



## **Four factor model in Indian equities market**

(Revised version of IIMA, W.P. No. 2013-09-05, Revised on September 5, 2014)

Sobhesh K. Agarwalla, Joshy Jacob & Jayanth R. Varma

**W.P. No. 2013-09-05**

September 2013

The main objective of the Working Paper series of IIMA is to help faculty members, research staff, and doctoral students to speedily share their research findings with professional colleagues and to test out their research findings at the pre-publication stage.

INDIAN INSTITUTE OF MANAGEMENT  
AHMEDABAD – 380015  
INDIA

## FOUR FACTOR MODEL IN INDIAN EQUITIES MARKET

(REVISED VERSION OF IIMA, W.P. No. 2013-09-05, REVISED ON SEPTEMBER 5, 2014)

Sobhesh K. Agarwalla, Joshy Jacob & Jayanth R. Varma\*

### Abstract

We compute the Fama-French and momentum factor returns for the Indian equity market for the October 1993 - December 2013 period using data from CMIE Prowess. We differ from the previous studies on this topic, in the Indian market, in several significant ways. First, we cover a greater number of firms relative to the existing studies. Second, we exclude illiquid firms to ensure that the portfolios are investible. Third, we have classified firms into small and big using a more appropriate cut-off considering the distribution of firm size. Fourth, as there are many instances of vanishing of public companies in India, we have computed the returns with a correction for the survival bias. During the period from January 1994 to December 2014, the average annual return of the momentum factor was 21.9%; the average annual return on the value portfolio (*HML*) was 15.3%; that of the size factor (*SMB*) nearly 0%; and the average annual excess return on the market factor (*MRP*) was 11.5%. This is a revised version of our earlier paper on this topic. The revision is carried out to primarily accommodate the data of firms which are retrospectively added to the prowess database by CMIE. The time series of daily, monthly and yearly returns on the factors and the underlying portfolios are made available at an [online data library](#). The authors would update the library on a monthly basis.

Keywords: Four factors, India, HML, WML, Momentum

JEL classifications: G12, C89

---

\*The authors are faculty members in the Finance & Accounting Area at the Indian Institute of Management, Ahmedabad (IIMA). All the authors have equally contributed to the paper. The authors can be contacted at [sobhesh@iimahd.ernet.in](mailto:sobhesh@iimahd.ernet.in), [joshyjacob@iimahd.ernet.in](mailto:joshyjacob@iimahd.ernet.in), and [jrvarma@iimahd.ernet.in](mailto:jrvarma@iimahd.ernet.in). Part of this research is supported by the R&P funding provided by IIMA. The authors acknowledge the excellent research support given by Mr. Ellapulli V. Vasudevan, Research Assistant at IIMA. We thank the CMIE for agreeing to disseminate the four-factor returns of the Indian market through the Prowess database. All errors are our own.

## 1 Introduction

This paper is part of our efforts toward making available the Fama-French and Momentum factors (four factor model) of the Indian equity market to academics and practitioners. The authors have collaborated with the Centre for Monitoring Indian Economy (CMIE) for regular update and timely dissemination of the factor returns. The collaboration allows computation of the factor returns based on the source data maintained by CMIE. This has led to differences in some of the factor returns as given in our earlier working paper on the four factor returns in the Indian market (IIMA, W.P. No. 2013-09-05) for two reasons. Firstly, in the earlier version, we had computed the various Fama-French and momentum portfolio returns based on the CMIE data available as of June 30, 2013. CMIE regularly updates the financial data of firms retrospectively as and when data are available. This often leads to the addition of more firms into the various portfolio later. The updation of the factor returns retrospectively with all the available financial data is also in the spirit of Fama and French (1993) who argued that in efficient markets, the pricing reflects the unpublished financial information of the firm. Secondly, portfolio construction and return calculations executed directly on the CMIE server ensures a more comprehensive coverage of the firms.

The estimation of four-factor returns accompanying this paper covers the October 1993 - December 2013 period. We would update the data on a regular basis while also attempting to extend the coverage backward in time. The objective is to provide data for the Indian market similar to what is provided for the US market at Kenneth French's website (French, n.d.). The time series of daily, monthly and yearly factor returns and the returns of the underlying portfolios are made available at <http://www.iimahd.ernet.in/~iffm/Indian-Fama-French-Momentum/>.

The starting of the estimates from 1993 is motivated by several considerations. First, interest rates in India were deregulated only in the early 1990s and therefore there was no market determined risk-free rate for earlier periods. As and when we extend the series backward, we will have to estimate the risk-free rate using some estimate of the magnitude of financial repression as discussed in Varma and Barua (2006). Second, the standard source of machine readable stock price and corporate financial data (the Prowess database published by the CMIE) begins only in the early 1990s. For this study, we have relied on data from Prowess and cannot therefore go back beyond the early 1990s. Data for the earlier periods has to be hand collected from multiple sources. We plan to perform this exercise and extend the data back to the early 1980s. While the major burst of economic reforms in India occurred in 1991, India had a vibrant equity market from at least the early 1980s, and we believe it is essential to extend the data back to cover this period.

Several authors including Connor and Sehgal (2001), Bahl (2006), Taneja (2010), Mehta and Chander (2010) and Tripathi (2008) have used or tested the Fama-French model or its variants in the Indian markets with a relatively small number of firms over relatively short periods of time. However, the study that comes closest to ours is Eun et al. (2010), who estimated the monthly size, value and momentum factors in India, for the period between July 1993 and December 2010. They used the data provided by Datastream and the factors were estimated based on total returns including dividends. We extend the analysis of Eun et al. in several ways. Firstly, our analysis covers a larger number of Indian firms provided by the CMIE Prowess database, the widely used database for academic research in India. Prowess covers more medium and small firms in the Indian market than Datastream. Secondly, we extend the factor estimates to daily frequency. Finally, while Eun et al. (2010) was a one-time exercise for a specific time period, we intend to provide these factors on an ongoing basis with regular updates.

## 2 Coverage of firms in the factors

We began with the list of all the firms listed in Bombay Stock Exchange (BSE)<sup>1</sup> covered in the CMIE Prowess database. Prowess had a total of 7,082<sup>2</sup> listed firms during the 1991 – 2013 period. However, of these 7,082 firms, only 6,943 firms had valid price and outstanding shares<sup>3</sup> data in Prowess.

The distribution of the market capitalisation of these 6,943 firms is given in Table 1. The number of firms covered, significantly increases from 1992 to 2012. The minimum and maximum number of firms covered during any one-year period is 2,156 (1992) and 5,304 (1995). The total market capitalisation of the firms during the same period (1992-2013) has gone up almost 30 times. It was around ₹67 trillion (around \$1.25 trillion) on June 2013. During the period, the median firm size has more than doubled and the average market capitalization has increased dramatically from around ₹1 billion in 1992 to ₹18 billion 2013. The average market capitalisation of the firms is very close to the market capitalization of the 90th percentile firm, indicating the presence of large number of small firms in India.

### 2.1 Liquidity Filter

All the firms that were traded on less than 50 days in a 12-month period prior to the portfolio creation date were excluded from the sample. The 50 trading days' criterion translates into roughly one trading day per

<sup>1</sup>The other leading exchange in India was the National Stock Exchange. However, the number of firms listed in BSE was substantially higher (more than 3 times) as compared to NSE. Further, almost all of the firms listed in NSE were also listed in BSE during the period covered in this study.

<sup>2</sup> CMIE Prowess database takes care of name changes and mergers and assigns a single firm identifier to the surviving entity before and after these events. We have used the CMIE identifier to distinguish the firms.

<sup>3</sup> In the remaining cases, either the price had a negative value or the outstanding shares were either missing or negative.

week. This ensures that the portfolios used for the estimation purpose are investible. The distribution of the firms based on their trading liquidity is given in Table 2. During the early years (1990s), when shares were traded in the physical form, there were more illiquid firms. The period from 1996-2000, which also corresponds with significant market decline in India, appears to have relatively poor liquidity. Between 2004 and 2010, the market enjoyed high liquidity and even firms in the first decile of liquidity traded nearly 100 days per year. The median number of trading days was 241 days out of about 250 trading days during the year 2011-2012.

The year-wise description of the firms eliminated by the liquidity criterion is provided in Table 2. Most of the firms, eliminated using the 50 trading days' filter, were small firms and belonged to the bottom 5 percentile, in terms of market capitalization. The liquidity filter eliminates a significant number of firms during 1997-2001 period. While more than 50% of the firms are excluded in the years 1998 and 2001, the market capitalisation of the excluded firms is very small. For instance, in the year where maximum number of firms are excluded (1998-1999) the market capitalisation of the excluded firms was only about 4.2%.

### 3 Estimation of size, value and momentum portfolios

#### 3.1 The Fama-French Size-Value portfolios and factors

The Fama-French methodology involves a cross classification of stocks on two dimensions – size, measured by market capitalization, and value, measured by the ratio of book value per share to market price per share –  $B/M$  ratio. This classification is tabulated below:

		Value as measured by $B/M$ ratio		
		Value ( $V$ )	Neutral ( $N$ )	Growth ( $G$ )
Size	Big ( $B$ )	$BV$	$BN$	$BG$
	Small ( $S$ )	$SV$	$SN$	$SG$

We follow Fama and French (2012) and use Value ( $V$ ), Neutral ( $N$ ) and Growth ( $G$ ) to denote the groups that Fama and French (1993) originally denoted as High ( $H$ ), Medium ( $M$ ), and Low ( $L$ ). Apart from being more descriptive labels, this notation also allows the letter  $L$  to denote the Losers group in the momentum analysis used later.

The portfolio  $BV$  can be regarded as the intersection of  $B$  and  $V$ , while  $BN$  is the intersection of  $B$  and  $N$ , and so on. Equally,  $B$  can be regarded as the union of  $BV$ ,  $BN$  and  $BG$ ; while  $V$  can be regarded as the union of  $BV$  and  $SV$ .

Following the literature, the Fama-French factors – size and value – were computed using the six disaggregated portfolios ( $BV$ ,  $BN$ ,  $BG$ ,  $SV$ ,  $SN$  and  $SG$ ) and not directly from the five aggregated portfolios ( $S$ ,  $B$ ,  $V$ ,  $N$  and  $G$ ). The reason for doing this was to make the size and value factors orthogonal to each other.

Fama and French (1993) described the construction of the size factor ( $SMB$ ) as follows:

“Our portfolio  $SMB$  (small minus big), meant to mimic the risk factor in returns related to size, is the difference, each month, between the simple average of the returns on the three small-stock portfolios ( $S/L$ ,  $S/M$ , and  $S/H$ )<sup>4</sup> and the simple average of the returns on the three big-stock portfolios ( $B/L$ ,  $B/M$ , and  $B/H$ )<sup>5</sup>. Thus,  $SMB$  is the difference between the returns on small- and big-stock portfolios with about the same weighted-average book-to-market equity. This difference should be largely free of the influence of  $B/M$ , focusing instead on the different return behaviors of small and big stocks.”

Put differently,  $SMB$  is the simple average of three return differences:  $SG - BG$ ,  $SN - BN$  and  $SV - BV$ , each of which is a difference between two portfolios that are matched in terms of value and differ only in size.

Similarly, the value factor  $HML$  (High minus Low)<sup>6</sup> is defined as the simple average of two differences:  $SV - SG$  and  $BV - BG$ , each of which is a difference between two portfolios that are matched in terms of size and differ only in value. The  $HML$  factor is thus designed to capture the effect of value while being largely free of the influence of size.

### 3.1.1 Size breakpoints (S & B portfolios)

Eun et al. (2010) bifurcated their size ranked portfolios into small and big based on the median size. However, we defined big firms ( $B$ ) as the top 10% by market capitalization and classified the remaining firms as small firms ( $S$ ). The naive approach of classifying all firms above the median as large and the rest as small was considered inappropriate for the Indian market given the size distribution of firms, because:

<sup>4</sup> $SG$ ,  $SN$  and  $SV$  in the Fama and French (2012) notation

<sup>5</sup> $BG$ ,  $BN$  and  $BV$  in the Fama and French (2012) notation

<sup>6</sup>VMG (Value minus Growth) would be a much more descriptive label for this factor, but the term  $HML$  is too well established to change. Fama and French (2012) while introducing the  $G/N/V$  notation for various portfolios, left the  $HML$  name for the value factor unchanged.

- The Indian market was dominated by a large number of small firms. For instance, the market capitalization of the 90th percentile firm was around ₹0.7 billion (approximately \$20 million) in 1997, ₹7 billion (approximately \$160 million) in 2004 and ₹16 billion (approximately \$300 million) in 2012. This is substantially lower than the NYSE size break-points published by [French \(n.d.\)](#).
- The average market capitalization of the firms over the years is close to the market capitalization of the 90th percentile firm.
- [Edwards and Cavalli-Sforza \(1965\)](#) suggested that the best split of observations into two clusters is one which minimizes the within-group sum of squares or maximizes the between-group sum of squares. We checked for various split-points starting from the 50th percentile to 90th percentile (based on market capitalization) in steps of 10 and found the within-group sum of squares to be the lowest at the 90th percentile in all the years.

It may be recalled that even though [Fama and French \(1993\)](#) used the median of NYSE listed stocks as the breakpoint for size, there were a disproportionate number of small stocks in their sample as most of the NASDAQ and AMEX stocks were smaller than the NYSE median.

### 3.1.2 Value breakpoints (V & G portfolios)

For the value breakpoints, we followed [Fama and French \(1993\)](#) and the stocks were grouped as below:

- High value group,  $V$ , consisted of the top 30% stocks in terms of the  $B/M$  ratio.
- Growth stocks (low value group),  $G$ , comprised the bottom 30% stocks in terms of the  $B/M$  ratio.
- The remaining stocks were grouped as neutral ( $N$ ) stocks.

### 3.1.3 Portfolio formation date

[Fama and French \(1993\)](#) formed their portfolios in June of each year after considering a 6-month gap from the fiscal yearends (December) to account for the time taken for the publication of accounting data. As the fiscal yearends for most Indian firms (89%) is March, assuming a 6-months gap<sup>7</sup> for publication of accounting data, we formed our portfolio in September of each year. In this, we have followed [Gregory et al. \(2009\)](#) who make the same argument for the UK, and have chosen to depart from [Eun et al. \(2010\)](#) who used the US formation date of 30th June. To summarize our methodology relating to portfolio formation date,

<sup>7</sup>The 6-months gap is more appropriate in the Indian context because Indian firms are required to hold their Annual General Meeting within six months of the fiscal yearend.

- At the end of September each year, the stocks were classified as Big ( $B$ ) and Small ( $S$ ), based on their market capitalisation at September-end.
- At the same time, the stocks were independently classified as Value ( $V$ ), Neutral ( $N$ ) and Growth ( $G$ ) based on their  $B/M$  ratio. There were two possibilities here depending on the financial yearend:
  1. If the firm's financial year ended in March, the  $B/M$  ratio was computed in September using the data as at the end of March of the same year.
  2. If the firm's financial year ended in any other quarter, the  $B/M$  ratio was computed in September of year  $t$  using the data as at the firm's financial yearend of year  $t - 1$ .

### 3.1.4 Number of firms in the size-value portfolios

In the size-value portfolio creation we have excluded all the firms with negative book value from the sample. The median number (over the years) of firms categorised into the different size-value intersection portfolios are given below.

		Value as measured by $B/M$ ratio		
		Value ( $V$ )	Neutral ( $N$ )	Growth ( $G$ )
Size	Big ( $B$ )	7	63	186
	Small ( $S$ )	666	821	494

The  $BV$  (Big-High value) portfolio is populated with fewer firms compared to the others. It indicates that most of the large Indian firms are also growth firms. In order to ensure that the portfolio returns are not driven by a few stocks, we did not consider the  $BV$  portfolio returns to estimate the  $SMB$  or  $HML$ , for years in which the number of stocks in the  $BV$  portfolio was less than five. This was the case for eight years. The choice of five stocks is based on the fact that a large part of the idiosyncratic risk is eliminated in a portfolio with as little as five stocks as may be seen in Figure 1 of [Evans and Archer \(1968\)](#) or Table 1 of [Statman \(1987\)](#).

## 3.2 Momentum Portfolios and Factors

As per the standard practice in literature ([Jegadeesh and Titman, 1993](#); [Carhart, 1997](#)), the classification of stocks as Winners ( $W$ ) and Losers ( $L$ ) was done based on their momentum returns at the end of each month.<sup>8</sup>

<sup>8</sup>The approach uses the daily total returns adjusted for dividends to estimate the 11-month holding period returns. We estimate monthly returns even when a stock is traded only for a part of its first month of trading using the available daily returns.



The momentum returns at the end of month  $t$  is the 11-month returns from the end of month  $t - 12$  to  $t - 1$ . By using the momentum returns, the stocks were grouped as below:

- $W$  – group consisted of the top 30% by the momentum return
- $L$  – group consisted of the bottom 30% by the momentum return

The buy-and-hold returns for month  $t + 1$  are calculated based on the above classification.

In line with the standard methodology (for example, [Fama and French \(2012\)](#)), the momentum portfolios were orthogonalized to the size factor. The size groups were created at the end of each month based on the size breakpoints as described in section 3.1.1. Based on the size and momentum groups, four size-momentum portfolios –  $WS$ ,  $WB$ ,  $LB$ ,  $LS$ , were formed every month, as below:

		Momentum	
		Winners ( $W$ )	Losers ( $L$ )
Size	Big ( $B$ )	$WB$	$LB$
	Small ( $S$ )	$WS$	$LS$

The median number of firms in the different size-momentum portfolios over the period are given below:

		Momentum	
		Winners ( $W$ )	Losers ( $L$ )
Size	Big ( $B$ )	105	31
	Small ( $S$ )	669	726

Similar to the method followed for size-value portfolios, we have excluded the portfolio in months where the number of stocks in the portfolio were less than five. As a result the  $BL$  portfolio was not considered in 10 months.

The momentum factor  $WML$  (Winners minus Losers) was computed as the simple average of the differences in the returns of  $WS - LS$  and  $WB - LB$ . The  $WML$  factor was thus designed to capture the effect of value while being largely free of the influence of size.

## 4 Survivorship Bias: Adjustment for Vanishing Firms

The literature documents many instances of the vanishing of public companies in India (Rao et al., 1999, for instance). In our dataset we have found that there were 3,184 firms that stopped trading during the period covered. Out of these, we could confirm that 439 firms had stopped trading due to mergers. Taking zero returns for all the remaining firms could have upwardly biased our return estimates as some of these firms could have disappeared (vanished) as an outcome of financial distress, leading to complete capital loss.

We have computed an alternative version of the factor portfolios assuming 100% capital loss for the firms vanishing due to distress<sup>9</sup>. Firms were identified as distressed if its last traded market price was below 50% of its face value. The year-wise distribution of these firms is given in Table 3. It can be seen that a large number of firms disappeared from the Indian market during the period 1996-2001. Most of these were small firms as they belonged to the bottom 2 deciles by market capitalization. The average market capitalization of these firms on their last trading day was only ₹0.2 million.

The change in the factor returns due to the above adjustment is somewhat trivial. Table 4 compares the portfolio returns with and without the adjustment. The difference in the annualized compounded returns over the 20-year period is about -0.1% for the *SMB* factor and -0.2% for the *HML* factor. This somewhat trivial outcome in terms of return occurs primarily due to the use of value weighted portfolios. Understandably, for the distressed firms, a significant portion of the loss in market capitalisation is already captured in the available trading data.

For future extension of the analysis, we intend to consider a lookahead period of 1-year for the purpose of classifying a firm as a vanishing firm. Therefore, the factor returns after adjusting for the vanishing firms could be computed only with a one-year lag.

## 5 Return on Size, Value, Momentum & Market Portfolios

### 5.1 Computation of Returns

The adjusted closing price (*Adjusted Close*) provided by CMIE Prowess is already adjusted for stock splits and other corporate actions but not for dividends.

The total return including dividends of day  $t$  was computed using prices from BSE for each unique firm identifier using the following formula:

---

<sup>9</sup>Some vanishing companies were not part of any portfolio on the last date because of other filters.

$$Total\ Return_t = \ln \left( \frac{Adjusted\ Close_t + DPS_t \frac{Adjusted\ Close_t}{Close_t}}{Adjusted\ Close_{t-1}} \right)$$

where  $DPS$  denotes the dividend per share. Using the above formula, we have computed the buy-and-hold returns for each size-value portfolio, as often employed in the factor return estimation (Roll, 1983). The weight of each stock in a portfolio was based on the market capitalization on the portfolio reconstitution date (the September yearend for the size and value portfolios, and the month-end for the momentum portfolios).

## 5.2 Estimation of daily four-factor returns

Daily four-factor returns were calculated using the portfolios created for the monthly 4-factors. As such on any particular day, stocks were classified on three different dimensions based on the following:

- The value-size intersections ( $BV, BN, BG, SV, SN, SG$ ) based on annual data.
- The momentum-size intersections ( $WB, WS, LB, LS$ ) based on monthly data.

## 5.3 Estimation of Market Risk Premium

The market portfolio is estimated as the value-weighted portfolio of all the stocks involved in the estimation of  $SMB, HML,$  and  $WML$  portfolios. The risk-free rate  $R_f$ , computed using the 91-days T-bill rate, is deducted from the return of the market portfolio to obtain the market risk premium  $MRP$  or  $R_m - R_f$ . The 91-day T-bill rate is sourced from the Reserve Bank of India's weekly auction data<sup>10</sup>. The implied yields have been converted to daily rates based on the number of trading days in the year following the issue.

## 5.4 Factor Returns

The cumulative logarithmic returns of the size, value, momentum and market portfolios are given in Figure 1. Over the period from January 1994 to December 2013, the cumulative market risk premium ( $MRP$ ) was about 59%. The cumulative return on the value factor ( $HML$ ) was about 216%. The size factor ( $SMB$ ) earned a negative cumulative return of about -58%). Our results suggest that the momentum earns significant positive returns (cumulative return of 341%) in the Indian market.<sup>11</sup> The correlations of the monthly factor

<sup>10</sup>URL: <http://dbie.rbi.org.in/DBIE/dbie.rbi?site=statistics>, under the main heading 'Financial Market' and sub-heading 'Government Securities Market'.

<sup>11</sup>The momentum factor return is not strictly comparable to the other two factor returns as it would involve a higher trading cost. This would happen as the momentum returns are estimated with monthly portfolio re-balancing whereas the other two factors have holding periods of one year.

returns are given in [Table 6](#). The correlations across the factor-returns are low and are in the lines of those reported from elsewhere in the world. A rigorous statistical analysis of the factor returns is required to arrive at reliable conclusions on the factors and their ability to explain the cross-sectional returns in India. This would require analysis over a longer period, which the authors intend to carry out in future.

## References

- Bahl, B. (2006), 'Testing the fama and french three-factor model and its variants for the indian stock returns', Available at SSRN 950899 .
- Carhart, M. M. (1997), 'On persistence in mutual fund performance', *The Journal of Finance* **52**(1), 57–82.
- Connor, G. and Sehgal, S. (2001), 'Tests of the fama and french model in india'.
- Edwards, A. and Cavalli-Sforza, L. (1965), 'A method for cluster analysis', *Biometrics* pp. 362–375.
- Eun, C. S., Lai, S., de Roon, F. and Zhang, Z. (2010), 'International diversification with factor funds', *Management Science* **56**, 1500–1518.
- Evans, J. L. and Archer, S. H. (1968), 'Diversification and the reduction of dispersion: an empirical analysis', *The Journal of Finance* **23**(5), 761–767.
- Fama, E. F. and French, K. R. (1993), 'Common risk factors in the returns on stocks and bonds', *Journal of financial economics* **33**(1), 3–56.
- Fama, E. F. and French, K. R. (2012), 'Size, value, and momentum in international stock returns', *Journal of Financial Economics* **105**(3), 457–472.
- French, K. R. (n.d.), 'Data library', [http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). Accessed: 2013-05-19.
- Gregory, A., Tharyan, R. and Christidis, A. (2009), 'The Fama-French and Momentum Portfolios and Factors in the UK', Available at SSRN: <http://ssrn.com/paper=1506960>.
- Jegadeesh, N. and Titman, S. (1993), 'Returns to buying winners and selling losers: Implications for stock market efficiency', *The Journal of Finance* **48**(1), 65–91.
- Mehta, K. and Chander, R. (2010), 'Application of fama and french three factor model and stock return behavior in indian capital market', *Asia Pacific Business Review* **6**(4), 38–56.
- Rao, K. C., Murthy, M. and Ranganathan, K. (1999), 'Some aspects of the indian stock market in the post-liberalisation period', *Journal of Indian School of Political Economy* **11**(4), 595–621.
- Roll, R. (1983), 'On computing mean returns and the small firm premium', *Journal of Financial Economics* **12**(3), 371–386.

Statman, M. (1987), 'How many stocks make a diversified portfolio?', *Journal of Financial and Quantitative Analysis* pp. 353–363.

Taneja, Y. P. (2010), 'Revisiting fama french three-factor model in indian stock market', *Vision: The Journal of Business Perspective* **14**(4), 267–274.

Tripathi, V. (2008), 'Company fundamentals and equity returns in india', Available at SSRN: <http://ssrn.com/abstract=1134651orhttp://dx.doi.org/10.2139/ssrn.1134651>.

Varma, J. R. and Barua, S. K. (2006), 'A first cut estimate of the equity risk premium in india', *Indian Institute of Management Ahmedabad, Working Paper* (2006-06-04).

**Table 1: Descriptive statistics of market capitalization of firms**

Year	Number of firms	Market capitalization - percentile (₹ million)					Total market cap. (₹ million)	Avg. market cap. (₹ million)
		10%	30%	50%	70%	90%		
1992 - 1993	2,156	31	70	144	332	1,593	2,257,449	1,047
1993 - 1994	3,103	34	78	153	391	1,986	4,276,565	1,378
1994 - 1995	4,484	36	71	132	308	1,482	5,529,141	1,233
1995 - 1996	5,304	21	39	68	156	874	4,718,224	890
1996 - 1997	5,044	9	19	41	102	679	4,772,805	946
1997 - 1998	4,160	7	16	36	102	728	4,537,186	1,091
1998 - 1999	3,793	7	17	39	118	1,000	4,923,274	1,298
1999 - 2000	4,048	12	26	58	175	1,429	8,242,648	2,036
2000 - 2001	3,495	8	21	51	155	1,206	6,001,943	1,717
2001 - 2002	3,084	10	25	58	190	1,520	5,809,285	1,884
2002 - 2003	2,899	10	25	66	242	1,914	6,803,081	2,347
2003 - 2004	2,839	11	30	89	402	3,692	12,100,902	4,262
2004 - 2005	2,975	23	76	229	854	6,956	17,864,922	6,005
2005 - 2006	2,940	26	96	344	1,427	11,885	28,027,785	9,533
2006 - 2007	3,087	29	118	442	1,729	15,609	40,924,175	13,257
2007 - 2008	3,207	40	148	523	2,034	19,475	56,391,429	17,584
2008 - 2009	3,221	30	96	290	1,087	10,997	39,750,024	12,341
2009 - 2010	3,412	40	149	495	1,939	19,299	62,666,128	18,366
2010 - 2011	3,512	43	160	509	2,032	20,421	68,027,024	19,370
2011 - 2012	3,698	32	114	345	1,406	14,853	60,371,950	16,326
2012 - 2013	3,698	31	110	338	1,492	16,143	66,900,915	18,091

The table shows the cross-sectional percentiles, total and average market capitalisation for various years for all listed firms. The market capitalization of a firm is taken as its average market capitalisation over the trading days of the firm during the period of 1-October to 30-September. The 2012-13 period covers only a 9-month period from 1 October, 2012 to 30 June, 2013.

**Table 2: Descriptive statistics of liquidity (Number of trading days per year)**

Year	Number of firms	No. of trading days - percentile								
		10%	20%	30%	40%	50%	60%	70%	80%	90%
1992 - 1993	2156	20	58	107	143	172	191	202	209	213
1993 - 1994	3103	11	36	72	111	153	190	207	215	219
1994 - 1995	4484	28	64	98	136	171	194	210	220	225
1995 - 1996	5304	24	67	115	151	179	199	214	226	233
1996 - 1997	5044	6	19	37	59	88	120	155	189	222
1997 - 1998	4160	3	6	14	29	50	86	128	173	224
1998 - 1999	3793	3	7	19	42	76	122	167	210	240
1999 - 2000	4048	6	21	47	81	117	152	188	220	245
2000 - 2001	3495	3	6	16	34	65	108	160	214	248
2001 - 2002	3084	4	16	39	76	125	171	213	240	250
2002 - 2003	2899	9	41	82	130	175	211	234	247	250
2003 - 2004	2839	16	73	135	191	226	246	254	256	256
2004 - 2005	2975	54	169	225	245	251	252	252	252	252
2005 - 2006	2940	95	190	233	245	249	250	250	250	250
2006 - 2007	3087	94	191	232	245	248	248	248	248	248
2007 - 2008	3207	97	192	232	245	249	250	250	250	250
2008 - 2009	3221	76	151	201	225	236	240	240	241	241
2009 - 2010	3412	96	197	236	247	249	249	249	249	249
2010 - 2011	3512	91	180	227	244	250	251	251	251	251
2011 - 2012	3698	54	127	187	227	241	248	249	250	250
2012 - 2013	3698	41	85	123	151	173	183	186	186	186

The table shows the cross-sectional percentiles (calculated using data of all listed firms) of trading days in Bombay Stock Exchange during 1-October to 30-September of various years. The 2012-13 period covers only a 9-month period from 1 October, 2012 to 30 June, 2013.



**Table 3: Number of firms that stopped trading over the years**

Calendar Year of last trading day	Number of Firms that stopped trading	Stopped trading due to mergers	Stopped trading for other reasons and had $P/FV \geq 50\%$ (no capital loss)		Stopped trading for other reasons and had $P/FV < 50\%$ (considered for 100% capital loss )	
			Number of firms	Formed part of any portfolio on the last trading day	Number of firms	Formed part of any portfolio on the last trading day
1992	9	2	7	0	0	0
1993	44	9	30	8	5	2
1994	43	10	25	11	8	2
1995	119	32	67	18	20	7
1996	260	28	130	27	102	54
1997	510	27	115	38	368	253
1998	247	25	56	8	166	57
1999	251	34	59	7	158	27
2000	402	29	85	23	288	104
2001	332	34	65	13	233	99
2002	145	23	33	11	89	32
2003	139	23	38	17	78	60
2004	115	14	25	9	76	36
2005	101	23	41	24	37	20
2006	81	36	19	9	26	22
2007	67	18	27	13	22	11
2008	52	17	29	12	6	1
2009	60	11	28	12	21	17
2010	74	23	35	10	16	7
2011	79	14	45	12	20	7
2012	54	7	28	9	19	16
<b>Total</b>	3,184	439	987	291	1,758	834

The table shows the number of firms that stopped trading over the years. Column 3 shows number of firms that stopped trading due to mergers. Columns 4-7 shows the number of firms that stopped trading for reasons other than mergers, showing separately the details of firms for which the *price/face value* on their last trading day was less than 0.50. The difference between columns 4 and 5, and columns 6 and 7, represents those firms which were not part of a portfolio due to various filters such as the liquidity filter.

**Table 4: Market and four-factors returns with and without survivorship bias adjustment**

Calendar Year	Four-factors with adjustment					Four-factors without adjustment				
	$R_m$	$SMB$	$HML$	$WML$	$R_m - R_f$	$R_m$	$SMB$	$HML$	$WML$	$R_m - R_f$
1994	19.4	29.0	15.1	-6.2	12.1	19.4	29.0	15.1	-6.2	12.1
1995	-44.0	-18.2	-15.0	11.2	-54.7	-44.0	-18.2	-15.0	11.2	-54.7
1996	-13.2	-38.9	-26.0	50.5	-23.2	-13.2	-38.9	-26.0	50.5	-23.2
1997	12.9	-52.2	-9.8	35.7	6.3	12.9	-52.1	-9.7	35.6	6.3
1998	-9.9	16.7	-9.3	-2.3	-17.4	-9.9	16.8	-9.0	-2.3	-17.4
1999	68.1	37.0	17.0	66.5	59.5	68.1	37.0	17.1	66.5	59.5
2000	-31.7	-30.0	15.8	-22.2	-40.4	-31.7	-29.7	16.2	-22.4	-40.4
2001	-24.1	-11.7	10.6	18.5	-31.3	-24.1	-10.4	12.3	18.1	-31.3
2002	18.9	-19.9	56.8	13.1	13.1	18.9	-19.9	56.7	12.9	13.1
2003	73.3	6.1	45.3	33.9	68.4	73.3	6.3	45.5	33.5	68.4
2004	19.6	16.9	35.2	22.5	14.9	19.6	17.0	35.4	22.2	14.9
2005	34.3	31.0	18.9	25.4	29.1	34.3	31.0	18.9	25.4	29.1
2006	32.4	2.4	1.3	29.8	26.2	32.4	2.4	1.3	29.5	26.2
2007	57.6	13.0	70.9	18.1	50.7	57.6	13.0	71.0	18.1	50.7
2008	-86.4	-42.2	-23.6	-9.7	-93.7	-86.4	-42.2	-23.6	-9.7	-93.7
2009	64.8	10.4	13.7	-24.8	61.3	64.8	10.4	13.7	-24.8	61.3
2010	11.7	5.0	4.2	16.6	6.4	11.7	5.0	4.2	16.6	6.4
2011	-31.6	7.2	-24.5	46.4	-39.2	-31.6	7.2	-24.5	46.3	-39.2
2012	25.2	1.8	7.1	-8.6	17.1	25.2	1.8	7.2	-8.6	17.1
2013	2.1	-21.8	12.9	26.3	-6.2	2.1	-21.7	13.0	26.2	-6.2
Cumulative	199.5	-58.3	216.5	340.8	59.2	199.5	-56.2	219.6	338.6	59.2
Mean	10.0	-2.9	10.8	17.0	3.0	10.0	-2.8	11.0	16.9	3.0
Max	73.3	37.0	70.9	66.5	68.4	73.3	37.0	71.0	66.5	68.4
Min	-86.4	-52.2	-26.0	-24.8	-93.7	-86.4	-52.1	-26.0	-24.8	-93.7
SD	41.1	25.4	26.3	24.1	42.1	41.1	25.3	26.3	24.1	42.1
Skewness	-0.4	-0.4	0.6	0.0	-0.4	-0.4	-0.4	0.6	0.0	-0.4

The table shows the annualised logarithmic market and four-factors returns (in percentage). The data covers the period from January 1994 to December 2013.

**Table 5: Size-Value and Size-Momentum portfolios' returns (adjusted for survivorship bias)**

Year	Size-value portfolios						Size-momentum portfolios			
	<i>BV</i>	<i>BN</i>	<i>BG</i>	<i>SV</i>	<i>SN</i>	<i>SG</i>	<i>WB</i>	<i>WS</i>	<i>LB</i>	<i>LS</i>
1994	23.4	12.1	11.6	54.0	50.2	36.9	12.2	56.6	42.0	41.9
1995	-50.7	-39.8	-33.0	-63.0	-61.3	-49.5	-39.4	-54.4	-47.1	-69.6
1996	-29.1	-10.9	-0.3	-59.3	-55.2	-35.5	0.7	-40.3	-66.9	-81.4
1997	-0.9	13.3	18.0	-40.6	-27.1	-22.8	23.1	-20.6	-37.0	-44.4
1998	-19.3	-22.7	-11.3	-7.0	-1.3	15.2	1.3	18.3	19.1	-4.5
1999	-13.6	36.8	68.7	107.9	79.0	83.9	101.4	117.4	17.5	64.5
2000	-5.3	-25.4	-31.6	-45.3	-36.8	-54.5	-35.3	-55.4	-10.7	-33.6
2001	-7.9	2.4	-28.9	-26.3	-3.7	-29.4	-17.5	-17.9	-65.5	-45.2
2002	87.1	71.5	2.5	53.1	34.9	22.8	22.8	33.1	-1.6	27.0
2003	80.2	111.7	66.5	115.1	98.5	75.0	94.4	94.0	60.0	55.1
2004		33.4	16.8	58.3	44.1	30.7	22.3	47.2	-0.1	25.2
2005		23.6	34.1	62.0	63.0	54.3	46.7	72.7	32.4	37.7
2006		17.3	38.1	39.3	16.4	37.2	41.7	34.7	28.7	2.3
2007		80.1	46.0	118.1	67.4	48.1	73.5	71.6	59.3	50.1
2008	-81.5	-74.5	-81.6	-117.0	-120.7	-108.1	-91.5	-116.4	-78.5	-117.4
2009	50.7	61.9	63.4	83.2	82.2	70.9	52.8	67.5	61.2	90.9
2010	14.4	17.4	10.2	20.4	21.5	19.0	14.2	25.6	1.0	5.6
2011	-64.0	-47.3	-28.7	-39.3	-47.5	-35.4	-17.0	-27.3	-68.0	-73.4
2012	15.4	27.5	24.0	32.5	29.0	29.6	27.0	32.7	42.4	27.2
2013	7.0	-8.9	6.2	-12.9	-12.0	-0.6	4.7	3.7	-24.8	-27.4
Cumulative	5.7	279.5	190.7	333.3	220.5	187.8	338.0	342.9	-36.5	-69.4
Mean	0.4	14.0	9.5	16.7	11.0	9.4	16.9	17.1	-1.8	-3.5
Max	87.1	111.7	68.7	118.1	98.5	83.9	101.4	117.4	61.2	90.9
Min	-81.5	-74.5	-81.6	-117.0	-120.7	-108.1	-91.5	-116.4	-78.5	-117.4
SD	46.5	45.5	38.4	66.3	57.1	50.2	45.9	58.2	46.3	56.0
Skewness	0.2	0.2	-0.4	-0.1	-0.5	-0.6	-0.2	-0.4	-0.3	-0.3

The table shows the annualised logarithmic returns (in percentage) of various size-value and momentum portfolios after adjustment for survivorship bias. The data covers the period from January 1994 to December 2013. The data of four-factor returns with adjustment in the last year should be taken as tentative.

**Table 6: Correlation matrix of monthly four-factors' returns (adjusted for survivorship bias)**

	<i>SMB</i>	<i>HML</i>	<i>WML</i>	$R_m - R_f$
<i>SMB</i>	100%			
<i>HML</i>	37%	100%		
<i>WML</i>	-7%	-17%	100%	
$R_m - R_f$	11%	22%	-16%	100%

Figure 1: Cumulative log-returns of the four factors (adjusted for survivorship bias)

