Understanding Investor behavior and it's implications on Capital Markets - The Indian Context

Bhaduri, Saumitra and Gupta, Saurabh

Madras School of Economics

21 November 2015

Understanding Investor behavior and its implications on Capital Markets – The Indian Context

Abstract

The paper aims to study the dynamic investor behavior and how it helps explain variation in stock returns. We propose a dynamic factor model to extract distinct latent factors representing fluctuations in asset returns due to changes in fundamentals and investor behavior. We study investor behavior under two broad categories, market-wide sentiment and herding. Our analysis suggests that both factors significantly impact the asset pricing and show varied volatilities across the sample. The model also ascertains empirical characteristics of the identified behavioral factors.

Dr. Saumitra Bhaduri

Professor, Madras School of Economics, Gandhi Mandapam Road,

Behind Government Data Center, Kottur

Chennai, Tamil Nadu, India - 600 025

Email : saumitra@mse.ac.in

Saurabh Gupta

Madras School of Economics

Email : gesaurabh@mse.ac.in

1

1. Introduction

Many studies have confirmed mispricing of assets and attributed it to the persistence of underreaction or overreaction of stock prices to public news. This has led to refutation of efficient market hypothesis over two underlying premises; security prices do not necessarily reflect all the information about a financial instrument and there are limits to arbitrage[1]. For example, moral-hazard problems help explain limits to arbitrage as it is argued that professional arbitrageurs manage other peoples' money (Shleifer and Vishny, 1997). Also, anomalies in asset pricing can be explained by behavioral biases of investor decisions. Investor decisions are at times irrational, driven by inherent behavioral tendencies of participants as opposed to rational line of thought. More often, the "sentiment" prevalent in the markets impacts the behavior of investors which is eventually reflected in asset prices. Such behavioral tendencies driven by sentiment and limits to arbitrage hinder the efficient process of eliminating mispricing in securities. Our paper empirically captures such behavioral phenomena and studies how it affects security prices.

Shleifer et al (1998) in their model of investor sentiment show that news associated with a security is only reflected in its price after a lag, due to behavioral reservations or sentiment. Even when there is evidence in the form of fundamental news, investors fail to incorporate the relevant information into security prices. As soon as the negative/positive sentiment is eliminated from the market, prices fall back to their fundamental values as investors correct their price expectations. However, if a large group of investors are influenced by one collective sentiment it can have severe repercussions over the market.

Therefore, it is often conjectured in the literature that the real markets consist of participants whose market decisions are often biased due to various psychological tendencies[2]. "Herding" is

---

[1] For a detailed study on why there are limits to arbitrage one can refer to Shleifer et al (1997)
[2] Tversky and Kahneman (1974) discuss the heuristics individuals resort to when they take decisions under uncertainty. Investor sentiment model by Barberis et al (1998) uses representativeness and conservatism heuristics to

one social consequence of such tendencies, where individual decisions tend to converge towards one collective decision. For instance, it is often the case that an investor is unwilling to take risks relying on his own private information which he thinks maybe insufficient. On the contrary, he will give more credibility to the decisions taken by his peers or hearsay in the market and follow the same. Shiller (1995) postulates that in a market environment as people interact regularly with each other, they mimic decisions of other investors, thereby form a "herd". A group of investors following a herd are driven by a collective herd-sentiment forming mis-informed price expectations and significantly contributing towards security mis-pricing.

In our model, we use a "top down" approach to study behavioral biases of investors, driven by sentiment, at an aggregate level. We categorize the factors influencing behavioral reactions of investors into two broad categories, market-wide sentiment and herding. Herding captures the reactions of investors in a group or a "herd", where individual decisions are correlated and converge towards a common juncture. In a market, several herd groups can co-exist, mutually exclusive of each other and driven uniquely by different sentiments. In our framework, emergence of signals leading to a sentiment in a herd is more localized, pertaining only to specific market segment. For example, a sentiment which leads a group of investors to come to a collective decision of picking growth stocks or value stocks would qualify for herding. However, public news and announcements are not always restricted to one market segment and more frequently propagate a collective sentiment over the market as a whole. Such investor sentiment which motivates deviation of every asset in the market from its fundamental value can be called a market-wide sentiment.

Several theoretical models have emerged explaining social mechanisms behind the herding phenomena which can be broadly categorized into rational or irrational models. Rational herding models argue that investors' price expectations, though rational, are influenced by the market environment. (Bikchandani and Sharma, 2000) introduce the basic idea of information cascades, which helps explain the inefficiency in investor decisions owing to reliance on actions of previous agents. (Scharfstein et al, 1990) extend similar ideas to managers and how they take

explain the mis-pricing of assets. A similar sentiment study by Daniel et al (1998) makes use of psychological tendencies such as self-attribution and over-confidence.

socially inefficient decisions due to various labor market conditions. Shleifer and Summers (1990) use the "noise trader" approach to explain irrational behavior of market participants and use the same to explain deviation of security prices from fundamentals.

The empirical studies have focused on testing under various events including cross country and cross market studies. Chan, Cheng and Khorana, (2000), provide empirical evidence analyzing herd behavior in the US, Hong Kong, South Korea, Taiwan, and Japanese stock markets. Most of the empirical literature stresses upon cross-sectional volatility of stock prices as a measure of herding. Christie and Huang, (1995) hypothesize that when individual returns herd around the market portfolio the dispersion in stock prices is relatively low. They use cross-sectional standard deviation to study the herd behavior; however, their results show that herding is prevalent only under extreme market conditions and not during regular market stress. Also, Chang et al, (2000) use an alternative measure, cross-sectional absolute deviation, to test for adverse herding across developed and emerging economies. They find out that only emerging markets happen to show significant evidence of herding. Hwang and Salmon, (2013) under similar guidelines extend their model to measure and capture the herding by studying dispersion in CAPM betas of assets. They separate adjustment to fundamentals, and herding due to market-wide sentiment by looking at variabilities in factor sensitivities.

It has remained a challenge to perform a holistic study where one can capture and distinguish various behavioral phenomena into an empirical framework and understand their effect on price dynamics. Further, given that such phenomenon arises out of individual actions it becomes difficult to quantify and proxy such factors with relevant information, without measurement errors. Although, it is worth looking at an alternative approach to study the investor sentiment as adopted by Baker and Wurgler in their recent studies. They argue that investor sentiment is a complex phenomenon at a market-wide level and cannot be precisely captured empirically as it involves different kinds of investors with varied behaviors and preferences. As a way out, they recognize sentiment proxies which represent the behavioral biases of investors. These proxies help explain effects of sentiment on various anomalies in the market.

Thus, we conclude that the question still remains as to how we precisely incorporate individual behavior in an empirical analysis. We explore the idea of using a state space framework to address this problem, which is ideal for capturing unobservable factors in a system. A state space framework has various advantages when it comes to dimensionality reduction, flexibility and the ease of estimation of the model. We assume that the three Fama and French factors[3] capture only the fundamental variances in stock prices i.e. when no behavioral aspects of individual investors are affecting the asset prices. As is evident from empirical and psychological evidence, there are various other factors which come into play while forming price expectations. We try to capture and define these factors by introducing two more latent variables in our model which account for the irrationality of the investors under stressful financial events and their reaction to public news and announcements.

Our study distinctively brings new perspective in the empirical study of investor sentiment. First, we distinguish between market-wide sentiment and herding by using two different representative factors constituting the investor behavior. The herding factor captures intrinsic price reactions pertaining to a specific industry whereas the market-wide sentiment studies the effect of global information on the same. Second, we consolidate our intuition into a dynamic factor model which allows for dynamic interaction of factors. The state space model produces relatively less correlated factors which help explain a specific phenomenon in the market more efficiently. Third, our model takes into account theoretical and psychological intuition underlying behavioral aspects of market participants and aggregates such phenomena at a macro level. Fourth, our model tries to capture the intuition presented by the Fama and French three factor model and the behaviorists' approach of irrational asset pricing on a single platform.

We develop and estimate this model to explain the security pricing for Indian Capital Markets during the post crisis period. Over the years, India has emerged as one of the most favored destinations for foreign investors among the developing markets with one of the highest market capitalization. Since the liberalization of capital market in 1991, FII's investment in Indian equity market has crossed $60 billion Bhaduri et al, (2013). The FII investment prospects for

---

[3] Fama and French (1993) discuss the three-factor model in detail. It has widely been accepted as an efficient model of rational asset pricing.

India are very bright considering the inherent advantages that the country has and its potential to absorb capital for its development and growth. Therefore, given the increasing importance of the Indian equity market as the most favored destination it is imperative for the Indian regulator to keep a constant vigil on herding in the market.

The paper follows with an explanation of methodology in the next section. Second, we mention the data used for our analysis, followed by estimation of results and their interpretation. Finally, the last section concludes the paper.

2. Methodology

As a consequence of investor decisions driven by "sentiment", market participants tend to over-react or under-react to information signals. Daniel et al (1998) in their model of investor sentiment categorize investors as informed or uninformed based on the information sets they use to form decisions. We assume that fundamental information accompanying a public signal is denoted by (S_t) which is necessary to price the assets correctly. However, uninformed investors often driven by sentiment and irrespective of their access to full information set S_t make use of information at hand; say information set (U_t). Where U_t is a subset of S_t i.e. $U_t \subset S_t$.

We therefore start by considering that the portfolio returns are biased owing to behavioral vulnerability of the market players. We assume that expectations of excess market returns $E_t^s(r_{mt})$ and individual asset portfolio returns $E_t^s(r_{it})$ are conditional upon the information set U_t and are biased in the following manner:

$$E_t^s(r_{it}/U_t) = E_t(r_{it}/S_t) + \delta_{it} \tag{1}$$

$$E_t^s(r_{mt}/U_t) = E_t(r_{mt}/S_t) + \delta_{mt} \tag{2}$$

where δ_{it} and δ_{mt} represent the bias. $E_t(r_{it}/S_t)$ and $E_t(r_{mt}/S_t)$ represent the fundamental values without any behavioral influences and conditional upon fundamental information S_t. The bias in equations (1) and (2) decide if the returns are over-valued or under-valued. A positive bias will signify decisions driven by positive sentiment whereas a negative bias will signify the opposite. We define the nature of bias by measuring it against the expected market returns, calling it degree of optimism when assets are over-priced and degree of pessimism when under-priced. Given by the following equation:

$$s_{it} = \frac{\delta_{it}}{E_t(r_{mt})}, \qquad s_{mt} = \frac{\delta_{mt}}{E_t(r_{mt})} \tag{3}$$

Further, we consider the behavioral influences of investors under two factors existing simultaneously in the market, market-wide sentiment and herd sentiment. Both factors impact the degree of optimism/pessimism in the market. A sentiment which affects assets across the market in equal degree is quantified by a market-wide sentiment factor (ms_{mt}), whereas a sentiment leading to herd formations is defined by (h_{it}), and ω_{it} captures any random behavioral deviations which cannot be captured by the former two. Thus, the degree of optimism/pessimism s_{it} constitutes the following

$$s_{it} = ms_{mt} + h_{mt} + \omega_{it} \tag{4}$$

Where it is expected that $E_t(s_{it}) = E_c(ms_{mt} + h_{mt} + w_{it}) = ms_{mt} + h_{mt}$, and we assume $E_t(w_{it}) = 0$. We substitute s_{it} from the above equation in equation (3) and write δ_{it} as,

$$\delta_{it} = [ms_{mt} + h_{mt} + \omega_{it}] * [E_t(r_{mt})] \tag{5}$$

segregating parametric terms on the basis of their characteristic nature. We substitute the above expression of bias in the expected returns of assets in equation (2) and finally obtain the returns equation as follows:

7

$$E_t^s(r_{it} / U_t) = E_t(r_{it}) + ms_{mt}E_t(r_{mt}) + h_{mt}E_t(r_{mt}) + \upsilon_{it} \qquad (6)$$

The equation (6) depicts the basic intuition underlying our empirical analysis. It denotes the expected portfolio returns as composed of expected fundamental changes in the characteristics of an asset ($E_t(r_{it})$), a factor capturing the market-wide sentiment ($ms_{mt}E_t(r_{mt})$) and a second factor capturing herding phenomena ($h_{mt}E_t(r_{mt})$).

The "behaviorists" have long debated the mispricing of assets as a consequence of irrational investor behavior. On the contrary, proponents of rational asset pricing models try to explain the same variation in stocks by recognizing structural patterns empirically and not relying on behavioral aspects. Fama and French (1993) three factor model is one significant example of the same which has enjoyed wider acceptance amongst researchers. It states that small cap and value stocks tend to exhibit significant positive returns as compared to others and thus propose three factors, market, size and value to explain variations in stock returns. In our model, we assume that changes in fundamentals ($E_t(r_{it} / S_t)$) are captured by the three Fama and French factors (market, size and value) since they represent rational pricing free from arbitrage. Thus, we write

$$E_t^s(r_{it} / U_t) = 3F.F.factors + Factor_1 + Factor_2 + \upsilon_{it} \qquad (7)$$

$3F.F.factors$ represent the three Fama and French factors i.e. We interpret $Factor_1$ and $Factor_2$ as fluctuations in the prevalent market sentiment and herding due to irrational behavior or bias in the asset returns due to the above mentioned factors. Also, we assume that the idiosyncratic variances are constant and time-invariant.

2.1 Dynamic Factor Model

We use a dynamic state space model to aggregate and study the behavior of the markets towards selected portfolios. It is well established that the unobserved factors extracted out of the state space framework have more desirable properties as expected out of explanatory factors, as is

empirically tested by He et al (2008) for the asset pricing context. We make use of the theoretical structure presented in the previous section and introduce a state space framework with five latent factors which are unobservable. The equation (7) motivates the basic structure of our state space framework.

Assuming a vector $R_t = \begin{bmatrix} R_{BH,t} R_{BM,t} R_{BL,t} R_{SH,t} R_{SM,t} R_{SL,t} R_{m,t} \end{bmatrix}$, representing the six excess demeaned portfolio returns sorted on size (B, S) and values (H, M , L) and one additional excess market return (m) . Let $F_t = \begin{bmatrix} F_{mkt} F_{size} F_{btm} F_{sent} F_{herd} \end{bmatrix}$ denote a vector of zero-mean unobserved state/latent variables. Where F_{mkt}, F_{size}, F_{btm} denote the rational three factors from the Fama and French three factor model, F_{sent} ($Factor_1$ in equation (7)) denotes the market-wide sentiment and F_{herd} ($Factor_2$ in equation (7)) represents the herding factor. The dynamic factor model is specified as: The measurement equation in its matrix form looks like:

$$
\begin{pmatrix} R_{BH,t} \\ R_{BM,t} \\ R_{BL,t} \\ R_{SH,t} \\ R_{SM,t} \\ R_{SL,t} \\ R_{m,t}^s \end{pmatrix} = \begin{pmatrix} \lambda_{1,1} & \lambda_{1,2} & \lambda_{1,3} & \lambda_{1,4} & \lambda_{1,5} \\ \lambda_{2,1} & \lambda_{2,2} & \lambda_{2,3} & \lambda_{2,4} & \lambda_{2,5} \\ \lambda_{3,1} & \lambda_{3,2} & \lambda_{3,3} & \lambda_{3,4} & \lambda_{3,5} \\ \lambda_{4,1} & \lambda_{4,2} & \lambda_{4,3} & \lambda_{4,4} & \lambda_{4,5} \\ \lambda_{5,1} & \lambda_{5,2} & \lambda_{5,3} & \lambda_{5,4} & \lambda_{5,5} \\ \lambda_{6,1} & \lambda_{6,2} & \lambda_{6,3} & \lambda_{6,4} & \lambda_{6,5} \\ \lambda_{7,1} & \lambda_{7,2} & \lambda_{7,3} & \lambda_{7,4} & \lambda_{7,5} \end{pmatrix} \begin{pmatrix} F_{mkt,t} \\ F_{size,t} \\ F_{btm,t} \\ F_{sent,t} \\ F_{herd,t} \end{pmatrix} + \begin{pmatrix} \upsilon_{BH,t} \\ \upsilon_{BM,t} \\ \upsilon_{BL,t} \\ \upsilon_{SH,t} \\ \upsilon_{SM,t} \\ \upsilon_{SL,t} \\ \upsilon_{m,t} \end{pmatrix}
$$

(8)

We assume that the unobserved state variables follow an autoregressive process of order 1 i.e. AR(1) process[4]. The transition equation when expressed in matrix form looks like:

---

[4] This assumption follows from the model used by Stock & Watson (1988) in their study. Where they use a dynamic factor model to formulate a co-incident index on inflation. Further, we put the same restrictions on latent variables as suggested by Stock & Watson (1998).

$$\begin{pmatrix} F_{mkt,t} \\ F_{size,t} \\ F_{btm,t} \\ F_{sent,t} \\ F_{herd,t} \end{pmatrix} = \begin{pmatrix} \phi_1 & 0 & 0 & 0 & 0 \\ 0 & \phi_2 & 0 & 0 & 0 \\ 0 & 0 & \phi_3 & 0 & 0 \\ 0 & 0 & 0 & \phi_4 & 0 \\ 0 & 0 & 0 & 0 & \phi_5 \end{pmatrix} \begin{pmatrix} F_{mkt,t-1} \\ F_{size,t-1} \\ F_{btm,t-1} \\ F_{sent,t-1} \\ F_{herd,t-1} \end{pmatrix} + \begin{pmatrix} \xi_{mkt,t} \\ \xi_{size,t} \\ \xi_{btm,t} \\ \xi_{sent,t} \\ \xi_{herd,t} \end{pmatrix}$$

(9)

We can represent the same matrix equations as a state space representation:

$$R_t = \lambda F_t + \upsilon_t \tag{10}$$

$$F_t = \phi F_{t-1} + \xi_t \tag{11}$$

Where υ_t and ξ both follow joint normal distributions, with the following restrictions:

$$\upsilon_t \sim i.i.d.N(0, \mathbb{Z})$$

$$\xi_t \sim i.i.d.N(0, \mathbb{Q})$$

$$E\left[\upsilon_t \xi_\tau'\right] = 0$$

for all t and τ. Where, $\mathbb{Z} \equiv diag\left[\sigma_{BH}^2 \sigma_{BM}^2 \sigma_{BL}^2 \sigma_{SH}^2 \sigma_{SM}^2 \sigma_{SL}^2 \sigma_M^2\right]$ is (7x7) co-variance matrix of idiosyncratic disturbances in portfolio and market returns. We identify the covariance matrix $\mathbb{Q}$ as a (5x5) identity matrix with no considerable implications on the results. λ denotes a vector of factor loadings of different factors on the asset portfolio returns, restrictions imposed on λ are discussed in the following sections.

2.2 Using the Kalman Filter to extract the latent factors

Given the dynamic state space framework, the most optimum estimation technique turns out to be the Kalman Filter. We use the Kalman Filter to estimate parameters contained in matrices λ

(factor loadings), $\mathbb{Z}$ (variances of the measurement equation) and ϕ (Auto-correlation coefficients of the latent factors) and then reiterate it recursively to extract the latent factors.

The estimation process follows a two staged process, prediction and update. We first predict the expected asset returns given the information set available and then update the same using Kalman gain to obtain optimal returns. At the beginning of the period we initialize the Kalman Filter with unconditional means and variances. At the start of time t, we first predict the returns R_t conditional upon the information available till time t-1. To achieve this we first extract the conditional means and variances of latent factors $F_{t/t}$ (unobserved latent factors).

1. Given information set I_{t-1}, the model will make its own predictions of the factors using the same. The unobserved dynamic factor $F_{t/t-1}$ is calculated using the following equation

$$F_{t/t-1} = \mu + \phi F_{t-1/t-1} \qquad \text{(12)}$$

2. The construction of our state space framework entails that the factor variances be constant and normalized at 1. The prediction of factor covariance matrix ($P_{t/t-1}$) conditional upon information available till t-1 is also predicted in a similar manner.

$$P_{t/t-1} = \phi P_{t-1/t-1} \phi' + \mathbb{Q} \qquad \text{(13)}$$

3. The factor estimates from previous steps help predict asset returns. The prediction error is thus calculated by taking the difference of the actual and predicted values of the portfolio returns in the given period, given by

$$\eta_{t/t-1} = R_t - R_{t/t-1} \qquad \text{(14)}$$

4. Following the prediction stage, we update the factors in order to obtain ex-post latent factors. The dynamic factors conditional upon information available at the beginning of time period t is

$$F_{t/t} = F_{t/t-1} + K_t \cdot \eta_{t/t-1} \qquad \text{(15)}$$

11

Where $K_t = P_{t/t-1} \cdot \beta_t{}' (f^{-1}_{t/t-1})$ represents the Kalman gain matrix, and $f^{-1}_{t/t-1}$ represents variance-covariance matrix of the forecast error.

2.3 Factor Identification

The model specified in the previous section is not exactly identified and is unfit for estimation. The unrestricted model defined above has 47 (40 factor loadings + 7 variances) free parameters for estimation[5]. However, as the independent variance-covariance terms of demeaned returns suggest, we can only estimate 28 parameters with the model[6]. Therefore, it becomes imperative to put restrictions on factor loadings and variances of our model. This has its own advantages; a restricted model will provide unique identity to the latent variables and free them from any random substitutions.

To obtain exact identification, some factor loadings are restricted for set portfolios depending upon the nature of the factors under study, especially for the behavioral factors. In the matrix equation (9), first column is determined by the market factor which will have a different impact on every individual portfolio and we assume the loading on market returns as one.

$$\lambda_{1,1} = \beta_{mBH} \, , \; \lambda_{2,1} = \beta_{mBM}$$
$$\lambda_{3,1} = \beta_{mBL} \, , \; \lambda_{4,1} = \beta_{mSH}$$
$$\lambda_{5,1} = \beta_{mSM} \, , \; \lambda_{6,1} = \beta_{mSL}$$
$$\lambda_{7,1} = 1.0$$

The second column pertains to the size factor. The impact of the size factor will have a common effect on all the individual portfolio returns of the same size. Hence, we have a common parameter for stocks with same size. As per our data we have two sets of portfolios on the basis of size i.e. big or small.

---

[5] With no implications on results we normalize disturbances of the transition equation to one as suggested by Stock & Watson, (1988).
[6] The number of variance-covariance terms is given by n(n+1)/2, where n represents the number of portfolios or number of observed series in state space framework.

$$\lambda_{1,2} = \lambda_{2,2} = \lambda_{2,2} = \beta_B, \ \lambda_{4,2} = \lambda_{5,2} = \lambda_{6,2} = \beta_S$$
$$\lambda_{7,2} = 0.0$$

Similarly, third column accounts for the value factor or portfolios sorted on the basis of book-to-market ratio. We select a common parameter for stocks with same value. The data marks three sets of portfolios segregated on book-to-market ratio, High, Medium and Low. Each category will have a different factor loading with no effects on the market returns.

$$\lambda_{1,3} = \lambda_{4,3} = \beta_{vH}, \ \lambda_{2,3} = \lambda_{5,3} = \beta_{vM}, \ \lambda_{3,3} = \lambda_{6,3} = \beta_{vL}$$
$$\lambda_{7,3} = 0.0$$

Fourth, we assume that market-wide sentiment has a global existence and all individual portfolio returns will have a similar sensitivity towards the market sentiment. Public signals which are of common knowledge impact every investor in a similar manner. Therefore, a market-wide sentiment will have similar repercussions on investor reactions across the market and related price expectations will also respond similarly. Also, the sentiment factor will have a lasting effect on the excess market returns. Hence, the restrictions will look like:

$$\lambda_{1,4} = \lambda_{2,4} = \lambda_{3,4} = \lambda_{4,4} = \lambda_{5,4} = \lambda_{6,4} = \beta_{sent}$$
$$\lambda_{7,4} = \beta_{msent}$$

The last factor accounts for the herding phenomena, when investors mimic decisions of other and suppress private information. We assume that the related sensitivities do not vary amongst the same sized stocks. Therefore, we adopt a similar pattern for factor loadings on herding factor as in the size factor. Further, we also assume that herding also impacts the returns on the market portfolio.

$$\lambda_{1,5} = \lambda_{2,5} = \lambda_{3,5} = \beta_{herdB}, \ \lambda_{4,5} = \lambda_{5,5} = \lambda_{6,5} = \beta_{herdS},$$
$$\lambda_{7,5} = \beta_{herdm}$$

3. Data

Our unique sample includes data from the financial crisis period 2007-2008 which helps us better understand and distinguish factors on investor behavior and fundamentals. We opt to choose six, size and BTM sorted, Fama & French (1993) portfolios pertaining to the Indian Context from 2007 to 2014. This helps us establish our results using data which has found a wider acceptance amongst researchers. Also it helps us contrast the effects of our factors over different market segments. The excess individual portfolio returns are collected from (Agarwalla, Jacob, & Varma, 2013). They use the same methodology as used by Fama & French, (1993) to construct similar portfolios dedicated to the Indian Capital Markets. To make our findings more robust we use daily returns data and a survivorship-bias adjusted data. Data adjusted for survivorship-bias helps us eliminate any companies which have shut down their business in the interim periods.

4. Results and Interpretation

A descriptive statistics of the latent factors is displayed in Table 1. The correlation between the factors is also reported in Table 2. As expected, the correlation amongst the latent factors is very minimal. Weak correlation helps better explain the variation in portfolio returns and immunes the model estimates to sensitivities relating to data. This also implies that the latent factors are independently more powerful with each specifying and capturing separate meaningful information.

Table 1: Descriptive Statistics of the Latent Factors from the State Space Model.

Factor	Median	Min	Max	Kurtosis	S.D
FF Factor 1 (F_{mkt})	0.02	-6.69	6.65	4.36	0.91
FF Factor 2 (F_{size})	0.07	-11.47	7.65	13.71	0.99
FF Factor 3 (F_{btm})	-0.01	-8.68	6.78	6.27	0.99
Sentiment Factor(F_{sent})	0.02	-6.07	9.10	7.45	0.99
Herding Factor (F_{herd})	-0.09	-4.63	8.06	5.08	1.42

Note: This table reports descriptive statistics of the five factors $\left[F_{mkt}, F_{size}, F_{btm}, F_{sent}, F_{herd} \right]$ extracted out of the state space model defined in equation (8) and (9). Where S.D represents the standard deviation of the factors.

Table 2:- Correlation between the Estimated Dynamic Factors

Correlation Coefficients	FF Factor 1 (F_{mkt})	FF Factor 2 (F_{size})	FF Factor 3 (F_{btm})	Sentiment Factor (F_{sent})	Herding Factor (F_{herd})
FF Factor 1 (F_{mkt})	1.00				
FF Factor 2 (F_{size})	0.01	1.00			
FF Factor 3 (F_{btm})	0.05	0.27	1.00		
Sentiment Factor (F_{sent})	-0.26	-0.47	0.10	1.00	
Herding Factor (F_{herd})	-0.05	-0.05	-0.10	-0.23	1.00

Note: This table reports the correlation coefficients between the estimates of the five latent factors $\left[F_{mkt}, F_{size}, F_{btm}, F_{sent}, F_{herd}\right]$ extracted out of the state space model defined in equation (8) and (9).

The estimated coefficients of the ex-post dynamic factors are presented under Table 3. We observe that most of the parameters are statistically significant. As expected, factor loadings for the fundamental Fama and French factors ($F_{mkt}, F_{size}, F_{btm}$) significantly impact the excess portfolio returns. These results are consistent with the conventional results exhibiting value and size effects. Our results show that returns in the small-stock group ($R_{SH,t}, R_{SM,t}, R_{SL,t}$) are more sensitive towards the size factor F_{size}, exhibiting positive coefficients, whereas negative coefficients for big-stock group ($R_{BH,t}, R_{BM,t}, R_{BL,t}$). The market factor shows a significant positive impact on all portfolio returns. Also, the value effect is very evident as high-valued stocks show a significant positive factor loading, denoting that stocks with a higher book-to-market ratio are more sensitive towards the value factor.

Table 3: Estimated Parameters of the Dynamic Factor Model

Estimated Parameter	Factor Loadings/Coefficients	T-Stat (S.E)
β_{mSL}	0.933	49.72*** (0.019)
β_{S}	0.694	59.61*** (0.011)
β_{vL}	-0.043	-20.08*** (0.002)
β_{sent}	1.422	58.42*** (0.024)
β_{herdS}	0.096	2.58*** (0.037)
β_{mSM}	1.054	52.36*** (0.020)
β_{vM}	0.190	25.58*** (0.007)
β_{herdm}	0.087	2.29** (0.038)
β_{mSH}	0.973	38.09*** (0.025)
β_{vH}	0.529	29.60*** (0.018)
β_{herdB}	0.084	2.25** (0.037)
β_{mBL}	0.866	284.67*** (0.003)
β_{B}	-0.080	-44.03*** (0.002)
β_{mBM}	1.641	139.91*** (0.012)
β_{mBH}	1.772	41.24*** (0.043)
β_{msent}	1.421	57.43*** (0.025)
σ^2_{SL}	0.183	-16.72*** (0.011)
σ^2_{SM}	0.261	42.51*** (0.006)
σ^2_{SH}	0.173	4.43*** (0.039)
σ^2_{BL}	0.000	0.00 (0.012)
σ^2_{BM}	0.299	52.82*** (0.006)
σ^2_{BH}	1.506	60.42*** (0.025)
σ^2_{M}	0.000	0.00 (0.007)

Note: This table reports the parameter estimates (factor loadings and variances) of the measurement equation (6) from the state space model defined in equation (8) and (9). The estimation results are reported for a daily sample data, spanning across 7 years (7 years: 01/10/2007 – 30/12/2014). The six Fama and French portfolios sorted on size and book-to-market ratio are notated as SL, SM, SH, BL, BM, BH and one additional market portfolio is denoted by M. Similarly, notation used for behavioral factors follows: a subscript "herd" is used for the herding factor and "sent" for the market-wide sentiment.

The behavioral factors, market-wide sentiment and herding, also show significant results. The excess portfolio returns are more sensitive towards the market-wide sentiment than the herding factor. Market-wide sentiment factor (F_{sent}) quantifies the extent to which investors misprice expectations based on a public signal which affects the market as a whole. Such signals, in the form of public news, can be encountered more frequently and over extended periods of time. Therefore, investors and eventually the asset prices are more prone to the global market-wide sentiment. Also, information revelation through such news doesn't adversely affect the stock prices as compared to the herding phenomena. This is because investors are quick to re-calibrate their expectations as the nature of information revealed is of common knowledge. This justifies a higher sensitivity in the market sentiment factor.

The fundamental information pertaining to a particular sector/market segment is usually more complex and hard to rely on, making it difficult for investors to re-calibrate their expectations, given their behavioral biases. Therefore, the adverse effects of herding can only be observed when its repercussions over stock prices become consequential. Also, this limits the sensitivity of portfolio returns to the herding factor F_{herd} relative to the sentiment factor. This is reflected in our results, where the factor loadings for the herding factor are almost half than that for the market-wide sentiment factor.

The auto-correlation coefficients for all the factors show statistically significant results. We expect a high magnitude and persistence for the herding factor since it emerges from a behavior when people mimic decisions of other investors over a period of time. The auto-correlation coefficient for the herding factor (ϕ_5) is of the magnitude of 0.9, relatively higher than for all the other factors. This result is expected and is consistent with theory of cascading information proposed by Bhikchandani et al, (1998). It states that information cascades are triggered when investors tend to defy their private signals and follow decisions already taken by other individuals in the herd. Also, cascades happen over a period of time, and investors' decisions are dependent on reactions of other investors in the past.
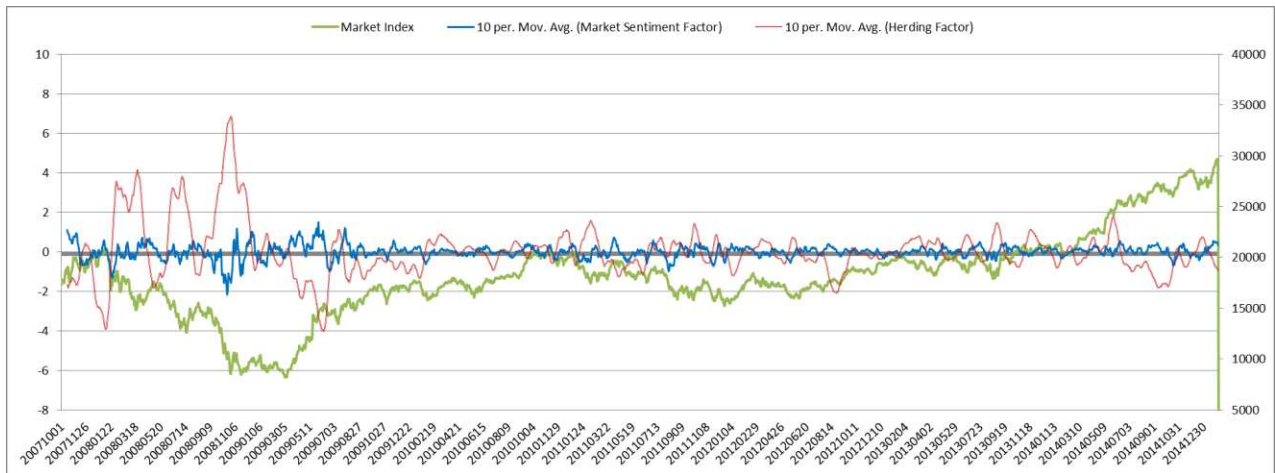
To check for the efficacy of our results we performed a residual analysis of the model and compared it with the three factor model proposed by He et al (2010). We estimated the model using only three factors, dropping the behavioral factors, and recorded the residual series for the same. We use the same restrictions on factor loadings as proposed for the fundamental Fama and French factors in our five-factor model. The residual analysis shows that mean squared error (MSE) does not show a significant change/gain between the two models. However, for certain portfolios (BL, BM) we see a considerable drop in MSE by 2-3%. This shows that our model of five latent factors best fits for a selected set of portfolios.

Table 4:- Estimated Parameters: Coefficients of Transition Equation

Parameter	Coefficient	T-Stat
Φ_1	0.095	3.82*** (0.025)
Φ_2	0.104	4.69*** (0.022)
Φ_3	0.329	11.69*** (0.028)
Φ_4	0.082	3.37*** (0.024)
Φ_5	0.945	31.89*** (0.03)

Note: This table reports auto-correlation coefficients of the transition equation (8) from the state space model defined in equation (8) and (9). The estimation results are reported for a daily sample data, spanning across 7 years (7 years: 01/10/2007 – 30/12/2014).

Figure 1 Herding and Sentiment against the Market Index

Note: The figure plots the ex-post behavioral dynamic factors for sentiment F_{sent} and herding F_{herd}, against the Indian Markets Index, BSE-SENSEX over a period of seven years (7 years: 01/10/2007 – 30/12/2014).

Further, we also plot the market-wide sentiment and the herding factor alongside market index to depict the episodes of herding across the sample period. We see an increased volatility of the behavioral factors during the period marked by financial crisis of 2008. However, the volatility subsides in the later period after 2009 when the world economy tries to recover from the crisis. As is evident in Figure 1, a very adverse form of herding is persistent during the crisis period, which is a cause for volatility in the herding factor. Since a financial event of such a magnitude tends to create an environment of panic and fear amongst the investors the rationality in their decisions while picking stocks is questioned.

Following this, to test the performance and robustness of our dynamic factors we divide our sample into two time frames, the crisis period and the post crisis period. Ideally, we expect the herding to be highly volatile and persistent during the crisis period and not so much during the post-crisis period whereas the market-wide sentiment must remain significant throughout. The first time frame, crisis period, takes into account the period from 2007 till late 2009. The second time frame, post-crisis period, uses sample after 2009. In order to check the robustness of our findings, we therefore estimate model parameters for these two time frames.

The Fama and French factors exhibit the same characteristics as demonstrated in the full sample and hence are not reported. However, the ex-post dynamic factors for herding (F_{sent}) turn out to be insignificant in the post crisis period whereas during the crisis period the factor loadings (β_{herdB}, β_{herdS}, β_{herdm}) still tend to be significant. The estimated factor loadings and auto-correlation coefficients for the herding and sentiment factor are reported in Table (5) and Table (6), respectively.

The crisis period was triggered by the collapse of housing bubble. Our model suggests that the bubble was aggravated by the convergence of investors' interest in securities related to the sub-prime mortgages. Since most of the major financial economies are linked to each other, repercussions of the crisis were seen all over the world. The crisis led market participants, globally, in a state of confusion and irrational behavior. This in turn gave rise to innovations and volatility in the financial markets. Our results very clearly record this volatile behavior as is evident in Figure (1).

However, post-crisis we see that the herding factor turns out to be insignificant. Also, as depicted using the full sample, the volatility in the factor also tends to subside as we move away in time from the crisis period.

Table 5:- Factor Loadings for the latent behavioral factors

Parameters	Crisis Period (2007-2009)	Post – Crisis Period (2010-2014)
	Coefficients and t-statistic	Coefficients and t-statistic
β_{herdS}	-0.09** (-2.78)	0.00 (0.60)
β_{herdB}	-0.08* (-2.34)	0.00 (.29)
β_{herdm}	-0.09* (-2.41)	0.00 (0.87)
β_{sent}	1.42*** (62.11)	0.87*** (51.52)
β_{msent}	1.42*** (60.84)	0.86*** (50.35)

Note: This table reports the factor loadings of the behavioral factors $[F_{herd}, F_{sent}]$ for equation (8). The sample daily returns data was split into two parts labeled as Crisis Period (2007-2009) and Post-Crisis Period (2009-2014).

Table 6:- Auto-correlation coefficients for the herding and sentiment factor

Parameters	Crisis Period (2007-2009)	Post – Crisis Period (2009-2014)
	Coefficients and t-statistic	Coefficients and t-statistic
Φ_1	0.09*** (4.06)	0.05*** (1.68)
Φ_2	0.10*** (4.50)	0.17*** (7.13)
Φ_3	0.33*** (12.77)	0.39*** (11.96)
Φ_4	0.08*** (3.37)	0.14*** (5.91)
Φ_5	0.94*** (37.67)	0.99*** (791.17)

Note: This table reports the auto-correlation coefficients of the latent factors $[F_{mkt}, F_{size}, F_{btm}, F_{herd}, F_{sent}]$ for equation (9). The sample daily returns data was split into two parts labeled as Crisis Period (2007-2009) and Post-Crisis Period (2009-2014).

5. Conclusion

We successfully devise a model to aggregate investor behavior, broadly categorized into market-wide sentiment and herding, empirically testing their influence over mispricing of asset prices. We construct a state space model and estimate parameters using the Kalman Filter, extracting the ex-post dynamic latent factors and study their behavior across time. We see that the Fama and French factors extracted from our model significantly impact the portfolio returns. Also, portfolio returns are restricted by exposure to sensitivities in the market sentiment and herding factor which also turn out to be significant for the whole sample. We also show that the behavioral factors exhibit properties of serial correlation. The herding factor in particular exhibits high correlation of the order of 0.9, denoting that information cascades and other behavioral mechanisms propagate over a distributed period of time and have adverse effects on portfolio returns.

To study the trend of market sentiment and herding in markets around the financial crisis of 2008 we divide our full sample into crisis and post-crisis period. Our analysis demonstrates that during the crisis period the factor loadings for herding factor are positive and significant whereas in the post-crisis period it is not. The market-wide sentiment on the other hand shows persistence across various samples tested for. This proves that keeping in view occurrences of financial events; investors tend to resort to adverse herding more than when the environment is relatively more stable. We successfully formulate a dynamic factor model and test it for the Indian context which shows favorable results taking into consideration the behavioral factors as well as the fundamental systematic risks.

References

1. Agarwalla, S. K., Jacob, J., & Varma, J. R. (2013). "Four factor model in Indian equities market". *Indian Institute of Management, Ahmedabad*, W.P. No. 2013-09-05.
2. Baker, M., & Wurgler, J. (2006). "Investor Sentiment and the Cross-Section of Stock Returns". *Journal of Finance*, 1645-80.
3. Baker, M., & Wurgler, J. (2007). "Investor Sentiment in the Stock Market". *NBER Working Paper Series*.
4. Barberis, N., Andrei, S., & Robert, V. (1998). "A model of investor sentiment". *Journal of Financial Economics*, 307-343.
5. Bhaduri, S. N., & Mahapatra, S. D. (2013). "Applying an alternative test of herding behavior: A case study of the Indian stock market". *Journal of Asian Economics*, 43-52.
6. Bikhchandani, S., & S., S. (2000). "Herd Behavior in Financial Markets". *IMF Working Paper WP/00/48*.
7. Bikhchandani, S., Hirshleifer, D., & Welch, I. (1998). "Learning from the Behavior of Others: Conformity, Fads and Informational Cascades". *Journal of Economic Perspectives*, 151-170.
8. Chang-Jin, K., & R.Nelson, C. (1999). *"State Space Models with Regime Switching"*. Cambridge, Massachusetts: The MIT Press.
9. Christie, W. G., & Huang, R. D. (1995). "Following the Pied Piper: Do Individual Returns Herd around the Market?" *Financial Analysts Journal*, 31-37.
10. Devenow, A., & Welch, I. (1996). "Rational herding in financial economics". *European Economic Review*, 603-615.
11. Fama, E. F., & French, K. R. (1993). "Common risk factors in the returns on stocks and bonds". *Journal of Financial Economics*, 3-56.
12. Fama, E. F., & French, K. R. (2004). "The Capital Asset Pricing Model: Theory and Evidence". *Journal of Economic Perspectives*, 25-46.
13. Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2000). "The Generalized Dynamic-Factor Model: Identification and Estimation". *The Review of Economics and Statistics*, 540-555.
14. He, Z., Huh, S.-W., & Lee, B.-S. (2010). "Dynamic Factors and Asset Pricing". *The Journal of Financial and Quantitative Analysis*, 707-737.
15. Hwang, S., & Salmon, M. (2004). "Market Stress and Herding". *Center for Economic Policy Research*, Discussion Paper No.4340.
16. Hwang, S., & Salmon, M. (2013). "Sentiment, Beta Herding, and Cross-sectional Asset Returns". *Korean Finance Association*.
17. Khorana, A., C., C. E., & W., C. J. (2004). "An Examination of Herd Behavior in Equity Markets: An International Perspective". *Journal of Banking & Finance*, 1651-1679.
18. Scharfstein, D. S., & Stein, J. C. (1990). "Herd Behavior and Investment". *American Economic Review*, 465-479.
19. Shiller, R. (1995). "Conversation, Information, and Herd Behavior". *Rhetoric and Economic Behavior*.
20. Shleifer, A., & Vishny, R. W. (1997). "The limits to Arbitrage". *The Journal of Finance*, 35-55.
21. Stock, J. H., & Watson, M. W. (2002). "Macroeconomic Forecasting using Diffusion Indexes". *Journal of Business and Economic Statistics*, 147-162.

22. Stock, J., & Watson, M. (November, 1988). "Probability model of coincident economic indicators". *National Bureau of Economic Research*, Working Paper No. 2772.
23. Tversky, A., & Daniel, K. (1974). "Judgement under Uncertainty: Heuristics and Biases". *Science*, 1124-113.