-run schism.

In case of *(1) equal weighting*, organisations assign the same importance to all objectives. So, if organisations apply the method of equal weighting, no preferences for objectives need to be articulated. In consideration of the incentive scheme (as outlined in section 2.1) the departments' utility function results as:

$$U_d^t = \sum_{z=1}^{|Z|} \left(r^{own^d} \frac{\sum_{i \in N^{own^d}} p_z^{i,t}}{|N^{own^d}|} + r^{res^d} \frac{\sum_{i \in N^{res^d}} p_z^{i,t}}{|N^{res^d}|} \right) . \tag{3}$$

The methods of *(2) long-run and (3) short-run schism* correspond to the concept of temporal separation [2]. Contrary to equal weighting, in schism approaches we map decision-makers to pursue only a subset of objectives Z at the same time without considering effects for future periods. Furthermore, these approaches are characterised by a high level of volatility in preferences. Changing preferences indicate that these methods of MCDM are especially applied in turbulent environments that often are characterised by changing stakeholder interests [1]. This turbulence may lead managers to change priorities of various stakeholder claims over time [12]. Especially, we map two different degrees of volatility, i.e. (2) long-run and (3) short-run schism. In case of (2) long-run schism, preferences change every ten periods, while in case of (3) short-run schism preferences are reversed in every single period.

In order to operationalise the methods of (2) long-run and (3) short-run schism, we introduce a weighting factor $w_z^t \in \{0,1\}$. In this simulation study weighting factors are exogenously given and they represent preferences for one objective from the headquarters perspective. Of course, these weigthting factors are also reflected in the decentral units' compensation system. So, for MCDM approaches (2) and (3) the departments' utility function results as

$$U_d^t = \sum_{z=1}^{|Z|} \left[w_z^t \left(r^{own^d} \frac{\sum_{i \in N^{own^d}} p_z^{i,t}}{|N^{own^d}|} + r^{res^d} \frac{\sum_{i \in N^{res^d}} p_z^{i,t}}{|N^{res^d}|} \right) \right] \tag{4}$$

where in case of (2) short-run schism $\forall w_z^t : w_z^t \neq w_y^t$ and $w_z^{t+1} \neq w_z^t$ and in case of (3) long-run schism $\forall w_z^t : w_z^t \neq w_y^t$ and $w_z^{t+10} \neq w_z^t$ with $t \in \{1,...,T\}; y,z \in Z$ and $y \neq z$

While departments aim at maximising their own utility function, the headquarters' objective is to maximise overall performance. The headquarters' utility function results as:

$$U_h^t = \begin{cases} \frac{1}{|Z|} \sum_{z=1}^{|Z|} P_z^t & \text{if (1) equal weighting} \\ \frac{1}{|Z|} \sum_{z=1}^{|Z|} \left(w_z^t P_z^t \right) & \text{if schism-approaches (2) and (3) .} \end{cases} \quad (5)$$

So, in case of equal weighting, from the headquarters' view all objectives contribute to overall performance equally, while in schism approaches preferences for objectives change during the observation time. The interval of reversal is given by the choice of MCDM-method (cf. equation 4).

3 Results

The present simulation study examines effectivity of MCDM methods in an organisational context. For this reason we map firms that are in charge of making ten decisions and pursuing two objectives at a time. The organisational lifetime and, hence, the observation time is 100 decision cycles. So, the constant key parameters for the organisational context are: $|N| = 10, |Z| = 2$ and $T = 100$. Optional design choices correspond to elaborations in section 2.1. We base our results on 450 different performance landscapes per objective each with 20 adaptive walks. Accordingly, each scenario is based on 9.000 adaptive walks.

We give two different measures for performance. First, as a snapshot of achieved performance at the end of the observation period, we report final performances $P_z^{t=100}$ per objective (cf. equation 2). Second, we report average performance P_z^{avg} over the whole organisational lifetime. P_z^{avg} also captures the speed of performance improvement per objective. Average performance is calculated as

$$P_z^{avg} = \frac{1}{9.000\,T} \sum_{s=1}^{9.000} \sum_{t=1}^{T} P_z^{t,s} \quad (6)$$

with index s standing for the number of simulation runs. As in case of the headquarters utility (cf. equation 5) we map all objectives to contribute to overall performance equally. So, overall final performance results as $P_{all}^{t=100} = 1/|Z| \sum_{z=1}^{|Z|} P_z^{t=100}$ and overall average performance results as $P_{all}^{avg} = 1/|Z| \sum_{z=1}^{|Z|} P_z^{avg}$.

We present results stepwise. First, we analyse evolution of performance separately for each MDCM mode. This is to answer the question which coordination mode appears to be benefical with respect to achieved performances. Second, in order to analyse effectivity of MDCM methods in case of given coordination mode, we analyse differences in achieved performances across MDCM approaches. Finally, implications for the choice of organisational design elements in hierarchical organisations can be derived from the presented results.

Table 1 Equal weighting

interdependencies obj 1 / obj 2	final performances $P_1^{t=100}$	$P_2^{t=100}$	$P_{all}^{t=100}$	average performances P_1^{avg}	P_2^{avg}	P_{all}^{avg}
Panel A: coordination mode: central						
low/low	0.8984	0.8994	0.8989	0.8941	0.8949	0.8945
low/medium	0.8777	0.8737	0.8757	0.8734	0.8694	0.8714
medium/medium	0.8515	0.8475	0.8495	0.8479	0.8437	0.8458
low/high	0.8515	0.8508	0.8512	0.8478	0.8474	0.8476
medium/high	0.8215	0.8334	0.8274	0.8186	0.8303	0.8245
high/high	0.8089	0.8084	0.8087	0.8070	0.8063	0.8066
Panel B: coordination mode: decentral						
low/low	0.8987	0.8975	0.8981	0.8957	0.8945	0.8951
low/medium	0.9004	0.8705	0.8855	0.8967	0.8638	0.8803
medium/medium	0.8599	0.8596	0.8598	0.8530	0.8525	0.8527
low/high	0.8961	0.8415	0.8688	0.8909	0.8310	0.8609
medium/high	0.8457	0.8323	0.8390	0.8369	0.8198	0.8284
high/high	0.8121	0.8153	0.8137	0.7989	0.8023	0.8006

Incentivisation: $r_z^{own^d} = 1$ and $r_z^{res^d} = 0.5$. Results are based on 450 landscapes each with 20 adaptive walks. obj = objective, confidence intervals vary from 0.002 to 0.003 on the 99.9% level.

3.1 Equal Weighting

Not surprisingly, we find that increasing complexity of interdependencies among decisions leads to decreasing final and average performances in the central as well as in the decentral coordination mode (cf. table 1). Furthermore, if organisations pursue objectives with the same nature of interdependencies, same final and average performances are achieved.

In case of objectives with different natures of interdependencies, results suggest overall performances to be significantly sensitive to the choice of coordination mode. The decentral mode of coordination appears to be superior to the central coordination mode from the final and average performance perspective.

On the single objective perspective and for the case of two objectives with different natures of complexity, results indicate that in most scenarios the spread between the two objectives is higher in the decentral coordination mode. In the central coordination mode, on the contrary, intervening headquarters appear to adjust the performance levels of the two objectives. So, in the central coordination mode results show no significant difference between the two single objective performances in most cases.

Table 2 Long run schism

interdependencies obj 1 / obj 2	final performances $P_1^{t=100}$	$P_2^{t=100}$	$P_{all}^{t=100}$	average performances P_1^{avg}	P_2^{avg}	P_{all}^{avg}
Panel A: coordination mode: central						
low/low	0.9795	0.6903	0.8349	0.8262	0.8335	0.8299
low/medium	0.9831	0.6613	0.8222	0.8527	0.7859	0.8193
medium/medium	0.9289	0.7968	0.8178	0.7990	0.8039	0.8014
low/high	0.9841	0.6479	0.8160	0.8777	0.7516	0.8147
medium/high	0.9358	0.6743	0.8050	0.8182	0.7617	0.7899
high/high	0.8748	0.6926	0.7837	0.7652	0.7668	0.7660
Panel B: coordination mode: decentral						
low/low	0.9826	0.6864	0.8345	0.8270	0.8266	0.8268
low/medium	0.9812	0.6478	0.8145	0.8279	0.7596	0.7938
medium/medium	0.9206	0.6666	0.7936	0.7673	0.7686	0.7680
low/high	0.9841	0.6455	0.8148	0.8345	0.7080	0.7712
medium/high	0.9231	0.6566	0.7899	0.7726	0.7115	0.7420
high/high	0.8261	0.6711	0.7486	0.7148	0.7184	0.7166

Incentivisation: $r_z^{own^d} = 1$ and $r_z^{res^d} = 0.5$. Results are based on 450 landscapes each with 20 adaptive walks. obj = objective, confidence intervals vary from 0.001 to 0.004 on the 99.9% level.

3.2 Schism Approaches

As in case of equal weighting, schism approaches lead to decreasing final and average performance with increasing complexity of interdependencies. Furthermore, in case of two objectives with the same nature of interdependencies the same levels of performances can be achieved on the overall as well as the single objective perspective and for both coordination modes.

In case of *long-run schism* (cf. table 2), in most scenarios results suggest the central coordination mode to be significantly superior to decentral coordination from the overall final as well as the overall average performance perspective. With increasing complexity of interdependencies among decisions, differences in achieved overall final and overall average performances increase too. Of course, this is also reflected in achieved performances on the single objective level.

Also in case of *short-run schism* (cf. table 3), the central coordination mode leads to higher overall final and overall average performances. With respect to performance, results suggest that the higher the complexity of interdependencies among decisions is, the more advantageous the central coordination mode appears to be.

3.3 Evaluation Across Multi Criteria Decision Making Methods

We find that increasing complexity of interdependencies among decisions leads to decreasing performances for all decision making modes and all coordination modes.

Table 3 Short run schism

interdependencies obj 1 / obj 2	final performances $P_1^{t=100}$	$P_2^{t=100}$	$P_{all}^{t=100}$	average performances P_1^{avg}	P_2^{avg}	P_{all}^{avg}
Panel A: coordination mode: central						
low/low	0.8878	0.8093	0.8486	0.8434	0.8429	0.8432
low/medium	0.9046	0.7465	0.8256	0.8624	0.7897	0.8261
medium/medium	0.8579	0.7790	0.8185	0.8062	0.8081	0.8072
low/high	0.9150	0.6954	0.8052	0.8751	0.7515	0.8133
medium/high	0.8664	0.7373	0.8018	0.8196	0.7683	0.7940
high/high	0.8382	0.7711	0.8047	0.7821	0.7823	0.7822
Panel B: coordination mode: decentral						
low/low	0.8859	0.7638	0.8248	0.8210	0.8246	0.8228
low/medium	0.8842	0.7055	0.7949	0.8177	0.7402	0.7790
medium/medium	0.7748	0.6979	0.7364	0.7311	0.7317	0.7314
low/high	0.8865	0.6766	0.7816	0.8164	0.6875	0.7520
medium/high	0.7750	0.6672	0.7211	0.7308	0.6788	0.7048
high/high	0.7133	0.6772	0.6953	0.6841	0.6846	0.6843

Incentivisation: $r_z^{own^d} = 1$ and $r_z^{res^d} = 0.5$. Results are based on 450 landscapes each with 20 adaptive walks. obj = objective, confidence intervals vary from 0.002 to 0.005 on the 99.9% level.

Hence, organisational design and especially departmentalisation with respect to cross-unit interdependencies among decisions is a fundamental parameter to affect organisational performance.

Results indicate that in the central as well as in the decentral coordination mode, for all combinations of different natures of interdependencies, equal weighting leads to the highest overall performances. In case of schism, intuition suggests that continuity in preferences affects performance positively. This could not be proved for the central coordination mode. Surprisingly, we find that with central coordination short-run schism leads to higher overall performances, while in the decentral coordination mode long-run schism appears to be superior to short-run schism.

Furthermore, differences between minimum and maximum of P_{all}^{avg} per MCDM method and coordination mode indicate that organisational performance in case of short-run and long-run schism is more sensitive to the choice of coordination mode than it is in case of equal weighting. While, with respect to the central coordination mode, in case of equal weighting differences solely increase slightly with decentral coordination, in case of short-run and long-run schism and decentral coordination differences approximately double.

4 Conclusion

Our results indicate that achieved performances subtly depend on organisational design elements. For the fact that, not at least due to increasing stakeholder orientation, multiple objectives seem to be immanent in organisational practice, advanced knowledge of effects of organisational design on performance is an important factor of success for management.

Results might give some new insights into organisational design. We show that equal weighting is superior to short-run and long-run schism with respect to organisational performance. Conventionally, in case of schism, continuitiy is benefical to performance. We show that this is not necessarily correct for all scenarios. In fact, in some setups discontinuity affects performance positively.

References

1. Can M. Alpaslan, Sandy E. Green, and Ian I. Mitroff. Corporate governance in the context of crisis: Towards a stakeholder theory of crisis management. *Journal of Contingencies and Crisis Management*, 17(1):38–49, 2009.
2. Richard Michael Cyert and James G. March. *A behavioral theory of the firm*. Blackwell, Oxford, 2. edition, 2005. J326 NPO-N-59.
3. Jason P. Davis, Kathleen M. Eisenhardt, and Christopher B. Bingham. Developing theory through simulation methods. *Academy of Management Review*, 32(2):480–499, 2007.
4. Sendil K. Ethiraj and Daniel Levinthal. Hoping for a to z while rewarding only a: Complex organizations and multiple goals. *Organization Science*, 20(1):4–21, 2009.
5. Jay Galbraith. *Designing complex organisations*. Addison-Wesley series on organization development. Addison-Wesley, Reading, Mass., 1973.
6. Nigel Gilbert. Simulation: an emergent perspective, 1995.
7. Michael C. Jensen and William H. Meckling. The nature of man. *Journal of Applied Corporate Finance*, 7(2):4–19, 1994.
8. S. A. Kauffman and S. Levin. Towards a General Theory of Adaptive Walks on Rugged Landscapes. *Journal of Theoretical Biology*, 128:11–45, 1987.
9. Stuart Kauffman. *The origins of order. Self-organization and selection in evolution*. Oxford Univ. Pr., New York, NY, 2010.
10. Marshall W. Meyer. *Rethinking performance measurement. Beyond the balanced scorecard*. Cambridge Univ. Press, Cambridge, 2002.
11. Henry Mintzberg. *The structuring of organizations. A synthesis of the research*. Prentice-Hall Internat., London, 1979.
12. Ronald K. Mitchell, Bradley R. Agle, and Donna J. Wood. Toward a theory of stakeholder identification and salience: defining the principle of who and what really counts. *Academy of Management Review*, 22(4):853–886, 1997.
13. Michael E. Porter. What is strategy? *Harvard Business Review*, 74(6):61–78, 1996.
14. Mitchel Resnick. *Turtles, termites and traffic jams. Explorations in massively parallel microworlds*. Complex adaptive systems. MIT Press, Cambridge, Mass., 5. print. edition, 1999.
15. Jan W. Rivkin. Imitation of complex strategies. *Management Science*, 46(6):824, 2000.
16. Jan W. Rivkin and Nicolaj Siggelkow. Balancing search and stability: Interdependencies among elements of organizational design. *Management Science*, 49(3):290–311, 2003.
17. Jan W. Rivkin and Nicolaj Siggelkow. Patterned interactions in complex systems: Implications for exploration. *Management Science*, 53(7):1068–1085, 2007.
18. Nicolaj Siggelkow and Jan W. Rivkin. Speed and Search: Designing Organizations for Turbulence and Complexity. *Organization Science*, 16(2):101–122, March-April 2005.
19. Herbert A. Simon. A behavioral model of rational choice. *Quarterly Journal of Economics*, 69(1):99–118, 1955.
20. Herbert A. Simon. The architecture of complexity. *Proceedings of the American Philosophical Society*, 106(6):467–482, 1962.
21. Geoffrey B. Sprinkle. Perspectives on experimental research in managerial accounting. *Accounting, Organizations and Society*, 28(2-3):287–318, 2003.

22. Brian Uzzi, Luis Amaral, and Felix Tschochas-Reed. Small-world networks and management science research: a review. *Europen Management Review*, 5:77–91, 2007.

The Problem of Emergency Department Overcrowding: Agent-Based Simulation and Test by Questionnaire

Roger A. McCain, Richard Hamilton and Frank Linnehan

Abstract Overcrowding of hospital emergency departments is a recognized public health problem in the United States. This paper develops and simulates a noncooperative game model of emergency department overcrowding. This hypothesis naturally takes the form of an agent-based computer simulation. The predictions of the model are validated by a questionnaire administered to patients with experience in an urban hospital emergency department. These results favor the Nash equilibrium hypothesis, implying that emergency department crowding is the equilibrium of the current state of the U.S. health care system. It is hoped that this paper contributes by 1) improving our understanding of the problem of emergency department overcrowding, and 2) providing an example of the potentiality for questionnaire studies as evidence for agent-based simulation studies.

1 The Problem of Medical Emergency Department Overcrowding

Overcrowding of hospital emergency departments is a recognized problem in the U. S. A. Data show that between 1994 and 2004, emergency department visits have increased by 26%, while the number of emergency departments has decreased by 9% and hospitals have closed 198,000 beds [7]. On its face, this might explain the

Roger A. McCain PhD
Lebow College of Business, Drexel University, Philadelphia, PA, 19104, USA
e-mail: mccainra@drexel.edu

Richard Hamilton M.D.
Department of Emergency Medicine, Drexel University College of Medicine, 245 N. 15th Street, Mailstop 1011, Phila. PA 19102, USA e-mail: richard.hamilton@drexelmed.edu

Frank Linnehan PhD
Lebow College of Business, Drexel University, Philadelphia, PA, 19104, USA
e-mail: linnehf@drexel.edu

S. Osinga et al. (eds.), *Emergent Results of Artificial Economics*, Lecture Notes in Economics and Mathematical Systems 652, DOI 10.1007/978-3-642-21108-9_8,

observed overcrowding in Emergency Departments. As a symptom of this overcrowding, wait times are increasing, diminishing quality and speed of care (Wilper et. al [14]; Derlet and Richards [5]; Burt and McCaig [2])

However, the problem may not be as simple as diminished capacity and increasing demand. While ED visits have increased and total number of EDs have diminished, the bed and treatment capacity of many EDs has actually increased. Between 2003 and 2004, 16.1 percent of all hospitals recently expanded their ED physical space, and approximately one-third of others plan to do so within the next 2 years. Over all, about 43.2 percent of all EDs had either recently expanded or had planned to do so. While it may be counterintuitive, the EDs most likely to be overcrowded were the ones which had recently expanded. Their characteristics were higher volume; located in a hospital classified as proprietary, voluntary, or nonprofit; affiliated with medical schools; experienced ambulance diversion; and finally demonstrated longer average visit durations [2]. One California study demonstrated that ED bed capacity increased between 1990 and 2001 in roughly 40% of EDs and that over time, the number of ED beds per 100,000 people actually increased by 5 to 10% [12].

Although ED overcrowding is a multifactorial problem, insured patients with less urgent conditions may be driving the volume increase [4]. This may account for why some hospitals have found it cost effective to expand ED capacity and others have not, perhaps depending upon the insurance status of the population they serve [12]. Additionally, across the health care system, patients report increasing difficulty in accessing office based care for both routine and urgent health care needs, driving patients to seek care in EDs [3].

Patients seek healthcare in the Emergency Department for a variety of reasons, but it suffices to say that a portion have an acute emergency and a portion seeks care because of lack of an acceptable alternative. In other words, a large portion of patients turn to Emergency Departments because they are diverted from other forms of medical care that is either at capacity or unavailable to them. Examples of this include but are not limited to: primary care physician not available within the desired time, specialist not available within the desired time, patients with no health care insurance or money to pay a specialist or primary care physician, or alternative health care source is inconvenient. In addition, it is a common practice for physicians themselves to use the ED as alternative access to the hospital for admission, lab work, evaluation, or procedures. Even the hospital itself diverts patients to the ED when the admissions are held in the ED until a bed becomes available. One final source of patients is those that are diverted from other closed EDs in the area. The capacity for handling patients in any given ED is limited by facility, number of physicians, and number of nurses, number of support staff and services.

One generally accepted solution to ED overcrowding and congestion is to increase capacity. But does this follow from an accurate diagnosis of the problem? Drawing on noncooperative game theory, we will argue that ED overcrowding is the equilibrium state of the current health care system. If this hypothesis is correct, then increasing capacity will merely reproduce the crowding problem on a larger scale.

2 Small-Scale Game-Theoretic Models

As a preliminary model consider Table 1. This is the simplest nontrivial game model of the problem, a 2x2 game in which the two players are two potential patients. Waiting time is one negative dimension of a medical care service, but some people might consider waiting 12 hours to be seen in an ED a worse payoff than waiting 24 hours to be seen in their PCPs office. In such a case the ED will be the best response only if it is uncrowded. We will express the payoffs in terms of a patient satisfaction scale, where 1 is the least satisfied and 10 is the most satisfied. In this example, we are using the following assumptions: satisfaction with quiet ED = satisfaction with quiet PCP office >satisfaction with crowded ED = satisfaction with crowded PCP office. Either service is crowded if both patients choose it.

Table 1 The Patients'Coordination Problem

		Patient 2[a]	
		Go to ED	Wait to see PCP
Patient 1	Go to ED	1,1	8, 8
	Wait to see PCP	8,8	1,1

[a] (First number is payoff to Patient 1 and second number is payoff to patient 2.)

In this case, no common strategy “Go to the ED” or “Wait to see PCP” is an equilibrium strategy combination. Whenever Patient 1 chooses to wait and see his PCP, Patient 2s best response is to choose take advantage of the ER, and conversely. In short, this is an *anticoordination game.*

While the two-person game captures an important aspect of the problem, it will be worthwhile to make our model a little more realistic by allowing for more than two patients, for different health states among patients, and for errors and uncertainty. As a first step the model is extended to N agents, as a numerical example. (Compare [10] pp. 227-230). Suppose we have ten potential patients and that all are alike (this is a counterfactual simplifying assumption) so that their benefits from medical care in the emergency department differ only to extent that their waiting times differ. Waiting time is roughly proportional to the position one has in line, but the position in line cannot be predicted there is some probability that I will be first, some that I will be last, depending on when I and others arrive. Suppose that the person first in line derives a satisfaction of 10 from the medical services of the ED, that being one place further back in the line reduces this satisfaction by two, and that the alternative to the ED provides a satisfaction level of five. We then have Table 2.

The number of ED patients will determine the payoffs for all players. The second column gives the benefit as it depends on the position in line. The third and fourth show average line position and average net benefit as they decline with the number choosing the ED. What we can see is that, if less than six patients choose the ED, then the ED is an individual's best response to the decisions of others; while if more than six choose the ED, the best response is to choose the alternative. Thus, for a Nash equilibrium to occur, just six choose the ED while the other four choose their

Table 2 The Patients Coordination Game with a Population of 10

Number of ED Patients and Position in line	net benefit	average position in line	average net benefit	benefit of alternative
1	10	1	10	5
2	8	1.5	9	5
3	6	2	8	5
4	4	2.5	7	5
5	2	3	6	5
6	0	3.5	5	5
7	-2	4	4	5
8	-4	4.5	3	5
9	-6	5	2	5
10	-8	5.5	1	5

alternative. This very simple example illustrates three conclusions which will extend (with a more formal treatment) to a realistic general case:

1. Just as in the two-person anticoordination game, equilibrium requires some agents to choose different strategies even if they themselves do not differ.
2. When the strategies are modes of service, the number choosing the different services in equilibrium will be such that the different services yield the same benefits, in expected value terms.
3. The equilibrium is not efficient, in general. In this example, the fourth person using the ED adds only 4 units of net benefits, while the alternative service yields 5. Thus, the efficient patient population for the ED in the example is 3. Queuing for service leads to inefficient congestion. We should add that queuing is difficult to avoid for a service, like emergency medical care, the mission of which is inconsistent with other forms of rationing that (while more efficient) would require that some people in need of medical care be turned away. However, it does mean that the emergency department is likely to be a relatively inefficient means of providing routine medical care, over and above its mission.

Nevertheless, we cannot expect that expanding the capacity of the emergency department will eliminate the overcrowding and inefficiency. Just the contrary: an increase in the capacity of the emergency department (so that there are two queues, for example, or that the queue moves twice as fast) will be offset by an increase in the number using the ED, to the point that the expected benefit of ED care is again the same as the expected benefit of the alternative service, i.e. 5 in Table 2. This is likely to mean increased waste, in that a larger proportion of the population chooses the less efficient mode of service. (This assumes that alternative medical care is not rationed by queuing, and so is less inefficient).

3 Agent-Based Computer Simulation

To further extend the model and allow for 1) much larger numbers of potential patients, 2) heterogeneity of health states, experience, and expectation, 3) boundedly rational learning, and 4) initialization effects, dynamic adjustment and transients, we undertook agent-based computer simulation [6], [13]) The formal structure of this larger model is expressed as computer program code. [1]

For these simulations, the agents are potential patients, while the ED is not a player in the game but a mechanism that mediates the interaction of the agents. It is assumed that (at each iteration of the simulation) agents are randomly sorted into four health states. The largest group, in state zero, have no health concerns that would lead them to consider seeking medical care either from the ED or the alternative mode. Agents in states 1 and 3 have health concerns such that treatment in the ED offers a higher benefit than the alternative in the absence of congestion. Recall that the difference in benefits may reflect the inconvenience of the alternative as well as a condition that requires emergency treatment. For agents in state 2, there is a health concern such that treatment through the alternative mode offers higher expected benefit than treatment by the ED, even in the absence of congestion. For all three types, experience of treatment in the ED depends on congestion and a random term. Each agent chooses the mode that offers the greater expectation of benefits. The results can best be illustrated by reviewing some representative simulations. Figure 1 shows the experience of agents of type 1 who choose the ED and those who choose the alternative over 500 successive iterations of the simulation. Expected benefit is at the top, with realized benefit below it converging toward the benefit of the alternative medical service. In this simulation agents of type 1 receive a benefit that averages 3 times as great from ED medical care than from the alternative, in the absence of crowding. However, as we see, crowding reduces the benefit for agents of type 1 to approximate equality with that from the alternative mode, as in simpler Nash equilibrium models. We note moreover that the expected benefit from ED care reported by agents who have chosen the ED remains above the true value. This occurs because of the heterogeneity of expectations. All of these pseudodata are reported only for agents who chose the ED. On the whole, the agents who choose the ED are the agents who have relatively favorable expectations for it, while those in the original population with less favorable expectations do not choose the ED and thus are not polled. Accordingly, we would expect to see their reported expectations biased upward, as here. The ability to allow for heterogeneity is a major advantage of agent-based computer simulation, and as we will see, can be important in interpreting empirical data.

In the simulations, there are 10,000 agents, and at each iteration they are sorted into health states such that about 60 % will seek health care from one source or another. Each agent makes the decision based on an expected benefit variable, with

[1] The compiled software is available in a more extensive working paper available from the authors, where the computer code is included as Appendix 1. The current version runs only on a Macintosh computer under System X.

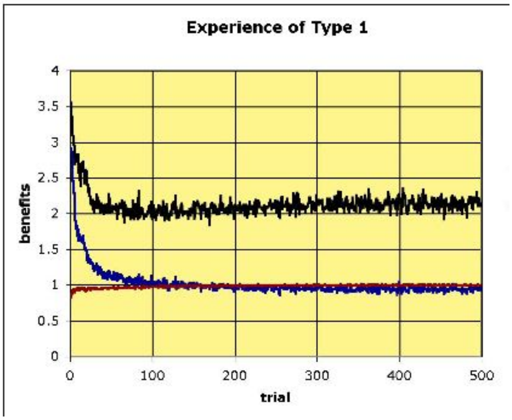

Fig. 1 Average Experience in a Single Simulation

a normally distributed pseudorandom error. (Qualification: For technical reasons having to do with the trial-and-error learning process, at least 5 % of those agents who seek health care choose the Emergency Department regardless of their expectations.) Each chooses the alternative that offers the better experience, on the basis of this estimate. After all these decisions have been registered, the congestion of the emergency department is computed by comparing its capacity parameter to the number of users. Congestion exists only when the number of users is greater than capacity; thus, for these simulations, capacity refers not to the maximum that the emergency department can do but to the design capacity to give optimal medical results. The experiences for all agents who choose the ED are then computed on the basis of congestion together with the parameters of their specific health states, with a pseudorandom variate to capture the uncertainty inherent in medical treatment. Numerical indicators of experience are roughly calibrated to the five-point Likert scale used in the questionnaire survey reported below. Expected benefits are then updated. The updating formula is the Koyck [8] lag formula, $E_t = \alpha X_{t-1} + (1-\alpha)E_{t-1}$. For the simulations reported $\alpha = 1/2$. Note that agents learn only from their own experience and not from imitation or communication. In order to achieve a correlated equilibrium in an anticoordination game, private signals of some sort are necessary; (McCain [11] Ch. 5; Aumann [1], Luce and Raiffa [9]) in this model the private signals derive from individual experience.

For this study 18 distinct simulations were recorded. Two simulations were run using each of 9 random number seeds. For one series of 9 simulations the capacity of the Emergency Department was set at 500, while for others it was set at 1000. The simulations were run for 200 iterations. Figure 2 shows the recorded Emergency Department congestion for the 18 simulations run. Congestion is indicated by an index that varies between zero and one, namely 1-(capacity/users), whenever the number of users is greater than the parametric capacity, and 0 otherwise. This congestion measure can be understood as a measure of the percent deterioration of the perceived quality of the experience from the point of view of the user. All simulations are initiated with user expectations such that there will be little or no congestion at the beginning. In Figure 2, simulations with capacity set at 500 are represented by grayed curves while those with capacity set at 1000 are shown in black.

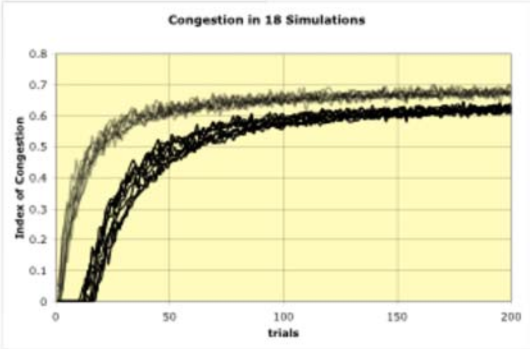

Fig. 2 Congestion in 18 Simulations

We see that in simulations of both types, there is an emerging tendency to congestion that stabilizes after about 100 iterations, and that represents 60% or more deterioration in the quality of user experience. (This 60% should not be taken as a quantitative prediction for real Emergency Departments as it may be sensitive to simplifying assumptions in the simulation program, but we can extract the qualitative prediction that there will be significant and important deterioration of patient experience). With the larger capacity of the Emergency Department, congestion is somewhat postponed and the stable level of congestion is somewhat less, but the existence of substantial overcrowding at equilibrium is a common result of all simulations.

Figures 3 and 4 show the reported quality of experience of agents of type 1 and 3, showing the simulations with capacity at 500 in gray and those with capacity at 1000 in black. These confirm that the experience of ED users of both types deteriorated in a qualitatively similar way, with quantitative differences that reflect the difference in the experience from the alternative mode of medical care. Reported experience for agents of type 2, whose emergency room experience would be less favorable than the alternative even in the absence of congestion, are not shown. For those agents, whose ED use converged to a minimum in each simulation, congestion did not influence their experience nor their decisions.

A key prediction of the small-N game theoretic models can be stated in this way: since the choice of modes of service is an anticoordination game, equilibrium requires that identical agents make different decisions. This self-sorting determines the noncooperative equilibrium of the game. This is why the solution in pure strategies (i.e. in which the individual agents do not randomize their decisions independently) requires some private information, information in this case derived from individual past experience. This is also a property of the emergent stable states in all simulations. Consider Figure 5, which shows the total number of users of the ED in each iteration of each of the 18 simulations, once again coding the simulations with

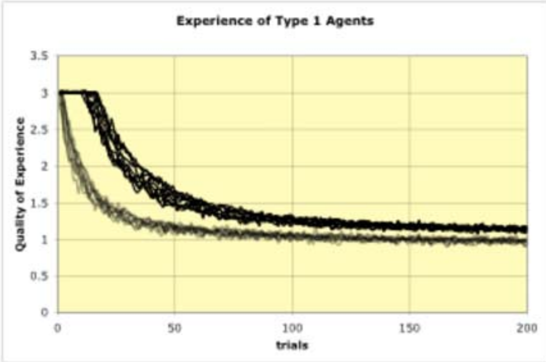

Fig. 3 The Reported Experience of Agents of Type 1

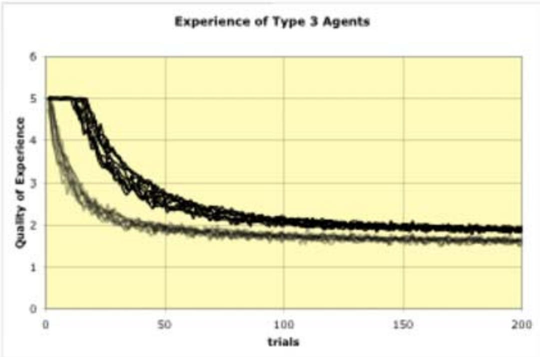

Fig. 4 The Reported Experience of Agents of Type 3

capacity = 500 in gray and those with capacity=1000 in black. Recall that, taking agents of types 1 and 3 together, there are on each iteration about 6000 agents who would benefit from Emergency Department treatment, relative to their alternative, in the absence of congestion. Thus, from 3200 to 4500 of the agents choose not to use the Emergency Department simply because their past experience there has been worse than average. Notice that doubling the capacity does not quite double the number of users of the Emergency Department, and it is this less than proportionate increase in use that follows the increase in capacity that permits the slight improvement of congestion and patient experience.

These more extensive examinations of the simulations support the following conclusions: 1) as in the small-N models, an equilibrium or stable state corresponds to congestion sufficient to reduce the benefits of users of the ED to approximate equality with the benefits from alternative service; 2) in these simulations with a large but finite number of agents and boundedly rational learning, the approximation to the alternative benefit may not be perfect and may vary somewhat with parameters and initialization, so that 3) an expansion of ED capacity can result in some slight improvement in congestion and patient experience, despite very substantial deterioration of the experience due to congestion, and finally 4) these results are uniform

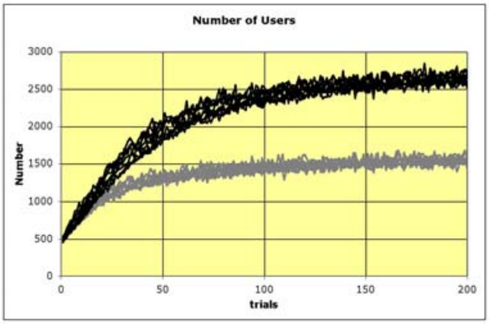

Fig. 5 Number of Users of ED in the Simulations

and predictable over simulations with a wide range of differing random inputs and detailed evolution.

4 Survey Method and Results

To test this model, a telephone survey was conducted of patients who visited the emergency department of the hospital of the Drexel University School of Medicine. [2] These telephone surveys were conducted by an independent research group who were given a list of all patients who had visited the ED during summer, 2007. Names and telephone numbers were randomly chosen from this list to complete 301 interviews. Interviewers made up to six calls, on different days and at different times, to reach and interview 301 patients.

Seven survey items were used to assess patient satisfaction. These items asked the patient to evaluate his/her satisfaction with the perceived quantity and promptness of care, the administrative and medical staffs effectiveness and consideration, staff time spent, and overall satisfaction with her or his care (sample item: How satisfied were you with the promptness of care you received at the ED?). A five point, Likert-type response scale was used for each item, ranging from 1 = Very satisfied to 5 = Very dissatisfied (the scores were reverse scored for the analyses, so higher numbers reflected greater satisfaction).

The same facets of satisfaction were also used to assess the patients expected experience in the ED and the patients expected experience with an alternative mode of care (sample items: How satisfied did you expect to be with the promptness of care you were going to receive in the ED? and Keeping in mind the facility you most likely would have gone to if you had not visited the ED: How satisfied would you have expected to be with the promptness of care?). In addition to these individual items, we also created three multi-item scales measuring overall satisfaction with the ED, expected satisfaction with the ED and expected satisfaction with the

[2] The complete script used for the survey accompanies the more extensive working paper available from the authors where is shown in Appendix 3.

alternative mode of care. The means of the seven items were used for each scale and each showed strong internal consistency (satisfaction with ED, $\alpha = .96$; expected satisfaction with ED,$\alpha = .95$; expected satisfaction with alternative mode of care, $\alpha = .95$). Note that since only patients who actually used the ED were surveyed, we have no data on the experienced satisfaction of the users of the alternative, nor on their expectations for the two modes .

Paired t-tests were used to assess differences in satisfaction levels between what the patients experienced at the ED and expected satisfaction with the alternative (Table 3), as well as the difference in the expected and experienced satisfaction with the ED (Table 4). In addition to the paired means for each item and the overall satisfaction scale, the tables show t-statistics and p-values The sign of the t-statistics is positive if the expectation for the alternative was greater (by however little margin) than the reported experience in the ED. Conventionally, the difference of means would be considered statistically significant if the p-value is less than 5%. The equilibrium hypothesis is that the there would be no statistically significant differences between the scores. The tests for difference of means are consistent with this hypothesis. It is true of every comparison except the promptness of care, in which significantly more satisfaction was expected of the alternative than of the ED. However, it is overall satisfaction that directly tests the equilibrium hypothesis, since overall satisfaction is the motive for choosing one alternative over the other. Moreover, the underlying mechanism that equilibrium is brought about by increasing congestion, which offsets other advantages of the ED is consistent with the statistically significant difference for promptness of care.

Table 3 Comparison of Experienced Satisfaction from ED with Expected Satisfaction from Alternative

Criterion	Expected from alternative	Experienced at the ED	t	p-value
Overall Quality of Care	3.83	3.77	,534	0.593
Quantity of Care	3.77	3.86	-.822	0.412
Promptness of Care	3.71	3.46	2.04	0.042
Administrative Staff Quality	3.79	3.87	-.728	0.467
Staff Ability to Resolve Your Medical Condition	3.89	3.91	-.164	.870
Personal Care and Consideration	3.96	3.93	.265	0.791
Amount of Staff Time Spent	3.69	3.69	.892	0.373
Overall satisfaction	3.80	.3.77	.379	0.134

Table 4 shows a similar series of comparisons for the expected and experienced satisfaction with ED care. The t statistic is positive if the expected satisfaction with ED care is greater than the experienced satisfaction. As we have noted, positive differences are to be expected if the expectations of the population polled were heterogeneous. These results are consistent with the assumption of heterogeneity.

Table 4 Comparisons of Expected and Experienced Satisfaction with ED Service

Criterion	Expected from the ED	Experienced at the ED	t	p-value
Overall Quality of Care	4.27	3.76	6.02	0.000
Quantity of Care	4.26	3.84	5.59	0.000
Promptness of Care	3.99	3.45	6.08	0.000
Administrative Staff Quality	4.15	3.86	4.05	0.000
Staff Ability to Resolve Your Medical Condition	4.41	3.93	6.15	0.000
Personal Care and Consideration	4.32	3.93	5.23	0.000
Amount of Staff Time Spent	4.06	3.60	5.34	0.000
Overall satisfaction	4.21	3.77	6.40	0.000

All in all, the results of the survey are consistent with the predictions arising from the agent-based simulation of the Nash equilibrium model for emergency department overcrowding.

5 Concluding Summary

The project reported in this paper was highly interdisciplinary, drawing ideas and techniques from several sources. There are novel contributions for each.

- For health care policy, we have specified, tested, and verified a Nash equilibrium hypothesis of the cause and nature of emergency room overcrowding. This hypothesis implies that increasing emergency room capacity may have little or no impact on overcrowding, in the absence of important changes in access to the alternative modes of medical care.
- For game theory, we have provided an example of testing a game-theoretic equilibrium model by questionnaire methods, using a realistically scaled agent-based computer simulation with boundedly rational learning to extend the insights of two- and small N-person game models to generate hypotheses for the survey.
- For questionnaire methods, we have provided an example of application to hypotheses from game theory and some evidence of the importance and consequences of heterogeneity, and the possibility of modeling heterogeneity explicitly by means of agent-based computer simulation.

For practical purposes, the policy implications would seem to be the most important, and a major limitation is that they reflect data from only one hospital. In future research, we hope to extend the questionnaire studies to a larger sample including hospitals in different regions and in different social contexts, e.g. rural and suburban as well as urban.

References

1. Aumann, Robert J, Subjectivity and Correlation in Randomized Strategies, Journal of Mathematical Economics v. 1, pp. . **1**, 67-96 (1974)
2. Burt CW, LF. McCaig (2006) Staffing, Capacity, and Ambulance Diversion in Emergency Departments: United States, 200304, Advance Data from Vital and Health Statistics no. 376 (27 September). http://www.cdc.gov/nchs/data/ad/ad376.pdf.
3. Burt, C.W, L.F. McCaig, and E.A. Rechsteiner (2004) Ambulatory Medical Care Utilization Estimates for 2004, http://www.cdc.gov/nchs/products/pubs/pubd/hestats/estimates2004/estimates04.htm (accessed 23 February 2007).
4. Cunningham P, May J. (2003) Insured Americans Drive Surge in Emergency Department Visits. Issue Brief no. 70 (Washington: Center for Studying Health System Change, 2003). http://www.hschange.com/CONTENT/613/.
5. Derlet RW, JR. Richards, Overcrowding in the nations emergency departments: complex causes and disturbing effects. Ann Emerg Med **35**, 63-68 (2000)
6. Holland, John and John Miller, Artificial Adaptive Agents in Economic Theory, American Economic Review (May) **81**, 365-370 (1991)
7. Kellermann, A.L. Crisis in the Emergency Department NEJM; 355:13001303 (2006)
8. Koyck, L. M. (1954), Distributed Lags and Investment Analysis (North-Holland Publishing Company, Amsterdam)
9. Luce, R. Duncan and Howard Raiffa (1957), Games and Decisions (New York: Wiley and Sons).
10. McCain, Roger A. (2010), Game Theory: A Nontechnical Introduction to the Analysis of Strategy, Revised Edition (Singapore: World Scientific).
11. McCain, Roger A. (2009), Game Theory and Public Policy (Cheltenham, UK: Elgar).
12. Melnick, GA, Nawathe AC, Bamezai A, Green L: (2004) Emergency Department Capacity And Access In California, 1990-2001: An Economic Analysis Health Affairs Web Exclusive, March 24, http://content.healthaffairs.org/cgi/content/abstract/hlthaff.w4.136v1
13. Tesfatsion, Leigh (2000), Agent-Based Computational Economics, (Iowa State University) available at: http://www.econ.iastate.edu/ tesfatsi/ace.htm
14. Wilper A, Woolhandler S, Lasser K, McCormick D, Cutrona S, Bor D, Himmelstein D. (2008) Waits To See An Emergency Department Physician: U.S. Trends And Predictors, 1997-2004 Health Affairs Web Exclusive, January 15, http://content.healthaffairs.org/cgi/content/abstract/hlthaff.27.2.w84v1

An Agent-based Model of Food Safety Practices Adoption

Tim Verwaart and Natalia I. Valeeva

Abstract Food processors and governments have an interest to motivate suppliers of animal products to implement food safety measures. This paper proposes an agent-based simulation model to compare possible consequences of alternative communication programs and incentive systems. The agent model combines economic incentive models with social-psychological survey results in an approach based on the theory of planned behavior. Food safety actions follow from producers' attitudes, social network influences, and perceived availability of resources and opportunities to implement the measures. The model allows for heterogeneity in the agent population, for instance with respect to openness to communications and factors that motivate producers to implement food safety measures. A sensitivity analysis can be performed for both aggregate outputs, such as food safety risk reduction downstream in the supply chain, and individual agent performance, such as the response to different incentives and communications. Conclusions are drawn about the model's feasibility for food safety policy support.

1 Introduction

Food safety is an issue in supply chains of animal products, such as dairy, meat, and eggs [14]. Food safety risk can be reduced by measures at each stage of the supply chain. In supply chains of animal products, many producers (farmers) deliver their products to a few processors. Producers and processors usually have long-term relationships. Processors have an interest to motivate their suppliers to implement measures that reduce food safety risk. Governments also have such interest, from a perspective of public health. A simulation model to evaluate possible consequences

Tim Verwaart and Natalia I. Valeeva
LEI Wageningen UR, Postbus 29703, 2502 LS den Haag, The Netherlands, e-mail: {tim.verwaart; natasha.valeeva}@wurl.nl

S. Osinga et al. (eds.), *Emergent Results of Artificial Economics*, Lecture Notes in Economics and Mathematical Systems 652, DOI 10.1007/978-3-642-21108-9_9,

of alternative communication programs and incentive systems *ex ante* can help these stakeholders to determine their food safety policies.

Models of economic incentives based on principal-agent theory and risk aversion are available and have been applied to design systems of penalties or premiums (e.g., [13, 16]). In addition to economic motives, the producers' intentions and actions depend on information, attitudes, and professional competences, and they are influenced through their social networks. In many research projects, field data is collected about efficacy of measures, about the cost of implementation, and about the factors that motivate producers to apply the measures (e.g., [10, 15]). This type of data can be combined with an economic modeling approach, to enhance the efficiency of simulation models in support of food safety policy.

In an agent-based approach economic and social factors can both be taken into account in a model of producers' decisions and actions. Such a model can be applied to estimate effects of alternative communication and incentive policies in a heterogeneous agent population. The effects on food safety risk and average quality emerge from the behaviors of individuals. Since the producers influence each other through their social networks, the effects cannot be viewed as a simple aggregate of individual decisions: social norms may withhold producers from acting. On the other hand, when thresholds of social acceptance are overcome and agents start following the examples of others, the implementation of measures may spread rapidly.

This paper focuses on the adoption of animal health practices by primary producers (farmers), in particular to reduce mastitis prevention in dairy cattle. The agent model of producer behavior proposed in this paper is inspired by the theory of planned behavior [1, 2]. According to this theory people's intentions to perform a behavior are based on their attitudes toward that behavior, on subjective norms (the extent to which important others are believed to approve or disapprove the behavior), and on their perceptions of behavioral control, i.e. to have the resources and opportunities to actualize the objectives of the behavior. The *intention* to perform a behavior is modeled to be a function of attitude A, subjective norm SN, and perceived behavioral control PBC, with

$$A \propto \sum_i b_i e_i, \tag{1}$$

$$SN \propto \sum_j n_j m_j, \tag{2}$$

$$PBC \propto \sum_k c_k p_k, \tag{3}$$

where e_i represent the effects that the behavior can bring about and b_i the subjective probabilities that the behavior will actualize the associated effects; m_j represent the motivation to comply with the jth referent and n_i the strength of the normative beliefs; p_k represent the perceived power of the control factors and c_k the believed control over the factors. According to the theory of planned behavior both the intention and the perceived behavioral control need to be sufficiently strong for a person to perform a behavior.

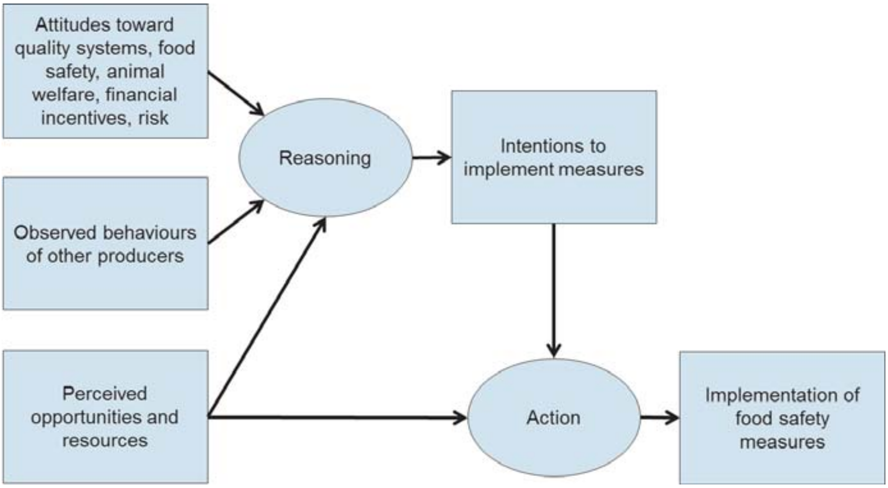

Fig. 1 Structure of the model of producers' intentions and actions

This paper proposes an agent model of producers' behavior as depicted in Fig. 1. The agents decide about food safety actions according to the theory of planned behavior. Their beliefs about efficiency and cost of possible food safety actions are modeled to be reinforced by professional communications, initiated by stakeholders like food processors and government authorities. The prevalence of food safety problems is modeled as a stochastic process. The expected value and variance of this process are reduced by food safety actions, but problems can still prevail. The producers receive feedback from the processor about actual prevalence and about premiums or penalties for actual performance. The producer agents use this feedback to update their beliefs about efficiency of actions. Furthermore, they observe the behaviors of other producers in their social network.

The research presented in this paper aims to assess the feasibility of agent based simulation to support food safety policy formulation. To that end a prototype has been developed for the case of mastitis prevalence in dairy cattle. Mastitis has farm level economic effects. Milk processors and government authorities are interested in policies to reduce mastitis prevalence. Premiums and/or penalties based on an analysis of *bulk milk somatic cell count (BMSCC)* in delivered milk are usually applied. Additional measures could further reduce mastitis prevalence. A host of data on mastitis prevalence and farmers' behaviors is available [5, 6, 7, 8, 10, 11].

The model proposed in this paper is abstract and is not calibrated to actual data. The latter is subject of current research and has to be well-founded on sensitivity analysis. The current prototype allows for such sensitivity analylis. Formal aspects of the model are discussed in Section 2. Section 3 presents the model's implementation and the approach to sensitivity analysis. Section 4 concludes the paper with a discussion of the model's feasibility to support food safety policy.

2 The Agent Model

This section decribes three key elements of the model. Firstly, information types available to the agents to found their decisions on. Secondly, the agent's information processing and decision mechanism to adopt measures and to act. Thirdly the process to generate prevalence data.

Information types. Information is a central issue in food safety policy. The following types of information are modeled in this multi-agent simulation.

(Information type A) Producers are informed by the food processor about penalties based on actual prevalence levels of contaminations in their products. Prevalence level itself may be considered a stochastic variable. In addition the test procedure has a measurement error. So, after taking food safety actions, some uncertainty about measured prevalence level will remain for the producers. In the simulation, contamination levels are randomly drawn from a log-normal distribution with expected value depending on the food safety actions taken (food safety actions reduce the expected value) and on prevalence level in the previous period (some autocorrelation is assumed). The coefficient of variation is kept constant, so variance is also affected by food safety actions. The generation of *BMSCC* levels $test_t$ is to be explained further on in this section. The values of $test_t$ are the basis for the penalties in period t.

The following example penalty system is applied in the simulation (note: this is an arbitrary penalty system, like it could be formulated by a milk processor):

$$prev_t = \sqrt{test_t prev_{t-1}}, \tag{4}$$

$$pen_t = 0 \text{ if } prev_t \leq \alpha_1, \tag{5}$$

$$pen_t = \min\{\alpha_2 + \alpha_3 pen_{t-1}, \alpha_4\} \text{ if } prev_t > \alpha_1, \tag{6}$$

where $prev_t$ is the basis for penality pen_t for period t; $test_t$ represents the last test result; $prev_0$ is randomly generated; α_1 is the penalty threshold; α_2 stands for the base penalty for first non-compliance, α_3 for an increase factor in case of repeated non-compliance; α_4 is the upper bound of penalties.

(Information type B) Producers are informed by food processors, the government, or other professional information services, about the consequences of possible actions. The information is modeled to be transfered in the form of tuples:

$$< s, a, prev\_reduction_a, E\_saving_a, add\_cost_a, ease\_of\_adoption_a >, \tag{7}$$

where s represents the sender, $prev\_reduction_a$ stands for the factor with which the prevalence expectation is reduced by action a, E_saving_a for the saving in animal health losses, add_cost_a for the implementation cost per period, $ease\_of\_adoption_a$ for the relative ease of adoption on a continuous scale from 0 to 1 (0: structural change and investment required, 1: fully reversible change of behavior; in the present paper we only study the case of fully reversible behavior).

The information is disseminated by s with some intensity of communication $\iota_a(s,t)$ for possible action a; $\iota_a(s,t)$ is interpreted in the model as the probability for a producer to receive the message in time period t.

Information type (C) The third source of information are other producers in a producer's social network. In the present version of the model the only information the producers exchange about food safety actions is whether they have taken them or not.

Information processing and decision making. Producers found their decisions on the information types presented above. Jansen et al. [12] discern four types of farmers with respect to openness to communications, using two dimensions: openness to communications from external sources (information type B) and openness to information from other farmers (information type C). The producer agent model reflects these types.

Information of types A and B may change producer agent's beliefs about potential effects of actions and about the probabilities that actions will actualize these effects. The effect of information type A is obvious: the producer will be aware of the current level of prevalence and penalty.

If an information type B is received in period t the value of $prev\_reduction_a$ is used to update the agent's belief about expected actual reduction $B\_prev\_reduction_{a,t}$ of $BMSCC$ by action a:

$$B\_prev\_reduction_{a,t} = (1-\gamma_s)\ B\_prev\_reduction_{a,t-1} + \gamma_s\ prev\_reduction_a, \quad (8)$$

where γ_s stands for the agent's openness to communications from sender s, with $0 \leq \gamma_s \leq 1$, and $B\_prev\_reduction_{a,0} = 0$. The values of other information communicated in types A and B are directly copied as agents' beliefs.

If no information about action a is received in period t the strength of belief is decayed with a small factor, e.g. $B\_prev\_reduction_{a,t} = 0.95\ B\_prev\_reduction_{a,t-1}$.

Agents observe infomation of type C in their social network. For each network relation i they have a motivation m_i to follow the example of i's behavior (in the simulation the values of m_i are initialized by a random generation process and to not change during the simulation). The agent's belief about the social norm with respect to action a is computed according to equation (2), with $n_i = 1$ for action a if i acts, $n_i = 0$ if i does not act.

The belief about the norm is input to the intention to perform a:

$$I_a = w_A A_a + w_{SN} SN_a + w_{PBC} PBC_a\ . \quad (9)$$

An agent's openness to information from other farmers is reflected by the value of parameter w_{SN}.

The other terms in the above equation refer to the agent's atitude toward a and its perceived behavioral control to perform a. With respect to the attitude, Valeeva et. al. have identified three clusters of producers with different motivations in a conjoint survey analysis regarding mastitis control [15]:

1. premium- or penalty-oriented motivation,
2. motivation to have a well-organized farm,
3. basic economic motivation.

From a viewpoint of economics, a producer is motivated to change behavior if the expected utility of the alternative behavior is greater than that of current behavior. The producer's utility is calculated as

$$U_t = 1 - \exp(-\lambda f_t),\ \lambda > 0, \tag{10}$$

where λ represents constant absolute risk aversion and

$$f_t = saving_t - cost_t - pen_t. \tag{11}$$

Producers can be more or less satisfied with their operational excellence (having a well-organized farm). We assume a producer's relative satisfaction with operational excellence as

$$S_t = \frac{AP_t - prev_t}{AP_t}\ , \tag{12}$$

where AP_t stands for the average $BMSCC$ prevalence in the producer population at time t.

Based on beliefs about recent prevalence and penalties, believed effects of action a on prevalence, and other beliefs resulting from information of type B, a producer agents can estimate U and S for alternative behaviors, and determine its attitude A_a towards action a:

$$A_{a,t} = \beta A_{a,t-1} + (1-\beta)(\tilde{U}_t(a) - \tilde{U}_t(\bar{a}) + \nu(\tilde{S}_t(a) - \tilde{S}_t(\bar{a}))) \tag{13}$$

where β is an agent parameter that represent persistence of the attitude, and ν represent the relative value which the producer attaches to operational excellence relative to economic utility.

The parameters λ and ν can be set to discern the motivational clusters identified by Valeeva et al. [15]:

1. premium- or penalty-oriented motivation: high value of λ, low value of ν,
2. motivation to have a well-organized farm: low value of λ, high value of ν,
3. basic economic motivation: low value of λ, low value of ν,

The third term in eqution (9) reflects an agent's perceived behavioral control: the ability to actually perform the action with the desired result. We determine perceived behavioural control of an agent as

$$PBC_a = B\_E\_saving_a \cdot B\_prev\_reduction_a - B\_add\_cost_a \tag{14}$$

where $B\_E\_saving_a$ reflects the agent's belief about expected saving per unit of reduction and $B\_add\_cost_a$ the believed additional cost of the action. In the present model we set $w_{PBC} = 0$ in equation (9).

A producer agent decides to performs action a if the intention $I_a > 0$ and perceived behavioral control has at least some threshold value: $PBC_a \geq PBCthreshold$.

After the decision, the new level of $BMSCC$ prevalence is randomly generated as described below, with

$$\mathrm{E}(prev_t) = M \text{ if no action is performed in period } t, \tag{15}$$

$$\mathrm{E}(prev_t) = M(1 - prev\_reduction_a) \text{ if } a \text{ is performed in period } t. \tag{16}$$

M represents the mean value of prevalence observed in a large dataset of $BMSCC$ analysis results; $prev\_reduction_a$ of actions can be obtained from literature or experts. Prevalence levels can be used to calculate actual penalties and savings that can be input to the next time step of the simulation. From the results of individual actions, aggregate outputs can be computed - such as average prevalence AP - and the effects of alternative or complementary communication and incentive policies can be analysed.

Random generation of prevalence data. We assume a lognormal distribution of $BMSCC$ prevalence values, with expected value $\mathrm{E}(prev_t)$ and coefficient of variation cv. The values of these parameters are known from a dataset of prevalence values. An actual value test result $test_t$ is calculated as

$$test_t = \exp Y_t, \tag{17}$$

with

$$Y_t = \delta Y_{t-1} + (1-\delta)X, \tag{18}$$

where X is drawn from a normal distribution, with

$$\mu = \ln \mathrm{E}(prev_t) - 0.5 \ln(1 + cv^2) \tag{19}$$

and

$$\sigma^2 = \frac{1+\delta}{1-\delta} \ln(1 + cv^2). \tag{20}$$

For each agent Y_0 is drawn from a normal distribution with $\sigma_0^2 = \ln(1 + cv^2)$ and μ according to equation (19).

3 Implementation and Results

The model described above has been implemented in Netlogo 4.1.2 [1].

Input data are based on published Dutch studies [5, 6, 7, 8, 10, 11]. For estimation of economic input parameters the base dairy farm is used. The farm is characterized by 65 milking cows with an average 305 day milk production of 8,500 kg/year, and

[1] The model and related information, such as sensitivity analysis procedures and results, are available at http:/www.verwaart.nl/FoodSafety

milking parlor with 12 stands (with a 12-place milking parlor under Dutch conditions), an average *BMSCC* of 200,000 cells/ml, an incidence of clinical mastitis of 30% per year (thus 20 cases for the base farm, given that 65% of the mastitis problems were caused by environmental mastitis pathogens and 35% by contagious mastitis pathogens).

Validation of an agent-based model is an ongoing proces. According to Gilbert sensitivity analysis is an essential step in the process of validation [4]. Sensitivity analysis can be performed, for instance, by Monte Carlo sampling of the parameter space and running the model for the parameter sets drawn [9]. A subsequent analysis of the outputs can identify the inputs and parameters to changes of which the model's outputs are most sensitive. That information can be used in the process of calibrating the model to real-world situations.

Table 1 presents an overview of the parameters of the present simulation.

Table 1 Model parameters of different categories, to be considered in sensitivity analyses

Category	Symbol	Parameter
Agents	λ	risk aversion
	ν	preference for operational excellence
	γ	openness to communication
	β	persitence of attitude
	d	network degree
	m_i	motivation to follow i (equal for all i)
	w_A	weight of attitude in intention determination
	w_{SN}	weight of social norm in intention detemination
	w_{PBC}	weight of PBC in intention determination
	$PBCthreshold$	perceived behavioral comtrol threshold to act
Penalty system	α_1	prevalence treshold for penalty
	α_2	penalty for first non-compliance
	α_3	increase factor for repeated non-compliance
	α_4	upper bound
Communication	ι_a	intensity of communication
	$prev\_reduction$	expected reduction of prevalence by action
	add_cost	additional cost of action
	E_saving	expected saving per unit of prevalence reduction
Prevalence	$\mathrm{E}(prev)$	expected level of prevalence
	cv	coefficient of variation
	δ	persistence parameter of prevalence

Like many agent-based models, this model has an extensive parameter set, in which agent parameters and global parameters of diverse types can be discerned. The agent parameters can be differentiated to compose a heterogeneous population. This is one of the purposes of the simulation (cf. the motivation types identified by Valeeva et al. [15]), but it complicates a sensitivity analysis. Due to the many interactions in an agent-based model, a sensitivity analysis is best performed for every situation the model is to be applied in [3]. This implies that a one-off sensitivity analysis is useful, but in addition procedures should be available to perform sensitivity analysis for particular situatons. Such procedures must allow for analysis of macro statistics and results of individual agents [3]. Procedures for sensitivity analysis of the present model and results are made available at the location specified in footnote 1.

This paper presents some results with a fixed setting of most parameters[2]. These parameters are not calibrated to actual data. The results are produced to demonstrate sensitivity of the model to communication intensity, network effects, and different motivations of the producers. Actual calibration remains for future research.

Table 2 presents results of simulations for the different motivation types of Valeeva et al. [15]. Network effects were excluded in the simulation The results show that information is an important factor for food safety adoption according to the model. Further, the reults dmonstrate the model's sensitivity to the different motivation types.

Table 2 Percentage of adoption for three motivation types at different levels of intensity of communication (ι) and penalty for first non-compliance (α_2), if no effects of social norms are assumed ($w_A = 1$; $w_{SN} = 0$)

ι	α_2	Adoption(%) $\lambda = 1$ $\nu = 0$	Adoption(%) $\lambda = 1$ $\nu = 1$	Adoption(%) $\lambda = 1.2$ $\nu = 0$
0.5	0	14	39	17
	2	35	52	36
	4	42	52	43
	6	47	52	45
0.8	0	17	51	23
	2	55	77	59
	4	72	85	77
	6	83	90	88

Table 3 presents results of simulations for a similar setting, but with network influence. The outcomes show that network effects reinforce the effects of communication, according to the model. A surprising effect is the reduced adoption for increasing penalty levels under low communication intensity in the simulations, if network effects are taken into account. High penalties reduce the relative effect of social norms in the decision functions, to the most for the penalty-oriented producers. Empirical validation or falsification of this effect remains for current research.

[2] average value of $\gamma = 0.75$; $\beta = 0.5$; $d = 4$; average value of $m_i = 0.75$; $w_{PBC} = 0$; $PBCthreshold = -0.2$; $\alpha_1 = 250$; $\alpha_3 = 1$; $\alpha_4 = 20$; $prev\_reduction = 0.07$; $add\_cost = 0.5$; $E\_saving = 5$; $\mathrm{E}(prev) = 210$; $cv = 0.3$; $\delta = 0.5$;

Table 3 Percentage of adoption for three motivation types at different levels of intensity of communication (ι) and penalty for first non-compliance (α_2), with network effects of social norms ($w_A = w_{SN} = 0.5$; network degree $d = 4$)

ι	α_2	Adoption(%) $\lambda = 1$ $\nu = 0$	Adoption(%) $\lambda = 1$ $\nu = 1$	Adoption(%) $\lambda = 1.2$ $\nu = 0$
0.5	0	87	86	86
	2	70	68	63
	4	59	59	52
	6	56	55	51
0.8	0	99	99	99
	2	98	98	98
	4	98	98	98
	6	98	98	96

4 Conclusion

The study develops an agent-based computational model of dairy farmers' decisions on adoption of measures to decrease the mastitis rates and consequently improve safety of the delivered raw milk. The current approach presents a way to model the food chain as a complex dynamic and more realistic system, taking human adaptation and learning (perceptions and behavior) into account next to traditional economic considerations (cost-effectiveness and incentives) while adopting certain food safety measures. In particular, the study shows how farmers' decisions on adoption of food safety measures are driven not only by economic considerations (by taking into account both actual and perceived by farmers economic parameters), but also by other motivations, cognitions and professional and personal networks. The study helps better understand interrelation of these aspects and their effect on success of the different tools used to motivate farmers to improve food safety (such as economic incentives and communication campaigns).

Under different parameter settings (e.g. risk-aversion, inclination to be an innovator, attitudes, network size, perceived availability of resources), model runs result in the emerging numbers of farmers adopting and not adopting food safety measures over time, given a certain incentive scheme and communication campaign parameters. Results include quantitative insights into consequences of several emerging collective behaviour patterns on the whole system of raw milk supply (such as milk safety and quality) and individual farmer performance (such as response to incentive schemes and communication campaigns).

The results reflect that individual farmers are more sensitive to costs of the measure, penalty, network influence, or communication according to personal characteristics like professional motivation and openness to communications from different sources. Simulations show that network influence can reinforce the adoption of food safety measures. Producers who would not adopt in response to incentive systems or communication programs, do so in the simulation under the influnce of social norms. The developed method can be used by chain participants (industry and primary production) to evaluate adoption of different food safety measures on the farm under a mix of approaches and to find out the most significant parameters affecting the adoption of food safety measures to reduce food safety risk under such a mix of approaches.

Results are helpful in evaluating tools aimed to decrease the mastitis rates and consequently improve safety of the delivered raw milk. The study provides valuable insights into the decisions to improve food safety management on the farm in a contemporary environment. More research would be however needed to find out the better estimates of the model parameters.

A further development of the model would be to include a possibility of exploring the adoption of other measures or different sets of measures. This would allow analysing the differences in effects of different aspects of decision-making on the decisions on adoption of different measures. This would give insights into information whether a particular economic incentive or a communication campaign is to be more effective for stimulation of adoption of a certain measure or a set of measures. In order to design effective tools (to improve food safety via better health of the animals), it is necessary to take these aspects into consideration.

References

1. Ajzen I (1991) The theory of planned behavior. Organizational Behavior and Human Decision Processes 50(2):179–211
2. Armitage CJ, Conner M (2001) Efficacy of the Theory of Planned Behavior: A meta-analytic review. British Journal of Social Psychology 40:471–499
3. Burgers SLGE, Hofstede GJ, Jonker CM, Verwaart T (2010) Sensitivity Analysis of an Agent-based Model of Culture's Consequences for Trade. In: LiCalzi M et al. (eds) Progress in Artificial Economics, Lecture Notes in Economics and Mathematical Systems 645, pp. 253–264. Springer, Heidelberg
4. Gilbert N (2008) Agent-based Models. Sage Publications, Thousand Oaks
5. Halasa T, Nielen M, Huirne RBM, Hogeveen H (2009) Stochastic bio-economic model of bovine intramammary infection. Livestock Science 124:295–305
6. Hogeveen H, Huijps K, Lam TJGM (2011) Economic aspects of mastitis: New developments. New Zealand Veterinary Journal 59:16–23
7. Huijps K, Lam TJGM, Hogeveen H (2008) Costs of mastitis: facts and perception. Journal of Dairy Research 75:113–120
8. Huijps, K, Hogeveen, H, Lam, TJGM, Oude Lansink, AGJM (2010). Costs and efficacy of management measures to improve udder health on Dutch dairy farms. Journal of Dairy Science 93:115-124.
9. Jansen MJW, Rossing WAH, Daamen RA (1994) Monte Carlo estimation of uncertainty contributions from several independent multivariate sources. In: Grasman J, van Straten G (eds) Predictability and Nonlinear Modeling in Natural Sciences and Economics, pp.334–343. Kluwer, Dordrecht
10. Jansen J, van den Borne BHP, Renes RJ, van Schaijk G, Lam TJGM, Leeuwis C (2009) Explaining mastitis incidence in Dutch dairy farming: The influence of farmers'attitudes and behaviour. Preventive Veterinary Medicine 92:210–223

11. Jansen, J, Renes, RJ, Lam, TJGM (2010). Evaluation of two communication strategies to improve udder health management. Journal of Dairy Science 93:604–612
12. Jansen J, Steuten CDM, Renes RJ, Aarts N, Lam TJGM (2010) Debunking the myth of the hard-to-reach farmer: Effective communication on udder health. Journal of Dairy Science 93: 1296–1306
13. King RP, Backus GBC, van der Gaag MA (2002) Incentive system for food quality control with repeated deliveries. European Review of Agricultural Economics 34:81–104
14. Kleter GA, Prandini A, Filippi L, Marvin HJP (2009) Identification of potentially emerging food safety issues. Food and Chemical Toxicology 47:932–950
15. Valeeva NI, LamTJGM, Hoogeveen H (2007) Motivations of Dairy Farmers to Improve Mastitis Management. Journal of Dairy Science 90:4466–4477
16. van Wagenberg CPA, Backus GBC, Kuiper WE, van der Vorst JGAJ, Urlings HAP (2010) Incentive mechanisms for liver lesion control in finishing pigs in the Netherlands. Preventive Veterinary Medicine 93:19–24

Part IV
Macroeconomics

Why Should the Economy be Competitive?

Hugues Bersini and Nicolas van Zeebroeck

Abstract Is a competitive free market the most efficient way to maximize welfare and to equally allocate rare resources among economic agents? Economists usually tend to think this is the case. This paper presents a preliminary attempt through an object-oriented multi-agent model to address this question. Agents which are alternatively producers, sellers, buyers and consumers participate in a market to increase their welfare. Two market models are compared: the *competitive market*, which is a double auction market in which agents attempt to outbid each other in order to buy and sell products, and the *random market*, in which products are allocated in a random manner. The comparison focuses on wealth creation (a common study) and above all on inequalities in resources allocation (much less frequently addressed). We compute and compare in both cases the money and utility generated within the markets as well as a Gini index that characterizes the degree of inequality in the allocation of these values among agents. In contrast with earlier works, we also compare welfare creation and distribution with more or less intelligent agents who use more or less information provided by the market. The results of hundreds of simulations are analyzed to find that, as is well known, the competitive market creates more value overall but, and generally less accepted and discussed by economists, at the expense of a much greater inequality. We further find that whereas the use of market information by the producers leads to less inequality in both types of markets, the inequality induced by the competitive market depends on the behavior of the agents, suggesting that it is both the institutions and the agents that foster inequality in competitive market structures.

Hugues Bersini
IRIDIA-CODE, Université Libre de Bruxelles, CP 194/6, 50 av. Franklin Roosevelt, 1050 Bruxelles, Belgium, e-mail: bersini@ulb.ac.be

Nicolas van Zeebroeck
ECARES and IRIDIA-CODE, Université Libre de Bruxelles, CP 114/04, 50 av. Franklin Roosevelt, 1050 Bruxelles, Belgium, e-mail: nivzeebr@ulb.ac.be

S. Osinga et al. (eds.), *Emergent Results of Artificial Economics*, Lecture Notes in Economics and Mathematical Systems 652, DOI 10.1007/978-3-642-21108-9_10,

1 Introduction

Economic tradition holds that competition and free markets are the best way to create value. Adherents to Adam Smith's exposition of the invisible hand and famous historical advocates of competitive or perfectly concurrent free market, such as Friedrich Von Hayek, Milton Friedman and many others, have usually seen in this type of decentralized and self-organized economical exchanges between competitive sellers and buyers the most efficient way to maximize social welfare. More recently, Gode and Sunder [2] have shown that highly competitive market structures such as double auctions lead to efficient product allocation and price setting, even when the agents have zero-intelligence (i.e. place offers and bids in a purely random way). They conclude that imposing a budget constraint to the agents is sufficient to maximize the allocative efficiency of these auctions, even with irrational (or Zero-Intelligence) agents.

However, a competitive system that is supposed to promote the winners to the expense of the losers may have little chances to equally distribute wealth. Competition in the economy is inherent to the nature of the law of supply and demand that rules the dynamics of these markets, calibrating the products price such as to adjust demand to supply. As a result of this competition among agents, the price of a rare product to acquire should increase, favoring the richest buyers, and the price of an abundant product to sell decrease, favoring the most performing sellers (who propose lower prices). One first naive counterargument to this apparent source of inequality is that, by competing, the sellers will finally provide products at cheaper price to the poorest buyers, and reciprocally, the buyers, by competing, will allow the unqualified sellers to also profit from the products they sell.

A succession of theoretical models has shown that it is possible to maximize the efficiency in product allocation through price adjustments. However, although all agents benefit from the exchanges at the aggregate level, individual welfare still depends on the agent's tastes distribution, budget constraints and production skills, and little is told about the welfare distribution among the agents. Allocative efficiency and Pareto optimality have nothing to do with equality among agents. As a matter of fact there is a long going philosophical dispute between the utilitarians (maximizing cumulated welfare) and the egalitarians (reducing welfare variability) [3], since both objectives (while equally ethically justified) appear to be antagonistic in many social contexts.

With regard to the egalitarian perspective, very little attention has been paid so far in the literature to the side-effects of competitive market structures on welfare distribution among agents. This is a difficult question to address from a structural point of view. In this paper, a stylized model is presented in order to compare two styles of market structures: competitive (a stylized double auction where the product goes to the most offering buyer and is sold by the most performing seller) v. random (or distributive) (where the matching between buyers and sellers is done in a purely random way insofar as certain constraints are met). The two market structures are compared along two main dimensions: the welfare produced as a result of the agents' interactions (in terms of utility and money), and the degree of inequality in welfare allocation between agents.

The paper that most closely relates to ours is Gode and Sunder [2], who look at the dispersion of profits among agents in the human v. zero-intelligence markets and conclude that contrarily to allocative efficiency that can be maximized by the market structure itself, profit dispersion may depend on the behavior of the agents. They however do not investigate the impact of human behavior on utility dispersion. The present article departs from Gode and Sunder's work in different ways. First, Gode and Sunder look at the difference in allocative efficiency of a single market structure when two different types of agents interact, human agents v. Zero-Intelligence machine agents who place bids and offers randomly. In contrast, we look at the same sets of machine agents who interact on two different types of market structures: competitive v. random. Second, we not only aim at looking at welfare creation (or allocative efficiency), but most importantly at welfare distribution among agents, in terms of both utility and money. Finally, we also investigate the impact of different agents behaviors in the different markets on welfare creation and distribution.

In practice, a computer program will be presented in the following, implementing simple rules for the agents to follow as well as a framework which allows them to interact, while logging facilities collect data on the transactions executed and the evolution of the different metrics in the model. The agents are then set free in numerous simulations, and the resulting metrics are compared across simulations. Our first sets of results show that a competitive market structure such as a double auction consistently leads to much higher welfare at the aggregate level than a random matching market, but they reveal considerable differences in the distribution of utility and money between the two market structures, with the competitive market leading to much more unequal distribution of welfare between agents. We then explore the impact of informed v. zero-intelligence producers and find that even in the most competitive market, ignorance of the producers leads to slightly more utility for the consumers but at the expense of much more money being wasted and therefore makes both markets much less efficient. We finally investigate different behaviors of our machine buyers and sellers in selecting the products they want to place bids or offers for and observe that although welfare creation is generally equal, the degree of inequality in utility and money distribution with the competitive market significantly depends on the rationality of the agents.

In the coming sections, the object-oriented model will be first described by detailing the constituting classes. Then the experimental results of the cumulated distribution of wealth will be compared across the two types of economy: competitive and random. As usual with this kind of artificial models (such as with classical game theory models and almost all artificial life models), the simulation presented here is not intended to depict any precise reality but needs to be construed as a software-thought experiment, namely the conception and the execution of virtual worlds helping to understand in outlines the behavior of a purposefully caricatured reality.

2 The Model

The model implemented in C# maps elegantly to an object oriented model with the distinct responsibilities distributed through the different classes. These classes and their relationships can be seen in the simplified class diagram in figure 1.

The model's different components all live within a structure called the *world*. The World contains all the agents as well as the market. Each world has one market, either a competitive double-auction market or a random market, a series of agents and some world specific settings such as the initial endowment of the agents, the number of different products the agents can make and trade,. . . A given number of products are bought and sold. The world is not limited in the number of units for each product, but each transaction concerns only one unit of the product. Each world corresponds to one simulation. Worlds always come in pairs with equal initial settings but one taking care of the competitive market and the other one of the random market.

The agents are the main actors of the model, they are the imaginary people who produce, consume and trade goods driving the model's markets. Each agent behaves alternatively according to its integrated Producer, Consumer, Buyer and Seller classes. So each agent plays the four roles in turn. Agents are defined by several key numbers, including:

Money, utility, stock: Every agent starts with the same amount of money at the beginning of the simulation. This allotment of money allows the agent to produce goods, in which case the money leaks out of the system, or to purchase them from one of the other agents during a transaction. Agents also have an amount of accumulated utility which they increase by consuming products. The way the utility increases depends on the agent's tastes. Agents also possess a certain amount of products which they have produced but not yet sold. For logging purposes, these stocks are valued at current market prices. At the beginning of the simulation, agents start with no inventory.

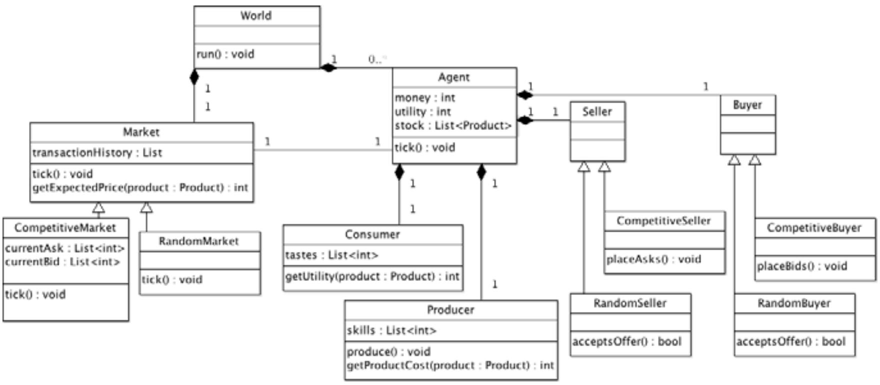

Fig. 1 The UML class diagram of the model

Skills and tastes: Agents are also characterized by two crucial vectors which are their skills and their tastes, one of each corresponding to each product. This is the departing point of agents differentiation during the simulation. Skills determine the cost of producing goods while tastes determine the amount of utility an agent will get from consuming a product. Skills and tastes are constant for each agent: once they are set, they will not evolve as the simulation proceeds. Furthermore, while individual skills and tastes vary, the total amount of skills ($\sum_i skill_i$) and tastes ($\sum_i taste_i$) is identical for every agent, hence agents only differ in the breakdown of their skills and tastes between products. At the initialization of the program, skills (tastes) are randomly set between 0 and 1 for each product, then they are all scaled so that the sum of all skills (tastes) for each agent is equal to 1.

Consumer behavior: When an agent purchases a product, it is immediately consumed and converted into utility. Tastes determine the amount of utility that will be produced: for each unit of product p consumed, the agent's utility will increase by $\lceil taste_p \rceil$.

Producer behavior: During each tick, an agent chosen randomly produces one unit of product. To produce a unit of product p, the agent will lose an amount of money determined by $\lceil skill_p \rceil$. This is the only process in the programme that dissipates money (all the other processes lead to money transfers between agents and/or to utility increases). A crucial part of the process is the selection of which product to produce. In a market designated as *random* (or Zero-Intelligence), the product will be selected randomly while in one designated as *informed*, the agent will search for the product which is expected to maximize his profit. As he knows his production cost for each product, he only needs to estimate the selling price. This is done by querying the market for estimated prices, or in case none are available yet (which is the case in the first ticks before any transaction has taken place) by guessing at random. Producers can therefore learn about the market supply and demand conditions. The market keeps a price estimate for every product, which is a moving average of the last n transaction prices for that product. Having an expected price and a given cost, the agent can calculate his expected profit ($expectedPrice - cost$) for each product and, knowing the product with the highest expected profit which he can afford to produce, select the kind of product to make on his turn.

The world moves forward through *ticks* which are discrete time-steps. During a tick, an agent is given a chance to produce one item (if the selected agent cannot produce anything, another agent is selected until one unit of a product has been produced), after what the market will execute one transaction. As a result, one product unit is exchanged between the seller and the buyer agents. Once acquired, the buyer agent immediately consumes the product and increases his utility by his taste for the consumed product.

Markets are the core of the model and represent the rules governing product exchanges between the agents. They also determine the buying and selling behavior the agents will play out. Two kinds of markets are studied in the model: the *competitive* one which is a double auction market and the *random/distributive* one.

Competitive Market: In the competitive market, agents bid to buy and sell goods. During a succession of steps, the market repeatedly invites two randomly selected agents to place asks and bids on one product they want to sell or purchase. This product selection can be done in different ways that are described in the results section below. In our baseline simulations, buyers choose the product that maximizes their expected net utility given the latest competing offer, and sellers choose the product that maximizes their expected net profit given the latest competing offer. In other settings, agents query the market to learn about actual prices, or use no market information but only base their choices upon their skills or tastes respectively. For each product, the market remembers only the highest bid and the lowest ask made in the current tick. As soon as these two numbers cross, the transaction is executed between the two winning agents for that specific product. If after a predetermined number of trials (arbitrarily high), no transaction can occur, the execution of the model stops and a market failure is reported. Agents place bids and offers in a manner inspired by Gode and Sunder [2]. Similarly to them, our agents are faced to their budget constraints, i.e. they cannot spend more than the money they have, cannot buy a product at a price above their utility, and cannot sell a product below their production cost.

Random/Distributive Market: In the random market, consumers are proposed a certain product to buy from a given producer. Provided that the taste of the agent for the proposed product is larger than the producer's cost (inverse of skill) for the given product and that the buyer is sufficiently endowed, the transaction is made and the price is randomly set by the market between these two bounds. Buyers and sellers therefore don't learn anything from the market in the random model and are therefore closer to the Zero-Intelligence agents of Gode and Sunder. Here Again, if following a predetermined number of trials, no transaction can occur, a market failure is reported.

3 Results

Key metrics: The first group of metrics quantifies the amount of welfare the agents accrued over time as a result of the transactions. There are two main dimensions of welfare: money and utility. Utility measures the amount of satisfaction the agents accumulated from their products consumption. It refers here to lifetime accumulated utility, so that it is a monotonic increasing function, both at the aggregate and at the individual level. Money is the most obvious metric. It can be seen as a form of *potential utility* as it can be used to buy and consume products. Money leaves the world when agents produce and there is no way to inject back more money into the system (i.e. our model does not allow for any endogenous growth). Therefore, at the aggregate level, money is a monotonically decreasing function. However at the individual agent's level, money can increase after a sale, though the overall trend will always be oriented downwards. Our welfare comparisons therefore look at the amount of utility that is obtained by the consumers, and the amount of money that is consumed by the production of goods. Total wealth is finally defined here as the sum of money and utility.

The second feature of the market to be examined is the amount of inequality it generates. The method used here to measure this inequality is the traditional Gini coefficient [1]. The Gini coefficient can be defined as twice the area between the Lorenz curve and the perfect equality line. As the data generated in this simulation give a polygonal Lorenz curve, a simplified method [4] is used to calculate this coefficient. It varies between 0 and 1 with 0 meaning perfect equality and 1 meaning perfectly inequality. So the closer to 1 the more unequal the market is. The Gini coefficient is used to measure inequality, not only for total wealth but also for money and utility separately.

Simulation Results: In our baseline simulations, the world was set with 50 agents, 10 products, an initial endowment of 500 money units per agent, and configured to keep a log of the 10 latests transactions for each product. Whatever the initial conditions we set in the model, the competitive market consistently and significantly produces more welfare at the aggregate level than the random market. This gain in welfare is mostly due to a gain in utility for consumers, in the order of 60% overall. Given that no money is produced in our model but only consumed (by the production of goods) or transferred between buyers and sellers (with no transaction costs), no difference significantly appears indeed in the amount of money that is left with the agents at the end of our simulations. This result, which confirms earlier results of Gode and Sunder (1993), is explained by a more efficient matching of consumers and products based on the preferences of the former. Each transaction therefore provides more utility to the consumer than what would statistically be achieved when the matching is purely random.

Nonetheless, our model allows for 2 possible modes of production choices by the producers: either *random* (agents choose which product to make essentially in a random way), or *informed* (agents choose which product to make based on the expected profit they could make with each product, itself computed based on their skills (which determine their costs of production) and the latest transaction prices observed on the market). When production choices are informed instead of random, considerably more money is left with the agents at the end of the simulations, indicating that the production is much more efficient (less money is wasted in producing goods for which the producer is less skilled). Overall, total cumulated production costs at the end of the simulations are 60% lower at the aggregate level when production choices are informed. In addition, one simulation out of 4 is interrupted before the end of the 50000 ticks due to the impossibility at some point to further match products with consumers' preferences. This result simply highlights the well-known benefits of specialization and comparative advantages theory.

Figure 2 reports statistics over the key indicators of the state of the economy at the end of the simulations (200 distinct simulations were run sequentially and the statistics reported in the figure are averages over the 200 simulations). The figure reports results for the two types of markets (random and competitive) and for the two distinct methods of production choices (random and informed). Note that in each run, the four possible combinations of markets and production choices strategies were tested successively with the exact same set of agents and products and the exact same initial conditions to ensure the comparability of the results.

What is more appealing in our results is the huge differences that appear between the two types of markets in terms of welfare distribution. At the end of each simulation, we compute the concentration of utility, money and their sum (representing the total wealth of the agents) across agents with a Gini index. A higher value of the index indicates that the utility, money or wealth is more concentrated (i.e. more unequal).

Figure 2 reports these indexes for each of the four combinations of markets and production choices strategies. It clearly appears that the competitive market leads to much more inequalities in the distribution of utility and money (hence wealth) than does the random market. With informed production, utility ends up 5 times more concentrated in the competitive than in the random market, revealing considerable inequalities in the distribution of utility between agents. Likewise utility, money is much more concentrated in the competitive than in the random market (the Gini index is 6.5 times larger). Given that these differences are robust to a variety of changes in the model and in its initial conditions as well as to a large number

INFORMED PRODUCTION

	Random	Competitive	% Difference
Total Utility	5857	9433	61%
Initial Endowment	25000	25000	
Total Money	23264	23242	0%
Total Money Spent	1736	1758	1%
Total Wealth	29122	32675	12%
Gini Utility	0,07	0,39	484%
Gini Money	0,01	0,10	638%
Gini Wealth	0,01	0,05	437%
Ticks per run	50000	50000	
Runs	200	200	
Failure rate	0%	0%	

RANDOM PRODUCTION

	Random	Competitive	% Difference
Total Utility	7092	9927	40%
Initial Endowment	25000	25000	
Total Money	20758	20772	0%
Total Money Spent	4242	4228	0%
Total Wealth	27849	30698	10%
Gini Utility	0,13	0,41	212%
Gini Money	0,04	0,15	279%
Gini Wealth	0,01	0,03	243%
Ticks per run	50000	50000	
Runs	200	200	
Failure rate	24%	23%	

% DIFFERENCE ACROSS PRODUCTION METHODS

	Random	Competitive
Total Utility	-17%	-5%
Total Money Spent	-59%	-58%
Gini Utility	-48%	-3%
Gini Money	-66%	-33%

Fig. 2 Table with simulation results

of successive simulations, we see these results as a proof that, caeteris paribus, a competitive matching of consumers and producers based on price competition leads to much more inequalities in welfare distribution than a purely random matching. With other words, a competitive market generates significantly more welfare but distributes it much more unequally between agents.

To help understand this phenomenon, recall that agents enjoy some utility from consuming products but they lose the money that corresponds to the price they had to pay for the products. Therefore, any transaction in the model distorts the distribution of utility and money by transferring some money from the consumer to the producer and some utility in the reverse direction. In the random market, consumers and producers are selected randomly. Therefore agents should statistically behave as producers and consumers with similar frequencies, so that transactions should statistically compensate for the distortions created by the previous ones, explaining why utility and money tend to be evenly distributed across agents in our random market. In contrast, in the competitive market, consumers and producers are matched based on their tastes and skills respectively. This implies that products have a propensity to go more frequently to the consumers that have more differentiated preferences across products as their taste for a given product is larger than that of most other agents. Provided that they do not hit their budget constraint, they will therefore be

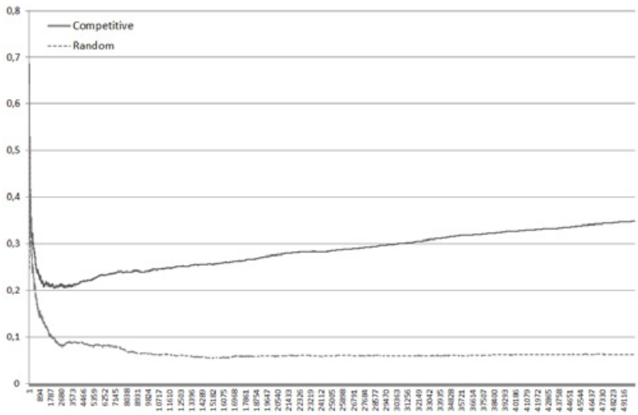

Fig. 3 Evolution of Gini index for the random and the competitive markets

	Baseline		Model 2		Model 3		Model 4		Model 5		Model 6	
Product selection	Latest bid		Skills/Tastes only		Latest transaction		Latest bid		Skills/Tastes only		Latest transaction	
Initial endowment	500		500		500		2		2		2	
Money increase	0.0		0.0		0.0		0.002		0.002		0.002	
	R	C	R	C	R	C	R	C	R	C	R	C
Utility	5857	9433	5863	8074	5865	7155	5877	9094	5878	8012	5875	7239
Money spent	1736	1758	1756	1687	1732	1639	1745	1689	1756	1658	1772	1683
Gini Utility	7%	39%	7%	24%	7%	13%	7%	29%	7%	21%	7%	14%
Gini Money	1%	10%	1%	5%	1%	2%	9%	44%	9%	29%	10%	18%

Fig. 4 Simulations with different endowments and production selection strategies

willing to pay a higher price than the other consumers to get the same product and will therefore be favored in the matching process (i.e. they will win their auctions more often and will therefore statistically spend more money and accrue more utility). Likewise, producers with more pronounced skills will enjoy lower production costs than most other agents for a given product and will therefore be willing to offer consumers a cheaper price for their products. As a result, they will be favored in the auctions and will tend to accrue more money. Therefore, distortions in utility and money tend to grow over time.

This is illustrated in Figure 3, which exhibits the evolution of the Gini index of utility along the 50000 ticks (iterations) in a simulation of the two markets with informed production. It shows that the Gini index quickly stabilizes over 0.06 in the random market, whereas in the competitive market it starts at about 0.2 (once some initial transactions have taken place) and then increases almost linearly until the end of the simulation, indicating that inequalities continue to grow over time. The differences are in fact even more pronounced with informed than with random production choices. A closer look at Figure 2 shows in fact that the Gini index of utility (money) is 50% (66%) smaller with informed than with random production in the random market, whereas Gini measures are less significantly impacted by production choice strategies in the competitive market. This suggests that all agents are better off with informed than with random production in a random market; probably because they all benefit from the advantage of specialization (i.e. goods are made by the most efficient producers and are therefore offered at a more affordable price, making them attractive to agents with less pronounced preferences). In the competitive market, informed production only generates benefits in terms of money distribution, which ends up less concentrated (i.e. less unequal) with informed than with random production, probably due to more efficient production leading to cheaper prices, so that voracious consumers need to spend less money to satisfy their preferences. Consumers with more differentiated tastes across products will still be favored in the auctions and will therefore continue to accrue more utility than the others, but they will need to spend less money to do so and the distortions in money distribution will therefore be smaller with lower production costs.

Finally, we investigated the impact of different behaviors (or degress of learning) of the agents in the competitive market by testing three different criteria for their selection of products to place offers for. In all the simulations reported above, sellers and buyers make their sale and purchase decisions respectively, based on the gap between the last bid that has been made over the same product by a competitor and their production costs or preferences respectively. This means that they choose a product that would maximize their expected benefit (in terms of net profit or net utility) if they were to win the auction by slightly outbidding the best competing offer so far. We explored two other product selection rules in the competitive model: The first alternative rule is the same as our default rule, except that the agents base their estimations of expected profit on the last transaction prices for the product rather than on the latest competing bid. These two strategies already point to different learning processes by the agents. The second alternative rule simply lets the producers select the product that they are best skilled at and can afford to produce, and the consumers select the product that they have the stronger inclination (taste) for and can afford to purchase. In this latter model, agents therefore don't use any information from the market to choose which product to sell or buy.

Figure 4 reports the key average metrics over 200 sets of simulations with each of these three production selection rules, and with two different money endowment strategies: our default one (a 500 unique endowment at the beginning), and a scheme in which agents receive a very small amount of money at the beginning (2 units) and are granted a 0.002 unit at the end of each tick. All of the reported simulations are based on informed production choices. Although our main findings about the superior performance of the competitive market over the random market and the higher inequality it leads to are rather robust, the magnitude of the differences between competitive and random markets vary from one simulation to the other. This is particularly due to inequalities in utility in the competitive market that significantly depend on the product selection strategy used. This result confirms that, as suggested by Gode and Sunder, the behavior of the agents can influence the distribution of profits (but more importantly of utility) in the competitive market. Combined with our first sets of results, this however indicates that institutions can generate inequalities (the competitive market always leads to more inequality than the random market design), but that the behavior of the agents can either reinforce or reduce these inequalities without affecting much the overall efficiency of the institutions in welfare creation.

A variety of tests has been performed to assess the robustness of our results to the main parameters of the model such as the agents' skills and tastes. These are randomly allocated at the start of each set of simulations (i.e. we always simulate the two markets (random and competitive) with the exact same parameters and agents). By running multiple sets of simulations, we test the sensitivity of the results to the allocation of tastes and skills. We further checked whether our results are robust to a variety of combinations of parameters, including the money endowment of the agents, the number of agents, and the number of products. They are. In the interest of space, the results of these robustness tests are available from the authors.

4 Conclusions

The objective of the present paper is to examine to what extent a competitive market compares to a random market in welfare creation and above all in welfare distribution among agents. Although the economic literature usually attributes a higher efficiency to competitive markets in maximizing social welfare, very little attention has been paid so far to the equality in welfare distribution resulting from the competition between agents, notably due to the difficulty in solving such problems analytically with a large number of products and agents. Agent-based simulations such as the one reported in this paper enables studying such emerging properties from individual interactions between agents. Various conclusions may be drawn from our results.

First, while it creates more welfare (utility and money) at the aggregate level, the competitive market distributes it much less equally. The competitive market structure is responsible for an inequality amplifying effect: goods become concentrated in the hands of greedy consumers and money in the hands of skillful producers. This result shows that institutions can in and by themselves generate inequality. Secondly, the behavior of the agents and the information they learn from the market (especially in choosing which products to bid on) have little effect on welfare creation (which is consistent with Gode and Sunder's results), but do significantly impact the distribution of welfare. This suggests that both institutions and agents share the responsibility for inequalities.

Various sources of randomness in real life are well-known to compensate for the positive feedback resulting from competition. For instance, among others, competing agents have limited time and cognitive resources to explore all possible offers, and many apparently irrational motivations undermine a lot of trading decisions. While a lot of casualties makes market to diverge from ideally competitive interaction in practice, how a fully random market could be practically designed in real life, although an interesting question, is out of the scope of this paper. However it is, for instance, quite plausible to imagine a computer-based market (such as ebay) where, after the consumers, hidden from each other, have indicated their preferences and the money they agree to pay and the sellers the price they are likely to accept, a transaction takes place based on one possible random pairing.

Despite the care we took in testing the sensitivity of our findings to arbitrary choices in design and in initial conditions, these results won't allow us to generalize our findings to any market design. These results describe a stylized exercise in which we compare a very aggressive competition-based market mechanism in the form of a double auction with a pure theoretical abstraction that represents a market in which producer and consumer allocation would be made purely on a random basis under a limited (and natural) set of constraints: the budget constraint, the no sale at loss rule for the producers, and the no purchase above utility value for the consumers. There are clearly large avenues for further analysis on how our results would resist to other market mechanisms and agent behaviors.

References

1. C. Gini. Measurement of inequality of incomes. The Economic Journal, 31(121):124126, 1921.
2. D.K. Gode and S. Sunder. Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. The Journal of Political Economy, 101(1):119137, 1993.
3. Binmore K. Natural justice. Oxford University Press, 2005.
4. Brian Slack and Jean-Paul Rodrigue. The gini coefficient. http://people.hofstra.edu/geotrans/eng/ch4en/meth4en/ch4m1en.html.

Economic Growth by Waste Generation: the Dynamics of a Vicious Circle

Benoît Desmarchelier, Faridah Djellal and Faïz Gallouj

Abstract The aim of this article is to explain the overall shape of the waste stream of durable goods in the US economy. For this purpose we construct an agent-based model of economic growth and waste generation and compare this latter variable to the amount of waste of durables generated in the USA between 1960 and 2008. The model highlights a particularly vicious circle in which waste generation becomes a positive factor of economic growth.

1 Background Literature and Issue

While several indicators of pollution exhibit a negative correlation with income per capita beyond a certain threshold of wealth [5, 9, 10], the waste stream from households keeps on growing, following the growth of income [13, 3, 12]. Thus, Panayotou [10] puts forth that ”the battle that can be won on the production side through structural change and technological progress can be lost on the consumption side through wasteful and unsustainable consumption patterns”. Given the principle that households cannot generate wastes without consuming, a good way to set the problem of waste generation by households is to ask how the continuous growth of consumption can be explained. Our attention will be particularly focused on the consumption of durables. This focus can be justified by the following reasons:

Benoît Desmarchelier
Lille 1 University, 59655 Villeneuve d'Ascq Cedex, France e-mail: benoit.desmarchelier@ed.univ-lille1.fr

Faridah Djellal
Lille 1 University, 59655 Villeneuve d'Ascq Cedex, France e-mail: faridah.djellal@univ-lille1.fr

Faïz Gallouj
Lille 1 University, 59655 Villeneuve d'Ascq Cedex, France e-mail: Faiz.Gallouj@univ-lille1.fr

S. Osinga et al. (eds.), *Emergent Results of Artificial Economics*, Lecture Notes in Economics and Mathematical Systems 652, DOI 10.1007/978-3-642-21108-9_11,

1. The consumption of durables is not essential for living.
2. The markets for durables keeps on growing in the developed economies despite a stable and high equipment rate of households.
3. Turned into waste, durables are particularly harmful for the environment and for human health.

According to Saviotti and Pyka [11], the continuous growth of consumption can be explained by the growing diversity of products produced by firms during the process of economic growth. This hypothesis is undoubtedly a part of the story, but it does not provide an explanation of the apparent non-satiety of consumers. According to Witt [16], this overall pattern of consumption can be explained by the following argument: ”with rising income, utilization rates [of durables] decline and depreciation rates [of durables] increase”. There are two interesting points here. First, consumers buy new durables without using up the old ones. According to Witt, their motivation is driven by the need of ”social approval” and ”aesthetic judgment”. Second, the life expectancy of durables decreases during the process of economic growth. These points underscore the fact that firms can deliberately choose to reduce the quality of their products and to push their fashion dimension. In response to this change in corporate strategy, consumers engage themselves in a race for buying the most up-to-date products, generating an ever larger amount of wastes. This explanation was already at the heart of Vance Packard’s work in the 1950’s [8] and it is at the heart of the throw-away society theory [15]. This explanation of waste generation is original because, unlike the usual view of pollution as a possible constraint for economic growth [7], it seems that waste generation is deliberatly used by economic agents as an engine for economic growth. The hypothesis of behavioral change of agents during the process of economic development makes it possible to construct an agent-based model in order to explore the consequences of this change for waste generation. Thus, the aim of this paper is to construct such a model and to observe if it is able to reproduce the general trajectory of waste streams of durables in developed economies.

2 The Model

The purpose of our economic growth model is to explain the growth of the waste stream of durables since the 1960’s. Following the hypothesis of the throw away society, it is possible to distinguish two main periods in the process of growth of an industrialized economy. In the first step, consumers are not equipped and the economic growth is pulled by the process of diffusion of durables. Then in the second step, a behavioral shift amongst firms and consumers is supposed to occur and to sustain the process of economic growth. We present these two periods of our model in the two following subsections.

2.1 A Simple Model of Economic Growth Pulled by Durables

The model includes three populations of agents: 10 firms of durables, 10 service firms[1] and 500 consumers. In the initial state of the model, 10% of the consumers have sufficient amount of money for buying a durable good.

At the beginning of each time step, consumers are mixed at random. Then, firms anticipate the demand they will receive during the period[2] and they consequently define their labor demand[3] and their level of production. Each consumer who is employed perceives a salary which is added to the amount of money he has not yet used up, and 70% of this budget is used for buying services. We assume that a consumer cannot buy more than one durable good per time. In addition, a consumer cannot consume two durables in the same time. Consequently, there is only two cases in which a consumer can buy a new good: when it is out of order or when its life time has expired[4].Unlike the rule applied to the purchase of durables, consumers try to exhaust their budget when they buy services. In other words, if a consumer has a sufficient amount of money for buying 3 services, he will demand 3 services to the provider he has chosen.

---

[1] The service firms are not necessary for the demonstration, but they are included in the multi-agent system in order to stabilize its dynamical behavior during the simulations. From this viewpoint, the model is consistent with the traditional conception of services as a sponge sector, which is likely to reduce the unemployment rate by offsetting employment losses in the manufacturing sector. (See the neo-industrial school in Gadrey [4]).

[2] Given D_{it-1}, the demand addressed to the firm i at time $t-1$, D^e_{it-1} the demand expected by i at time $t-1$, and β the firm's sensitivity to errors of anticipation, the demand expected by i at time t is given by the equation (1).

$$D^e_{it} = D^e_{it-1} + \beta[D_{it-1} - D^e_{it-1}] \tag{1}$$

[3] Given a_{it} the labor productivity in the firm i at time t and given S_{it-1} the level of stocks inherited from previous periods, the equation (2) gives the number of workers demanded by an industrial firm.

$$L^d_{it} = \frac{D^e_{it} - S_{it-1}}{a_{it}} \tag{2}$$

In the case of a service firm, there are no stocks. Thus, the labor demand of a service firm is given by the equation (3).

$$L^d_{it} = \frac{D^e_{it}}{a_{it}} \tag{3}$$

[4] The industrial firms are described by a parameter of "Quality" which is identical for each firm. The firm i has a chance to be chosen by the consumer j according to a competitivity index. Given q_{it} the quality of the good produced by the firm i and P_{it} its price, the competitivity index of the firm i is given by: q_{it}/P_{it}. the life time of the good is then fixed by a random draw in a Normal law $N(q_{it}, 1)$ but is constrained to be at least equal to one period. In addition, the good has a probability to be broken up. The price P_{it}, is determined by adding a mark up μ upon the average cost of production. Thus, $P_{it} = (1+\mu)(W_t/a_{it})$, with W_t, the wage of the time step t. We assume that the wage is equal to the average labor productivity in the economy.

After these processes of production, employment, selling and buying, industral or service firms have made positive or negative profits, which increase or reduce the amount of money they have accumulated. When a firm has a negative amount of liquidities, it is removed from the market and replaced by a new firm, which is a copy of one of the most profitable surviving firms. When an industrial firm has made positive profits, a share δ of these is invested in R&D. Concretely, given π_{it} the positive profit of the industrial firm i, $RND(0,1)$ a random draw in a Uniform law defined in the compact $[0;1]$ and ξ a scale parameter, an innovation occurs when the condition (4) is satisfied.

$$RND(0,1) < 1 - e^{-\xi \times \delta \times \pi_{it}} \tag{4}$$

If the innovation is successful, the productivity of the firm, a_{it}, is increased by the amount of the random draw $RND(0;1)$. Because the level of wage is fixed at the level of the average labor productivity of the economy, this process of innovation within industrial firms is able to generate a process of economic growth. Indeed, such an innovation can rise the level of wage and it produces a decrease of the level of price of the innovative firm (see footnote 4). The innovative firm becomes more attractive than the other non-innovative firms for customers because of its decreasing price. However, this hypothetic virtuous circle encounters rapidly its limits: as we assume that service firms cannot innovate (see Baumol [1, 2]), their prices increase at the same rate than the level of wage, making impossible any growth of real output in services. This phenomenon, also known as ”cost disease” (Baumol [1]), is generated in the part (b)[5] of the figure 1 where we see that the growth of the outpout in the service sector is entirely explained by inflation. Consequently, the process of economic growth is only based on the industrial sector. But this latter encounters also its own limit. Indeed, we see on the part (a) of the figure 1 that the equipment rate of the consumers in durables reaches a high level pretty early (79% after only 200 time steps), and because consumers cannot use two durables at the same time, the growth of the real output in durables is bounded. Thus, we obtain a situation of asymptotic stagnancy of the GDP (Figure 1 part (c)) and the stream of waste of durable (figure 1 part (d)) is also constant in the long run.

2.2 Beyond the Limits: the Throw Away Society

The model produces an asymptotically stagnant economy. This result is relevant because it was expected and also because historically, the firms of such an economy pulled by durables have encountered the problem of a bounded level of activity. Faced with a more competitive environment, industrial firms have to change the way they produce and sell goods. The basic idea is, for a firm, to expand its market by differentiation of its products and by a decrease of the lifetime of the product in order

[5] Each simulated time series presented in the figures of this article are the average of 10 simulation runs, each of these runs uses a different seed for the random draws in probability laws.

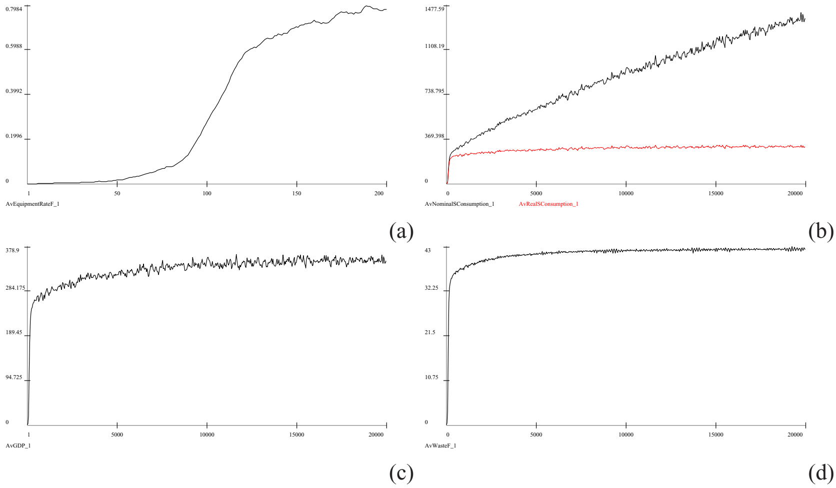

Fig. 1 (a)Diffusion process of durables, (b) cost disease. The **red black** represents the nominal service consumption and the **red curve** represents the real service consumption, (c) real GDP, (d) waste stream of durables

to increase the replacement rate of the consumers : the quantitative growth of the firm is henceforth based on the rise of the frequency of the buyings of the consumers. We can argue that the US firms in the post war period were already perfectly aware of the economic interest for them to produce such ”planned obsolescence” (Packard [8]). For example the General Motors designer, Harley J. Earl, stated in 1955 that ”our job is to hasten obsolescence. In 1934 the average car ownership span was 5 years; now it is two years. When it is one year, we will have a perfect score” (in Whiteley [15]).

Our model includes and connects these two kinds of planned obsolescence (style and lifetime). When 90% of the consumers are equipped with a durable good, industrial firms invest a part ϕ of their positive profits in R&D[6]. This new R&D seeks to improve the value of a parameter of style of the product. The ensuing innovation process is identical to the one implemented for productivity improvement. We consider that the style M_{it} of the good provided by the firm i is negatively related to the quality q_{it} by the logistic function represented by the equation 5.

$$q_{it} = q_{Min} + \frac{q_{Max} - q_{Min}}{1 + e^{\tau \times M_{it}}} \tag{5}$$

q_{Max} and q_{Min} are respectively the maximum level and the minimum level for quality attainable by a firm in our simulated economy, τ is a scale parameter. In other words, we assume that the improvement in the style dimension of the product is accompanied by a decrease in the quality of the product and thus, by a decrease of its lifetime expectancy (see footnote 5 for the determination of the product’s lifetime).

[6] We precise that the R&D expenditures for labor productivity improvement are also done

This product innovation brings a new source of waste generation: the disposal of stocks of obsolete products.

In response to this new marketing strategy of firms, the consumers use the style of their durable good ”as a social language to broadcast [their] status in society” (Whiteley [15]). Consequently, they can choose to throw away and to replace their durable good by a new one before the end of its lifetime. With M_{jt} the style of the product owned by the consumer j and with T the threshold at which the consumer believes that his product is outdated, a new durable good is bought when $M_{it} - M_{jt} > T$[7] .

The consequence of these two behavioral changes is the rise of the buying frequency, and thus the rise of the production of durables and of the production of waste. Thus, the GDP of our simulated economy (Figure 2 (a)), after a period of relative stagnation when the equipment rate of consumers approaches 90% (Figure 2 (b)) begins to grow again ”thanks” to these behavioral changes. The simulated GDP shows a logistic time path because the level of production is still upper bounded. Indeed, we have 500 consumers, and when there is a replacement of durable per period and per consumer, there is no way for growing further for industrial firms in our model. Of course, this is not the case in the real world, because there is always a process of creation of new sectors (Saviotti and Pyka [11]). But despite of this limit of our model, we get a logitic time path of the waste stream of durables similar to the real one illustrated by the US economy[8] (figure 2 (c) and (d)).

Our model suggests that waste generation has become an engine for economic growth since the advent of the mass consumption in the post-war period. The logistic time path of the waste stream of durables in our model seems to sustain the hypothesis of the throw away society as the main explaination for the growth of the real world’s waste stream. By now, all the firms and all the consumers of the model have adopted the throw away behaviors. But we can consider what would happen to the waste stream of durables if there was heterogeneity amongst firms or amongst consumers, that is to say a coexistence of throw away consumers and not throw away ones, and/ or such a coexistence amongst firms. Similarly, what kind of time path of waste stream would have emerged if only firms or consumers had adopted the behavior of the sthrow away society ? We will use the model for responding to these questions in the following section.

[7] Before the birth of the ”throw away spirit” (Packard, 1960[8]), an industrial firm was chosen by a consumer according to its performance index, q_{it}/P_{it}. But now, in the throw away period, this choice is based on a new index, M_{it}/P_{it}. thus, the quality of the product does not enter anymore in the product choice of a consumer. Indeed, this latter know that he will throw away the durable good before the end of its lifetime.

[8] Source for the US datas: urlhttp://www.epa.gov/epawaste/nonhaz/municipal/pubs/msw2008data.pdf A similar logistic pattern for waste generation is observed for France: urlhttp://www2.ademe.fr/servlet/KBaseShow?sort=-1&cid=96&m=3&catid=17571 But we can underline that the French data are agregates of all the types of wastes coming from households. To our knowledge, the only organization which provides long time series for waste stream coming specifically from durable goods is the US Environmental Protection Agency.

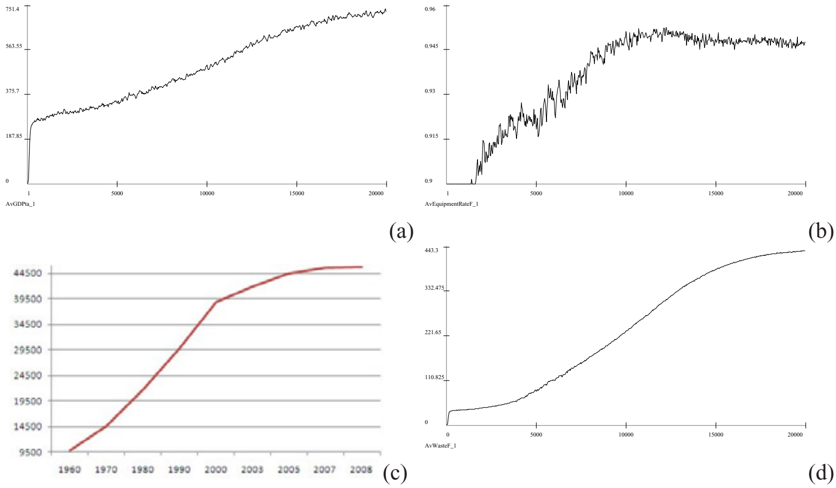

Fig. 2 (a)Real GDP with throw away society, (b) Equipment rate upper than 90% in the throw away scenario, (c) Waste stream of durables in the USA (in thousands of tons) since 1960's, (d) waste stream of durables generated by the model in the throw away scenario

3 Heterogeneous Agents and the Waste Stream of Durables

We have conducted various simulations with different proportions of throw away agents in our two populations of interest : indutrial firms and consumers. The figure 3 presents the waste streams of durables generated in the cases were (a) 10%, (b) 50%, (c) 80% and (d) 100% of industrial firms are throw away ones. For each of these situations, we have conducted simulations with 100, 200, 300, 400 and 500 throw away consumers (for a total population of 500 consumers) [9] .

These results allow us to sketch a number of conclusions about the respective roles of firms and of consumers in the generation of wastes.

1. The more widespread is the throwing away behavior amongst the agents, the more the production of wastes is important. This conclusion is true both for firms and for consumers, but we can underscore that we need at least one throw away firm for having an active role played by consumers in the production of wastes (see curve E0_C500 in the part (a) of the figure 3). In this sense, we can say that firms are the ones who have the initiative in the vicious circle of increasing wastes.

[9] The labels of the curves represented on the figure 3 have to be understood in the following way. E represents the number of throwing away firms (on a total of 10 firms) and C repsresents the number of throwing away consumers (on a total of 500 consumers). Thus, the label *AvWaste-E10-C100* means the average level of waste when there is 10 throw away firms and 100 throw away consumers.

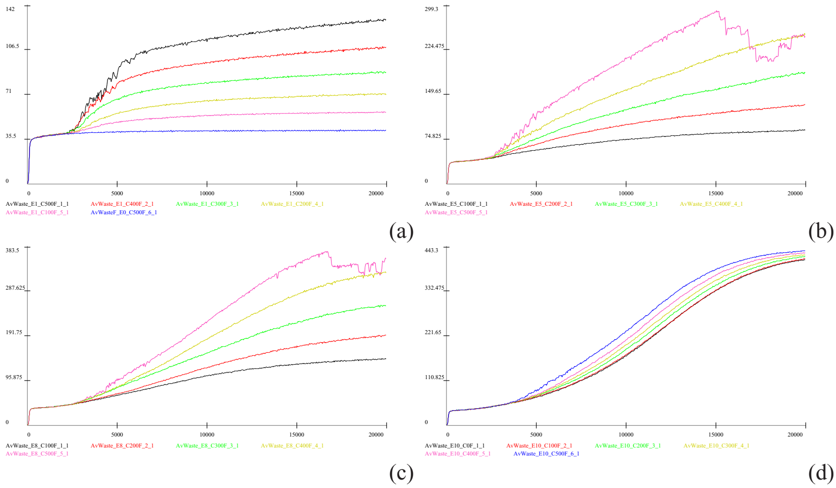

Fig. 3 (a)Waste stream with 10% of throw-away firms and with 0 throw away firm but 500 throw away consumers, (b) Waste stream with 50% of throw-away firms, (c) Waste stream with 80% of throw-away firms (d) Waste stream with 100% of throw-away firms

2. The parts (c) and (d) of the figure 3 reveal that we need to have an important proportion of throwing away firms for generating a logistic waste stream like the one in the USA. But in the same time, the part (d) of the figure (curve E10_C0) reveals that we can obtain a logistic time path without any throwing away consumer if all the firms have adopted the destructive behavior. This observation means that the model doesn't support the assumption that the consumers are the main responsibles for the production of wastes. Rather, this result allows us to question the relevance of the consumer information policies (products labelling...).

3. We also find that when all the consumers are throw away consumers (parts (b) and (c) of the figure 3) the waste stream becomes unstable. The reason for this unstability may lie in the process of entry and exit of firms from the market. Indeed, we find that the non throwing away firms are also those with the most important level of labor productivity. When all the consumers are throw away ones, a selection process occurs : the goods of these high productive firms are not sufficiently appealing for consumers. Thus, these firms eventually disappear and are replaced by new firms of the throwing away type, which are less productive. As far as in our model the level of wage is equal to the level of the average labor productivity, the disappearence of the most productive firms causes a rough decrease in wage and therefore in consumption. Of course, a less important level of consumption leads to a lower level of waste generation. This result is interesting because it shows a possible feedback of the behavior of consumers on the behavior of firms. In other terms, after having initiated the throw away shift, firms can be influenced in return by the consumers. Thus, the process of diffusion of the throwing away behavior becomes self-sustained.

4 Conclusion

Our simple model seems to be able to generate a time path of the waste stream of durables compatible with the one generated by the US economy since the 1960's. The interest of the agent-based approach here is to exhibit the importance of waste generation for the economic growth. Indeed, without the adoption of the throw away behavior by the agents, the process of economic growth is rapidly stopped in the model. Another interest of the agent-based approach is to show that it is the firms which have the initiative of the throw away cultural shift, but also to show that there is a possible feedback from consumers to firms wich induce a self-sustained diffusion process of the throw away culture amongst firms.

The vicious circle of an economic growth based on the production of wastes seems to sustain the call for a post-growth society (Meadows [6], Speth [14]). However, we do not take part in this debate because in its present formulation our model doesn't represent a current developped economy based on the tertiary sectors.

Several remarks are useful to qualify the interpretation of our results. First, the empirical datas we used are in thousands of tons while the datas coming from the model are in units of durables, there is therefore a problem of measurement variable. Second, there is a problem of temporal unit since we are not able to determine what is the correspondance between the real time and a time step in our model.

References

1. William J. Baumol. Macroeconomics of unbalanced growth: the anatomy of urban crisis. The American Economic Review, 57:415426, 1967.
2. William J. Baumol, Sue Anne Batey Blackman, and Edward N.Wolf. Productivity and American leadership: the long view. The MIT PRess, 1994.
3. M. A. Cole, A. J. Rayner, and J. M. Bates. The environmental kuznets curve: an empirical analysis. Environment and Development Economics, 2:401416, 1997.
4. Jean Gadrey. L'économie des services. La Découverte, 1992.
5. Gene M. Grossman and Alan B. Krueger. Economic growth and the environment. The Quaterly Journal of Economics, 110:353377, 1995.
6. Donella H. Meadows, Dennis L. Meadows, Jorgen Randers, andWilliamW. Behrens. Rapport sur les limites la croissance. In Halte la croissance ? Fayard, 1972.
7. William D. Nordhaus. Managing the global commons. The economics of climate change. The MIT PRess, 1994.
8. Vance Packard. The waste makers. Penguin Books, 1960.
9. Theodore Panayotou. Demystifying the environmental kuznets curve: turning a black box into a policy tool. Environment and Development Economics, 2:465484, 1997.
10. Theodore Panayotou. Economic growth and the environment. Economic Survey of Europe, 2:4572, 2003.
11. Pier Paolo Saviotti and Andreas Pyka. Economic development by the creation of new sectors. Journal of evolutionary economics, 14:135, 2004.
12. Nemat Shafik and Sushenjit Bandyopadhyay. Economic growth and environmental quality. time series and cross-country evidence. World Bank Policy ResearchWorking Paper WPS904, june 1992.
13. Magnus Sjöström and Göran Östblom. Decoupling waste generation from economic growth - a cge analysis of the swedish case. Ecological Economics, 69:15451552, 2010.

14. James Gustave Speth. The Bridge at the Edge of the World. Yale University Press, 2008.
15. Nigel Whiteley. Toward a throw-away culture. Consumerism, ’style obsolescence’ and cultural theory in the 1950s and 1960s. Oxford Art Journal, 10:327, 1987.
16. Ulrich Witt. Learning to consume-a theory of wants and the growth of demand. Journal of Evolutionary Economics, 11:2336, 2001.

Using Agentization for Exploring Firm and Labor Dynamics

A Methodological Tool for Theory Exploration and Validation

Omar A. Guerrero and Robert L. Axtell

Abstract Agentization is the process of rendering neoclassical models into computational ones. This methodological tool can be used to analyze and test neoclassical theories under a *more flexible* computational framework. This paper presents agentization and its methodological framework. We propose that, by classifying the assumptions of a neoclassical model, it is possible to systematically analyze their influence in the predictions of a theory. Furthermore, agentization allows the researcher to explore the potentials and limitations of theories. We present an example by agentizing the model of Gabaix (1999) for the emergence of Zipf laws. We show that the agentized model is able to reproduce the main features of the Gabaix process, without holding neoclassical assumptions such as equilibrium, rationality, agent homogeneity, and centralized anonymous interactions. Additionally, the model generates stylized facts such as tent-shaped firm growth rates distributions, and the employer-size wage premium. These regularities are not considered in the neoclassical model. Thus, allows the researcher to explore the boundaries and potentials of the theory.

Introduction

Agentization is the process of rendering neoclassical models into computational ones. This allows the researcher to set different models into a common technical arena where they can be compared with more flexibility than equation-based models. It facilitates the understanding of how microeconomic behavior affects system-wide dynamics at different scales.

This paper presents a systematic method for agentizing neoclassical models. Section 1 introduces the conceptual framework of agentization. In Section 2 we present

Omar A. Guerrero and Robert L. Axtell
Department of Computational Social Science, George Mason University e-mail: {oguerrer, rax222}@gmu.edu

S. Osinga et al. (eds.), *Emergent Results of Artificial Economics*, Lecture Notes in Economics and Mathematical Systems 652, DOI 10.1007/978-3-642-21108-9_12,

an example by agentizing a popular neoclassical model of firm dynamics. Section 3 shows how the flexibility of agentized models facilitates the exploration of additional empirical regularities that are not treated in the neoclassical model. Section 4 presents a summary of the results and the conclusions.

1 Agentization as a Methodological Tool

Agentization is the process of turning a neoclassical model into an agent-based one. In it, the modeler faces the problem of turning the main neoclassical assumptions into more realistic ones. Axtell [2] has referred to these assumptions as *the neoclassical sweet spot*: rationality, agent homogeneity, equilibrium, and non-interactiveness. These assumptions live in the core of all neoclassical models, and without them, mathematical tractability becomes impossible in most cases. Additionally, there exist other assumptions that are important to particular models. We can classify these assumptions based on how necessary they are to agentize the model.

First-order: These assumptions have to be necessarily confronted by the modeler. Without re-thinking them, the agent-based version cannot be implemented. Non-interactiveness is the best example. Without designing an explicit interaction mechanism, agents have no means to behave.
Second-order: They live in the core of neoclassical models and modifying them is important to explore the robustness of theories. However, they are not necessary for mapping the neoclassical model into a crude agent-based one. Equilibrium, agent homogeneity, and rationality are examples.
Third-order: These are assumptions that are specific to some models, but are not general to all neoclassical models. Some examples are time and space representations.

When agentizing a model, we move from the crudest possible computational implementation (one that only deals with first-order assumptions) to a richer one that considers more realistic or meaningful concepts (deals with second and third-order assumptions). This can be done systematically by studying the impact of first, second, and third-order assumptions on the evaluated output. Comparing the output of the agentized version with the neoclassical one yields information about the economic implications of holding certain assumptions. We summarize this process in three steps:

1. Implement a crude agent-based model with an explicit interaction mechanism, trying to keep all other assumptions unchanged.
2. Test the crude model for parameters sensibility and output differences with respect to the neoclassical model.
3. Repeat steps 1 and 2 for each of the second and third-order assumptions.

2 Agentization Example

We exemplify agentization by using a popular microeconomic model about the emergence of the Zipf distribution. The Zipf distribution and power laws are pervasive statistical regularities that occur in different social systems. We are rather interested in power-laws since the Zipf distribution is only a special case of them. In economics, it is well-known that they appear in the distribution of firm sizes. The most popular models are stochastic processes [7, 11, 12, 8]. Although these models have been important for understanding firm dynamics, they lack of deep microeconomic meaning. Gabaix's model [6] has received great attention due to its capability of producing a power-law distribution from microeconomic principles, which is normally not possible in neoclassical models due to the lack of heterogeneous agents.

2.1 Micro-Foundations

Although Gabaix [6] originally created the model for explaining the distribution of cities sizes, it has been applied to firms [10]. The model has a microeconomic specification of utility-maximizing representative workers, and firms with homogeneous production functions that offer a wage and a benefits plan. If the system is in equilibrium, there exists a master equation that describes the growth process of every firm in terms of benefits, utilities, and probabilities of exiting the market. The master equation defines an identical growth process for every firm that is independent from their initial size: Gibrat's law [7]. Finally, by explaining the variance of such process in terms of exogenous shocks, the author proves that the distribution is Zipf [14]. The main assumptions of the model are summarized in Table 1.

Recent studies of Gabaix's model [10] are concerned with the stochastic process of the master equation. Although this tells something about the properties of the distribution, it does not say much about the implications of economic behavior. We are interested in agentizing Gabaix's model and studying how different specifications of behavioral rules can impact the stylized fact of a skewed firm size distribution.

Table 1 Microeconomic assumptions.

Number	Assumption	Order
1	Firm i offers random benefits $b_{i,t}$ at time t	3rd
2	Workers that consume c get utility $u(c) = b_{i,t}c$	2nd
3	Workers consume all their income, thus $c = wage \equiv w$	3rd
4	Workers exit the market each period with probability δ	3rd
5	Workers and firms interact through a tatonnement process	1st
6	Workers choose the firm that solves: $\max u = b_{i,t} w_{i,t}$	2nd
7	Workers can choose firms only once	1st
8	Firm i's production function is $F = (N^o_{i,t})^\alpha (N^y_{i,t})^{1-\alpha}$	3rd
9	$N^y_{i,t} \equiv$ young workers that moved to firm i in period t	3rd
10	$N^o_{i,t} \equiv$ old workers that were already in firm i before period t	3rd
11	$w^y_{i,t} = \alpha \left(\frac{N^o_{i,t}}{N^y_{i,t}} \right)^{1-\alpha}$	3rd
12	In equilibrium, all the utility-adjusted wages of period t are equal	2nd

2.2 Crude Agentization

The crude version of the model implies that we keep all the assumptions of the original model, and only make the necessary adjustments to implement it into its agent-based form. For this, we define an interaction mechanism in which firms post their vacancies in a newspaper. A post contains the benefits plan and the wage offered by the corresponding firm. Agents look at the offers, and take the job that maximizes their utility. When a worker picks a job, the firm hires her and updates the post in the newspaper with the new wage as specified by assumption 11 in Table 1. The crude model has the following structure:

1. All agents are instantiated.
2. Workers are initially assigned to firms.
3. Workers exit the market with probability δ and replaced by new ones.
4. For each worker that has not changed jobs yet:
   a. Read the newspaper and pick the best job offer.
   b. The chosen firm hires the worker and updates her post.
5. Repeat the previous step until an equilibrium is reached.
6. Repeat from step 3.

2.3 Equilibrium Impossibility

In its agentized form, this equilibrium is achieved through an iterated search process:

- For each worker that has not migrated between firms in the previous periods:
  - Inspect the newspaper and pick the best job offer.
  - The chosen firm hires the worker and re-adjusts the wages of all the new workers that have been hired in the current period.
- Repeat as long as there are workers that change jobs in the current period.

Notice that the original model assumes an equilibrium in which all wage offers will be the same (adjusted to utilities). Mathematically, this works with continuous populations, since there is always an ε small enough such that the population of employees of firm i plus ε produces a wage that will balance the effect of the random benefits plan $b_{i,t}$. In the agentized case this is a problem since the population is discrete, which implies ε might not exist. This turns into an infinite cycle in which workers are continuously swapping jobs due to the offers imbalance. Therefore, we replace the equilibrium concept for an attainable one.

In equilibrium, workers that are hired by a different firm at period t get to keep their offers. This means that if a worker chose a firm, and then a new worker joined the same employer, the firm will only update the job offer for future candidates, and not those that already joined.

Notice that we have already modified a second-order assumption. This means that, although necessary, first-order assumptions are not always sufficient for producing outputs in an agentized model. We observe an interesting feature regarding the stochasticity of the agents behavior that helps to understand the nature of the firm size distribution: with higher probability δ of exiting the market, the firm size distribution is less degenerate[1]. Figure 1 shows that with more exit and entry of workers, the firm size distribution is closer to a power law.

We can understand this property in terms of the firms' marginal product, specified in assumption 11 in Table 1. The more $N^{y}_{i,t}$, the smaller $w_{i,t}$ becomes. All the young workers that arrived to firms in step t will become old workers in $t+1$. Assume that no worker exits the market, then those firms that hired more workers in period t, will offer higher wages in $t+1$. Therefore, if δ is low enough, larger firms do not reduce their wage offers substantially, which gives them an advantage to attract new young workers. In the long run, this degenerates into a few giant firms, and the remaining population of firms with a couple of employees. If δ is large enough, a significant

[1] In this case, degeneration takes the form of a few large firms and several firms with a couple of workers.

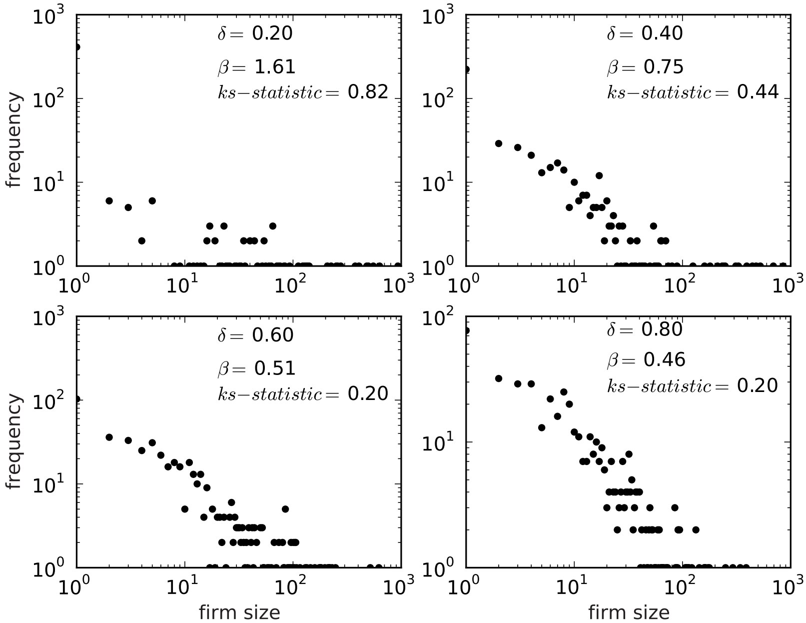

Fig. 1 Effect of exit probability on the firms size distribution output. Experiments were ran for a population of 500 firms and 12,500 workers. We ran 50 simulations for each $\delta \in [0.2, 0.4, 0.6, 0.8]$ and averaged the results for each case. We used maximum likelihood estimation for the functional form: $p(x) = x^{-\beta}$. We used the Kolmogorov-Smirnov (ks) statistic as a measure of how degenerate is the distribution. The plots are in log-log scale.

amount $N^o_{i,t}$ will leave the largest firms, producing a drop in their wage offers, which in turn makes the offers of the smaller firms more competitive.

This tells us a great deal about the role of labor mobility in the dynamics of firms. The system need enough noise to produce imbalances that will allow labor to flow, and prevent firms from establishing permanent advantages in their offers. The noise can come from $b_{i,t}$ and δ. On one hand, $b_{i,t}$ is not enough source of randomness because it requires that workers have low enough income so that they have incentives to change jobs. On the other, the effect of δ does not depend on incentives. It represents an exogenous source of randomness that, if is large enough, can produce the necessary imbalances to generate a stable firm size distribution. Therefore, the next steps towards agentization is to find what are the economic principles that produce enough endogenous fluctuations so that the output remains close to the one already obtained.

2.4 Labor Mobility and Time

We can think of the fluctuations produced by δ as labor mobility. Therefore, mobility of workers is crucial to the model. First, we notice that by releasing assumption 7 in Table 1 labors have free mobility, which does not modify the previous results.

A simple approach to endogenize the probability of exiting the market is by making it depend upon workers' age. However, time does not have a meaningful role in the neoclassical model, other than differentiating between periods. Therefore, we redefine the role of time, and adjust the related assumptions.

- Each time step t represents one month.
- Workers age as the model runs.
- Workers exit the market when they retire.
- When a worker exits the market, she is replaced by a younger one.
- Probability δ of exiting the market is super linearly proportional to the worker's age.
- Firms update benefits plans every year.
- A worker is considered young by firm i if she has been working there for less than a year.

Endogenizing retirement behavior can be the result of complex behaviors [3], which is out of the scope of this paper. Instead, we are interested in showing how simple changes in the nature of the assumptions affect the outcome of the model. Empirical evidence from labor markets [9] will serve to motivate the super linear relation between age and probability of retirement. This enables the model to generate a non-degenerate firm size distribution similar to the those in Figure 1 for high δ.

2.5 Heterogeneity and Local Interaction

We are left with two remaining second-order assumptions from the neoclassical sweet spot: agent homogeneity and rationality. Notice that in Section 2.4 we gave the workers the possibility of aging, which makes them heterogeneous. We can make firms heterogeneous assigning different α to their production functions. The effect of this modification is a degenerate firm size distribution shown in Figure 2.

Workers' rationality in assumption 6 in Table 1 requires that they maximize over all the population of firms. We can modify this by engaging the agents into local interactions. This is the k-lateral exchange mechanism proposed by Axtell [1], which is globally stable and has polynomial computational complexity.

Every month, workers read their local newspaper in order to obtain information about jobs and choose the best offer. The combination of advertisements each worker can see is different. Therefore, workers are heterogeneous and rationally bounded.

Local interaction produces outliers in the firm size distribution. This means that the shape of the distribution remains similar to the one from the previous results, but now there exist several large firms that makes the tail heavier. Interestingly, when we combine local interaction with heterogeneous firms, the effect of the former assumption is large enough to correct the degenerative effect produced by the latter one. Figure 2 shows the different distributions produced by these changes.

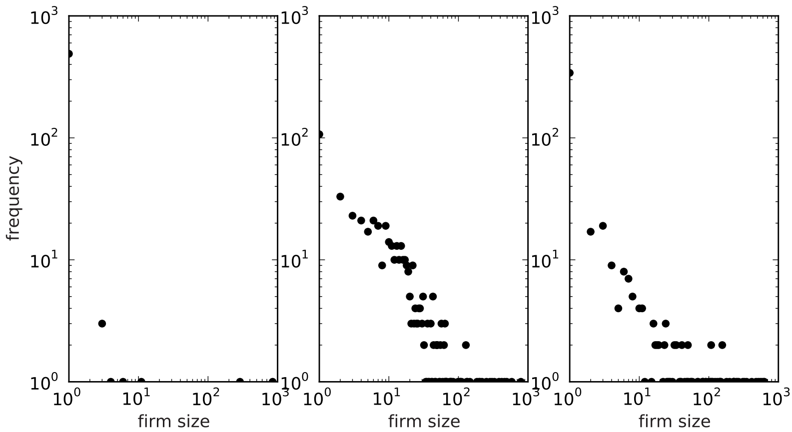

Fig. 2 The left plot is the firm size distribution when firms are heterogeneous and workers interact globally. The center plot is the firm size distribution when firms are homogenous and workers interact locally. The right plot is the firm size distribution when when firms are heterogeneous and workers interact locally. Simulations were ran for a population of 500 firms and 12,500 workers.

We can do one last refinement to the neoclassical sweet spot. The local equilibrium mechanism is still unrealistic because it assumes that some auctioneer places the workers in the best possible jobs until no further job changes can be made. In reality, a worker takes the best job offer that is available *at the moment* she is searching. This equilibrium refinement also eases the computational complexity:

When a worker searches for a job, she signs a contract that binds her to that employer for the duration of period t. Breaking the contract would cause the worker to incur in higher penalties than any alternative offer in the market. This mechanism guarantees that workers do not have incentives to relocate multiple times during period t. This *contractual equilibrium* also reduces the computational complexity to polynomial time $O(MN)$, where N is the number of firms and M is the number of workers.

The model is robust enough to preserve the previous results. Notice that now workers and firms are completely autonomous, since there is no central matching mechanism or auctioneer. The results emerge from pure decentralized local interactions. In summary, we have an agentized model that does not hold any of the assumptions of the neoclassical sweet spot, and with realistic time representation. Now we can proceed to explore other stylized facts that are not treated in the neoclassical model.

3 Limits Exploration

Two well known empirical regularities regarding firm dynamics are the growth rates tent-shaped distribution and the power-law relation between growth rate volatility and firm size [13]. Figure 3 shows that the model is able to emerge a tent-shaped firm growth rate distribution and a power-law relation between growth volatility and firm size. We must point out that these results were not possible in the model presented in Sections 2.3 and 2.4. It was only when we incorporated heterogeneity and local interactions that they became clear.

One of the oldest results in economics is the skewed income distribution. The agentized model is able to produce some skewness in the wage distribution, but with no clear signs of following a Pareto distribution. We can take advantage of the model's flexibility and incorporate a simple behavior that would explain such phenomena. It is well documented that wages increase with tenure inside firms [5]. We can use this fact to motivate a mechanism through which firms increase the wage of their employees on a yearly basis. Investigating how these agreements are achieved is out of scope of this study. Instead, we add a simple behavioral rule to the model.

Every year, firms offer a percent increase in the wage of all their employees. Therefore, workers with longer tenures have less incentives to change jobs.

Another well known stylized fact in labor economics is the employer-size premium [4]. This indicates that average wage is higher in larger firms. Figure 4 shows that both empirical regularities emerge while the previous are preserved. This are

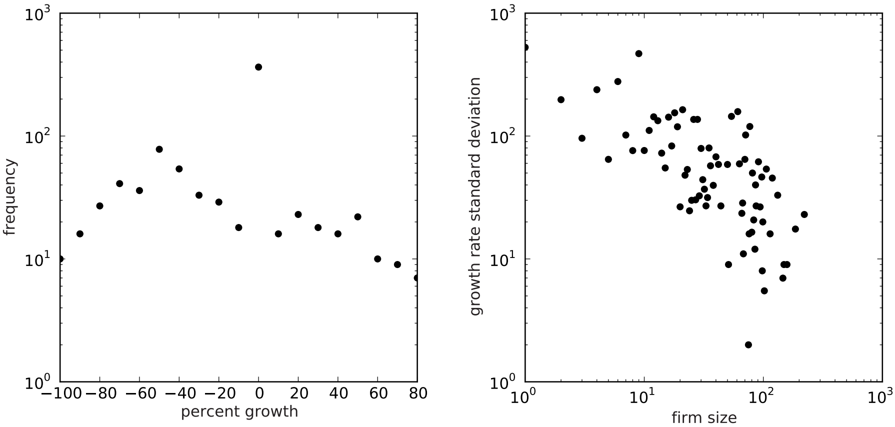

Fig. 3 The left plot is the annual growth rates tent-shaped distribution. The y-axis is in logarithmic scale. The right plot is the power-law relationship between annual growth volatility and firm size. The plot is in log-log scale. Simulations were ran for a population of 1,000 firms and 25,000 workers.

remarkable results because they show how with a simple behavioral addition, the model is able to produce clear results that were not contemplated in the original model. Furthermore, it sheds some light on the importance of contracts and bargaining in generating these stylized facts.

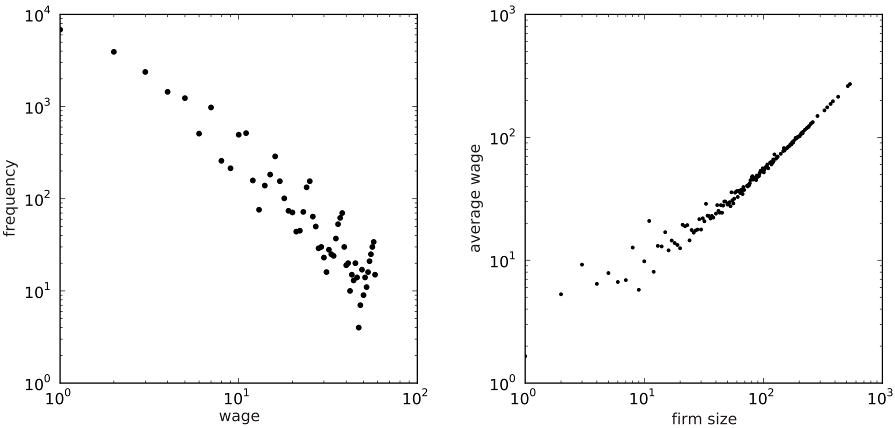

Fig. 4 The left plot is the wage distribution of workers. The right plot is the relationship between employer size and average wage. The plot is in log-log scale. Simulations were ran for a population of 1,000 firms and 25,000 workers.

4 Summary and Conclusions

This paper presented a systematic way in which neoclassical models can be rendered into agent-based ones. For this, it is necessary to understand what assumptions live in the core of every neoclassical model and which ones are particular to the specific model. The researcher can study how modifying each assumption changes the predicted results. Furthermore, due to the flexibility of the agentized model, the researcher can explore if the theory that is captured in the model is able to explain additional results that are not contemplated in its neoclassical form.

We provided an example of this process by agentizing Gabaix's model. We showed that the crude agentized model generates similar results to the ones of the original one. By modifying each of the four assumptions of the neoclassical sweet spot and providing a realistic time representation, the model is robust enough to generate a skewed firm size distribution. Furthermore, with a few behavioral additions, the model emerges additional stylized facts that are well documented in the empirical literature (e.g. firm growth rates distribution, income distribution, and employer-size premium). Future exploration can be extended to the inclusion of bargaining between firms an workers, vacancies information transmission through networks, and processes of firms' birth and death.

The main lesson form this study is that the methodological tool of agentization serves the researcher as a mean of testing and exploring economic theories that exist in the form of neoclassical models. Furthermore, it provides a systematic guide for building agent-based economic models that depart from the existing literature. We expect that with further developments, agentization can be generalized to other kind of economic models, and serve as reliable system for advancing scientific discovery.

References

1. Axtell, R. (2005). The complexity of exchange. *The Economic Journal*, 115(504):F193–F210.
2. Axtell, R. (2007). What economic agents do: How cognition and interaction lead to emergence and complexity. *The Review of Austrian Economics*, 20(2-3):105–122.
3. Axtell, R. and Epstein, J. (2006). Coordination in transient social networks: an agent-based computational model of the timing of retirement. *Generative social science: studies in agent-based computational modeling*, page 146.
4. Brown, C. and Medoff, J. (1989). The employer Size-Wage effect. *National Bureau of Economic Research Working Paper Series*, No. 2870.
5. Brown, J. (1989). Why do wages increase with tenure? On-the-Job training and Life-Cycle wage growth observed within firms. *The American Economic Review*, 79(5):971–991.
6. Gabaix, X. (1999). Zipf's law for cities: An explanation. *Quarterly Journal of Economics*, 114(3):739–767.
7. Gibrat, R. (1931). *Les Inégalités Economiques*. Sirey, Paris.

8. Levy, M. and Solomon, S. (1996). Power laws are logarithmic boltzmann laws.
9. Peracchi, F. and Welch, F. (1994). Trends in labor force transitions of older men and women. *Journal of Labor Economics*, 12(2):210–242.
10. Saichev, A., Malevergne, Y., and Sornette, D. (2009). *Theory of Zipf's Law and Beyond*. Springer, 1st edition. edition.
11. Simon, H. (1955). On a class of skew distribution functions. *Biometrika*, 42(3-4):425–440.
12. Simon, H. (1960). Some further notes on a class of skew distribution functions. *Information and Control*, 3(1):80–88.
13. Stanley, M., Amaral, L., Buldyrev, S., Havlin, S., Leschhorn, H., Maass, P., Salinger, M., and Stanley, H. (1996). Scaling behaviour in the growth of companies. *Nature*, 379(6568):804 – 806.
14. Zipf, G. (1949). *Human behavior and the principle of least effort*. Addison-Wesley Press.

Part V
Market dynamics

Firm Entry Diversity, Resource Space Heterogeneity and Market Structure

César García-Díaz and Arjen van Witteloostuijn

Abstract Evolutionary explanations of market structures have usually focused on the selection pressures impacted by a number of factors such as scale economies, niche width, firm size and consumer heterogeneity. How selection processes work in markets is highly dependent on the available firm type variation at entry. In order to explore the implications of different degrees of firm diversity at entry on posterior market selection processes, we develop an agent-based computational model. Results indicate that a proper understanding of market selection processes should, indeed, involve understanding the effects of firm type variation at market entry.

1 Background

Organizational ecologists have studied the evolution of organizational populations in a large number of industries, driven by the interplay of founding and exit rates [5]. However, little has been done on trying to understand how pre-producers organize resources until successfully entering the market [6], and even less on how specific entry patterns affect the configuration of organizational populations. Exceptions are the works of [19] and [16], who explore how delays in organizing resources and pre-production before market entry affect firm population trajectories. Carroll and Khesina [6] address components that affect founding rates, like age dependence and entry mode, and argue that diversifying (*de alio*) firms have significantly lower mortality rates than new start-ups (*de novo* firms) in many industries. It has also been argued how diversifying entrants outcompete new start-ups in turbulent envi-

César García-Díaz
Management Department (ACED), Faculty of Applied Economics, University of Antwerp, Prinsstraat 13, BE-2000 Antwerp (Belgium), e-mail: cesar.garcia-diaz@ua.ac.be

Arjen van Witteloostuijn
Management Department (ACED), Faculty of Applied Economics, University of Antwerp, Prinsstraat 13, BE-2000 Antwerp (Belgium), e-mail: arjen.vanwitteloostuijn@ua.ac.be

S. Osinga et al. (eds.), *Emergent Results of Artificial Economics*, Lecture Notes in Economics and Mathematical Systems 652, DOI 10.1007/978-3-642-21108-9_13,

ronments, but such an outcome is reversed when new start-ups have strong learning capacities [8].

Organizational ecologists have used density-dependent entry processes in an attempt to explain market evolution (e.g., [1, 3]), but have acknowledged their limitations to explain industry evolutionary patterns [2, 15]. Moreover, organizational ecologists have emphasized the role of market selection processes through the investigation of the determinants of hazard rates [14], which we could call the "second" selection process. Nonetheless, there has not been given the same attention to the "first" selection process as to how attempts of founding materialize in successful entrants in the marketplace (e.g. [6]), and much less as to how this relates to specific market configurations. The relevance of the second selection process depends on how much variation is generated in the first selection process. Explanations of this source of variation have been given with reference to market adaptation. As Levinthal [13] puts it: "For selection processes to be meaningful, organizations must exhibit stable sources of heterogeneity. Organizational adaptation may contribute to these stable sources of organizational level differences that, in turn, form the basis of differential selection" (p. 934). However, the organizational ecology work referred to above ignores the implications of firm variation at entry on the evolution of market populations. Hence, an account of how entrant diversity affect market structure is still largely unexplored. Here, we do so by adopting industrial organizations strategy of modeling the interaction among individual agents directly in the context of an organizational ecology analysis of market evolution, as computational models in industrial organization (e.g., [7]) have incorporated individual-level behavior to explain industry dynamics.

Specifically, our objective is to study the joint effects of entry variation and consumer heterogeneity on the evolution of market structures. We define *firm entry diversity* as the firm type variation at entry. We aim to explore the effect of entry diversity in settings with different consumer distributions over the set of available socio-demographic taste preferences, which we define as the *resource space* [18, 4, 9]. Following [21], we consider three types of resource spaces: a uniform distribution of consumers over the whole set of taste possibilities (a rectangular space), a space with a clear abundant market center and scarce resource peripheries (tailed space), and a space in which consumers are concentrated around a small set of taste preferences (condensed space). We develop an agent-based computational model that attempts to understand the determinants of market structure evolution from different combinations of firm entry diversity and consumer demand heterogeneity.

2 The Model

The model refers to a varying set of firms that compete for consumers in a resource space. The number of entrants is defined according to a density dependence model [5]. Firms leave the marketplace if their cumulative profits become negative. The number of entrants is drawn from a negative binomial distribution [12]. For the entry rate parameters values, we adopt those used in [9]. The resource space corresponds to a distribution of consumers b_k, $k = 1,2,,N$, where N represents the total number of consumer preferences (taste positions) in this space. The resource space is set according to a discretized beta distribution with parameters $\alpha' = \beta' = \eta$. The total number of consumers $\sum b_k$ is set around 5,600 consumers, but the specific number differs as to the specific value of η. The higher the η value is, the lower is the associated resource space heterogeneity, as is visualized in Figure 1.

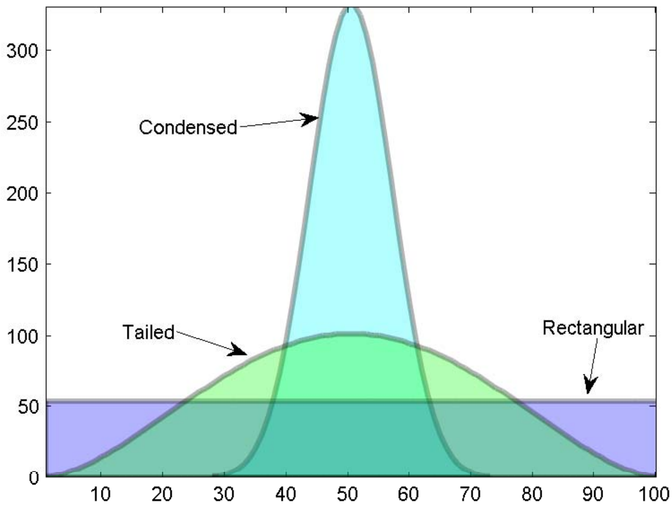

Fig. 1 Resource spaces with different degrees of heterogeneity. The x-axis represents the different consumer tastes. Rectangular spaces are the most heterogeneous, while condensed spaces are the most homogeneous.

Firms that enter the market select their location stochastically according to the probability distribution of non-served consumers, $\rho_{k,t}$, which is given by:

$$\rho_{k,t} = \frac{(1 - CBP_{k,t-1})b_k}{\sum_{i=1}^{N}(1 - CBP_{i,t-1})b_i}, \tag{1}$$

where $CBP_{k,t-1}$ represents the active consumer base percentage at position k at time $t-1$. This implies that, in unimodal spaces, firms will first attempt to target the consumers located near the market center. As the resource space gets populated, firms will gradually switch their focus towards the peripheries. At the beginning of the simulation ($t = 0$), there is no active consumer base. Hence, $\rho_{k,0} = b_k / \sum b_i$.

Firms have a two-piece cost function. One piece relates to the production costs, $C^i_{PROD,t}$, and the other accounts for the niche-width costs, $C^i_{NW,t}$. We assume that firms operate under one single technology, but differ in their fixed investment, which concedes different levels of scale advantages. Thus, we assume that firms derive their cost function from the same underlying industry-level long-run average cost (LRAC) curve.

Thus, following standard microeconomics [17], production levels for the firm, Q, are quantified through a Cobb-Douglas production function ([9]):

$$Q_{i,t} = F_i^{\alpha_1} V_{i,t}^{\alpha_2}. \tag{2}$$

Coefficients α_1 and α_2 correspond to production volume elasticities with respect to production factors F and V. Firms derive their production costs, $C^i_{PROD,t}$, from the LRAC curve ($\alpha_1 + \alpha_2 > 1$ for downward-sloping average costs). Production costs for the firm are calculated according to the usage of production factors F and V. That is, assuming that production factor prices are W_F and W_V, respectively, $C^i_{PROD,t} = W_F F + W_V V$ and the LRAC curve is calculated by solving the following optimization problem:

$$\begin{array}{ll} \min & C^i_{PROD,t} \\ \text{s.t.} & Q_{i,t} = F_i^{\alpha_1} V_{i,t}^{\alpha_2}. \end{array}$$

The factor F is fixed for each firm, but varies across firms according to their type at entry. Each F value corresponds to a unique efficient point Q of the LRAC curve (i.e., the solutions of Equation 3) and, consequently, to a different scale advantage potential. Choices depend on the assumed distribution of the Q possibilities, according to the resulting LRAC efficient cost points. The LRAC curve is normalized so that the minimum possible unit cost is 1 when $Q = \sum b_k$ ($W_V = 4.15$, $W_F = 2W_V$, $\alpha_1 = \alpha_2 = 0.7$). We then investigate to what extent the specific distribution of Q values influences the configuration of market structures, as depicted in Figure 2.

Niche-width costs represent the negative effect of covering a market with a large scope of consumer preferences. Niche-width costs are defined as a function of the upper and lower limits of firm i's niche [9], and a proportionality constant NWC:

$$C^i_{NW,t} = NWC \| w^u_{i,t} - w^l_{i,t} \|, \tag{3}$$

where $\|.\|$ represents the (Euclidean) distance between the two niche limits, and $w^i_{i,t}$ and $w^u_{i,t}$ represent the firm i's lower and upper niche limits. At every time t, firms move in the direction that is likely to increase their scale advantage. Thus, the niche center of each firm is updated at every time step. Firm i's niche center is defined as $nc^i_t = \|w^u_{i,t} - w^l_{i,t}\|/2 + w^l_{i,t}$. A consumer evaluates the offerings at her or his location k. The consumer buys from the firm that offers the lowest compound cost (price plus product dissimilarity) from the set of options $U^i_{k,t}$ (that is, considering all the possible i firms that belong to the set of offering firms at k, $S_{k,t}$):

$$U^*_{k,t} = \min_{i \in S_{k,t}} \left\{ U^i_{k,t} \right\} = \min_{i \in S_{k,t}} \left\{ P^i_t + \gamma \frac{\|nc^i_t - k\|}{N-1} \right\}, \tag{4}$$

where $P_{i,t}$ is firm i's price at time t, and γ is a constant that quantifies the effect of offer's distance from the firm's niche center to the consumer taste location (product

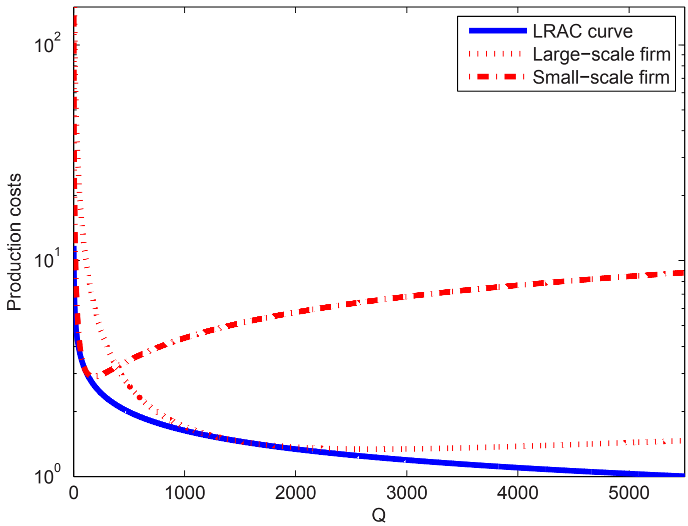

Fig. 2 LRAC curve with two different firm production costs. A small-scale firms curve reflects limited scale advantages, contrary to the large-scale firms curve.

dissimilarity). The distance-related effect is normalized to the maximum possible Euclidean distance in the model, $N-1$. For the sake of simplicity, we do not use demand functions, but define a limit price value for firm offers. The maximum price a consumer is willing to pay corresponds to a markup value over the smallest efficient firm in the industry, $P_{max} = (1+\varphi)\mathrm{LRAC(Q)}|_{Q=1}$, where φ is a markup factor.[1] Coefficient φ ranges between 0 and 1. In other words, consumers are allowed to bear a maximum cost, U_o, so that $U^i_{k,t} \leq U_o = P_{max}$. The amount U_o defines a cost-related participation constraint for consumers. If firm i expects to serve consumer k, its price P^i_t has to comply with

$$P^i_t \leq P_{max} - \frac{\left\|nc^i_t - k\right\|}{N-1}. \tag{5}$$

In order to reflect scale advantages, firms use a markup price over average costs (i.e., $(1+\varphi)C(Q)/Q$), provided that the markup price complies with Equation 5:

$$P^i_t = \min\left\{P_{max} - \gamma\frac{\left\|nc^i_t - k\right\|}{N-1}, (1+\varphi)C(Q)/Q\right\}. \tag{6}$$

Using the criteria of Equation 6, and the rivals' observed prices in the last time step, firms determine expected incremental demand and compute expected incre-

[1] This means that the only reason why a consumer will buy from a larger firm is that such a firm is more cost-efficient than the smallest possible firm in the industry.

mental profit, which, if positive, determines expansion to a new niche position (horizontal expansion). Firm expansion can also be vertical (i.e., directed at unserved demand in the current niche), following a similar expansion mechanism. Horizontal expansion implies targeting new consumer tastes. We assume that a firm will not attempt to target new consumer segments at every time step. We do so by associating horizontal expansion with an expansion probability coefficient, $ExpCoef$. Both expansion mechanisms are taking into consideration at every time step. Since firms take time to grow and find their operational point, they might undergo losses during the first time periods. For such a reason, firms are assumed to have an initial endowment for a number of E periods. The endowment is defined as a proportion Q_i/Q_{max} of the firm's fixed costs, where Q_i is firm i's quantity efficient point, and Q_{max} is the largest quantity efficient allowed for any firm in the model.

Entry diversity depends on a beta distribution with parameters $\alpha' = \beta' = \beta$ over the Q efficient points. We consider three alternatives: a median firm representation with a typical efficient capacity (with $\beta > 1$), a uniform distribution over the full Q spectrum ($\beta = 1$), and a bimodal case ($\beta < 1$), which resembles firm types in organizational ecology's resource-partitioning empirical studies (i.e., the generalist-specialist dichotomy; see Refs. [5, 10]).The lower the β value is, the more diverse the set of entrants becomes.[2] It is worth noticing that while the diversity among entrants varies according to β values, the density dependence entry parameters are kept constant.

We consider nine different scenarios, which result from the combination of three different entry diversity values ($i.e., \beta = 0.5, 1, 3$) with the three different resource spaces ($\eta = 1, 3, 30$). For each scenario, we run 30 simulations. Every simulation run was set for 10,000 time steps. The set of constant values in the model are $N = 100$, $x = 3$, $NWC = 195$, $\gamma = 60$, $\phi = 0.2$, $E = 12$ and $ExpCoef = 0.04$. Additionally, we impose some limits on the Q values. A firm not only needs a minimum cost efficient point (i.e., a minimum fixed investment) to enter the market, but also its most efficient operation point cannot imply covering more than half of the total market, $Q \in [Q_{min}, Q_{max}] = [10, \sum(b)/2]$.

3 Results

Results indicate that under high entry diversity ($\beta < 1$), the number of firms is maximized in tailed spaces, as is revealed in Figure 3. As expected, additionally, results show that highly heterogeneous resource spaces impede the formation of concentrated markets (see Figure 4). Organizational ecology's resource-partitioning theory indicates that unimodal spaces, taste heterogeneity, market-center scale economies and center-periphery scope diseconomies are required conditions for the emergence of dual market structures: then, generalist firms populate the market center, whilst

[2] In order to quantify diversity (D), we use the measure proposed by [20]: $D = \sum_{ij, i \neq j} d_{ij} p_i p_j$, where d_{ij} represents the distance (or disparity) between the Q efficient points of firms i and j, and p_i and p_j correspond to their proportions.

specialists survive at the market fringe [5, 21]. These dual markets are characterized by having a high market concentration and a non-declining number of firms [21]. Our results indicate that markets with high concentration and non-declining number of firms (as in resource-partitioning theory) can only emerge under high entry diversity. So, our results identify yet another condition for resource partitioning, one which has been ignored in the literature, to date.

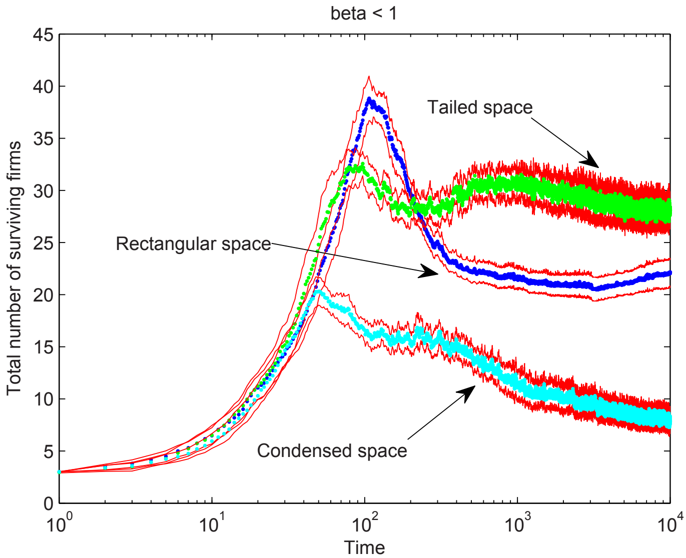

Fig. 3 Total number of firms according to the different resource spaces at high entry diversity. Border lines indicate confidence intervals at 95% significance

Property 1: Tailed spaces give conditions for resource-partitioning outcomes only when entry diversity is high.

As we might have expected, results also indicate that the more condensed the space, the closer the market structure outcome is to an oligopoly / monopoly configuration. This outcome is reinforced with decreasing entry diversity. In terms of market concentration, a closer look reveals that both high and low heterogeneity resource space conditions tend to be dominant over the entry diversity condition. Under low entry diversity, high resource space heterogeneity (see Figure 4) is associated with a clear market shakeout, where the firm population reaches a peak and then sharply declines. The market shakeout pattern has been observed in many industries ([7, 11]).

Property 2: Low entry diversity coupled with highly heterogeneous resource spaces generated a market shakeout characterized by an increasing number of firms that first reaches a peak, to be followed by a sharp decline.

As observed in Figure 5, rectangular spaces tend to generate unconcentrated market structures. Consistent with the above-mentioned explanations, in the case of

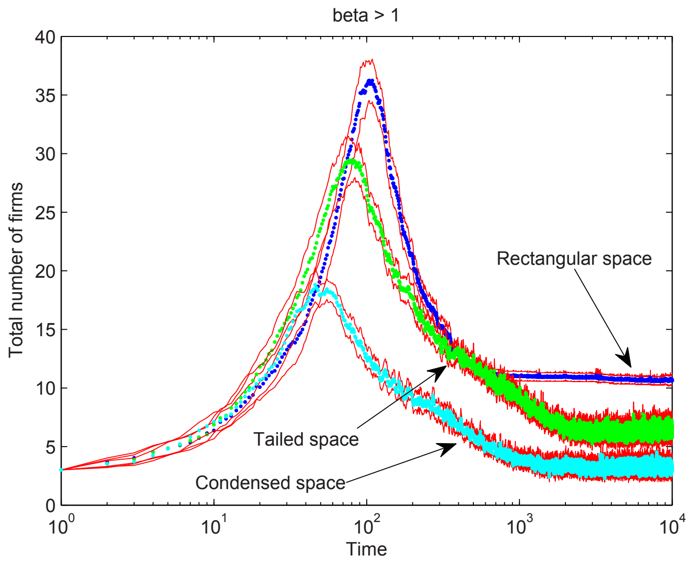

Fig. 4 Total number of firms under low entry diversity. Border lines indicate confidence intervals at 95% significance

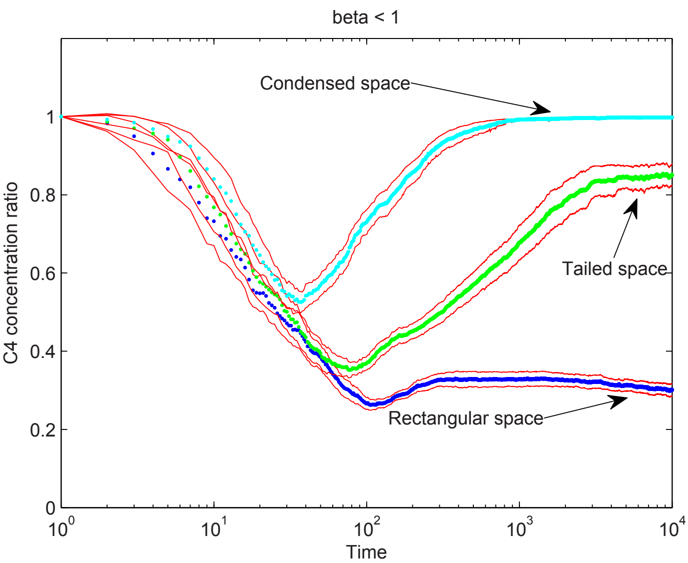

Fig. 5 C_4 market concentration ratios according to the different resource spaces at high entry diversity. Border lines indicate confidence intervals at 95% significance

rectangular spaces, the survivor population of firms decreases with decreasing entry diversity. Firms with medium and low-scale capacity survive mostly under high entry diversity ($\beta < 1$), while large and medium-scale firms survive under low entry diversity ($\beta > 1$). Therefore, rectangular spaces produce fragmented market structures with rather small firms under $\beta < 1$, and markets with rather homogeneous firms that split the market equally under $\beta > 1$.

Property 3: Rectangular spaces generate fragmented market structures composed of small firms under high entry diversity, and homogeneous markets with medium-sized similar firms under low entry diversity.

Tailed spaces present the highest variation of results in terms of the evolution of market concentration and firm population. A closer look at tailed spaces reveals how entry diversity relates to the speed of market consolidation. We consider this an important aspect since, as mentioned by [19], expectations of market actors are affected by the speed at which market conditions change. For instance, delays in capital appropriation for entrepreneurs and / or delays in organizational learning capabilities affect firm response to market variations, and hence market dynamics. Figure 6 illustrates how entry diversity affects the speed at which market concentration emerges. As we might expect, entry diversity slows down the speed of market concentration convergence.

Property 4: The rate of market concentration change increases as entry diversity decreases.

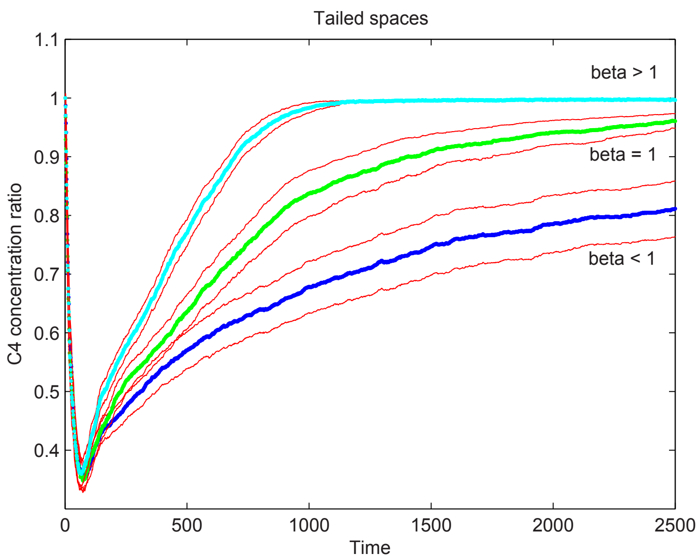

Fig. 6 C_4 market concentration ratios for tailed spaces. Border lines indicate confidence intervals at 95% significance

Finally, we refer again to tailed resource spaces. We explore how entry diversity undermines the selection forces of scale-driven competition. A common intuition would be that only firms with a large-scale capacity should triumph in the market center where the most abundant resources are. We categorize all surviving firms in three different classes: small-scale, medium-scale and large-scale firms, according to their efficient production points at the LRAC curve (ranges, respectively: $Q \leq 100$, $100 < Q \leq 1000$ and $Q > 1000$). In Figure 7, we observe that under low entry diversity, the size–distance-to-market-center inverse relationship is confirmed. That is, the closer to the market center the firm locates, the larger its size is. Although the medium entry diversity case ($\beta = 1$) allows for the existence of some small-scale firms, the bimodal entry case illustrates that the center is taken over mostly by medium-scale firms, and the periphery by both large-scale and small-scale firms. Large-scale firms at the periphery are dying firms that survive for a while due to accumulated profits, and have no option than to conquer additional demand to cover their large fixed costs. This also shows that not only the largest firms survive at or near the center: while selection favors the large-scale firms under the unimodal entry case, this is clearly not so in the bimodal entry case.

Property 5: In tailed resource spaces, the dominance of the largest-scale firms at the market center is weakened as entry diversity increases. In all cases, most firms at the center have a larger scale than most survivors at the periphery, but proliferation of small-sized firms is only observed under high entry diversity.

4 Conclusions

The simulation study presented in this paper illustrates how the outcome of market-selection forces heavily depends on the variation source at entry. Although the exact processes that influence firm entry diversity are beyond the scope of this work, we show how such an exogenously defined inflow variety to the market might influence the outcomes, and consequently the interpretations related to consumer heterogeneity-based explanations.

Results indicate that resource-partitioning processes require more than scale economies and a unimodal space with consumer heterogeneity alone. According to our findings, a resource-partitioning outcome could not even be obtained with a uniform distribution of entry types along all the possibilities of the LRAC curve ($\beta = 1$), and strictly needs a highly diverse entry inflow. Also, shakeout patterns seem clearly salient when both low entry diversity ($\beta > 1$) and rectangular space are present. Likewise, condensed spaces seem to always lead to oligopoly / monopoly configurations, irrespective of the degree of entry diversity.

Extensions of this work might consider comparative analysis of mortality rates according to the different scenarios, and additions to include entrepreneurial behavior that influences the earlier-mentioned “first selection” process of entrants, its diversity and subsequent effects on the market or “second selection” process.

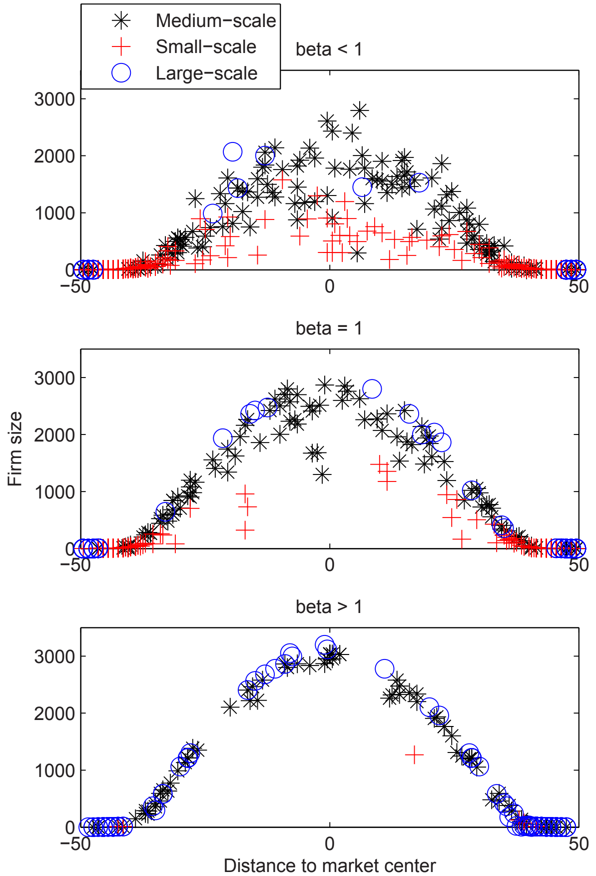

Fig. 7 Aggregated data on all simulation runs that illustrates the relative position (i.e., niche center) of each surviving firm with respect to the market center, its size (quantified in terms of sales) and its scale advantage according to the production cost function

Acknowledgements The authors acknowledge financial support from the Antwerp Centre of Evolutionary Demography (ACED) through the Odysseus grant, an FWO (Flemish Scientific Organization) funding initiative.

References

1. Barron, D. N. (1999). The structuring of organizational populations. *American Sociological Review*, 64(3), 421–445.
2. Barron, D. N. (2001). Organizational ecology and industrial economics: a comment on Geroski. *Industrial and Corporate Change*, 10 (2), 541–548.

3. Barron, D. N. (2001). Simulating the dynamics of organizational populations. In: Lomi, A., Larsen E. R.(Eds). *Dynamics of Organizations: Computational Modeling and Organization Theories*, Cambridge, MA: MIT Press, p. 209–242.
4. Boone, C., van Witteloostuijn, A., Carroll, G. R. (2002). Resource distributions and market partitioning: Dutch daily newspaper organisations, 1968 to 1994. *American Sociological Review* 67(3), 408–31.
5. Carroll, G. R., Hannan, M. .T. (2000). *The Demography of Corporations and Industries*. Princeton: Princeton University Press.
6. Carroll, G. R., Kessina O. M. (2005). The ecology of entrepreneurship. In: Alvarez, S., Agarwal, R., Sorenson, O. (Eds). *Handbook of Entrepreneurship Research: Disciplinary Perspectives (International Handbook Series on Entrepreneurship)*. New york: Springer, p. 167–200.
7. Chang, M-H. (2009). Industry dynamics with knowledge-based competition: a computational study of entry and exit patterns. *Journal of Economic Interaction and Coordination*, 4(1), 73–114.
8. Ganco, M., Agarwal, R. (2009). Performance differentials between diversifying entrants and entrepreneurial start-ups: a complexity approach. *Academy of Management Review*, 34(2), 228–252.
9. García-Díaz, C., van Witteloostuijn, A., Péli, G. (2008). Market dimensionality and the proliferation of small-scale firms. *Advances in Complex Systems*, 11(2), 231–247.
10. Kovacs, B., Carroll G. R. (2010). Niche width and scale in organizational competition: a computational approach. *Computational and Mathematical Organization Theory*, 16(1), 29–60.
11. Klepper, S., Simons, K. L.(2000). The making of an oligopoly: firm survival and technological change in the evolution of the U.S. tire industry. *Journal of Political Economy* 108(4): 728–760.
12. Lee, J., Harrison, J.R. (2001). Innovation and industry bifurcation: the evolution of R&D strategy. *Industrial and Corporate Change* 10(1): 115–149.
13. Levinthal, D. A. (1997). Adaptation on rugged landscapes. *Management Science*, 43(7), 934–950.
14. Lewin, A. Y., Weigelt, C. B., Emery, J. D. (2004) Adaptation and selection in strategy and change: Perspectives on strategic change in organizations. In: Poole, M. S., Van de Ven, A. H. (Eds). *Handbook of Organizational Change and Innovation*, p. 108–160. New York: Oxford University Press.
15. Lomi, A., Larsen, E. R, Freeman, J. H. (2005). Things change: dynamic resource constraints and system-dependent selection in the evolution of organizational populations. *Management Science*, 51(6), 882–903.
16. Lomi, A., Larsen, E. R, Wezel, F. C. (2010). Getting there: exploring the role of expectations and pre-production delay in processes of organizational founding. *Organization Science*, 21(1), 132–149.
17. Mas-Colell, A., Whinston, M. D., Green, J. R. (1995). *Microeconomic Theory*. USA: Oxford University Press.
18. Péli, G., Nooteboom, B. (1999). Market partitioning and the geometry of the resource space. *American Journal of Sociology*, 104(2), 11321153.
19. Ruef, M. (2006). Boom and bust: the effect of entrepreneurial inertia on organizational populations. In: Baum, J. A. C., Dobrev, S. D., van Witteloostuijn, A. (Eds). *Ecology and Strategy, Advances in Strategic Management*, Emerald Group, p. 29–72.
20. Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society Interface*, 4 (15), 707–719.
21. van Witteloostuijn, A., Boone, C. (2006). A resource-based theory of market structure and organizational form. *Academy of Management Review*, 31(2), 409426.

Time-Dependent Trading Strategies in a Continuous Double Auction

Shira Fano and Paolo Pellizzari

Abstract We model a continuous double auction with heterogenous agents and compute approximate optimal trading strategies using evolution strategies. Agents privately know their values and costs and have a limited time to transact. We focus on equilibrium strategies that are developed taking into account the number of traders that submitted orders previously, as well as the number of who will submit subsequently. We find that it is optimal to place increasingly aggressive orders, according to a roughly linear schedule, and test the resulting equilibrium for robustness and accuracy.

1 Introduction

The Continuous Double Auction (CDA) is one of the most commonly used trading platforms in financial markets. Trading optimally in such an environment is hard due to the sheer complexity of the decision options faced by traders, that choose quantity, type, limit price and timing of their orders based on past history and/or beliefs about others' behavior.

Given the analytical and computational problems involved in the analysis, models of CDAs usually make a number of simplifying assumptions, like unit-orders, no cancellation and little heterogeneity of traders. Foucault 1999 [5], Foucault et al 2005 [6] and Rosu 2009 [8] solve analytically models of a CDA with a flow of patient/impatient traders, under several additional assumptions related to the exogenous arrival of market orders and the presence of a "trading crowd" that absorbs any quantity outside of a fixed spread. A robust feature of these models is that impatient

Shira Fano
Dept. of Economics, Bocconi University Milan e-mail: shira.fano@gmail.com

Paolo Pellizzari
Dept. of Economics, University Ca' Foscari of Venice e-mail: paolop@unive.it

S. Osinga et al. (eds.), *Emergent Results of Artificial Economics*, Lecture Notes in Economics and Mathematical Systems 652, DOI 10.1007/978-3-642-21108-9_14,

traders always submit market orders whereas patient agents always prefer to issue limit orders.

We consider many types of buyers and sellers, with different evaluations and costs, respectively, that proxy varying degrees of impatience. We focus, in particular, on the time constraint faced by traders and assume that there is a finite amount of time to complete a transaction. Traders enter the market in random order and face the risk of missed execution that never occurs in the previously mentioned models. In equilibrium, they adapt their bids or asks to the number of traders that submitted orders previously, as well as the number of who will submit subsequently. This appears to be a rather realistic feature as presumably the trading behavior with hours to go is different from the one within few minutes from the closing bell. The seminal tournament among trading strategies in [9] suggested that "timing [...] is crucial for successful performance" and the two best programs, Kaplan and Ringuette, both took advantage of the current elapsed time. Interestingly, while simple timing information is useful, only a handful of programs tried more sophisticated approaches like predicting the responses of the opponents.

Our model has similarities with [7] where a two-tick model of a CDA is presented. In Parlour's paper, heterogenous agents can submit a market order or queue at the fixed best bid B or best ask A. In the first case, they transact immediately at a less favorable price, while a limit order will be executed (at better terms) with some probability that endogenously depends on the accumulated orders at the quotes and on the behavior of the traders that are to appear before the end of the session.

We extend the previous model in that traders are allowed to develop a time-dependent submission strategy and can submit limit orders *at any price*. We do not explicitly model market orders but, instead, allow marketable limit orders to be immediately executed. Hence, impatient agents seeking for immediacy can optimally post aggressive limit orders that will result in outcomes similar to the ones obtained by market orders.

To the best of our knowledge, such a model cannot be solved analytically and we resort to computational techniques to determine the optimal strategies for all traders (with different types). Formally speaking, we look for a Nash equilibrium of a game with stochastic payoffs, played by agents and by nature: the former pick a bid/ask when it is their turn to partecipate in the trading session; the latter samples the participants to the market and selects a queue that randomly affects the gains of the agents.

We compute the optimal actions using an evolutionary algorithm known as Evolution Strategies (ES) and find that, in equilibrium, agents submit more aggressive bids or asks as time elapses, in a roughly linear fashion.

The paper is organized as follows. The next section describes the CDA and the bidding functions used by traders. together with an illustration of ES. Our results are showed in Sect. 3 and the robustness of the findings is analyzed in Sect. 4, where we estimate the accuracy of the computed equilibrium strategies. We then conclude with some final remarks.

2 The Model

Agents interact using a continuous double auction as a market platform. Traders on both sides of the market sequentially submit offers to buy and sell one item of a traded asset and can accept offers at any time during a trading session. Whenever a buyers' bid is greater or equal than a sellers' ask, the agents conclude a transaction; otherwise, their offers are stored in two separate limit order books and remain valid for future transactions.

In the *bid book*, offers are sorted from the highest to the lowest limit price. The order is reversed in the *ask book*, so that the most competitive offers, called *best bid* and *best ask*, are always the first in the respective side. Both books are reordered whenever a new offer arrives. When a transaction occurs, the offers involved in the exchange are removed. A standard rule of precedence decides the transaction price, namely the price is set by the quote that was submitted earlier.

The protocol can be implemented in different ways and real markets exhibit a large number of subtle variations. To simplify the model, as customarily done in the literature, we do not allow agents to cancel or resubmit orders during a single session and assume no informational asymmetries in that all traders access the same time-related information.

An equal number N buyers and N sellers are sampled from a set $\mathscr{P}$ of traders. The ith buyer, $i = 1, \ldots, N$ can redeem one unit of the exchanged asset for $v_i \in V$. Hence, his profit is $v_i - p$ if he trades at the price p. No profit is cashed in the case of no execution. Analogously, the jth seller has a cost $c_j \in C$. The profit is $p - c_j$ if a transaction occurs at price p and zero elsewhere. Buyers and sellers can't exchange roles and a buyer (seller) who acquires (transfers) a unit of the asset can't sell (buy) it later in the same session.

At the beginning of each day all agent are active, meaning they have the right to post offers. We allow traders to make only unit orders, so they have to choose prices, but not quantities. Agents have a unique chance to submit a limit order and, hence, they can transact at most once per day.

We assume that the order of arrival of traders is randomly chosen by nature, according to a uniform distribution over all possible queues. This reduces the complexity of the strategy space, eliminating issues related to the choice of the time to act and orders' cancellation.

The main source of uncertainty, in this framework, is the position $t \in \{1, \ldots, 2N\}$ in the queue. Moreover, as the types of the agents (values or costs) are random, the profits of each agent depend, in general, on both his/her position in the queue *and* on the set of the other traders taking part in the auction.

Agents can be active for two reasons: either they are waiting for their turn to post an offer, or their offer is already on one of the books. The trading day ends when all $N + N$ agent have submitted their offer and the queue of traders waiting to place an order is exhausted. At the end of each day, all the remaining unfilled orders are cancelled and the books are erased. Differently from other models, traders have a limited time to transact, i.e., there are relatively few agents involved in each trading session. This can be interpreted as a thin market or, alternatively, as a device to model a short span of time in a longer day. In fact, realistically, at each point in time, only a handful of traders are competing on the books at viable prices. There is also a technical reason to restrict the number of traders: unless mechanisms to cancel orders are in place, offers will accumulate and the bid-ask spread will decrease to the point where there is a unique transaction price. In such a steady state, only intra-marginal agents trade, bidding and asking exactly the same equilibrium price. The resulting dynamics, like in [3], is extremely flat when a large number of traders is considered.

A strategy (or bidding function) for a buyer with value v is the vector $w(v) = (b_1^v, b_2^v, \ldots, b_{2N}^v) \in \mathbf{R}^{2N}$, whose t-th component is the bid to post when the agent has to act at position t in the queue. When it is not needed, we drop the value v (or c) and simply write w for a strategy. For instance, let the strategy used by a buyer be the vector $(0.4, 0.4, 0.4, 0.4, 0.4, 0.5, \ldots, 0.5, 0.95, 0.95) \in \mathbf{R}^{20}$. Then, the trader bids 0.4 if he acts as fifth or earlier in the queue; if not, he always submits 0.5, except if he is last or next to last in the queue, when he bids 0.95 (say, because the time is almost over.) We skip for brevity a detailed description of sellers' strategies, that clearly requires only trivial changes with respect to buyers.

Agents maximize the expected profit, that is dependent on their type (value or cost) and on their bidding function. Intuitively, a good strategy should be individually rational, that is, no bid should exceed the value of the trader. Moreover, it should strike the right balance in the so called immediacy-efficacy trade-off that is typical of a limit order market: the more aggressive the bid, the larger is the probability to trade but, conversely, the smaller is the gain.

Let $\pi(w_i, \mathbf{w}_{-i})$ be the expected profit of a buyer using the strategy w_i when all the other agents behave according to $\mathbf{w}_{-i}$. An equilibrium is a set $\mathscr{E} = \{w_i, i \in \mathscr{P}\}$ of strategies such that

$$\pi(u_i, \mathbf{w}_{-i}) \leq \pi(w_i, \mathbf{w}_{-i}),$$

for all i and all alternative strategies u_i. Loosely speaking, in an equilibrium, no agent has the incentive to change its strategy to increase his profits.

The determination of the equilibrium strategies is a hard computational task. Estimates of the expected profits can, to the best of our knowledge, be obtained resorting to simulation only. In other words, given a strategic profile $(w_i, \mathbf{w}_{-i})$, the quantity $\pi(w_i, \mathbf{w}_{-i})$ is an average over (many) sampled positions in the queue and over (many) sampled sets of trading partners.

2.1 Evolution Strategies

Evolution strategies (ES) are population-based optimization algorithms. Superficially, they look similar to the well-known genetic algorithms (GA), sharing broad ideas related to evolutionary adaptation, selection, mutation and cross-over of the members of the population. A modern treatment of GA is in [1] and an insightful introduction to ES is provided in [2]. ES are tailored to maximize real-valued functions of continuous variables, use deterministic selection in the creation of a new population and cleverly endogenize the mutation rate using specific evolvable meta-parameters in the chromosome. In contrast, the archetypical (and, perhaps, old-fashioned) GA uses binary encoding, stochastic selection and exogenously given mutation strength.

ES are conveniently described using the $(\mu/\rho \overset{+}{,} \lambda)$-ES notation, to be read "mu slash rho plus or comma lambda". μ denotes the number of parents that generate λ offsprings which are then evaluated and selected to form the next generation (again, with μ individuals). The value of λ usually exceeds μ to create selective pressure on the population. The recombination parameter $\rho \leq \mu$ refers to number of parents that are used to breed an offspring, so that, say, $\rho = 1$ means that descendants are clones of the (unique) parent and, at the other extreme, $\rho = \mu$ implies that an offspring is a function of all the μ parents. Finally, the "+" or "," in the notation refer to the way deterministic selection based on fitness is performed: in comma-ES, the μ parents for the next generation are deterministically selected among the λ offsprings picking the ones with highest fitness. In other words, *all* the individuals of the previous generation are discarded. In plus-ES, instead, the μ parents of the next generation are deterministically selected in the wider set of $\mu + \lambda$ solutions formed by both (old) parents and (fresh) offsprings. Strong elitism[1] is clearly at work in a plus-ES, as the best solution is never discarded and, indeed, only improving offsprings are included in the next generation.

In more detail, we describe how our $(\mu/\rho,\lambda)$-ES algorithm works for buyers having a specific value, assuming the strategies played by other agents are fixed in a given batch of s simulated sessions. Obvious modifications apply to sellers. In the following, z denotes a real standard normal random draw variable and $\mathbf{z}$ is a λ-dimensional standard multivariate normal vector. Each time z and $\mathbf{z}$ are used, they are resampled independently from the previous draws.

1. Set $g = 0$.
   Initialize the population $\mathscr{A}^{(g)} = \{(w_i^{(g)}, \sigma_i^{(g)}), i = 1, \ldots, \lambda\}$ of λ individuals ($w_i \in \mathbf{R}^{2N}, \sigma_i \in \mathbf{R}$). The vectors $w_i^{(g)}$ dictate the bid to issue for each of the $2N$ possible positions in the queue, while $\sigma_i^{(g)}$ is a meta-parameter related to mutation.
2. Repeat ("years" of s sessions, where $N + N$ agents trade in a CDA).
   a. Compute average profit $\pi_i^{(g)}, i = 1, \ldots, \lambda$, and let $\mathscr{F}$ be the set of indexes of the μ agents (out of λ) with the largest average profit in year g (comma-ES).
   b. For $l = 1, \ldots, \lambda$ do:
      i. Sample (without replacement) a random subset $\mathscr{F}' \subseteq \mathscr{F}$ with ρ elements. Set

[1] *Elitism* refers to algorithms where the elite, namely some set of best solutions, is guaranteed to survive to the next generation.

$$w = \frac{1}{\rho} \sum_{k \in \mathscr{F}'} w_k^{(g)}$$
$$\sigma = \frac{1}{\rho} \sum_{k \in \mathscr{F}'} \sigma_k^{(g)}$$

ii. Set $\sigma_l^{(g+1)} = \sigma \exp(\tau z)$.
iii. Set $w_l^{(g+1)} = w + \sigma_l^{(g+1)} \mathbf{z}$.

c. Set $\mathscr{A}^{(g+1)} = \{(w_l^{(g+1)}, \sigma_l^{(g+1)}), l = 1, \ldots, \lambda\}$ (this is the next generation)
d. Set $g = g + 1$

3. Until termination ($g \leq 250$).

Step 1 initialize a random population of λ individuals at generation $g = 0$. Observe that each member of a population is a couple of one strategy vector $w_i^{(g)} \in \mathbf{R}^{2N}$, with tentative bids to be issued in position $1 \leq t \leq 2N$ in the queue, and one real-valued meta-parameter $\sigma_i^{(g)}$, that controls the size of the mutations of the components of $w_i^{(g)}$. Successful evolution will, at the same time, produce individuals with effective w_i (in terms of average profit) and low σ_i (to reduce the chance of catastrophic mutations that preclude survival in future generations).

Step 2 is the main loop in the ES algorithm, repeatedly evaluates the average profit in s sessions and select the best μ individuals to form the parent population $\mathscr{F}$ in 2a. Then, λ offsprings are created (2b) by randomly picking ρ parents, taking averages of both the sampled strategies and meta-parameters (i), mutating first σ (ii) and then w (iii). In 2c, the just generated recombined and mutated individuals form the next generation $\mathscr{A}$, and the index g is incremented (2d).

Step 3 check whether some (more or less arbitrary) terminal condition is satisfied. In this paper, we stopped the ES after 250 years of s days, i.e., trading sessions.

The vigilant reader may notice that our description slightly differs from the one given in Figure 1 of [2], line 3, in that we start with a population of size λ to generate μ parents only *after* the average profits in a CDA are simulated. The two formulations are perfectly equivalent but allow us to keep constant the number of agents involved in the market. For the same reason, comma-ES were preferred to plus-ES, where, at times, $\mu + \lambda$ buyers with the same values would have to be taken into account. It is also noteworthy that all the (important) minutiae related to the trading in a CDA are under the hood in 2a, where we succinctly say that "average profits are computed". In particular, profits in 2a are computed in trading sessions with buyers and sellers of *any* type, whereas evolution occurs always within individuals with same values or costs: s sessions are used to compute the profits of all groups; based on these fitness measures, parallel ES algorithms evolve the strategies of all the agents of each type; such strategies are used in the next s sessions and the whole process repeats.

If desired, different strategies adopted by agents with the same values or costs can be interpreted as competing trading rules used by the same physical agent in several sessions. The agents then learn the best rules by evolving the most effective behavior within the set of their own trading strategies.

3 Computational Results

In this Section, we discuss our results for a representative set of parameters' values. Table 1 summarizes our choices:

Table 1 Parameters of the model. The upper part shows the choices relative to the agents, the lower part the ones related to ES.

Parameter	Value	Description
N	10	Number of buyers and sellers in a session
V	$\{0.05, 0.1, \ldots, 0.9, 0.95\}$	Set of values for buyers
C	$\{0.05, 0.1, \ldots, 0.9, 0.95\}$	Set of costs for sellers
$\mathscr{A}$		Entire population (20 agents for each $v \in V$ and $c \in C$)
s	200	Number of trading session in one generation
λ	20	Number of offsprings
μ	10	Number of parents
ρ	1,5,10	Recombination parameter
τ	$1/\sqrt{4N}$	Meta-mutation rate in ES

Each sessions involves $N = 10$ buyers and sellers. Hence, 20 agents are randomly selected and act in a session according to a random queue. As discussed before, a relatively low N must be used to avoid trivial dynamics and we used the intermediate size of the market described in [4].

A strategy w (bidding function) is then encoded in a vector with 20 components representing the bid/ask to post in each of the possible positions $t = 1, \ldots, 20$ in the queue. The realized payoff is a random variable whose mean is $\pi(w_i, \mathbf{w}_{-i})$, estimated taking the average profit over $s = 200$ sessions.

The values and costs of traders are sampled in the discrete and equal sets V and C. A standard approach for the computation of equilibria would require to solve a global optimization problem with 760 real variables: in fact, each strategy requires 20 entries and there are $19 \times 2 = 38$ different types.

As far as the ES parameters are concerned, we pick $\lambda = 20$ offsprings and $\mu = 10$ parents. We tried different ρs, to investigate the robustness of the search results with respect to different recombination mechanisms.

As outlined in Sect. 2, we decompose the task to find an equilibrium in 38 subproblems, iteratively maximizing the profits of each type in isolation over s sessions, keeping constant the actions of the other types. In such a way, we solve the 38 disjoint problems

$$\max_{w_i} \pi(w_i, \mathbf{w}_{-i}),$$

relative to agents having the same $v \in V$, where $\mathbf{w}_{-i}$ denotes the strategies played by all remaining agents.

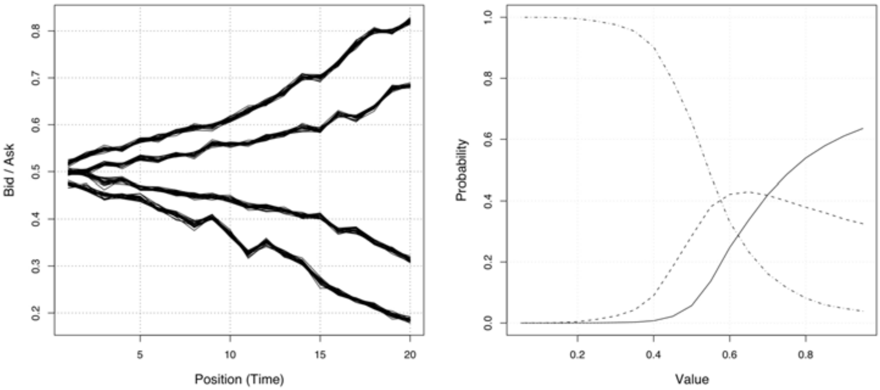

Fig. 1 Left panel: bunches of optimal bids for buyers with values $0.95, 0.75$ (top) and equilibrium asks for sellers with costs $0.05, 0.25$ (bottom). The results are relative to $\rho = 5$. Right panel: probabilities of immediate (solid), delayed (dashed) and missed execution(dash-dot) as function of the values of buyers. The results are relative to $\rho = 10$.

Figure 1, on the left, depicts a bunch of 20 equilibrium strategies (averaged over 50 runs of the ES) for some intramarginal buyers and sellers. In equilibrium, buyers increase their bid as time elapses and sellers, symmetrically, lower their ask. At the beginning of the session, when only the first traders in the queue have posted bids or asks, the orders tend to cluster around 0.5. This is done as plenty of time is available and delayed execution is likely, due to high number of agents that are still to come. Close to the end of the session, however, orders are much more aggressive, trading off gains for a higher execution probability. We numerically computed the equilibrium strategies of all agents in 50 independent runs, for each $\rho = 1, 5, 10$, totalling 150 experiments. In all cases, it is optimal to place increasingly aggressive orders, according to a roughly linear schedule and, hence, we have reasons to think that this is a robust feature of the evolved equilibrium strategies.

The right part of Figure 1 graphically illustrates some interesting features of the equilibrium outcomes. We plot the probability of immediate execution (solid), of delayed execution through a limit order (dashed) and of missed execution (dash-dot) as functions of the value of the trader (a symmetric plot for sellers is skipped.) The solid line shows that the most impatient traders, with high values, issue market(able) orders that results in immediate execution more than 60% of times. The remaining orders are filled with some delay even in the case of deeply intramarginal buyers. This is a novel and, we believe, realistic aspect of the model that is not present in some of other models of CDAs. Most of deeply intramarginal buyers, say when $v \geq 0.8$, transact almost certainly in equilibrium.

Moderately intramarginal agents trade less frequently with market orders, very often experience delayed executions (with a peak when $v \approx 0.6$) and there is an increased probability of no trade. Marginal traders issue market orders very rarely but still have a significant probability to have their orders filled at a later time. As expected, very few transactions involve deeply extramarginal traders that have a probability of missed execution that approaches 100%.

Overall, this picture is consistent with a world where impatient traders often (but not always) ask for immediate liquidity that is often provided using limit orders by patient agents. Buyers and sellers of intermediate strength roughly obtain the same proportion of immediate and delayed executions.

4 Quality of the Equilibrium and Robustness Test

We investigate in this Section the accuracy of the claim that the output previously described is an equilibrium set of strategies, focusing on the lack of any incentive to profitably deviate.

To evaluate the quality of the equilibrium found with ES, we use the notion of ε-equilibrium. A set $\mathscr{E}$ of strategies is an ε-equilibrium if, for all players i, for every strategy w_i and for every deviation u_i we have

$$\pi_i(w_i, \mathbf{w}_{-i}) \geq \pi_i(u_i, \mathbf{w}_{-i}) - \varepsilon,$$

where $\varepsilon \geq 0$ and there is no $\varepsilon' > \varepsilon$ such that the previous equation holds. We can interpret ε as measure of the distance of $\mathscr{E}$ from a (true) Nash equilibrium, meaning that at least one agent can obtain an amount ε of additional profits by abandoning his equilibrium strategy w_i to switch to u_i. That is to say that an ε-equilibrium restricts the profit attainable using a strategic deviation to be smaller than ε and coincides with a Nash equilibrium only if $\varepsilon = 0$.

Given $\mathscr{E}$ we can, in principle, compute the profits' differentials $\pi_i(w_i, \mathbf{w}_{-i}) - \pi_i(u_i, \mathbf{w}_{-i})$ for all possible deviations, keeping constant the strategies of all the other players $\mathbf{w}_{-i}$.

However, checking for all realizable deviations of vectors in $\mathbf{R}^{20}$ is unquestionably too demanding and an exhaustive comparison would just be practically impossible. Therefore, to solve this problem, we somewhat arbitrarily decided to reduce the complexity of the task by considering a subsample of deviations for some specific agents. We are aware that, by definition, we will underestimate the real ε but, at the same time, we provide useful simulation-based measures of the accuracy reached by ES in approximating the equilibria.

Without loss of generality, we will focus in what follows only on intramarginal buyers, sampling agents with probability proportional to their value. Deep intramarginal traders, in fact, obtain higher gains from trade and, on the other hand, can experience greater losses if they are not using an optimal strategy. Thus, we expect them to potentially benefit from strategic deviations.

Regarding which deviations to analyze, we consider three alternative ways to sample candidates, leading to three different families of ε. First, as in [Fano et al., 2011], we examine simple constant strategies where the bid does not vary with the position in the queue.[2]

Second, we create a deviation by adding a random quantity to each of the components of an existing strategy $w \in \mathbf{R}^{20}$. We, therefore, consider the vector w' whose jth component is $w'_{(j)} = w_{(j)} + \tilde{u}_{(j)}$, where $\tilde{u}_{(j)}$ takes the values -0.05, 0, 0.05 with equal probability.

Third, taking into account that we have 50 runs for each parameters' setup, we randomly replace an equilibrium individual strategy in *one* run with a deviation from *another* run. The rationale for this procedure is that bids that were optimal in one environment may also be able to provide good performance in another instance of the problem.

Table 2 Estimates of ε and ε^w, obtained using different methods to sample deviations, for different values of ρ. The p-value of the test that ε or ε^w are equal to zero are also shown, together with the value v of the buyer that deviates from the equilibrium. All entries, as explained in the text, are based on 5000 simulated deviations.

		ε	p-value	Value v	ε^w	p-value	Value v
$\rho = 1$	1	0.0057	0.416	0.85	0.133	10^{-14}	0.95
	2	0.0048	0.083	0.90	0.066	0.0004	0.95
	3	0.0021	0.062	0.95	0.110	10^{-5}	0.85
$\rho = 5$	1	0.0034	0.091	0.95	0.036	0.0006	0.80
	2	0.0019	0.027	0.95	0.022	0.0023	0.65
	3	0.0018	0.051	0.90	0.037	0.0008	0.80
$\rho = 10$	1	0.0025	0.340	0.85	0.032	10^{-5}	0.60
	2	0.0029	0.092	0.95	0.016	0.0004	0.90
	3	0.0010	0.014	0.85	0.073	10^{-5}	0.75

For each run, we have computed by simulation the profits arising from 100 independent deviations for each of the three methods described above. Table 2 shows the results, as a function of the value of the recombination parameter $\rho = 1, 5, 10$. In this and in the following table, every row begins with 1, 2 or 3, depending on the way deviations are generated. As an example, the row labelled "2" for $\rho = 5$ demonstrates that an additional mean profit of 0.0019 can be gained by the strongest buyer whose type is 0.95. With a p-value of 0.027, this amount can be considered different from zero at the 5% but not at the 1% level of statistical significance.

A look at the left part of the table, reveals that the size of ε, regardless of the type of deviations, is rather small and of the order of 10^{-3}. This means that, on average, minute profits can be obtained diverging from the equilibrium. Statistically,

[2] In detail, we require the deviations to be individually rational and sample the constant from a uniform distribution in $[0.45, v]$, where v is the value of the agent that deviates. Bids smaller than 0.45 would remain unmatched with high probability and, hence, would surely produce tiny, if any, additional profit.

the amounts are often significant at the 10% level, but never at the 1% significance level and, in general, the values are larger for deviations of the first type and smaller when generated using the third method. Moreover, it is apparent from the sizes of the various εs that the equilibria computed using $\rho = 1$ are less accurate that the one obtained with $\rho = 5$ or $\rho = 10$. Indeed, deviations from the equilibrium computed with $\rho = 1$ are rewarded with additional gains that more than double the one achievable the ones relative to $\rho = 10$.

The right part of the Table 2 reports the largest difference between the equilibrium profits and the ones coming from a deviation. Such a worst-case measure ε^w, that quantifies the largest (over $50 \times 100 = 5000$ simulations) single increment in profits after departing from the equilibrium, is at least one order of magnitude larger than any ε. This clearly shows that there are instances in our 50 runs in which profits can be substantially increased by deviating from the strategy computed by ES in that specific experiment. This is, again, particularly true for $\rho = 1$, confirming that the accuracy of the equilibrium for this value of the recombination parameter is rather low.

Judging from the sizes of ε and ε^w, the quality of the equilibrium solutions are similar when $\rho = 5$ and $\rho = 10$, with a preference for the latter recombination parameter that often produces smaller additional profits.

Table 3 Fractions of cases, expressed as percentages, in which deviations are profitable and significant at various significance levels α.

		DEV>ES	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.001$
	1	38.86	4.7	0.7	0.4
$\rho = 1$	2	25.36	2.0	0.2	<0.1
	3	7.58	0.7	0.1	<0.1
	1	30.00	3.0	0.4	0.2
$\rho = 5$	2	21.94	2.2	0.2	0.1
	3	15.11	1.0	0.1	<0.1
	1	24.68	2.4	0.3	0.2
$\rho = 10$	2	25.36	1.5	<0.1	<0.1
	3	6.59	0.5	<0.1	<0.1

Table 3 fine-tunes the information on the quality of the equilibrium evolved by our ES algorithm, displaying the fraction of cases in which deviations yield higher profits and how often the additional gain is statistically significant. Take again the row labelled “2” with $\rho = 5$: deviations are profitable in 21.94% of the simulations but the additional gains are significant at the 0.1, 0.01 and 0.001 levels only in 2.2%, 0.2% and 0.1% of the cases, respectively. Loosely speaking, with the exception of $\rho = 1$, that has already been shown to be a poor choice, most of the deviations do not increase the profits with respect to the optimal ES strategy and, even in the few cases in which mutations are profitable, the gains are only very rarely significant. There are, however, a small number of beneficial deviations that appears to produce significant gains, thus confirming the comments relative to the worst-case ε^w in Table 2.

5 Conclusion

We have analyzed in this paper the equilibrium strategies of heterogenous agents in a CDA where limited time was available to trade. This form of time-related pressure induces agents to submit more aggressive orders as time goes by, according to a roughly linear schedule.

The equilibrium strategies were numerically computed decomposing the task to solve separate problems (one for every type of trader) and using a Evolution Strategies algorithm. Satisfactory results are robustly obtained when the recombination parameter is $\rho = 5$ and $\rho = 10$, but not when $\rho = 1$.

We evaluated the accuracy of our results using the concept of ε-equilibrium that allows to gauge to what extent strategic deviations are profitable. As it is impossible in our framework to exhaustively test all possible variations of the bidding functions, we resort to randomization sampling agents proportionally to their strength and selecting strategies to be mutated in three different ways. Even taking into account some potential underestimation, the incentive to deviate quantified by ε appears to be minute in all occasions, emphasizing the overall quality of the approximated equilibrium strategies calculated by Evolution Strategies.

References

1. Ashlock, D. (2006). Evolutionary computation for modeling and optimization. Springer.
2. Beyer, H.-G. and Schwefel, H.-P. (2002). Evolution strategies: a comprehensive introduction. Natural Computing, 1:352.
3. Dawid, H. (1999). On the convergence of genetic learning in a double auction market. Journal of Economic Dynamics and Control, 23(9-10):1545 1567.
4. Fano, S., LiCalzi, M., and Pellizzari, P. (2011). Convergence of outcomes and evolution of strategic behavior in double auctions. Journal of Evolutionary Economics, pages 126. 10.1007/s00191-011-0226-4.
5. Foucault, T. (1999). Order flow composition and trading costs in a dynamic limit order market1. Journal of Financial Markets, 2(2):99134.
6. Foucault, T., Kadan, O., and Kandel, E. (2005). Limit order book as a market for liquidity. Review of Financial Studies, 18(4):11711217.
7. Parlour, C. A. (1998). Price dynamics in limit order markets. Review of Financial Studies, 11(4):789816.
8. Rosu, I. (2009). A dynamic model of the limit order book. Review of Financial Studies, 22(11):46014641.
9. Rust, J., Miller, J. H., and Palmer, R. (1994). Characterizing effective trading strategies : Insights from a computerized double auction tournament. Journal of Economic Dynamics and Control, 18(1):61 96.

An ACE Wholesale Electricity Market Framework with Bilateral Trading

Davide Provenzano

Abstract In this paper, an agent-based simulation model for a hybrid power market structure is presented. A bilateral transaction mechanism is combined with a uniform-pricing auction settlement in order to isolate the impact of medium-term bilateral contracts on market power and spot prices in a competitive wholesale market setting. First we describe the negotiation method for bilateral trading of energy and then introduce a new approach for bidding in the DA market based on the load duration curve. We find that, despite the conventional concerns, the foreclosure effect produced by the bilateral agreement between a generation and a retail business will not necessarily lead to higher prices, and will be manifested only according to the specific market characteristics.

1 Introduction

During the last 20 years, the traditional vertically integrated structure of the electricity market has undergone a worldwide transformation as many countries have begun to restructure and/or liberalize the sector. The new electricity markets have been characterised by an oligopoly of generators, very little demand-side elasticity in the short term, and complex market mechanisms designed to facilitate both financial trading and physical real-time system balancing. In many countries, therefore, the cost-based fully centralized dispatch has been abandoned towards a bid-based dispatch and/or a bilateral trading.

In order to figure out in advance the effects of these restructuring policies, more and more researchers have been developing electricity market models. In this research field, thanks to its capability of modelling large-scale complex systems better than the more traditional optimization and equilibrium models, the agent-based

Davide Provenzano
Dipartimento di Scienze Statistiche e Matematiche "Silvio Vianelli", University of Palermo, Italy
e-mail: provenzano@unipa.it

S. Osinga et al. (eds.), *Emergent Results of Artificial Economics*, Lecture Notes in Economics and Mathematical Systems 652, DOI 10.1007/978-3-642-21108-9_15,

(AB) paradigm has become very popular. Moreover, increasingly powerful computational resources as well as the development of toolkits that facilitate the implementation of agent based models in object-oriented programming languages have further pushed the development of the Agent-based Computational Economics (ACE) [1].

An AB model of the England and Wales electricity market [2] is implemented by Bower and Bunn [2] to compare different market mechanisms, i.e. daily versus hourly bidding and uniform versus discriminatory pricing. A very detailed discussion of the results of this model can be found in [3] where the basic model is validated against classical models of monopoly, duopoly, and perfect competition. The same model is also applied to the case of German electricity sector in [4] to analyze the impact of four mergers of large German utilities that were probable at the time of the study (and have actually taken place shortly after). A more detailed model of the New Electricity Trading Arrangements of England and Wales (NETA) is discussed in [6] where, differently from the precedent study, the authors explicitly model an active demand side and the interactions between the bilateral market and the balancing mechanism as a call market with pay-as-bid settlement. An extension of the same model and further analysis is presented in [7] where it is analyzed whether two specific generation companies in the England and Wales electricity market can manipulate market prices to increase their profits. The same model has also been applied in [5] to the analysis of the market power on the electricity market in England and Wales.

In the present work a simplified power generating sector has been modeled. A medium-term bilateral transaction mechanism is combined with a uniform-pricing auction settlement. Generation Companies (GenCos, i.e. companies possibly owning several plants with different generation technologies) and Suppliers (Supps, i.e. the agents purchasing from wholesale market in order to supply end-use customers) are not mandated to submit energy bids in the auction mechanism as they are allowed to sign bilateral medium-term (6 months or more) contracts. In this framework, we envisage a wholesale market with the demand side being price-taking and GenCos offering above the marginal cost.

An Independent System Operator (ISO) daily controls and coordinates the power exchange market whereas it does not have any role in the financial negotiations and settlements of a bilateral contract. In the auction market, clearing is given by the point in the (Euro, MW) space where demand and supply curves meet, resulting in a single price for the whole market, the marginal system price (MSP), assuming no operating and transmission constraints are violated.

Bilateral (B) exchange contracts are included in the day-ahead (DA) market, receiving a scheduling priority, and are concluded after a negotiation process based on a similarity measurement paradigm.

Within this agent-based framework, the proposed computer simulation model allows us to isolate the impact of medium-term bilateral trading on market power and spot prices in a competitive wholesale market setting.

---

[1] A complete survey of tools for agent-based simulation of electricity markets can be found in [9].

[2] A complete survey of agent-based electricity market models can be found in [8].

The hybrid structure of the model presented in this paper is innovative in two key regards. First, to our knowledge, no prior research has focused on a wholesale energy market where the uniform-pricing auction settlement is combined with medium-term bilateral trading. Second, this model implements a price bidding for the auction market based on the load duration curve.

The remaining structure of the paper is organized as follows. Section 2 outlines the composition of the proposed power market. Section 3 discusses the bilateral algorithm and the trading behavior in the B market. Section 4 describes the demand side and the supply side in the DA market. Section 5 defines the simulation settings. Results are presented in Section 6. Section 7 concludes.

2 Market Composition

The electricity production process can be divided into four different stages:

- Generation;
- Transmission along the high voltage network;
- Distribution along the medium and low voltage network;
- Supply to final customers.

While the network fixed sunk costs make the transmission and the distribution stages natural monopolies, at the national and regional level respectively, generation and retailing are instead potentially competitive, as the technology allows more than one firm on the market.

In the generation-retailing model under study N GenCos sell energy to M Supps in an electricity market where medium-term bilateral contracts are negotiated before a uniform-pricing auction takes place. Suppliers, in turn, re-sell the bought energy in the end-user market.

Firms are assumed to be capacity constrained. The generation capacity of the i-th GenCo at time t, $\overline{G}_i(t)$, is the maximum amount of power available to conclude B contracts, $G_i^B(t)$, and/or for bidding in the DA market, $G_i^{DA}(t)$[3].

$\Theta = \sum_{i=1}^{N} \overline{G}_i(t)$ is, therefore, the market capacity that includes a percentage β of reserve margin over the expected peak-demand, consistent with the normal operations of the most de-regulated energy markets.

Let q $(0 \leq q \leq 1)$ be the decision parameter to express the ratio of bilateral contracts with respect to the total trading activity of the generic agent; $\overline{S}_j(t)$ be the total quantity of energy supplier j is willing to buy at time t; $S_j^B(t)$ the total quantity of energy supplier j buys on the B market at time t; $S_j^{DA}(t)$ the total quantity of energy supplier j buys on the DA market at time t.

Then, the two markets together are described by the generation constraint

[3] Because of the short-term analysis the model assumes no investment over the whole simulation horizon and, therefore, $\overline{G}_i(t) = \overline{G}_i$.

$$\overline{G}_i(t) \geq qG_i^B(t) + (1-q)G_i^{DA}(t) \tag{1}$$

and the demand constraint

$$\overline{S}_j(t) \geq qS_j^B(t) + (1-q)S_j^{DA}(t) \tag{2}$$

$\forall i = 1,2,\ldots,N;\ j = 1,2,\ldots,M;\ t = 1,2,\ldots,T.$

Marginal costs are assumed to be constant throughout and no transmission constraints are assumed. Our model excludes blackouts due to extreme weather of technical failure.

The demand elasticity is set to zero (i.e. the demand bidding is considered independent from the market clearing price). This assumption is quite reasonable, since in the day-ahead markets worldwide loads have usually shown so far small changes with respect to energy prices. Demand is not subject to any type of curtailment and, therefore, loads are completely supplied both when they are involved in a bilateral transaction, and when they are traded in the DA.

3 The Match-Making of Agents in the Bilateral Market of Energy

B contracts are signed way ahead of time compared to the DA energy auctions and their agreed price could be, therefore, higher than the average DA market clearing prices (MCP). Yet, B contracts are financially safer for market participants because they can hedge against the high price volatilities of the real-time energy markets and their price, possibly higher than DA MCP, could be however convenient for the buyer during on-peak demand hours (during off-peak hours the convenience could work in the opposite direction).

Following [1] agents are assumed to represent their preferences in the electricity market by making use of a describing tree. Each node is labelled with a key factor a buyer considers in selecting its power supplier (i.e. price tariff, power quality, reliability, and customer service) and a set of weights, taken from the real interval [0,1], reflects the importance of branches on all levels of the tree. Therefore, node labels represent attributes of the energy system and branch weights represent their relative importance. The higher the weight, the higher the importance of that issue for the agent. Branches will always be labelled in lexicographic (alphabetical) left-to-right order while branch weights on the same level of any subtree are required to add up to 1.

Once the seller and buyer agents have constructed their trees, an algorithm measures the similarity bottom-up and recursively between every buyer-seller pair based on the weighted similarity of nodes belonging to the compared trees.

A similarity function $A(\sigma_{i,j}^{h,k})$ is defined to measure the similarity of nodes.

In the lowest level of the trees, $k = 1$, $\sigma_{i,j}^{h,k}$ is computed as follows:

$$\sigma_{i,j}^{h,k} = 1 - \frac{|x_i^{h,k} - x_j^{h,k}|}{max(x_i^{h,k}, x_j^{h,k})} \tag{3}$$

$\forall i = 1,2,\ldots,N;\ j = 1,2,\ldots,M;\ h = 1,2,\ldots,H;\ k = 1;$
where, $x_i^{h,k}$ and $x_j^{h,k}$ represent the value of nodes entitled to the h-th electricity factor at level k in the GenCo's and Supp's compared tree respectively.
The similarity of nodes, $\sigma_{i,j}^{h,k}$, is then adjusted to $A(\sigma_{i,j}^{h,k})$ by an arc function $A(\sigma_{i,j}^{h,k}) = \sqrt{\sigma_{i,j}^{h,k}}$.
The similarity of nodes at level $k+1$ is then computed by summing the similarities of nodes at level k, $A(\sigma_{i,j}^{h,k})$, weighted using the arithmetic mean of branches' weight, $(w_i^{h,k} + w_j^{h,k})/2$. The equation looks as follows:

$$A(\sigma_{i,j}^{r,k+1}) = \frac{\sum_{h=1}^{H} A(\sigma_{i,j}^{h,k})(w_i^{h,k} + w_j^{h,k})/2}{\sum_{h=1}^{H} (w_i^{h,k} + w_j^{h,k})/2} \tag{4}$$

$\forall k = 1,2,\ldots,K;$
where $r = 1,2,\ldots,R$ is the number of nodes at level $k+1$.

The similarity value of the compared trees ranges from zero to one. A value of zero (one) means that the seller-buyer pair under consideration is totally dishomogeneous (homogeneous).

3.1 The Bilateral Transaction Mechanism

At the end of the similarity measurements each buyer lists its potential sellers in decreasing order of similarity value and starts the negotiation process with its top ranked seller in the priority list. If the bilateral agreement or the time constraint for negotiation is not met, the agent starts a negotiation with the next agent in the list. The ISO does not have any role in the financial negotiations and settlements of a bilateral contract and no broker fees are assumed for bilateral contracting.

Each agent has minimum and maximum reference values for each of the negotiation subjects, meaning for each of the key electricity factors.
Let a ($a \in \{i, j\}$) represents a generic agent and h the issue under negotiation, then $x_a^h \in [min_a^h, max_a^h]$. When negotiating, the value agent a offers to its client/server at time t for issue h, depends on time as shown by the following function:

$$x_a^{h,\cdot}(t) = \begin{cases} min_a^h + \lambda_a(t)(max_a^h - min_a^h) \\ min_a^h + (1 - \lambda_a(t))(max_a^h - min_a^h) \end{cases} \tag{5}$$

with the first or the second equation to be used if the agent's offer is increasing or decreasing respectively.
Many functions can be used to define $\lambda_a(t)$. In this paper we use an exponential function for $\lambda_a(t)$ formulated as:

$$\lambda_a(t) = \left(\frac{t}{\bar{t}_a}\right)^{\frac{1}{\beta_a}} \quad (6)$$

where β_a defines the convexity of the function and $\bar{t}_a$ defines the upper limit of the negotiation time acceptable for agent a.
If the offer of agents i and j, $x_i^{h,\cdot}(t)$ and $x_j^{h,\cdot}(t)$, intersect at a time equal or less than $min(\bar{t}_i, \bar{t}_j)$, then an agreement is met for the issue h. A bilateral contract for energy supply is concluded when an agreement is met for all the attributes of the energy system involved in the negotiation process.

4 The DA Market

Once the time for bilateral contracting expires, the bidding activity starts in the energy market. The power capacity offered by any GenCo on the DA market will be up to the residual of its maximum capacity if a medium-term bilateral contract has been previously concluded.
On a hourly basis, each agent is allowed to submit a single offer for each hour of the next day. In the auction market the ISO collects bids from all generators and suppliers, sorts these offers in merit order (plants, starting from the cheapest to the most expensive one, are scheduled to generate until demand is met for each hourly period) and, matching the demand and supply curves into the (Euro, MW) space, clears the market at the price offered by the marginal unit on the merit order schedule: the marginal system price (MSP). Given the equilibrium quantity, Q^*, the ISO assigns full capacity, $q_i = G_i^{DA}$, to the n GenCos with bids below the MSP; the remaining capacity $q_i = Q^* - \sum_{i=1}^{n} G_i^{DA}$, to the GenCos with bids equal to the MSP[4]; whereas plants that have offered above the marginal plant's price are not scheduled to generate, and receive no payment. At the same time, Supps with bids above the MSP receive the exact quantity of energy demanded; Supps with bids equal to the MSP receive the remaining quantity of energy, and Supps submitting below the MSP reveice no energy at all. We assume that agents in the DA market estimate the quantity of energy to trade for each hour of the next day looking at the expected load shown by the energy load curve. In particular, the expected load $\overline{L}(t)$ is drawn, independently in each round, from a uniform distribution in $[L(t) - \varepsilon, L(t) + \varepsilon]$, where $L(t)$ is the value read on the energy load curve and ε accounts for the small uncertainty typical in day-ahead forecasting. Since load varies over time, the MSP

[4] In case of a tie, the GenCo is selected randomly.

fluctuates as long as more or less expensive generators become successful bidders and set the competitive price.

4.1 The Supply Side

We assume that any GenCo in the electricity market estimates its bidding price by making use of the load duration curve. Accumulating the time intervals for which load has a certain value during a period T (a year, a month, or a day) and plotting the ordered values of the load versus time will produce the load duration curve (LDC). Then, normalizing the values on time axis and reversing the axes, the resulting curve $F(X) = Pr(p \geq X)$ can be used for probability purposes. In fact, this curve estimates the probability that the average hourly demand takes on values greater than or equal to X in an hour of period T[5].

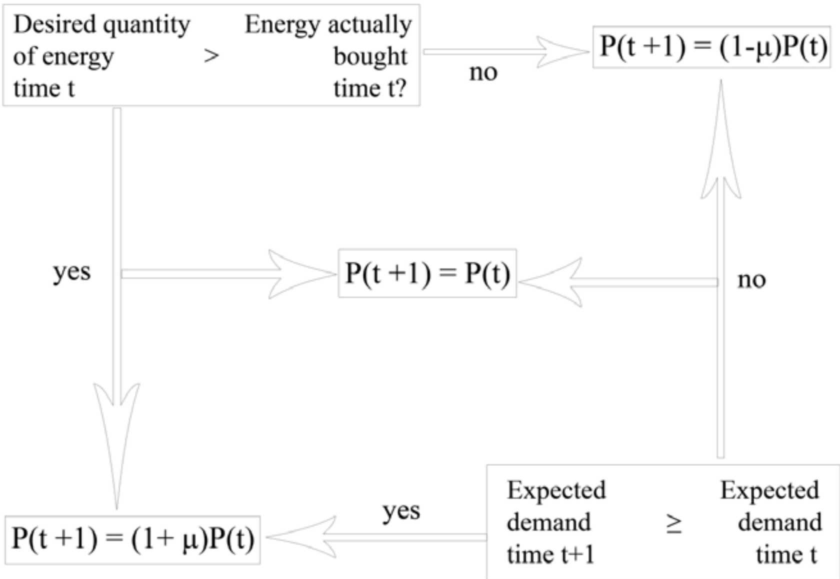

Fig. 1 The algorithm for agents compeating through price in the DA market.

[5] An example of this curve is given in Fig.3 for the case under study.

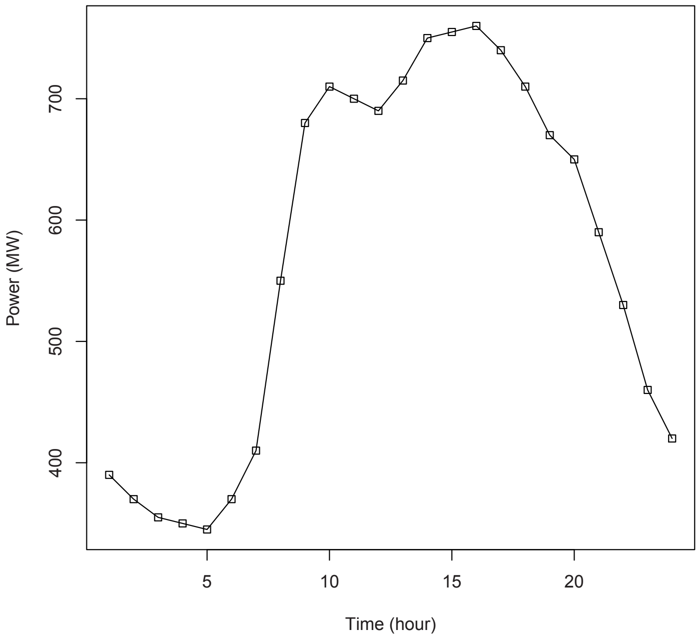

Fig. 2 A typical load curve for a Summer demand day.

4.2 The Demand Side

Supps decide the quantity of energy to bid by taking the pattern of variation of the electricity demand for each hour of the day into consideration.
The bidding price, instead, is fixed to the marginal price for the base load and then is it updated by following the algorithm shown in Fig. 1.

Given the marginal cost of the i-th GenCo at time t, $MC_i(t)$, its bidding price $p_i(t)$, is then defines as:

$$p_i(t) = MC_i(t)/g(F(X)) \tag{7}$$

where $g(F(X))$ is the mark-up function for the i-th generator, expressing how much the bidding price is increased above the marginal cost.
We assume a consistency rule such that the pricing strategies do not alter the sequence of plants in the marginal cost merit order: a more expensive plant will never undercut the bids of a less expensive one. The algorithm allows Supps competing through price to decide on their bids, at each iteration, looking at the quantity of energy actually bought for the bidding hour t (compared with the quantity of energy desired for the same hour), and considering the expected demand of energy for the next period of time (hour t+1).

If, for instance, the desired quantity of energy at time t is greater than the quantity actually bought, the difference might depend on a bidding price lower than or equal to the MSP. Therefore, if the demand at time t+1 is expected to be higher than the demand at time t, the agent increases the chances of being a successful bidder by increasing the bidding price. μ is the parameter used to update the price in accordance with the expected increase in the energy demand.

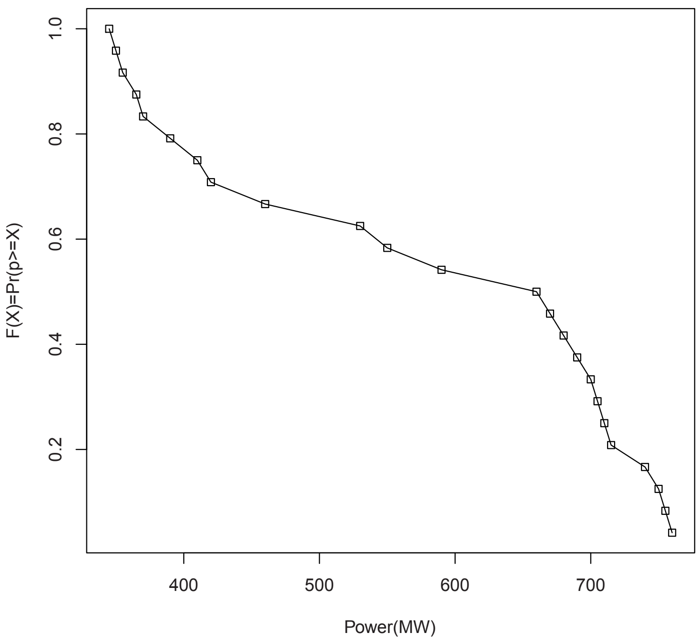

Fig. 3 The reversed load duration curve for probabilistic applications.

5 Simulation Settings

Simulations are defined in terms of iterations of trading days, each one for a set of 24 hourly periods. To keep things simple, we only consider one-level trees where the energy attributes to be negotiated are energy price and energy quantity. We have assumed firms in each segment (generation or supply) to have the same market capacity. Hence, the individual GenCo capacity is $\theta^{GenCo} = \Theta/N$ and the individual Supp capacity is $\theta^{Supp} = \Theta/M$. Agents, however, hold different technologies

and, therefore, GenCos have different marginal costs, while Supps have different marginal benefits. Marginal costs and marginal benefits have been used to set the values of min_i^h and max_j^h respectively for the bilateral transaction mechanism.
β, meaning the percentage of reserve margin, has been set equal to 20%. Therefore, assuming the peak load around 760 MW, the total capacity installed in the power market is equal to about 900 MW. The variation of daily energy load ε has been fixed to 3%.
Simulations have been run with two different market structures. The first composed of 2 GenCos ($N = 2$) and 3 Supps ($M = 3$), and the second with 3 GenCos ($N = 3$) and 2 Supps ($M = 2$).
Loads agents consider to define their strategies are depicted in Fig. 2, whereas Fig. 3 shows the reversed load duration curve built as described in §4.1.
The ACE wholesale electricity market framework developed in this paper has been implemented in Java using AnyLogic.

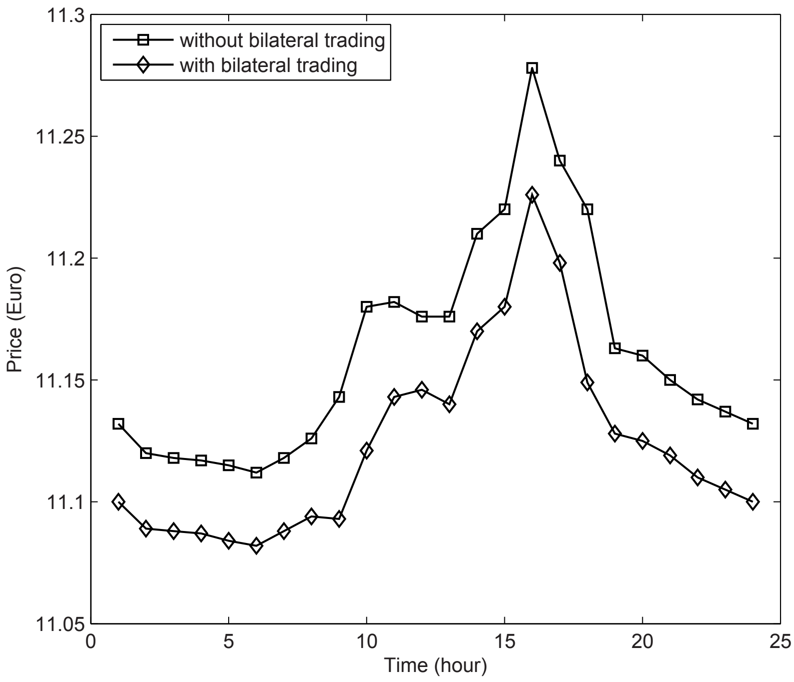

Fig. 4 The equilibrium price in the first DA market structure with and without bilateral negotiations.

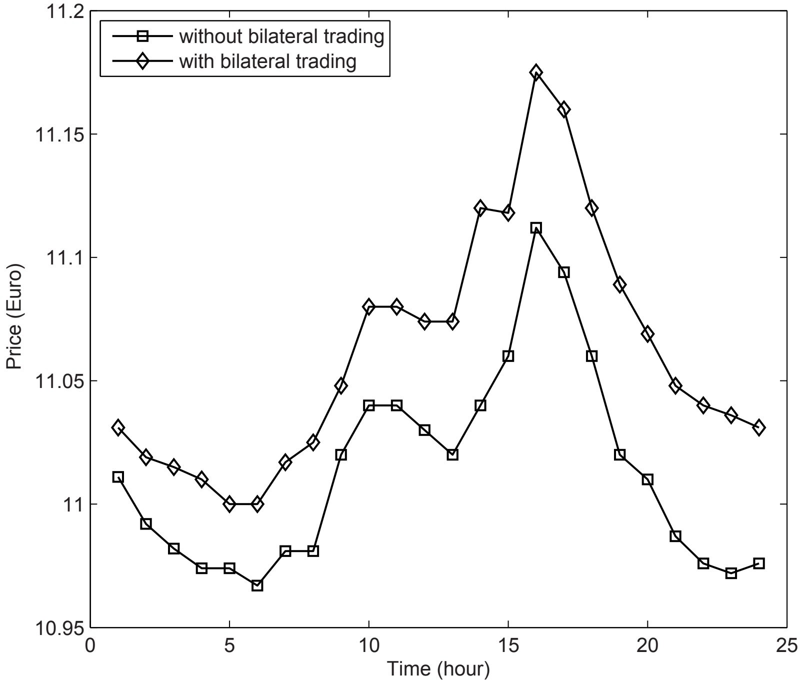

Fig. 5 The equilibrium price in the second DA market structure with and without bilateral negotiations.

6 Results

We have run our simulations assuming different percentages of bilateral contracts with respect to the total demand of electricity. In particular simulations have been run for $q = 0.0$ (no bilateral contracting), $0.6, 0.7, 0.9$, $\forall i = 1, 2, \ldots, N$; $t = 1, 2, \ldots, T$. Results are shown with reference to a typical trading day (24 hrs).

For the first market structure, where the number of Supps is higher than the number of GenCos, bilateral contracts have shown to have the effect of reducing prices in the DA market (Fig.4). This happens all the times the GenCo with bids equal to the MSP is involved in the bilateral agreement.

When the number of GenCos is higher than the number of Supps, simulations show that energy bought in the bilateral market may result in a potential increase of the energy prices in the DA market (Fig.5). As expected, this happens every time the bilateral agreement involves the most efficients agents.

In the very simple structure of the proposed model, the different percentages q of bilateral contracts do not produce any effect on the equilibrium price. Very probably, in an electricity market model with a greater number of agents the quantity of bilateral agreements would play a more active role.

7 Conclusions

The application of a computational agent-based model to the very simple energy market presented in this paper has provided some useful general insights.
It is plain that, a bilateral contract between a generation and a supplier effectively disengages two active players from each side of the wholesale market. Yet, this foreclosure effect will not necessarily lead to higher prices, despite the conventional concerns, and will be manifested only according to the specific market characteristics. In an energy market with a number of generation companies lower than the number of end-user suppliers, the bilateral contracs may produce lower equilibrium prices when the agreement involves the marginal unit in the merit order. The typical situation of increasing prices is instead reproduced in a market structure where the number of generation companies is greater than the number of suppliers and the bilateral agreement involves the most efficient agents.
We do believe that the proposed model contributes to the existing literature of power markets with new arguments about the effects of bilateral contracting and presents a new approach for bidding in the uniform-pricing auction settlement.

References

1. Bhavsar VC, Harold B, Yang L. (2003) A Weighted-Tree Similarity Algorithm for Multi-agent System in E-Business Environments. Proceedings of the Business Agents and the Semantic Web (BASeWEB) Workshop, Halifax, National Research Council of Canada, Institute for Information Technology, Fredericton:53–72
2. Bower J, Bunn DW (2000) Model-based comparisons of pool and bilateral markets for electricity. Energy Journal 21(3):1–29
3. Bower J, Bunn DW (2001) Experimental analysis of the efficiency of uniform-price versus discriminatory auctions in the England and Wales electricity market. Journal of Economic Dynamics & Control 25:561–592
4. Bower J, Bunn DW, Wattendrup C (2001) A model-based analysis of strategic consolidation in the German electricity industry. Energy Policy 29(12): 987–1005.
5. Bunn DW, Martoccia M (2005) Unilateral and collusive market power in the electricity pool of England and Wales. Energy Economics 27:305– 315
6. Bunn DW, Oliveira FS (2001) Agent-based Simulation: An Application to the New Electricity Trading Arrangements of England and Wales. IEEE Transactions on Evolutionary Computation 5(5):493–503
7. Bunn DW, Oliveira FS (2003) Evaluating individual market power in electricity markets via agent-based simulation. Annals of Operations Research 121(1-4): 57–77
8. Weidlich A, Veit D (2008) A critical survey of agent-based wholesale electricity market models. Energy Economics 30:1728–1759
9. Zhou Z, Chan WK (Victor), Chow JH (2007) Agent-Based simulation of electricity markets: a survey of tools. Artif Intell Rev, doi 10.1007/s10462-009-9105-x

Part VI
Games

Dynamics of Cooperation in Spatial Prisoner's Dilemma of Memory-Based Players

Chenna Reddy Cotla

Abstract In a population of extremely primitive players with no memory, interaction with local neighbors in a spatial array can promote the coexistence of cooperators and defectors, which is not possible in the well mixed case (Nowak, Bonhoeffer, and May, 1994). However, the applicability of this insight is unclear in the context of a social system where memory plays a significant role in the conscious decision-making of the members. In this paper, the problem of cooperation is analyzed in a population of players with the memory model embodied in the ACT-R cognitive architecture (Anderson and Lebiere, 1998). Using agent-based simulations, it is shown that in a population of memory-based agents, spatial structure supports higher levels of cooperation in comparison to the well mixed paradigm.

1 Introduction

The Prisoner's Dilemma (PD) has long been used as a paradigm to study the problem of cooperation faced by unrelated individuals in the absence of central authority [4, 16]. Even though, the PD offers an invaluable framework to study the problem of cooperation in the context of two self-regarding players, realistic investigation of cooperation problems at societal level involves a population of more than two players [18]. Consideration of a population of more than two interacting players leads to the further assumptions concerning the structure of the interaction. Mean-field approximation and rigid spatial structures are often considered as the two limiting cases of interaction topologies [9]. In the simplest mean-field approximation, each player interacts with every other player with equal probability. This represents the well-mixed scenario and here the interaction has no structure at all. The other extreme, the rigid spatial structure, represents the case where players are situated on a regular lattice and interact only with their local neighbors. The seminal work by

Department of Computational Social Science, George Mason University.
e-mail: ccotla@gmu.edu

S. Osinga et al. (eds.), *Emergent Results of Artificial Economics*, Lecture Notes in Economics and Mathematical Systems 652, DOI 10.1007/978-3-642-21108-9_16,

Nowak and May [16] has established that this crude approximation of interaction topologies observed in the real world could be used to explain the emergence and maintenance of cooperation in a population of extremely simple players with no memory. Nowak and May [16] have employed a spatial version of evolutionary PD, which is commonly referred to as the "Spatial Prisoner's Dilemma" (SPD), to show that cooperators and defectors can coexist in a chaotically shifting balance when the interactions are restricted to local neighbors on a square lattice. Computational experiments with numerous variations of such an evolutionary game confirmed the robustness of the claim that interaction with local neighbors in a spatial array can promote coexistence of cooperators and defectors in a population of memory-less players which is not possible in the well-mixed case [15] .

The evolutionary spatial framework considered in Nowak and May [16], and in Nowak et al. [15] provides an apt template for modeling strategic interactions and explaining the maintenance of cooperation in a spatially structured population of simple biological or physical entities that lack memory. However, straightforward adaptation of these results to reason about the dynamics of cooperation in a social system may not be appropriate. Human decision-making in repeated strategic interaction appears to be more sophisticated than the pure imitation in the evolutionary models. Moreover, some researchers argue that the evolutionary template may not be an appropriate framework to study learning and adaptation processes at the cognitive level [3]. This paper studies the problem of cooperation in a spatial setting by explicitly taking into account the adaptive character of the human memory. By using the memory model embodied in the ACT-R cognitive architecture [1], Lebiere et al. [13] were able to reproduce important experimental observations in the context of the iterated PD. The current paper extends this successful two-person memory-based game playing model to the context of SPD and studies dynamics of cooperation in such a framework.

The paper is organized as follows: Section 2 presents the representational details of the decision-making model of players in the model. Section 3 presents details about computational simulations with SPD involving memory-based players described in Section 2 and corresponding results about the dynamics of cooperation in such a framework. A discussion about the decision-making process of players in the model and some perspectives for future research is given Section 4. Finally, concluding remarks are pointed out in Section 5.

2 Model

Similar to the evolutionary spatial framework considered in Nowak and May [16], players in the current model are located on a regular lattice and play PD game with other players in their neighborhoods to receive the corresponding payoff. The principal divergence from the evolutionary SPD model is that the players in the current model use decision-making mechanism offered by ACT-R memory model to choose a strategy rather than imitating the strategy of the best scoring neighbor. In ACT-R, higher level decision-making is embodied in the declarative memory of facts and the procedural memory of production rules []Taatgen2006. The production rules in the procedural memory are condition-action rules that encode the potential actions to be taken when certain conditions are met. The facts are encoded in declarative memory using items called *chunks*. Chunks encode knowledge as structured, schema-like configurations of labeled slots. Each chunk has a level of activation that depends on its previous usage, its relevance to the current context, and a noise component. ACT-R memory model retrieves the chunk with highest activation and applies the relevant production rule to achieve a goal [19].

The representations of a player's declarative and procedural memory components in the present model are largely derived from the memory based account of two-person Prisoner's Dilemma game proposed in Lebiere et al. [13] . In this transition from two-person game playing context to the spatial game, the procedural component of the player is kept intact: a player looks at its two possible moves, determines the most likely outcome given each move, and makes the move associated with the best likely outcome. This logic is captured in the following production rule:

IF	the goal is to play Spatial Prisoner's Dilemma
	and the most likely outcome of making move C is outcomeC
	and the most likely outcome of making move D is outcomeD
THEN	make the move associated with the larger of outcomeC and outcomeD
	Note the actual outcome and push a new goal to make the next play

The number of possible outcomes for each player in a spatial game depends upon the definition of the neighborhood under consideration. If each player has n neighbors then there are 2^{n+1} possible outcomes per player. For simplicity, a totalistic representation of the outcomes is adopted that is inspired by the totalistic approach used in Ishida and Mori [12] to represent spatial strategies. The symbol kC is used to represent a configuration where k neighbors of a given player have chosen to cooperate and $n-k$ neighbors have chosen to defect, where n is the size of the neighborhood and $0 \leq k \leq n$. Using this notation, the outcome where the player under consideration has cooperated, and the configuration of neighbors' moves is kC, is denoted with C-kC, and the outcome where the player has chosen to defect for the same configuration of neighbors' moves is denoted with D-kC. In addition to simplifying the representational matters, such a totalistic representation of outcomes explicitly takes into account the spatial phenomena that is an important characteristic of spatial games, in contrast to the well known SPD frameworks that use either very simple strategies (e.g. Nowak and May [16]) or use conventional two person strategies in repeated PD games which depend upon the past actions of a single opponent (e.g. Axelrod [4]). With this totalistic representation scheme, there are $(n+1)$ possible outcomes for each possible move when a neighborhood of size n is considered. All these possible outcomes are represented in a player's declarative memory as chunks of type *outcome* with three slots: *p-move* that encodes the player's action; *N-config* that encodes the choice of moves by the neighbors; and, *payoff* that encodes payoff received by the player for that particular outcome calculated from Reward (R), Temptation (T), Sucker (S), and Punishment (P) payoffs of the PD game. The $2(n+1)$ chunks necessary to encode all possible outcomes for a given player in the model are given below:

(C-nC isa outcome p-move C N-config nC payoff nR)
(C-$(n-1)$C isa outcome p-move C N-config $(n-1)$C payoff $(n-1)R+S$)
.
.
(C-kC isa outcome p-move C N-config kC payoff $kR+(n-k)S$)
.
.
(C-1C isa outcome p-move C N-config 1C payoff $R+(n-1)S$)
(C-0C isa outcome p-move C N-config 0C payoff nS)
(D-nC isa outcome p-move D N-config nC payoff nT)
(D-$(n-1)$C isa outcome p-move D N-config $(n-1)$C payoff $(n-1)T+P$)
.
.
(D-kC isa outcome p-move D N-config kC payoff $kT+(n-k)P$)
.
.
(D-1C isa outcome p-move D N-config 1C payoff $T+(n-1)P$)
(D-0C isa outcome p-move D N-config 0C payoff nP)

For a given player, the first clause of the production rule will retrieve one of the $(n+1)$ chunks associated with the player making the move C, and the retrieved chunk is denoted as *OutcomeC*, and the second clause will retrieve one of the $(n+1)$ chunks associated with the player choosing to defect, and the retrieved chunk is denoted as *OutcomeD*. The payoffs associated with these two outcomes, *OutcomeC* and *OutcomeD*, are compared and the *p-move* associated with the chunk with the highest payoff is taken.

The production rule retrieves the most likely outcome for each move by retrieving the outcome chunk with the highest activation. For simplicity, relevance factor is not considered in calculating the activation of a chunk in the current model. The activations of a declarative chunks are calculated using the following equation:

$$B = \ln\left[\sum_{i=1}^{k} t_i^{-d} + \frac{(n-k)(t_n^{1-d} - t_k^{1-d})}{(1-d)(t_n - t_k)}\right] + N\left(0, \frac{\pi . s}{\sqrt{3}}\right) \tag{1}$$

The first part of the sum accounts for the adaptive nature of the human memory observed in various psychological experiments reported in [2]. This quantity, known as "Base level activation", increases with each reference and decreases with time justifying the power of learning and forgetting [2]. t_j in the sum refers to the time since jth reference, n is the total number of references, and d is the forgetting rate. This computationally efficient approximation of the original formula proposed in Anderson and Schooler [2] is due to Petrov [17]. Petrov has shown that by keeping the most recent k references, the base level activation can be approximated with great accuracy. In the actual implementation we used $k = 1$ for computational efficiency. The second part of the equation accounts for the stochasticity and is calculated as noise that is normally distributed with the mean of zero and the standard deviation determined by the activation noise parameter s [13]. The same default values are considered as in Lebiere et al. [13] for the forgetting rate of $d = 0.5$, and the activation noise parameter of $s = 0.25$. The initial references of declarative chunks are uniformly distributed such that on average each chunk would get 100 references. It has been observed from computational experiments that results are qualitatively unchanged when we varied the number of initial references from 10 to 100.

3 Simulation Results

The first computational experiment is carried out on a square lattice of the size 50×50 with periodic boundary conditions. Memory-based players with the procedural and declarative memory components described in the previous section are placed at each lattice site and each of them interacts with its von Neumann Neighbors (self interaction is not considered here). The standard PD payoff matrix [4] with $R = 3, S = 0, T = 5$, and $P = 1$ is considered. In each generation (time step), all the players simultaneously make their choice of moves using the production rule, receive payoffs determined by the corresponding outcomes, and update their declarative memories. As such the underlying dynamics of the model are synchronous. To characterize the macroscopic dynamics of the model, the fraction of cooperators (f_C) in the population at each generation is considered. Since the model involves stochastic elements, simulation output from a single realization may be misleading and some statistical treatment would be more appropriate. The simulation is carried out 30 times with a different random seed each time to ensure statistical independence across the runs. In each run, it was observed that f_C in the model asymptotically stabilizes and fluctuates around a constant after some number of generations. The model is considered to be asymptotically stable in a given run when the difference in the mean values of f_C over two consecutive windows of 10^4 generations is less than 10^{-3} in magnitude []gomez2007. After the model is considered as asymptotically stable, the mean value of f_C over next 10^4 generations is taken as the asymptotic f_C of the run. Figure 1 depicts the behavior of the mean value of f_C over the consecutive time windows of 10^4 generations in a sample run of the model that is run for a total of 400,000 generations. It can be easily seen that slope of the curve is continuously diminishing in magnitude indicating that mean f_C is approaching an asymptotic value. In this run, after 120,000 generations, the model

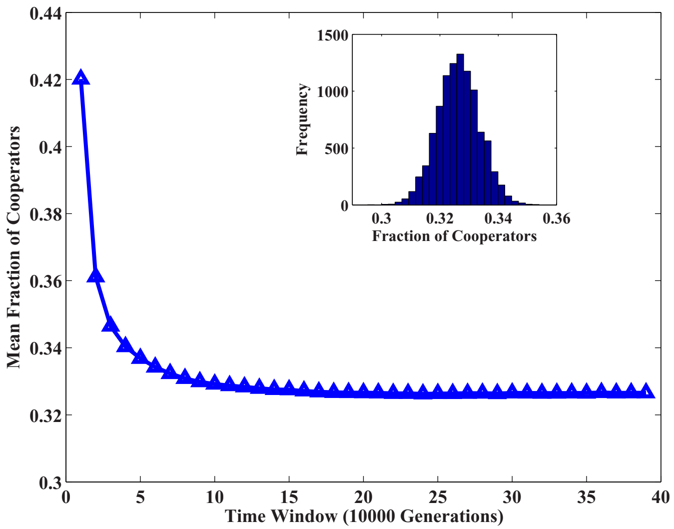

Fig. 1 Mean f_C over time windows of 10^4 generations in a simulation run of $400,000$ generations. The *insert* depicts the histogram of frequencies of different f_C values over the 10^4 generations after the model is considered as asymptotically stable.

meets the asymptotic stability criteria and the asymptotic f_C is the mean fraction of cooperators over the next 10^4 generations, which is 0.3226 for this case. The *insert* in Figure 1 depicts frequencies of different values of f_C in the window of 10^4 generations after the model is considered as asymptotically stable. Almost normally distributed frequencies suggest that f_C is fluctuating around the mean. The 95% confidence interval obtained for asymptotic f_C from 30 different runs is obtained as $[0.3220, 0.3228]$. Such a narrow confidence interval indicates that f_C is fluctuating around a constant. Also the experiments were repeated for lattice sizes from 20×20 to 400×400 and it was observed that asymptotic cooperation levels are almost independent of lattice size. These emergent stable cooperation levels independent of the initial configuration of the model are very interesting.

The second experiment considered the mean-field interaction scenario for a population size of 2500 players so that the effect of spatial structure on cooperation levels can be evaluated. Since the interactions in well-mixed case are bilateral, game playing model proposed in lebiere et al. [13] is directly used. A generation in this case consists of $N/2$ micro time steps, where N is the size of the population. In each micro time step, two randomly selected distinct players play a bilateral PD game. Thirty statistically independent runs of this well mixed scenario were carried out with stopping times determined by the asymptotic stability criteria discussed earlier. The 95% confidence interval of the asymptotic f_C was calculated as [0.2446 0.2448]. Here too the asymptotic f_C values were almost independent of size of the population. Also, this is in contrast with evolutionary PD where cooperation is not possible in mean-field interaction case. Comparison between cooperation lev-

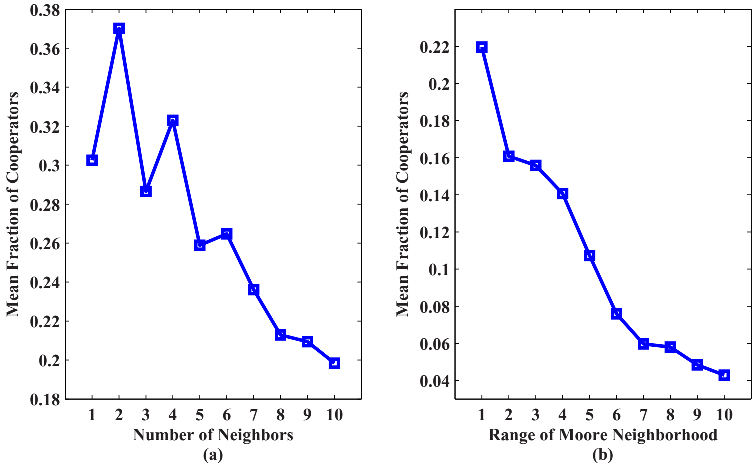

Fig. 2 (a). Mean asymptotic f_C of 30 independent runs of the model for different sizes of neighborhood. Players are located on a 50×50 square lattice. (b) Mean asymptotic f_C of 30 independent runs of the model for different ranges of Moore neighborhood. Players are located on a 50×50 square lattice.

els in the spatial context involving von Neumann neighborhoods and well-mixed case suggests that spatial structure can support higher levels of cooperation than that of well-mixed case in a population of memory-based players. However, further experiments with different notions and sizes of neighborhoods suggested that spatial structure does not always lead to the higher asymptotic cooperation levels than that of the well-mixed scenario. For example, the mean value of asymptotic f_C obtained from 30 independent runs of the spatial model considering Moore neighborhood was 0.2196, which is less than the mean asymptotic f_C observed in well-mixed scenario. To further characterize the relation between the neighborhood size and asymptotic cooperation levels, two more computational experiments were carried out. The first of these explored the effect of number of neighbors on the asymptotic cooperation levels. In this experiment, the spatial game is run with neighborhood sizes from 1 to 10. For a given neighborhood size, the corresponding number of distinct neighbors are randomly initialized for each player. Figure 2(a) plots the mean of the asymptotic f_C of the 30 independent runs against the neighborhood size from 1 to 10 for SPD with memory-based players located on a 50×50 lattice. Even though, neighborhood sizes between 1-6 support higher asymptotic cooperation levels than well-mixed case, neighborhood sizes bigger than 6 support asymptotic cooperation levels smaller than that of the well-mixed case. The other experiment explored the effect of the range of Moore neighborhood on the asymptotic cooperation level in the model. Figure 2(b) plots the mean asymptotic f_C observed in 30 independent runs against the range of Moore neighborhood for SPD with memory-based players located on a 50×50 lattice. Again, it can be observed that asymptotic cooperation levels decrease with larger neighborhood sizes and they are all smaller than the asymptotic cooperation levels observed in the well-mixed case. These simulation results suggest that in a SPD with memory-based players as described in this paper, neighborhoods larger than 6 in size lead to lower asymptotic cooperation levels compared to the well-mixed case. Thus in a population of memory-based players spatial structure is beneficial for cooperative behavior only for smaller neighborhood sizes. Furthermore, it is interesting to note that for all these experiments, the asymptotic cooperation levels were observed to be almost independent of the size of the population.

4 Discussion

In each generation, players in the current model make a decision to cooperate or defect through the use of accumulated experience. The particular decision-making model used in this paper is referred to as Instance based decision-making model in the cognitive science literature [8]. This model of decision making is based on two principles: a. storage of instances of experiences and b. decision making logic that involves selecting the most promising action by reviewing past experiences [8]. In the current model instances are encoded as the outcome chunks in the declarative memory and decision-making logic is encoded as the production rule in the procedural memory of a player. Since these two principles are implemented using memory model of a validated cognitive architecture, in this case ACT-R, this model is also cognitively plausible. Coming back to the decision making process, a player considered in the model makes use of its memories of outcome instances in the past generations to construct an expectation about the neighbors' moves conditional upon its own choice of move. Such an expectation is entirely experience based and is facilitated in a unique manner by the adaptive memory of a player. In this way, the adaptive nature of the memory of a player captures the observed pattern of neighbors' play using the past occurrences of the outcomes. The maximizing move taken by a player in a given generation is a best response to such an observed pattern of the neighbors' play. A given player's choice of a move in turn affects the adaptive memories of the player's neighbors and their future moves. Due to the presence of such causal loops that couple players and their neighborhoods analysis by reduction may not be practical in this model. Agent based models are very useful to model and analyze the emergent characteristics arising out of interactions among the elements of a complex adaptive system like the one discussed here [14]. Agent based models facilitate explicit representation of individual members of a complex adaptive system and the direct interactions among them to systematically analyze various emergent aspects [5]. In the current model, attainment of constant asymptotic cooperation levels for a given neighborhood independent of initial configuration of the model can be considered as such a system level emergent property. Before deriving conclusions from agent models, it is often important to be cautious about certain representational matters. Importantly, the consideration of different updating schemes proved to elicit significantly different emergent behaviors from certain agent based models [6, 11].

To examine the effect of different updating schemes, multiple runs of the previous experiments were carried out for each of the two asynchronous updating schemes of Random Activation(RA) and Uniform Activation (UA) [6]. It was observed that the previous conclusion that spatial structure is beneficial for cooperative behavior only for smaller neighborhood sizes remains valid for these asynchronous updating schemes too.

Behavioral experiments on spatial games have been very recent. Results from these recent experiments call for alternative efforts are needed to explore the mechanisms underlying the dynamics of decision-making in these games [10]. Especially, experiments on spatial Prisoner's dilemma with human subjects have established that strategy change of a player in these experiments is different from unconditional imitation of the best scoring neighbor as assumed in many evolutionary game theoretical models [20]. We believe that cognitive framework proposed in this paper to analyze spatial games may be an important candidate among the possible alternatives that can complement the evolutionary framework in understanding the dynamics of strategic puzzles in social sciences. Furthermore, it is quite straightforward to extend this model to study other social dilemmas by changing the relevant payoff quantities in the model. In our future studies we intend to extend the model considering various other social dilemmas, interaction topologies, and validate the model output with the experimental data on spatial games.

5 Conclusion

This paper has investigated the effect of the adaptive nature of human memory on the dynamics of cooperation in Spatial Prisoner's Dilemma. Computational experiments showed that fraction of cooperators in such a framework stabilizes around a constant almost independently of initial configuration of the model for a given definition of neighborhood and updating scheme. Furthermore, it is shown from the simulation results that spatial structure is beneficial for cooperative behavior only for smaller neighborhood sizes in a population of memory-based players. This work may be relevant in understanding the dynamics of cooperation in a social system where the memory processes that facilitate and constrain decision-making of individuals may not be ignored.

References

1. Anderson, J. and Lebiere, C. (1998). *The atomic components of thought*. Lawrence Erlbaum Associates.
2. Anderson, J. and Schooler, L. (1991). Reflections of the environment in memory. *Psychological Science*, 2(6):396.
3. Aunger, R. (2001). *Darwinizing culture: The status of memetics as a science*. Oxford University Press, USA.

4. Axelrod, R. (1984). *The evolution of cooperation*. Basic Brook, New York.
5. Axtell, R. (2000). Why agents? on the varied motivation for agent computing in the social sciences. *Brookings Institution CSED Technical Report*.
6. Axtell, R. (2001). Effects of interaction topology and activation regime in several multi-agent systems. *Multi-Agent-Based Simulation*, pages 33–48.
7. Gómez-Gardenes, J., Campillo, M., Floría, L., and Moreno, Y. (2007). Dynamical organization of cooperation in complex topologies. *Physical Review Letters*, 98(10):1–4.
8. Gonzalez, C. and Lebiere, C. (2005). Instance-based cognitive models of decision making. In D., Z. and Courakis, A., editors, *Transfer of Knowledge in Economic Decision Making.* New York: Palgrave McMillan.
9. Hauert, C. (2002). Effects of space in 2×2 games. *Int. J. Bifurcat. Chaos*, 12:1531–1548.
10. Helbing, D. and Yu, W. (2010). The future of social experimenting. *Proceedings of the National Academy of Sciences*, 107(12):5265–5266.
11. Huberman, B. and Glance, N. (1993). Evolutionary games and computer simulations. *Proceedings of the national academy of sciences of the United States of America*, 90(16):7716.
12. Ishida, Y. and Mori, T. (2005). Spatial strategies in a generalized spatial prisoner?s dilemma. *Artificial Life and Robotics*, 9(3):139–143.
13. Lebiere, C., Wallach, D., and West, R. L. (2000). A memory-based account of the prisoner's dilemma and other 2×2 games. In *Proceedings of the 3rd International Conference on Cognitive Modeling, Groningen, Netherlands*, pages 185–193.
14. Miller, J. and Page, S. (2007). *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton University Press.
15. Nowak, M., Bonhoeffer, S., and May, R. (1994). Spatial games and the maintenance of cooperation. *Proceedings of the National Academy of Sciences of the United States of America*, 91(11):4877.
16. Nowak, M. and May, R. (1992). Evolutionary games and spatial chaos. *Nature*, 359(6398):826–829.
17. Petrov, A. (2006). Computationally efficient approximation of the base-level learning equation in act-r. In *Proceedings of the Seventh International Conference on Cognitive Modeling*, pages 391–392.
18. Szilagyi, M. (2003). An investigation of n-person prisoners' dilemmas. *Complex Systems*, 14(2):155–174.
19. Taatgen, N., Lebiere, C., and Anderson, J. (2006). Modeling paradigms in act-r. In Sun, R., editor, *Cognition and multi-agent interaction: From cognitive modeling to social simulation*, pages 29–52. Cambridge University Press.
20. Traulsen, A., Semmann, D., Sommerfeld, R., Krambeck, H., and Milinski, M. (2010). Human strategy updating in evolutionary games. *Proceedings of the National Academy of Sciences*, 107(7):2962.

Indian Food Supply Chains: a Game and Model to Study Economic Behavior

S.A. Meijer, J. Raghothama, R. King and B. Palavalli

Abstract Food supply chains are vital for the societal welfare in India. Given the tremendous issues on food safety and security, more insight in the function of Indian food markets is needed. This paper introduces the multi-agent model on which the Mango Mandi Gaming Simulation is build. The model holds the major roles present in the Indian mango markets, and is in both it structure and processes built upon case studies of these markets. The interaction with human players, who can take over a role in the simulation model, is specified. Applications are identified, and a road map for validation is laid out.

1 Introduction

In India, increasing population and urbanization has led to concerns about adequate food supplies. The recent price fluctuations in the world markets have led to crises on commodity food prices like unions and tomatoes. While a 50 to 100% increase in prices may not be a problem to consumers in the sections of the society with higher welfare, it is a problem for about 2/3 of Indias population of 1.15 billion people [1]. The Green Revolution of the 1960s increased yields in the North of India applying scientific knowledge [2], and now new approaches are being applied to provide additional insight. The attention is not only geared towards production techniques,

Sebastiaan. A. Meijer
TUDelft, Netherlands e-mail: Sebastiaan.Meijer@tudelft.nl

Jayanth Raghothama
CSTEP, Bangalore e-mail: jayanth@cstep.in

Robin King
CSTEP, Bangalore e-mail: robin@cstep.in

Bharath Palavalli
CSTEP, Bangalore e-mail: bharath@cstep.in

S. Osinga et al. (eds.), *Emergent Results of Artificial Economics*, Lecture Notes in Economics and Mathematical Systems 652, DOI 10.1007/978-3-642-21108-9_17,

but also on the institutional arrangements and economic incentive system, as India has faced scandals with rotting food and hunger existing simultaneously [3]. Recent studies show that waste in agricultural production in India reaches as high as 20 percent while the perception reflected in government pronouncements and the media often extends to 40% [4] [5].

The late 1990s and start of the 21st century showed an increasing popularity of supply chain gaming simulations in the domain of supply chains and networks. Gaming simulations (games, in short) in this domain are becoming a standard component of business school curricula [6], and researchers use them as a research method for both quantitative and qualitative research [7]. Supply chains and networks are considered complex adaptive systems, as they meet all the aspects in the definition by Holland [8]. Bekebrede and Meijer [9] demonstrated that complex adaptive systems could be simulated in a gaming simulation. The latest development in gaming simulation is the integration of better simulators as technical backbone and more detailed representations of the aspect system under study in the game. Multi agent simulation is one of the techniques used for this integration.

In this paper we have selected the mango supply chain as representative of semi-perishable fruits in India. Mangos are grown in many areas of India, increasingly exported, and represent many of the supply chain challenges present in other Indian horticultural products as well. The multi-agent simulation design presented models the various roles that exist in the supply networks. Human players can ‘take over’ each of the roles that are in the model. The play is computer-based with optional face-to-face interaction between participants. The design facilitates distributed sessions over the Internet or other computer networks.

2 Indian Mango Supply Chain

India has been the leading producer of mangos in the world for many years, and since 2005 has become the leading exporter as well. In 2007, India produced about 40% of total world production, and about 25% of export volume [10] [1] Indian production is dominated by domestic markets, which are highly segmented by location, cultivar, and quality.

In order to better understand the dynamics of the market, we conducted a small survey during the 2009 and 2010 mango seasons in Bangalore, the capital of the southern state of Karnataka, one of the principal mango producing states in India [11]. Our research showed that distribution in Karnataka is similar to that docu-

[1] Figures differ between the international FAO and Domestic APEDA (Indian Agricultural and Processed Food Products Export Development Authority), probably because of different dates (calendar year vs. Indian fiscal year (April-March)) and differing definitions (the broader mango, mangosteen and guava category used by the FAO). We will use the FAO figures here as they allow for international comparisons and the time series extends further. Naidu and Naidu discuss this matter further, and add the strange case of mango pulp, for which APEDA shows substantial exports, while the FAO shows none.

mented in the neighboring state of Andhra Pradesh [12]. In the South, most production occurs on small-scale farms, but medium and large-scale farms also produce mangos. While some farmers directly market their mangos themselves, especially large scale producers selling to food processors, most mangos are sold through pre-harvest contractors (PHCs), referred to in the Bangalore market as dealers, who benefit from economies of scale by gathering the produce of many farmers. These dealers then sell the produce in the wholesale markets, or mandis, and earn commissions from both sellers and buyers, generally totaling between eight and twelve percent. These dealers provide advance payment to the farmers, and then bear both the risk and opportunity of the market.They arrange for transportation and storage in the journey to the mandi. The market itself lasted for around 30-45 days. Dealers sourced their products mainly from farmers around Bangalore unless they were re-selling special varieties of mangoes from another region. Dealers paid the farmers a portion for their mangoes only if the mangoes were sold in the market. Formal written contracts of any form appeared to be rare. Problems with quality and consistency continue to persist. Based on our field studies, the structure of the Mango supply chain emerged shown in figure 1.

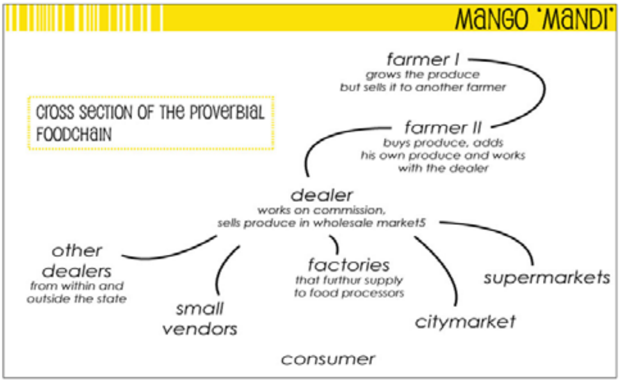

Fig. 1 Market Structure

A wide variety of problems in the supply chain exist. A comprehensive value-chain survey-based study of Indian horticulture conducted by the World Bank concluded that high variation in prices reflect market segmentation, that logistics and intermediation costs are much higher than production costs, and that high wastage reflects inefficiencies in the delivery chain. Mangoes showed an average 20% wastage, with 8% of waste occurring from the farm gate to the local mandi, and 2% each within the district, state, and outside the state [4]. The Asian Development Bank also notes the problems with getting produce in acceptable conditions in a timely fashion to consumers, and the need to upgrade transport, irrigation, marketing and financial sector infrastructure to address them. Our primary research also revealed the challenges in transport, processing, storage, and financial linkages.

Based on this analysis, we believe that key elements of the Indian mango supply chain that must be addressed include:

1. heterogeneous production methods;
2. segmented markets across geography, time, cultivar, and quality;
3. severe credit constraints, with informal credit often provided by middlemen;
4. inadequate transportation support services;
5. inadequate cold storage and warehouse facilities; and
6. incomplete information across the supply chain, worsening the closer one gets to the farmer/producer.

In a previous publication the authors showed that no existing gaming simulation currently addresses these issues, mainly because most work done has been on Western contexts where the seven constraining variables mentioned above do not apply.

The design reflected below focuses on the trading element once production is complete, and thus abstracts away from the production process itself. Elements 1, 4 and 7 are included in the design presented below. Future work is expected to include the other elements.

3 A New Design: Mango Mandi Gaming Simulation

The Mango Mandi Gaming Simulation (MMGS) is afirstinstantiation of a multi-player gaming simulationframework on supply chain management for semi-perishable fruits and vegetables. The MMGS models the mango supply chain in India, with all the actors modeled as software agents, with relatively simple models of behavior, preferences and payoff functions. A game is being built on top of the multi-agent simulation to explore and elicit more complex real human behavior.

3.1 Description of the Roles / Agents in the MMGS

FARMER I This section of farmers grows the produce, and sells it to another farmer.

FARMER II This section of farmers collect the produce from other farmers, add it to their own produce, and sell it via the 'dealer', or the distributor. An advance is given to him by the dealer, and then balanced out through the season. This locks in the farmers to the dealers.

DEALER The dealer works on a commission basis. When the fruit matures, the farmers get in touch with the dealer and the wholesale markets are set up. The dealer pays the farmer an advance, which is paid back through the three months of sale. This advance assures loyalty of the farmer. The dealer takes a certain percentage of the price on the produce sold.

OTHER DEALERS From here, the produce goes to other dealers from other states and the network is repeated.

SMALL VENDORS The source of fruit for these types of vendors is both the wholesale market as well as the city market. Cart Vendors have routes they follow in order to sell their mangos. The Roadside Basket Vendors often buy the waste fruit from the wholesale market and sell it right outside.

FACTORIES The factories are another large client for the wholesale market. They buy the fruit in bulk, convert it into pulp, and store it in cans so as to then sell it to other food processors such as juice makers, cold drink factories, etc.

SUPERMARKETS The supermarkets also buy from the wholesale markets and mark up the prices and sell the mangos. They typically buy only quality mangos or specific varieties.

In the current version, Farmers, Dealers, Supermarkets, Small Vendors have been modeled. Exporters are also an important agent type that has been modeled as well. There is enough flexibility and modularity to model other actors in the future. Human players take on the role of the dealer when they play the game.

The first implementation of the game is a simple trading model, based on preference and utility functions. The object traded is mango. Based on the producer states, agents are located: Bangalore(Karnataka), Chennai(Tamil Nadu), Pune(Maharashtra), Ahmedabad(Gujarat) and Srinagar(Jammu and Kashmir). Agents in these cities can both buy and sell mangos, except for Srinagar where agents can only buy mangos. Both road and train connect all cities.

A round simulates one sales season, a period of forty-five days. Each day is compressed in real time, and lasts for approximately three to four minutes. This is configurable based on the speed of the players. Several negotiations/deals can happen between the agents within one day.

3.2 Game Design

Software agents can play the roles of all the actors in the system. The MMGS was designed in a manner that allowed us to substitute any of these agents with human players. Three attributes were chosen for the product(mango), being: type, price and time to decay. Quality has been ignored for the current version, since it is intrinsically tied to both type and ripeness. Quantity is a fixed item. Trader trade packages of mangoes. An actor maintains an inventory of items and can track the status of items that have been bought from other actors at various places. The produce ripens in transit and there is a random probability that a small percentage of the cargo might be lost due to reasons such as mishandling, theft, etc. Figure 2 shows the legal transactions between actors (both humans and agents) in the game. Actors trade mangos based on their preferences. The preferences/biases that influence an agent's decisions are as follows[2],

- Preferences for commodities,

$$f(\omega_{attribute}, \theta) \tag{1}$$

  where $0 \leq \omega \leq 1$ and attribute $\in (type, ripeness, price), \theta$ is order of the attribute for that particular commodity and $\theta in(1, \ldots noofattributes)$
- Scrooge factor ψ, which defines the elasticity of an agent's spending where $0 \leq \psi \leq 1$
- Preference for history or past experience ω_{hist} where $0 \leq \omega_{hist} \leq 1$
- An actor rates the other actor at the end of a transaction based on the following parameters,
  - Ease of negotiations α_{buyer} or α_{seller} and quality of goods received for the buyer[3] β_{buyer} where $0 \leq \alpha_{buyer} + \beta_{buyer} \leq 1$ and $0 \leq \alpha_{seller} \leq 1$
  -

$$\tau_{i.average} = \frac{\sum \frac{\alpha_{buyer}+\beta_{buyer}}{\alpha_{seller}}}{NumberOfTransactions} \tag{2}$$

  where each actor i maintains a list for all other actors[4].

At the beginning of each day, the following actions happen:

- Inventory is handed out the city. This inventory is usually a large amount of mangos of various types. These numbers are based on production data for the cities in the game obtained from various sources. This forms the mechanism through which data is fed in to the simulation, liquidity of mangos maintained, and cultivation of mangos simulated. Variations in the inventory for each day will lead to heterogeneity in the markets.
- The inventory for each day is distributed equally to all farmers in that particular city. This assumption is made to keep the model simple initially. The function to allocate inventory to farmers within a city can be changed to reflect reality (based on data from markets), and to provide for true heterogeneity in supply of mangos.

$$I_i = \frac{I_{city}}{N_{city}} \tag{3}$$

[2] In all the formulae being described above please note that 0 is the least value and 1 being the highest value unless specified otherwise.

[3] A seller would have rated the buyer based on the efficiency of the transfer of money but in the system the transfer of currency is instantaneous and a buyer is allowed to buy only if he has enough money.

[4] In order to highlight certain issues brought out during the interviews such as preferences based on relationships/caste it is possible to pad τ to make it extremely high.

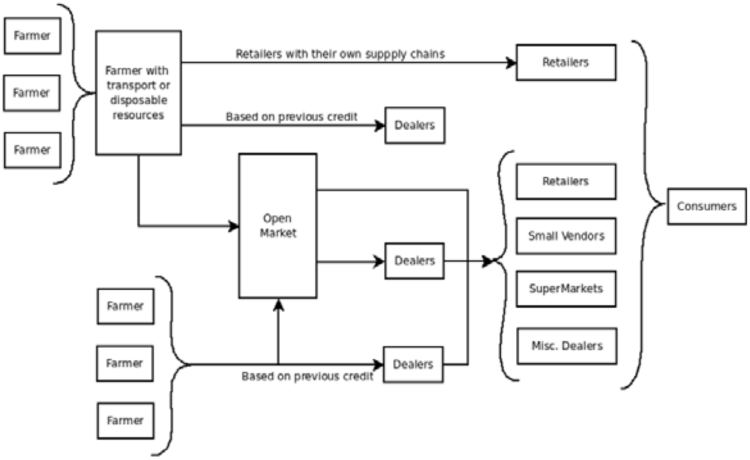

Fig. 2 Legal transactions in the game

where I is an inventory lot, defined by $attribute \in (type, ripeness, price)$ and $quantity(kgs/tons), N$ is the number of farmers in the city and $city \in (Bangalore, Pune, Ahmedabad, Chennai, Srinagar)$

- The attributes of all mangos change at the beginning of every day. For example the price of mangos may change and the mangos become riper:

$$ripeness = ripeness + 1 \tag{4}$$

- Because of the changed attributes of mangos and arrival of new mangos into an agents inventory, the agent has to compute a new price for his inventory lots and offer them in the market again. This new price is a function of the attributes and the average price of mangos. The average price of mangos is information available to all agents and is updated regularly as and when agents trade mangos. The average price of mangos can be calculated based on multiple parameters, for example the type of mango, the city where it is being traded and so on.

$$I_{price} = f(I_{attributes}, AveragePrice_{type,city}) \tag{5}$$

- Once the price is set, all agents offer their inventory lots at the new price. In the current version of the model, all these offers are visible to every agent.
- Contracts and credit between farmers and dealers can be simulated by limiting the network of a farmer, i.e, ensuring that offers from a farmer is visible only to one or a set of dealers. Additionally, the values of Scrooge Factor ψ and Ease of Negotiations for a particular agent τ can also be manipulated to ensure that a farmer deals with only one or a set of dealers.
- Agents now search for offers based on their preferences. If certain offers fit an agents criteria, the agent can use the same function to decide on a *quotedprice* and propose a deal to the selling agent.

All software agents are assumed to operate under conditions of bounded rationality and hence take rational decisions based on the perceived utility that a certain product gives them according to their preferences or biases. To reduce the number of decisions human players have to make, the negotiation protocol is a simple propose-accept/reject mechanism.

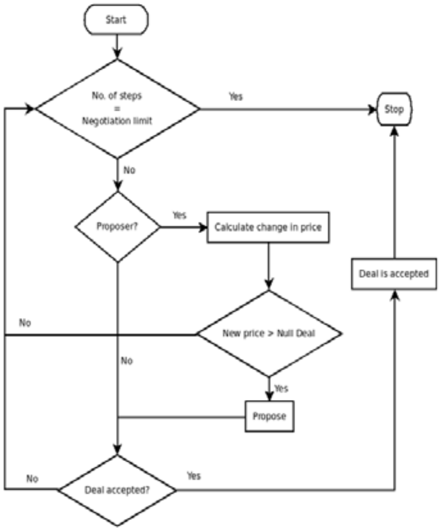

Fig. 3 Negotiation Protocol followed by the agents

When the proposal for a new transaction is received, agents calculate the payoff with the price provided and proceed based on the other biases they have. Agents can however proceed with a more complex negotiation protocol that mimics behavior observed in markets as described in figure 3. This mechanism is followed only for stand-alone simulations and not used when agents play the game with humans.The change in the new price is calculated as follows,

$$\frac{(1-(\omega_{hist} * \tau_{actoravg})) * quotedprice * (1-\psi_{Scrooge})}{noofstepsfornegotiation} \tag{6}$$

The scrooge factor determines which mode of transport will be used by the agent as the transport will cost both time and money. An extremely greedy agent may find it hard to bargain with other agents and may be blacklisted by other agents as negotiations always end in a null deal. A null deal would essentially result in $\alpha = 0$ and hence would mean that $\tau = 0$. This would deter other agents from approaching an agent with such a bad transaction history unless the agents are bound to each other based on credit or other mechanisms. Information about the reputation of agents is public.

For the buyer in a transaction, the attributes of the inventory lot bought will be decided as follows:

- The price is determined by

$$price = quotedprice + \delta_{newprice} + transportprice \tag{7}$$

where $\delta_{newprice}$ is the change in price as decided through the negotiation, and $transportprice$ is the cost of transporting from the sellers location to the buyers'.
- The ripeness of the mango when it arrives at the buyers location will be

$$ripeness = ripeness_{bought} + noofdaysfortransport \tag{8}$$

where $noofdaysfortransport$ is a function of the distance between the locations of the buyer and seller and the type of transportation chosen
- The quantity of mangos that actually arrives at the buyers location is defined by:

$$quantityreceived = quantitybought - (\gamma * quantitybought) \tag{9}$$

where γ is a factor that determines the amount lost in transit, and $0 \leq \gamma \leq 1$

Both $transportprice$ and γ depend on the type of transportation chosen by the buyer, which in turn depends on his Scrooge Factor ψ. There are several types of transportation available[5]. Each type has a cost for one unit of mangos, and a certain capacity. The total quantity of mangos wasted is the quantity lost in transit in all transactions, as well as all unsold inventory[6].

For the seller, the profit from each transaction is a percentage of the price, which is his commission. This percentage can vary depending on whether the farmers are bound to the dealer through credit or other mechanisms. The dealer incurs no loss for unsold inventory.

At the end of every day, the following actions happen:

- All agents remove all of their unsold offers from the markets, to facilitate the change in attributes for the next day.
- All mangos which are past ripeness are removed from agents inventories and counted as waste.

$$\begin{aligned} if I_{type,ripeness} &= timetodecay_{type} \\ waste &= waste + I_{quantity} \end{aligned} \tag{10}$$

Other agent types that inherit from the generic model can be easily added. There are some differences between the various agent types, illustrated in the table below:

The network of each agent type varies; with farmers being accessible only to dealers (or only a few specific dealers, to mimic contracts and credit mechanisms), and other agent types can only access dealers within their city. Supply and demand information is localized, and agents can only access the information they need to process the various mango offers.

[5] It is assumed that all transportation choices are always available to all agents.

[6] Mangos cannot be sold by the dealer once they are ripe

Table 1 Agent Preferences

Agent Type	Mango Attributes	Scrooge Factor	History	Transport Type
Dealer	Prefers to look at ripeness of the mangos first, type of the mangos second and price last	Very high scrooge factor	Very high preference for history	No/Equal preference for all transport types
Exporter	Prefers to look at type of the mango first (usually only one type), and price second. Has no preference for ripeness	Moderate Scrooge Factor	High preference for history	Transport not considered
Small Vendor	Prefers to look at type in conjunction with ripeness, and price last	Very low scrooge factor	Low preference for history	Transport not considered
Super Market	Prefers to look type first, ripeness second and price last	Moderate to high scrooge factor	Low preference for history	Transport not considered
Farmer	Looks at type first, and then price history of mangos next, to set the price	Very low scrooge factor	Moderate to high preference for history	Cheapest transport option preferred

4 Validation

In the world of gaming simulations there is no commonly accepted method of evaluation, and the most well-known are geared towards learning applications [13]. Raser [14] gave four criteria, re-used by Peters et al [15] for evaluating gaming simulations. These are:

- psychological reality
- structural validity
- process validity
- predictive validity

The game on top of the multi-agent simulation is still in development, but we can say something about the validity of the underlying MAS.

The structural validity of the MAS modelcan been validated by a test to put in primary data collected in surveys in mandis around Bangalore, secondary data regarding production and prices of mangos from various ministries of the government. anddata from interviews with landowners and farmers in the states of Karnataka and Andhra Pradesh. An initial test using this data led to values on the key variables that are at least within a reasonable range of the patterns in power distribution, margins and the central role that the dealers play in the supply chain as corroborated in all the interviews, and in reports by other agencies. Limited by computing power and infrastructure in the Indian lab, further advancements in these are planned for Summer 2011 when new facilities will be in place.

For the process validity, great care has been taken that processes used by the agents in the model match those derived from the interviews and various government regulations that are in force in all parts of the county. For example, the APMC(Agriculture Produce Market Committees)Act that regulates how agricultural products can be marketed and sold specifies that food processors and other consumers should not approach farmers directly to procure goods. The same process has been followed in the simulation. Qualitative checks have been performed to see whether the agents actually follow the routines and behave according to the rules built into the system. Similar to the structure validity, quantitative validation will be performed once the infrastructure to do so is in place.

For the predictive validity, data is collected from various years of mango production and growth. Once the structure and process validity are checked, this historical data will be used to check the predictive validity of the MAS, starting in the past.

5 Conclusions and Discussion

This paper introduced a new design of a gaming simulation in the field of supply chains and networks, called the Mango Mandi Gaming Simulation. The design of the MMGS is a multi-agent simulation on top of which a gaming simulation is being built. This allows for combining simulated agents with real human players, each representing actors in the supply chain. Key issue in this is the way that the multi-agent simulation is structured. In our design great care has been taken to mimic the structure and processes that occur in the real markets. Special attention in this market has the issue of wastage, which will be the first topic to study using the MMGS after validation has been completed. Before the game can be used validation studies need to be done, for which we have a list of criteria.Currently projected applications of the MMGS include, but are not limited to:

- Institutional economics experiments: similar to the work of Zuniga et al [16] and Meijer et al [17] researchers can set up sessions with controlled parameter settings to discover the factors driving the course of actions in the supply chain.
- Improved understanding of actor behavior: by tuning and improving the models in the software agents, the behavior of the simulated supply chain and real supply chain can be compared, leading to a better understanding of the behavior of real actors, similar to Tykhonovet al [18].
- Understanding interactions between software agents and human players: by allowing human players to play with software agents, the behavior of both agents and players can be compared leading to a better understanding of the limitations or efficacy of these interactions., related to the work of for instance Kamphorst al [19].

Given the societal issues in the Indian food chain context, and the role of scientific knowledge in the previous Green Revolution, the authors hope that approaches like the MMGS can contribute to solve the current lack of insight on what happens in the food markets.

References

1. Asia Development Bank, Special Report, Food Prices and Inflation in Developing Asia: Is Poverty Reduction Coming to an End?.
2. Arunachalam, V.S., and R. King (2004). Knowledge Management and the Indian Industrialization Process. In Davis, Subrahmanian and Westerberg, editors, Knowledge Management: Organizational and Technological Dimensions. Springer-Verlag.
3. Kaushik Basu. India's Foodgrain Policy: An Economic Theory Perspective. Economic & Political Weekly, 46(17).
4. Mattoo, A.,D. Mishra, and A. Narain(2007). Press Lauch presentation of *From Competition at Home to Competing Abroad: A Case Study of Indias Horticulture.*Washington: World Bank.
5. India Today. Agricultural Losses, 2009.
6. Corsi, T. M., S. Boyson, A. Verbraeck, S. P. Van Houten, C. Han, and J. R. Macdomald (2006). The real time global supply chain game: New educational tool for developing supply chain management professionals. Transportation Journal 45(3): 61-73.
7. Meijer, S.A. (2009). The organisation of transactions: Studying supply networks using gaming simulation. Wageningen Academic, ISBN.
8. Holland, J.H.(1995), "Can There Be A Unified Theory of Complex Adaptive Systems?", in: Harold J. Morowitz, Jerome L. Singer, editors. The Mind, The Brain, and Complex Adaptive Systems. Addison-Wesley.
9. Bekebrede, G., Meijer, S. (2009). Understanding complex infrastructure systems: the case of SimPort-MV2. IEEE. Second International Conference on Infrastructure Systems and Services, Developing 21st Century Infrastructure Systems, Chennai.
10. FAOSTAT. FAO. http://faostat.fao.org Accessed January 1-4, 2010.
11. Interviews and survey, conducted by CSTEP staff, Bangalore, 2009.
12. Sudha, M. and F. Kruijssen (2008). Varietal Differences in the Supply Chains of Two Mango Varieties in South India. ActaHortaculturae 787: 379-386.
13. Kriz, W.C., and Hense, J.U. (2006). Theory-oriented evaluation for the design of and research in gaming and simulation. Simulation & Gaming, 37(2), 268-283.
14. Raser, J. (1969). Simulations and society: An exploration of Scientific Gaming.
15. Peters, Vincent and Vissers, Geert and Heijne, Gerton, (1998). The Validity of Games. Simulation & Gaming, 29(1), 20-30.
16. Zuniga-Arias, G., S. Meijer, R. Ruben and G.J. Hofstede (2007) Bargaining power and revenue distribution in the Costa Rican mango supply chain: a gaming simulation approach with local producers. Journal of Chain and Network Sciences 7:2 pp 143-160.
17. Meijer, S.A.; Hofstede, G.J.; Omta, S.W.F.; Beers, G. (2008) The organization of transactions: research with the Trust and Tracing game. Journal on Chain and Network Science 8 (1). pp. 1 - 20.
18. Tykhonov, D.; Jonker, C.M.; Meijer, S.A.; Verwaart, D. (2008) Agent-Based Simulation of the Trust and Tracing Game for Supply Chains and Networks. Journal of Artificial Societies and Social Simulation 11 ((3) 1). - p. 1 - 30.
19. Kamphorst,, B. , A. van Wissen, and V. Dignum. Incorporating BDI agents into human-agent decision making research. In: H. Aldewereld, V. Dignum, G. Picard (Editors): Engineering Societies in the Agent World X, LNAI 5881, Springer, 2009.