UNIVERSITÀ DI PISA

DIPARTIMENTO DI ECONOMIA E MANAGEMENT

Corso di Laurea Magistrale in

BANCA, FINANZA AZIENDALE E MERCATI FINANZIARI

Herd Behavior: An Estimate for the Italian Stock Exchange

Author:

Caterina Santi

Advisor:

Prof. Davide Fiaschi

— ACADEMIC YEAR 2013-2014 —

Herd Behavior:
An Estimate for the Italian Stock Exchange

Caterina Santi

Advisor: Prof. Davide Fiaschi

Academic Year: 2013-2014

Abstract

Herd behavior is widely believed to play a crucial role in financial markets and particularly when the market is in stress. This work analyses the phenomenon of herd behavior from both a theoretical and an empirical point of view. We apply the approach by Hwang and Salmon (2004), based on the cross-sectional standard deviations of the betas, to analyse herd behavior in the Italian Stock Exchange in the period January 1998 - December 2012. We find that herd behavior towards the market portfolio is significant and persistent, independently from and given the particular state of the market, and it shows a positive correlation with the FTSE MIB. Another remarkable result, given that herd behavior can lead to significant mispricing, is that herd behavior is never greater than the 40% of its maximum potential value during the sample period. Further, we examine herd behavior towards SMB and HML factors and find evidence of significant periods of herd behavior towards SMB and HML.

Acknowledgments

I would like to take this occasion to express my very great gratitude and appreciation to Prof. Davide Fiaschi for his valuable and constructive suggestions during my study and during the planning and development of this thesis. Its critical approach in analysing and interpreting economic phenomena has helped me to develop a critical thinking.

I wish also to thank Prof. Carlo Luigi Bianchi and Prof. Mauro Sodini for their precious advices and patience. They have been really important in my education.

A special thanks goes to my course mate, Annarita, Viola, Adele and Silvio, to my dearest friends Veronica, Marco, Giulia, Blerim and Giusy, and my cousin Chiara who supported me in these years.

Last but not least, I wish to thank my parents and Angelo for their support and encouragement throughout my study.

Preface

Herd behavior occurs when individuals imitate more or less blindly the decisions of others. There are many social and economic situations in which we are influenced in our decision making by what others do. One of the most famous anecdotal evidence of fads is the "beauty contest"[1]. Keynes described the action of rational market participants using an analogy based on a fictional newspaper contest, in which participants have to choose between six photographs of women the most beautiful face. Those who picked the most popular face are then eligible for a prize. Keynes wrote:

> "It is not a case of choosing those [faces] that, to the best of one's judgment, are really the prettiest, nor even those that average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practice the fourth, fifth and higher degrees." (Keynes. The General Theory of Employment Interest and Money.).

Herd behavior is widely believed to be a crucial element of behavior in financial markets. Indeed, as Shiller wrote in his book *Irrational Exuberance*, if the investors were all independent of each other prices would not have been affected by a faulty thinking. However, if a large number of people shares a faulty thinking, then this can lead to stock market booms and busts[2].

In the past decades there has been an increasing interest in herd behavior in financial markets. In fact, many researchers have attempted to find theoretical explanations and empirical evidence of herd behavior. In the thesis we provide an overview of the theoretical and empirical literature of herd behavior, then we use the measure proposed by Hwang and Salmon (2004) in order to detect herd behavior in the Italian stock market in the period January 1998 - December 2012. This approach to measuring herd

[1] Keynes, J. M. (1964). The general theory of employment, interest and money. New York: Harcourt, Brace and World.

[2] Shiller, R. J. (2005). Irrational exuberance (2nd ed.). Broadway Books, p. 157.

behavior is based on deviations from the CAPM prices and it captures market-wide herding. In other words, this measure aims to analyse the collective behavior of all market participants towards the market views. For the three factor model where the factors are the market excess return, the SMB and HML factors (SMB and HML stand for “small [cap] minus big” and “high [book/price] minus low”, these factors measure the historic excess returns of small caps and high book-value-to-price ratio stocks over the market as a whole) the herding measure is simply calculated from the relative dispersion of the betas on each factor for all the assets in the market. For instance, when herd behavior towards the market portfolio arises, the cross-sectional standard deviation of the estimated betas will diminish so that traders herd around the market consensus. Further, the measure captures adjustments in the cross-sectional standard deviation of the betas caused by herd behavior rather than adjustments due to fundamentals. In fact, it takes the underlying movement in the market as given.

We have found that herd behavior towards the market portfolio is significant and persistent independently from and given the particular state of the market and it shows a positive correlation with the FTSE MIB. We have also examined herd behavior towards SMB and HML factors and found significant evidence of herd behavior towards SMB, while herd behavior towards the HML factor is significant and positive correlated with the FTSE MIB only in the period from January 1998 to December 2005.

The thesis has been organised in the following way. Chapter 1 gives a brief overview of the evolution of decision theory. Chapter 2 and Chapter 3 review the theoretical research on rational and non rational herd behavior. Chapter 4 examines some empirical measures of herd behavior in financial markets. Then, Chapter 5 reports the results of our empirical analysis of herd behavior in the Italian Stock Exchange. Chapter 6 contains our conclusions. Lastly, in the appendix in Chapter 7 a brief treatment of the Kalman filter and smoothing is rendered.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Economics, like any other social science, has as ultimate aim to develop theories in order to help us better understand the world we live in[1]. In order to explain the complexity of economic phenomena economic theories proceed on the basis of a number of assumptions or premises.

One of the most important assumption of economic theory is the rationality of economic agents. In everyday speech, people are called reasonable if it is possible to reason with them, if their preferences are in line with their interests and their values. On the other hand, for economists and decision theorists, rationality relates to the internally consistency of a person's beliefs and preferences[2]. In order to guarantee the satisfaction of the hypothesis of rationality, economists assume that the preferences satisfies certain standard properties.

Consider X as the consumption set, x, y and z as basket of goods in X. We read $x \succeq y$ "x is weakly preferred to y". An ordering $\succ$ of strict preference can be defined simply by defining $x \succ y$ to mean not $x \succeq y$. We read $x \succ y$ as "x is strictly preferred to y". Similarly, indifference is defined by $x \sim y$ if and only if $x \succeq y$ and $y \succeq x$. The consumer is assumed to have preferences on the consumption bundles in X. Economists want the preferences to order the set of bundles. Here, the assumptions on these preferences:

COMPLETENESS. For all x and y in X , either $x \succeq y$ or $x \preceq y$ or both.

REFLEXIVITY. For all x in X , $x \succeq x$.

TRANSITIVITY. For all x, y and z in X , if $x \succeq y$ and $y \succeq z$, then $x \succeq z$.

[1] Wilkinson, N. and Klaes, M. (2012). An introduction to behavioral economics (2nd ed.). Palgrave Macmillan, p. 2.

[2] Kahneman, D. (2011). Thinking, fast and slow. Penguin Books, p. 411

MONOTONICITY. If $x \geq y$, then $x \succeq y$.

CONVEXITY. Given x, y and z in X. If $x \sim y$ and $z = \lambda x + (1-\lambda)y$ with $0 < \lambda < 1$, then $z \succ x$ and $z \succ y$.

CONTINUITY. For all y in X, the sets $\{x : x \succeq y)$ and $\{x : x \preceq y)$ are closed sets. It follows that $\{x : x \succ y)$ and $\{x : x \prec y)$ are open sets.

The first two assumptions state that the consumer is able to express his preferences about basket of goods, and the third assumption is necessary for preference maximization. If preferences were not transitive, there would be sets of bundles with no best elements.

Monotonicity asserts that if consumers could dispose of unwanted goods, without sustaining further costs, they would prefer basket of goods with a greater quantity of at least one good where the quantity of the other good remains unchanged.

The assumption of convexity implies that economic agents prefer averages to extremes quantity of a certain good. Finally, the assumption of continuity allows to eliminate certain discontinuous behavior.[3]

1.1 Lottery and the Expected Value Criterion

Previously, we have briefly treated the consumers' preferences under conditions of certainty. However, many choices made by economic agents take place under conditions of uncertainty.

When there are risky outcomes, agents could make their decisions according to the expected value criterion. In other words, they choose higher expected value investments. For instance, suppose there is a gamble in which the probability of getting a \$100 payment is 2% and the alternative is getting nothing. Then the gamble expected value is \$2. If we allow an agent to choose between this gamble and a certain payment of \$1.50, according to the expected value theory the agent will choose the \$100-or-nothing gamble.

Unfortunately, the expected value criterion can lead to some paradoxical results.

1.1.1 St. Petersburg paradox

Consider a game in which a fair coin is tossed at each stage. The pot is started at 1 dollar and is doubled every time a head appears. The game ends and the player wins

[3] Varian, H. R. (1992). Microeconomic analysis (3rd ed.). Norton & Company, pp. 94-96.

the amount of money in the pot when the first tail appears. Hence, the player wins \$1 if a tail appears on the first toss, \$2 if a tail appears on the second toss, and so on. In other words, the player wins 2^{k-1} dollars, where k indicates the number of heads that are tossed before the first tail appears. Now, we compute the expected value of the game.

$$E[X] = \frac{1}{2}\cdot 1 + \frac{1}{4}\cdot 2 + \frac{1}{8}\cdot 4 + \frac{1}{16}\cdot 8 + ... = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + ... = \infty$$

Assuming the game can continue till infinite. According to the expected value criterion, people should pay any price to enter the game. The paradox is the discrepancy between the price people seem willing to pay to participate to the game and the price suggested by the expected value criterion[4].

A solution to the St. Petersburg Paradox was proposed by Bernoulli. He argued that the utility function used in real life considers the expected utility of a gamble as finite, even if its expected value is infinite. Thus he hypothesized decreasing marginal utility of increasingly larger amounts of money[5].

1.2 Expected Utility Theory

In Economics, Decision Theory, and Game Theory the expected utility theory is related to people's preferences with respect to choices that have uncertain outcomes. The Von Neumann-Morgenstern utility theorem provides necessary and sufficient "rationality" axioms under which the expected utility hypothesis holds. This theory claims that if certain axioms are satisfied, the subjective value associated with a gamble by an individual is equal to the statistical expectation of that individual's valuations of the outcomes of that gamble.

Theorem. *Expected Utility Theorem*

A preference relationship ($\succeq$) satisfies the axioms of reflexivity, transitivity, continuity, independence[6] and the Archimedean property[7] if and only if $\exists\ U : X \to \mathbb{R}$ such that $p \succ q$ if and only if

[4]Bernoulli, D. Originally published in 1738; translated by Dr. Louise Sommer. (1954). Exposition of a new theory on the measurement of risk. *Econometrica*, 22, pp. 33-36.

[5]*Ibid.*

[6]**Independence Axiom.** If $p \succ q$, then $\forall r \in P$ and $\forall \alpha \in (0, 1]$
$\alpha \cdot p + (1-\alpha)\cdot r \succ \alpha \cdot q + (1-\alpha)\cdot r$

[7]**Archimedean Property**. If $p \succ q \succ r$, then there exists a probability $\alpha \in (0, 1)$
$(1-\alpha)\cdot p + \alpha \cdot r \succ q \succ \alpha \cdot p + (1-\alpha)\cdot r$

$$\sum_{x_i \in supp(p)} p(x_i) \cdot U(x_i) > \sum_{x_i \in supp(q)} q(x_i) \cdot U(x_i)$$

where:

- $p(x_i)$ is the probability of occurrence of prize x_i in lottery P;
- $supp(p)$ is the set of all possible prizes of lottery P;
- $q(x_i)$ is the probability of occurrence of prize x_i in lottery Q;
- $supp(p)$ is the set of all possible prizes of lottery P;
- $U(x_i)$ is the utility associated with the prize x_i.

This theory has helped to explain some popular choices that contradict the expected value criterion. The expected utility theory states that rational agents act as though they were maximizing expected utility. Further, it takes account for the possibility that people may be risk averse, that is individual would not accept a fair gamble (a fair gamble has an expected value of zero). Risk aversion implies that economic agents utility functions are concave and show decreasing marginal wealth utility. The risk attitude is directly related to the curvature of the utility function. In fact, the utility function is linear for risk neutral agents, convex for risk seeking individuals and concave for risk averse individuals[8].

1.3 Prospect Theory

Prospect theory was developed by Daniel Kahneman, a professor at Princeton University's Department of Psychology, and Amos Tversky in 1979.

Prospect theory describes how individuals choose between probabilistic risky alternatives, where the probabilities of outcomes are known. The theory describes the decision processes in two stages: editing and evaluation. During the first stage, the possible outcomes of a decision are ordered with respect to certain heuristic. In particular, agents set a reference point, decide which outcomes they consider equivalent, and then consider lesser outcomes as losses and greater ones as gains. In other words, the editing phase wants to alleviate any framing effects and to resolve isolation effects that are the propensity of individuals to isolate consecutive probabilities rather than treating

[8]Neumann, J. and Morgenstern, O. (1953). Theory of games and economic behavior. Princeton, NJ. Princeton University Press.

them together. In the evaluation phase, agents compute a value based on the potential outcomes and their associated probabilities, and then choose the alternative that gives a higher utility. Under Expected Utility Theory, people can not change risk preference unless there would be a violation of axioms and thus the predicted outcome would be wrong. In contrast, Prospect Theory does not assume agents to behave always in the same way, but to behave accordingly to their preferences when facing gains or losses.

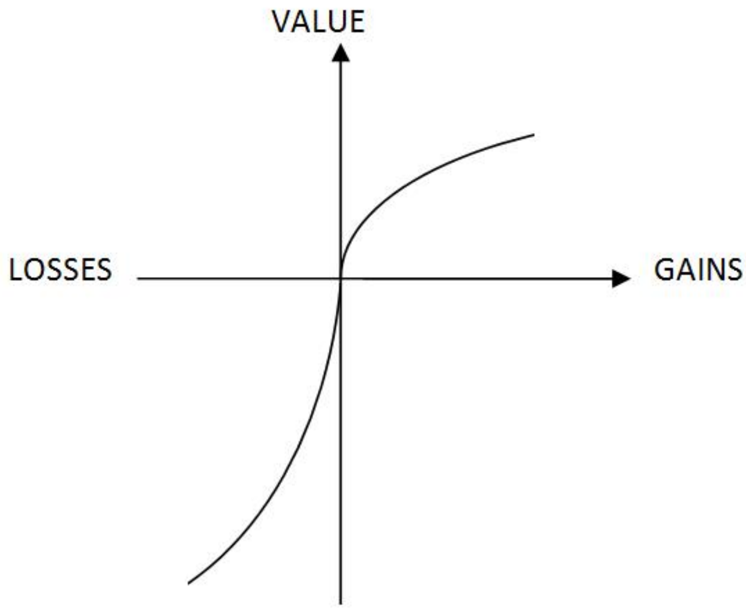

Figure 1.3.1: Value Function

Source: Kahneman, D., and Tversky, A. (1979). pp. 263-91.

The value function (Figure 1.3.1) is s-shaped, it is defined on deviations from the reference point, and it is generally concave for gains and convex for losses. This shape of the value function implies that individuals are loss averse. Differently, according to the expected utility theory, rational agents are indifferent to the reference point. In fact, individuals only care about absolute wealth, not relative wealth in any given situation[9].

Figure 1.3.2 presents an individual's "indifference map" for two goods.

All locations on an indifference curve are equally attractive, this is literally what indifference means. Hence, we can say $A \sim B$ (A is indifferent to B). However, the indifference curves do not indicate individual's current income and leisure (the reference point).

[9]Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), pp. 263-91.

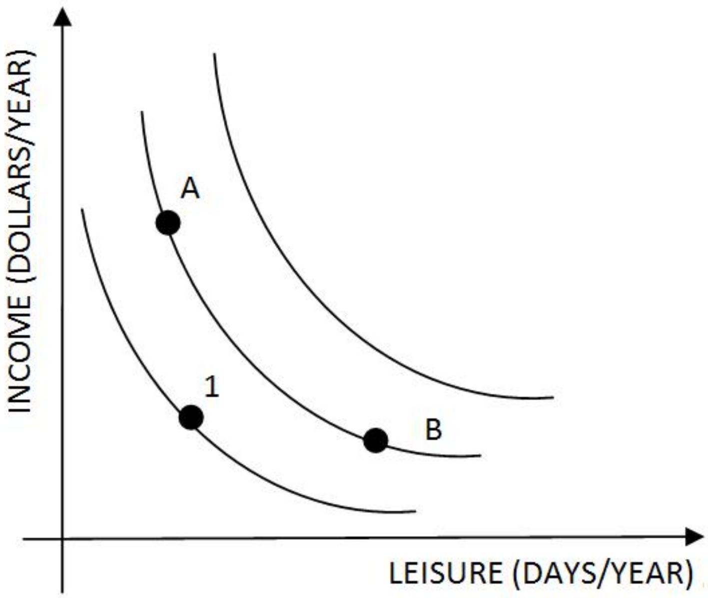

Figure 1.3.2: Indifference Map

Source: Kahneman. (2011). p. 289.

In order to better appreciate the power that the reference point exerts on choices, Kahneman in his book *Thinking Fast and Slow*, gives an interesting example.

> Consider Albert and Ben, "hedonic twins" who have identical tastes and currently hold identical starting jobs, with little income and little leisure time (point 1 in figure 1.3.2). The firm offers them two improved positions, A and B, and lets them decide who will get a raise of $ 10,000 (position A) and who will get an extra day of paid vacation each month (position B). As they are both indifferent, they toss a coin. Albert gets the raise, Ben gets the extra leisure. Some time passes as the twins get accustomed to their positions. Now the company suggests they may switch jobs if they wish.

The standard theory represented in the figure assumes that preferences do not change over time. Position A and B are equally attractive for both twins. Thus, Albert and Ben are indifferent to switch position or remain where they are. In sharp contrast, prospect theory states that both twins will definitely prefer do not change job. This preference for the status quo is due to loss aversion.[10]

[10]Kahneman. (2011). *op.cit.*, pp. 289-91.

1.4 Econs and Humans

In 2008 the economist Richard Thaler and the jurist Cass Sunstein published a book, "Nudge", which introduced several new words into the language, including the terms Econs that is referred to the *homo economicus*, described by economic theories, and Humans. Unlike the Econs, Humans have a limited view of the world because of the information available at a given moment, and therefore they cannot be as consistent as Econs. To qualify as Econs, people are required to make unbiased forecasts, which means that the forecasts does not have to be necessarily correct. However, it cannot be systematically wrong in a predictable direction. Instead, Humans predictably err. Consider for example, the "planning fallacy" that is the systematic tendency toward unrealistic optimism about the time it takes to complete projects.[11]

In addition, Econs communicate with others only if they can gain something, they care about their reputations, and they will learn from others if actual information can be obtained. Humans, on the other hand, are frequently nudged[12] by other Humans.

Social influences can occur because other persons' actions and thoughts reveal information about what might be best to do or think, or because people prefer conformity in order to avoid negative judgement or curry other people favour.[13]

1.5 Behavioral Finance

In finance, the efficient-market hypothesis (EMH) claims that financial markets are "informationally efficient". This means that one cannot consistently achieve returns in excess of average market returns, given the information available at the moment the investment is made. There exists three major versions of this hypothesis: "weak", "semi-strong", and "strong". The weak-form EMH asserts that prices on traded assets already reflect all past publicly available information. The semi-strong-form EMH states both that prices reflect all publicly available information and that they instantly change when new public information arrive. On the other hand, the strong-form EMH states that prices instantly reflect even hidden or "insider" information[14].

[11] Thaler, R. H. and Sunstein, C. R. (2008). NUDGE: Improving decisions about health, wealth, and happiness. Yale University Press, p. 7.

[12] A nudge, as Thaler and Sunstein use the term in their book "NUDGE: Improving decisions about health, wealth, and happiness", is any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives.

[13] *Ibid.*, pp. 53-54.

[14] Fama, E.F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2), p. 383.

The efficient-market hypothesis has been disputed by investors and researchers both empirically and theoretically. From the empirical side, one of the most discussed facts is that stock prices exhibit more volatility than fundamentals or expected returns do. Some non traditional model for return determination has been introduced in order to explain stock price volatility. In "fads" interpretations of the excess of volatility, noise trading by naive investors plays a significant role in stock price determination[15]. Shiller (1984) and DeBondt and Thaler (1985) argue that psychological and sociological evidence shows that investors follow "irrational" trading rules and overreact to the news. Potentially, this behavior generates wide variations in expected returns and traditional models for return determination becomes inadequate.

Another interpretation is that while some fraction of trading is done by the noise traders (naive traders that make their decisions following some rules of thumb), another fraction of trading is done by sophisticated investors who ensure that there are no extraordinary expected returns once risk is accounted for[16].

1.6 Herd Behavior

Observing human society it is easy to note that people who communicate regularly with one another think similarly[17].

Herd behavior occurs when people mimic what others do even when, their private information suggests doing something quite different[18]. This behavior can explain some economic phenomena like price bubbles. In fact, if the investors were all independent of each other, any faulty thinking would have no effect on prices. However, if a large numbers of people share a faulty thinking, then this can indeed be the source of stock market booms and busts.[19]

There are a lot of social and economic situations in which we are influenced in our decision making by what others do. Consider for example the choice between two restaurants (A and B) that are in front of each other and that are both more or less unknown to us. Assume that there is a population of 100 people who are facing such a choice and 99 of them have received signals that B is better and only one person has

[15] West, K. D. (1988). Bubbles, fads, and stock price volatility tests: A partial evaluation. *Journal of Finance*, 43, 639-60.

[16] *Ibid.*

[17] Shiller, R. J. (1995). Conversation, Information and Herd Behavior. *Rhetoric and Economic Behavior*, 85(2), p. 181.

[18] Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics,* 107(3), p. 798.

[19] Shiller, R. J. (2005). *op.cit.*, p. 157.

received a signal that favors A. Moreover, the prior probabilities are 51% for restaurant A being the better and 49% for restaurant B being the better.

People arrive at the restaurants in sequence and the first person who face the choice has signal A. Clearly, she will go to A. The second person will now know that the first person had signal A, while his signal is B. Since the signals are of equal quality, she will choose by prior probabilities and go to A. The third person is in the same situation as that of the second person, hence she will make the same choice and so on. At the end everyone go to restaurant A even if, considering the aggregate information, restaurant B is almost certainly better[20].

Herd behavior can occur for several reasons. Firstly, others can have more or more accurate information and their actions reveal their information. Secondly, and this is relevant only for money managers who invest on behalf of others, imitation could be rewarded by the compensation scheme or and terms of employment. Thirdly, individuals may prefer conformity. In addition, it is important to distinguish between intentional and spurious herd behavior. The former is the result of the intent by investors to imitate the actions of other's, and it may lead to inefficient market outcomes. Spurious herding, on the other hand, is a situation where groups face similar information sets and decision problems and then they make similar decisions. Thus, this is an efficient outcome[21].

In the following we will focus on herd behavior generated by imperfect information. Other causes of herd behavior include behavior that is not fully rational. This is the case of noise traders, who decide on the base of what they observe in the market. In fact, in the absence of any piece of such information they necessarily have to rely on what others do[22].

[20] Banerjee, A. V. (1992). *art.cit.*, pp. 798-99.

[21] Bikhchandani, S. and Sharma, S. (2000). Herd behavior in financial markets: A review. IMF Working Paper, pp. 3-4.

[22] Lux, T. (1995). Herd behavior, bubbles and crashes. *The Economic Journal*, 105, p. 882.

Part I

Literature Review on Herd Behavior Models

Chapter 2

Rational Herd Behavior Models

2.1 Introduction

This chapter provides an overview of the theoretical research on rational herd behavior that started with the seminal papers by Bikhchandani et al. (1992) (section 2.2), and Banerjee (1992) (section 2.3). These papers discussed herd behavior in an abstract framework, where agents with private information make their decisions in sequence. They pointed that, after a finite number of agents have chosen their actions, all following agents will ignore their personal information and herd. This is an important result as they provided a rational interpretation for the herd-like behavior we observe in consumers' and investors' decisions. Then, section 2.4 presents the paper by Zhiyong et al. (2010). This paper dealt with the role of ambiguity in herd behavior, and showed that herd does not occur when informed traders and market makers have the same ambiguity aversion. Finally, section 2.5 discusses the paper by Avery and Zemsky (1998) who explained herd behavior in a model where stock prices are endogenous. This model showed that when informed traders have private information on only a single dimension of uncertainty, herd behavior do not occur. Herd behavior arises when there are two dimensions of uncertainty, and it can lead to a significant short-run mispricing when a third dimension of uncertainty occurs[1].

2.2 Informational Cascade

Bikhchandani et al. (1992) analysed a sequential decision model where individuals will rapidly converge on one action on the basis of only little information.

[1] Avery, C. and Zemsky, P. (1998). Multidimensional uncertainty and herd behavior in financial markets. *The American Economic Review*, 88(4), p. 724.

An informational cascade occurs when, after having observed the actions of the preceding individual, it is optimal for an individual to imitate the actions of others without regard to his private information[2].

Assume that there is a sequence of individuals, each deciding whether to adopt or reject some behavior. The cost of adoption is the same for all individuals, $C = \frac{1}{2}$. The gain of adopting V, is also the same for all individuals and $V \in [0, 1]$, with equal prior probability $\frac{1}{2}$.

Each individual can observe the actions of others (the ordering of individuals is exogenous and it is common knowledge), and a conditionally independent signal (either Good, G or Bad, B) about value. If $V = 1$, signal G is observed with probability $p > \frac{1}{2}$, and B is observed with probability $1 - p$. Similarly, if $V = 0$, signal G is observed with probability $1 - p$, and B is observed with probability $p > \frac{1}{2}$[3]. Applying Bayes rule, the posterior probability of $V = 1$ after observing a signal G is[4]

$$Prob[V = 1|G] = \frac{Prob[G|V = 1] \cdot Prob[V = 1]}{Prob[G|V = 1] \cdot Prob[V = 1] + Prob[G|V = 0] \cdot Prob[V = 0]}$$

$$\frac{p \cdot 0.5}{p \cdot 0.5 + (1 - p) \cdot 0.5} = p > 0.5$$

Thus, the first individual adopts if his signal is G and rejects if it is B. The second individual can observe the first individual's action. If the first individual adopted, the second individual adopts if his signal is G. However, if his signal is B he adopts with probability $\frac{1}{2}$. The third individual is faced with three situations: (1) both predecessors have adopted, the third individual will adopt and an UP cascade will starts, (2) both predecessors have rejected, the third individual will reject and a DOWN cascade will starts, (3) one has adopted and the other rejected. The third individual is in the same situation of the first individual and he will take his decision on the base of his signal. Then, the fourth individual will be in the same situation as the second individual, the fifth as the third, and so forth.

Cascades tend to start sooner when individuals have more precise signals. More precisely, accurate signals increase the probability to end up in the correct cascade. Moreover, the probability of not being in a cascade falls exponentially with the number of individuals[5].

[2]Bikhchandani, S., Hirshleifer, D. and Welch, I. (1992). A theory of fads, fashion, custom, and cultural changes as informational cascades. *Journal of Political Economy*, 100(5), p. 994.

[3]*Ibid.,* p. 996.

[4]Bikhchandani, S. and Sharma, S. (2000). *art.cit.*, p. 6.

[5]Bikhchandani, S., Hirshleifer, D. and Welch, I. (1992). *art.cit.*, p. 997.

In this model herd behavior always implies an informational cascade.

Definition 1. An informational cascade occurs in period t when

$$P[h_t|V, H_t] = P[h_t|H_t] \, \forall V, h_t$$

where h_t is the action taken by the trader who arrives at time t and H_t denotes the publicly observable history of trades up until time t.

In an informational cascade, no new information reaches the market because the distribution over the observable actions is independent of the state of the world. In particular, this happens when the actions of all informed traders are independent of their private information[6].

The problem with cascade is that they prevent the aggregation of information. When an individual takes an action that is uninformative to others, it creates a negative externality. If a number of early individuals would make the altruistic choice of following their signal, this behavior would ultimately lead to almost perfectly accurate decisions. Instead, individuals act in their own self interest, imitate what others do creating negative externality that leads to an inefficient outcome[7].

2.3 A Simple model

Banerjee (1992) discussed a sequential decision model in which each decision maker observes the action adopted by the previous decision makers and then decide to follow his private information or ignoring it and mimic what others do. This behavior is rational as previous decision makers could have some important information. Banerjee showed that the decision rules that are chosen by optimizing individuals will be characterized by herd behavior.

Consider a population of N agents each of whom maximizes the identical risk-neutral utility function Von Neumann Morgenstern defined on the space of asset returns.

The decision is between a set of assets indexed by numbers [0, 1]. Call the ith asset $a(i)$. Let us assume that there is a unique asset i^* that has positive return, while the others have return equal to zero. Of course, everybody, given these payoffs, would want to invest in i^*[8].

There is a probability α that each person receives a signal telling her that the true i^* is i'. The signal could be false with probability $1-\beta$. If it is false, then it is assumed

[6] Avery, C. and Zemsky, P. (1998). *art.cit.*, p. 728.
[7] Bikhchandani, S. and Sharma, S. (2000). *art.cit.*, pp. 8-9.
[8] Banerjee, A. V. (1992). *art.cit.*, pp. 802-03.

that it is uniformly distributed on [0,1] and therefore gives no information about what i^* really is. The individuals take their decisions sequentially; each person can observe the choice made by the previous person.

After everybody has made her choice, all the alternatives that have been chosen are tested and those who have chosen the alternative that is true receive their rewards.

Here, some assumptions are made to minimize the possibility of herd behavior.

ASSUMPTION A. If a decision maker does not have a signal and everyone else has chosen $i = 0$, she will choose $i = 0$.

ASSUMPTION B. Decision makers will decide to follow their own signal, whenever they are indifferent between following their own signal and following someone else's choice.

ASSUMPTION C. A decision maker will choose to follow the decision maker who has the highest value of i, when she is indifferent between following more than one of the previous decision makers[9].

If the first decision maker's has a signal, she will certainly follow her signal. Otherwise, if she has no signal, by ASSUMPTION A she will choose $i = 0$.

If the second investor has no signal, she will imitate the first decision maker and invest in the same asset. However, if she has a signal and the first person has not chosen $i = 0$, by ASSUMPTION B she will follow her own signal.

If the third decision maker has no signal by ASSUMPTION C she chooses to follow the previous decision maker who has the highest value of i. However, if she has a signal i' she will follow it, unless both people before her have chosen the same option and this option is neither $i = 0$ nor $i = i'$.

Whenever some person's signal matches the choice made by one of her predecessors, she should always follow her signal. This follows from the fact that the probability that two people should get the same signal and yet both be wrong is approximately equal to zero[10].

Under assumptions A, B and C, the unique Nash equilibrium decision rule that everyone will adopt is the following:

1. If the first decision maker has a signal, she will follow her own signal. Otherwise she will choose $i = 0$.

[9] *Ibid.*

[10] *Ibid.*, pp. 804-05.

2. For $k > 1$, if the kth decision maker has a signal, she will choose to follow her own signal either if and only if her signal matches some option that has already been chosen or if no option other than $i = 0$ has been chosen by more than one person.

3. Assume that the kth decision maker has a signal. If any option other than the one with the highest i has been chosen by more than one person, the decision maker will choose this option, unless her signal matches one of the other options that has already been chosen. In this case she chooses the latter option.

4. Assume that the kth decision maker has a signal. If the option with the highest i has been chosen by more than one person and no other option (except $i = 0$) has been chosen by more than one person, she will choose this option unless her signal matches one of the options already chosen. In this case she chooses the latter option.

5. Assume that the kth decision maker has no signal. Then she will choose $i = 0$ if and only if everyone else has chosen $i = 0$. Otherwise, she will choose the option with the highest value of i that has already been chosen unless one of the other options (excluding $i = 0$) has been chosen by more than one person. In this case, she chooses the latter option[11].

The equilibrium decision rule in the above model is characterized by extensive herd behavior; agents abandon their own signals and follow others even when they are not completely sure that the other person is right. In addition, it can happen that no one in the population chooses the right option, while if all the decision makers took their decisions without observing the choices made by others, some people will always end up choosing the correct option[12].

2.4 Ambiguity Aversion

Zhiyong et al. (2010) investigated the relationship between ambiguity aversion and herd behavior.

The concept of ambiguity is quite different from that of risk. The distinction between them is that risk refers to situations where the perceived likelihoods of events can be represented by a unique probability distribution, whereas if there is ambiguity, not only

[11] *Ibid.*, pp. 806-07.
[12] *Ibid.*, p. 808.

is the outcome of an act uncertain but also the expected payoff of the action, since the probability distribution of the possible events is unknown.

People differ in their attitudes towards ambiguity. The majority of decision makers are ambiguity averse, evaluating an act by the minimum expected value that may be associated with it and thereby being more cautious when the probability distribution is undefined, differently a significant minority of decision makers appear to be ambiguity-loving that is exactly the opposite[13].

In order to understand the effect of ambiguity aversion on herd behavior, the Bikhchandani, Hirshleifer, and Welch (BHW) assumptions are extended in the following ways:

1. V follows the binary distribution of $\{b_j, g_j\}$, where $g_j > b_j \geq 0$, $j = 1, 2, ..., K$ and $C = \frac{(g_j + b_j)}{2}$;
2. the binary distributions are different if the distribution types are different, then we have $g_{j_1} \neq g_{j_2}$ and $b_{j_1} \neq b_{j_2}$ when $j_1 \neq j_2$
3. the probability for the j-th distribution type to occur is q_j; the probability for V to be b_j or g_j is 0.5 at time 0. The probability for V to be b_j or g_j can be any number between 0 and 1 after time 0. As long as this probability deviates from 0.5, different distributions will have different means due to the fact that $g_{j_1} \neq g_{j_2}$ and $b_{j_1} \neq b_{j_2}$, $\forall\ j_1 \neq j_2$. Thus the model will not lose generality by assuming C is constant across j[14].

It is also assumed that asset prices are determined by competitive market makers and an infinite number of traders who trade sequentially. As before, $t = 0, 1, 2, ...$ represents the t-th trading period, and H_t represents the historical public information prior to period t. There are two groups of traders in the market, the informed traders and the noise traders. The behavior of noise traders is considered exogenous. Further, informed traders can be divided into N categories according to their attitudes towards ambiguity.

As in BHW model, every informed trader receives a signal (note that the signal cannot be observed by market makers or other traders) that could be either G (Good news) or B (Bad news). When $V = g_j$, G is observed with probability $p > \frac{1}{2}$, and B is observed with probability $1 - p$. Similarly, if $V = b_j$, G is observed with probability $1 - p$, and B is observed with probability $p > \frac{1}{2}$. Each informed trader can observe his signal and previous traders actions.

[13]Zhiyong Dong, Qingyang Gu, and Xu Han. (2010). Ambiguity aversion and rational herd behaviour. *Applied Financial Economics*, 20, p. 332.
[14]*Ibid.*, p. 334.

Definition 2. The i-th trader with private information x_i engages in herd behavior at time t if he buys when $V_{i,0}(x_i) < V_{m,0} < V_{m,t}$ or if he sells when $V_{i,0}(x_i) > V_{m,0} > V_{m,t}$; and buying (or selling) is strictly preferred to other actions.

Where $V_{i,0}(x_i)$ is the estimate of the asset return at time t of an informed trader i with private information x_i, $V_{m,0}$ is the expectation of asset value by market maker at time 0, and $V_{m,t}$ is the expected value of the asset return by market maker in period t.

Herd behavior by a trader satisfies three properties. First, the signal received by the informed trader $V_{i,0}(x_i) < V_{m,0}$ incentive to sell. Second, the history of trading must be positive, $V_{m,0} < V_{m,t}$. Finally, given the historical and the private information, informed traders are willing to buy the asset, which is $V_{m,t} < V_{i,t}(x_i)$.

The signal contains bad news that incentive to sell. However, after having observed the trading history, the signal constitutes positive information, causing their evaluation to exceed asset price $V_{m,t}$ in period t, such that they follow the historical trend ignoring their own information[15].

In this model a function Φ that characterizes traders' attitudes towards ambiguity is introduced. Φ is assumed to be increasing ($\Phi' > 0$) and concave ($\Phi'' < 0$) in order to focus on the case of ambiguity aversion.

Proposition 3. *Given $\Phi' > 0$, if all informed traders and market makers share the same value function and have the same level of ambiguity aversion, herd behavior will never occur*[16].

Proof. Denote $\pi_{v_j,t}$ as the probability of $V = v_j$, given the historical transaction information and distribution type j, ($\pi_{v_j,t} = P(V = v_j|H_t)$).

Let g and b denote different values of v_j, which implies $\pi_{b_j,t} = 1 - \pi_{g_j,t}$. Recalling the assumption that the probability for V to be b_j or g_j is 0.5 at time 0, this implies that $\pi_{g_j,t} = \pi_{g,t}$, and $\pi_{b_j,t} = \pi_{b,t}$, regardless of the distribution type j, with $0 < \pi_{g,t} < 1$ and $0 < \pi_{b,t} = 1 - \pi_{g,t} < 1$. Moreover, since market makers and informed traders have the same degree of ambiguity aversion, it is applicable to use Φ to denote their value functions. When traders receive bad news in period t, they estimate the asset value to be

$$V_{i,t}(B) = \Phi^{-1}\left\{\sum_j q_j \Phi[E_j(V \mid H_t,\, x = B)]\right\}$$

[15]Avery, C. and Zemsky, P. (1998). *art.cit.*, p. 728.
[16]Zhiyong Dong, Qingyang Gu, and Xu Han. (2010). *art.cit.*, p. 335.

$$= \Phi^{-1}\left\{\sum_j q_j \Phi\left[g_j \frac{(1-p)\pi_{g,t}}{(1-p)\pi_{g,t}+p(1-\pi_{g,t})} + b_j \frac{p(1-\pi_{g,t})}{(1-p)\pi_{g,t}+p(1-\pi_{g,t})}\right]\right\} \quad (2.4.1)$$

When traders receive bad news in period t, they estimate the asset value to be

$$V_{i,t}(G) = \Phi^{-1}\left\{\sum_j q_j \Phi[E_j(V \mid H_t,\, x = G)]\right\}$$

$$= \Phi^{-1}\left\{\sum_j q_j \Phi\left[g_j \frac{p\pi_{g,t}}{p\pi_{g,t}+(1-p)(1-\pi_{g,t})} + b_j \frac{(1-p)(1-\pi_{g,t})}{p\pi_{g,t}+(1-p)(1-\pi_{g,t})}\right]\right\} \quad (2.4.2)$$

The expectation of asset value by the market maker in period t is

$$V_{m,t} = \Phi^{-1}\left\{\sum_j q_j \Phi[g_j\pi_{g,t} + b_j(1-\pi_{g,t})]\right\} \quad (2.4.3)$$

To prove the absence of herd behavior it is only necessary to show that $V_{i,t}(B) < V_{m,t} < V_{i,t}(G)$, which guarantees that traders will follow their private information instead of the historical trend. As Φ and Φ^{-1} are increasing functions, it is sufficient to compare $E_j(V \mid H_t)$ and $E_j(V \mid H_t,\, x = B)$ in order to verify $V_{m,t} - V_{i,t}(B) > 0$

$$E_j(V \mid H_t) - E_j(V \mid H_t,\, x = B) = [g_j\pi_{g,t} + b_j(1-\pi_{g,t})] - \left[\frac{g_j(1-p)\pi_{g,t} + b_j p(1-\pi_{g,t})}{(1-p)\pi_{g,t}+p(1-\pi_{g,t})}\right]$$

$$= \frac{\pi_{g,t}(g_j - b_j)(2p-1)(1-\pi_{g,t})}{(1-p)\pi_{g,t}+p(1-\pi_{g,t})} \quad (2.4.4)$$

Given $g_j > b_j$, $0.5 < p < 1$, $0 < \pi_{g,t} < 1$, it is clear that $V_{m,t} - V_{i,t}(B) > 0$. Thus $V_{m,t} > V_{i,t}(B)$.

Similarly it can be proved that $V_{i,t}(G) > V_{m,t}$

$$E_j(V \mid H_t,\, x = G) - E_j(V \mid H_t) = \left[\frac{g_j p\pi_{g,t} + b_j(1-p)(1-\pi_{g,t})}{p\pi_{g,t}+(1-p)(1-\pi_{g,t})}\right] - [g_j\pi_{g,t} + b_j(1-\pi_{g,t})]$$

$$= \frac{\pi_{g,t}(g_j - b_j)(2p-1)(1-\pi_{g,t})}{(1-p)\pi_{g,t}+p(1-\pi_{g,t})} > 0$$

[17] □

[17] *Ibid.*, p. 336.

The situation is different if the degree of ambiguity aversion vary amongst traders.

The coefficient for absolute ambiguity aversion $-\Phi''/\Phi'$ is used to characterize traders and market maker's attitudes towards ambiguity.

Lemma 4. *Functions Φ_1 and Φ_2 are defined within real domain and have continuous first-order and second order derivatives. If $\Phi_1' > 0$, $\Phi_2' > 0$, $\Phi_1'' < 0$, $\Phi_2'' < 0$ and $-\Phi_1''/\Phi_1' < -\Phi_2''/\Phi_2'$, then for any different real numbers x_1, x_2, ω_1 and ω_2 where $x_1 < x_2$, $\omega_1 > 0$, $\omega_2 > 0$ and $\omega_1 + \omega_2 = 1$, the following inequality holds:*

$$\Phi_1^{-1}(\omega_1\Phi_1(x_1) + \omega_2\Phi_1(x_2)) > \Phi_2^{-1}(\omega_1\Phi_2(x_1) + \omega_2\Phi_2(x_2)) \tag{2.4.5}$$

For $x_1 < x_2 < ... < x_n$, $\omega_1 > 0$, $\omega_2 > 0$, ..., $\omega_n > 0$, where $\sum_{i=1}^{n}\omega_i = 1$, the following inequality holds[18]*:*

$$\Phi_1^{-1}\left(\sum_{i=1}^{n}\omega_i\Phi_1(x_i)\right) > \Phi_2^{-1}\left(\sum_{i=1}^{n}\omega_i\Phi_2(x_i)\right) \tag{2.4.6}$$

Herd behavior can occur when informed traders and market makers have different degrees of ambiguity aversion.

Proposition 5. *Assume that there is one type of informed traders whose value functions are denoted as Φ_1. The market maker's value function is denoted as Φ_2. Both Φ_1 and Φ_2 have continuous first-order and second-order derivatives. If $\Phi_1' > 0$, $\Phi_2' > 0$, $\Phi_1'' < 0$, $\Phi_2'' < 0$ and $-\Phi_1''/\Phi_1' < -\Phi_2''/\Phi_2'$, which means the absolute ambiguity aversion coefficient of informed traders is always less than that of the market maker, buying herd may occur in the market*[19].

Proof. Denote $V_{1,t}$ as the informed traders' certainty equivalent asset value in period t based on the private signal x and $V_{2,t}$ as that of the market maker's based on the public information prior to period t. As long as the probability for $V_{2,t} < V_{1,t}(B)$ is positive under the condition $V_{1,0}(B) < V_{2,0} < V_{2,t}$, buying herd can occur. First, it can be shown that $V_{1,0}(B) < V_{2,0} < V_{2,t}$ is possible. The proof of Proposition 3 shows that

$$\begin{aligned} V_{2,0} &= \Phi_2^{-1}\left\{\sum_j q_j\Phi_2[0.5g_j + 0.5b_j]\right\} \\ &= \Phi_2^{-1}\left\{\sum_j q_j\Phi_2[C]\right\} = C \end{aligned} \tag{2.4.7}$$

[18] *Ibid.*

[19] *Ibid.*, pp.336-37

From Equation 2.4.4 we have

$$V_{1,0}(B) = \Phi_1^{-1}\left\{\sum_j q_j\Phi_1[E_j(V|H_0, x = B)]\right\}$$

$$= \Phi_1^{-1}\left\{\sum_j q_j\Phi_1\left[g_j\frac{0.5(1-p)}{0.5(1-p)+0.5p} + b_j\frac{0.5p}{0.5(1-p)+0.5p}\right]\right\} < C \qquad (2.4.8)$$

Therefore, it is true that $V_{1,0}(B) < V_{2,0}$. Also, as long as there are more buy-in traders than sell-out traders in the market prior to period t, the price in period t would be expected to increase, which is $V_{2,t} > V_{2,0}$. Given $\pi_{g,0} = 0.5$, this implies $\pi_{g,t} > 0.5$

The second step is to prove the existence of $V_{2,t} < V_{1,t}(B)$. Recall the assumption $g_{j_1} \neq g_{j_2}$, $b_{j_1} \neq b_{j_2}$when $j_1 \neq j_2$. Under the condition $\pi_{g,t} > 0.5$, $\xi = g_j\pi_{g,t} + b_j(1-\pi_{g,t})$ are positive numbers that differ with j. Arrange ξ in an increasing sequence, $C < \xi_1 < \xi_2 < ... < \xi_k$. From Equation 2.4.1, it is clear that $E_j(V|H_t, x = B) > C$, $\forall j$, as long as $\pi_{g,t} > p$. Define $\psi = E_j(V|H_t, x = B)$ and arrange ψ in an increasing sequence $C < \psi_1 < \psi_2 < ... < \psi_k$. Proposition 3 leads to the conclusion that $\xi_j > \psi_j$, $\forall j$. Applying Lemma 4 we obtain

$$\Phi_1^{-1}\left(\sum_j q_j\Phi_1(\xi_j)\right) > \Phi_2^{-1}\left(\sum_j q_j\Phi_2(\xi_j)\right) \qquad (2.4.9)$$

To prove the possibility of herd behavior, the only thing remaining is to show the existence of positive ψ_j $(\xi_j > \psi_j, \forall j)$, such that

$$\Phi_1^{-1}\left(\sum_j q_j\Phi_1(\psi_j)\right) > \Phi_2^{-1}\left(\sum_j q_j\Phi_2(\xi_j)\right) \qquad (2.4.10)$$

Since both Φ_1 and its inverse function are strictly increasing and continuous functions, there always exists $\delta > 0$ such that $\Phi_1^{-1}\left(\sum_j q_j\Phi_1(\xi_j - \delta)\right) > \Phi_2^{-1}\left(\sum_j q_j\Phi_2(\xi_j)\right)$. As long as there are $\psi_j$such that $max(\xi_j - \psi_j) < \delta$, inequality (2.4.10) will hold. By Equation 2.4.4 in Proposition 3,

$$\xi_j - \psi_j = E_j(V \mid H_t) - E_j(V \mid H_t,\, x = B)$$

$$= \frac{\pi_{g,t}(g_j - b_j)(2p-1)(1-\pi_{g,t})}{(1-p)\pi_{g,t} + p(1-\pi_{g,t})} \qquad (2.4.11)$$

The last step is to find some parameters such that $max\left[\frac{\pi_{g,t}(g_j-b_j)(2p-1)(1-\pi_{g,t})}{(1-p)\pi_{g,t}+p(1-\pi_{g,t})}\right] < \delta$. Since g_j, b_j are exogenous parameters and can be any positive numbers, so appropriate g_j and b_j can always be found[20]. □

It is sufficient only a small number of informed traders having different coefficients for absolute ambiguity aversion from that of market makers to generate herd behavior[21].

2.5 Multidimensional Uncertainty

Avery and Zemsky (1998) proposed a model in which herd behavior arises only in the presence of two dimensions of uncertainty, and it can lead to a significant short-run mispricing when a third dimension of uncertainty occurs.[22]

In the preceding discussion the cost of taking an action is fixed ex-ante and it remains so. This assumption is relaxed in Avery and Zemsky (1998).

Suppose that after a trader has made a decision to buy or sell, the asset price adjusts to take into consideration the information revealed by this action. In a setting with competitive market makers, the stock price will always be the expected value of the investment conditional on all publicly available information. Therefore, investors without private information will be indifferent between buying or selling. Further, the action of any informed trader will reveal their own information. As a consequence herd selling or buying never occurs.

As BHW, this model considers an asset which value V can either be 0 or 1. Informed traders observe a signal which can be G or B. When $V = 1$, G is observed with probability $p > \frac{1}{2}$, and B is observed with probability $1 - p$. Similarly, if $V = 0$. Prices are set by a competitive market maker. If the first trader buys, the asset price will increase to $E(V \mid x = G) = 2p - 1$. Then, the second investor will deduce the previous trader signal from his action. If the second trader has private signal B then his posterior expected value is 0 which is less than $2p - 1$. If, instead he observes G, then his posterior expected value of V is $\frac{2p-1}{p^2+(1-p)^2}$ which is greater than $2p - 1$. Hence, the second investor will follow his private signal. Consequently, herd behavior will not occur when the asset price adjusts to reflect available information[23].

Now, it is considered the case that investors have private information about two dimensions of uncertainty. In addition to information related to *value uncertainty*, there

[20] *Ibid.*
[21] *Ibid.*
[22] Avery, C. and Zemsky, P. (1998). *art.cit.*, p. 724.
[23] Bikhchandani, S. and Sharma, S. (2000). *art.cit.*, p. 9.

is a second dimension of uncertainty called *event uncertainty*. That is the trader has private information that there has been a shock to the underlying value of the asset, while the market maker does not. The information asymmetry gives the investors an advantage in interpreting the history of trades. When *event uncertainty* appears together with *value uncertainty*, herd behavior arises and resembles an information cascade in fact the market does not learn the efficient price. However, the effect of this herd behavior is bounded and the impact on pricing may be small if the bound is small[24].

Then, Avery and Zemsky consider a third dimension of uncertainty that is *composition uncertainty*. There is *composition uncertainty* when in the market there are investors of different types and their probability is not common knowledge.[25] Suppose there are investors of type H and L. Type H has very accurate information and type L has very noisy information. Further, it is supposed that the two types of traders in the population is not common knowledge. A sequence of identical decisions may arise naturally in a well-informed market. In addition, a sequence of identical decisions is also natural in a poorly informed trader because of herd behavior by type L investors who mistakenly believe that most of the other investors are of type H. Thus, informationally inefficient herd behavior may occur and can lead to price bubbles and mispricing when the accuracy of the information with market participants is not common knowledge. Traders may imitate the behavior of an initial group of investors in the erroneous belief that this group knows something[26].

[24]Avery, C. and Zemsky, P. (1998). *art.cit.*, p. 731-32.
[25]*Ibid.*, 735.
[26]Bikhchandani, S. and Sharma, S. (2000). *art.cit.*, p. 9-10.

Chapter 3

Non Rational Herd Behavior Models

3.1 Introduction

This Chapter is mainly concerned with the determination of the behavior of noise traders who do not have access to information about fundamental values. Following others' opinion in this case is not irrational as in the absence of any piece of information they necessarily have to rely on what others do[1].

In the following section an application of epidemic models to interpersonal communication is discussed. These models can help us to better understand the transmission of attitudes between traders. Then, in section 3.3 we analyse the paper by Lux (1995), which dealt with a formalization of contagion.

3.2 Epidemic models applied to interpersonal communication

Interpersonal and interactive communications have a powerful impact on our behavior. The mathematical theory of the spread of disease has been used by epidemiologists to predict the course of infection and mortality.

In the simplest epidemic model, it is assumed that the disease has a given infection rate (i), which is the rate at which the disease spreads from contagious people to susceptible people, and a given removal rate (r), which is the rate at which an infected people cease to be contagious.

[1] Lux, T. (1995). *art.cit.*, p. 882.

If $r = 0$, the number of infected people after the introduction of one contagious person increase till the entire population is infected.

If $0 < r < i$, the course of epidemics will be bell shaped: the number of infective will at first rise from zero, peak and then drop back to zero.

If $r > i$, the epidemic will never get started.

When epidemic models are applied to predict the course of interpersonal communication; the infection rate is the rate of communication of ideas and the removal rate is the rate of forgetting or of losing interest[2].

Economist Alan Kirman (1993) used epidemic models to study the behavior of ants in exploiting food sources, and he noted that the model is also relevant to stock market changes. It has been found experimentally that ants, when presented with two identical food sources near their nest, tend to exploit them in an asymmetric way. Kirman observed that ants individually recruit other ants by contact and following or by laying a chemical trail, which is the ant equivalent of word-of-mouth communication. Kirman showed that if there is randomness in the recruitment process, the experimentally observed phenomena can be explained in terms of a simple epidemic model. Although epidemic models and ant behavior are of theoretical interest, they appear to be less accurate for modeling social processes. One reason is that interpersonal communication is more imprecise and variable than the spread of disease or other biological processes. However, epidemic models are still helpful in understanding the kinds of things that can bring about changes in market prices[3].

3.3 An elementary formalization of contagion

The behavior of noise traders was formalized by Lux (1995) by referring to the concept of synergetics. Synergetics' basically consists of a probabilistic, macroscopic approach to the analysis of the dynamics of multi-component systems with interactions among the units constituting the system[4].

Lux assumed that there exists a fixed number of speculative traders $2N$. These can either be optimistic or pessimistic about the future development of the market.

Moreover, n_+indicates the number of optimistic traders and n_- the number of pessimistic traders. Obviously, $n_+ + n_- = 2N$. The "neutral" subject is not considered. Defining $n \equiv \frac{n_+ - n_-}{2}$ and $x \equiv \frac{n}{N}$. Where x is an index describing the average opinion of

[2]Shiller, R. J. (2005). *op.cit*, pp. 164-65.
[3]*Ibid.*, pp. 165-66.
[4]Lux, T. (1995). *art.cit.*, p. 883.

speculative investors, $x \in [-1, +1]$ and $x = 0$ means that there exists an equal number of optimistic and pessimistic individuals. On the contrary, $x > (<) \, 0$ exhibit more or less predominant optimism (pessimism).

In case of a high portion of optimistic traders, it would be very probable that the few remaining pessimistic ones would also change their attitude. Hence there exists a probability p_{+-} for a pessimistic trader to become optimistic and *vice versa* $p_{-+}(p_{+-} > 0, p_{-+} > 0)$. With contagion both probabilities should depend on the actual distribution of attitudes[5].

$$p_{+-} = p_{+-}(x) = p_{+-}\left(\frac{n}{N}\right), \quad p_{-+} = p_{-+}(x) = p_{-+}\left(\frac{n}{N}\right). \tag{3.3.1}$$

The existence of financial gurus is excluded by this formulation, as all individuals can influence one particular trader in the same way. Another simplifying assumption is that each speculator can change his opinion only once at any one time.

From this it is possible to determine the change in time of the number of optimistic traders

$$\frac{dn_+}{dt} = n_- p_{+-} - n_+ p_{-+}$$

and pessimistic speculators

$$\frac{dn_-}{dt} = n_+ p_{-+} - n_- p_{+-}$$

Then,

$$\frac{dx}{dt} = \frac{1}{2N}\left(\frac{dn_+}{dt} - \frac{dn_-}{dt}\right) = \frac{1}{N}\left(n_- p_{+-}(x) - n_+ p_{-+}(x)\right)$$

Noting that $N - n = n_-$ and $N + n = n_+$, we can write

$$\frac{dx}{dt} = \frac{1}{N}\left[(N-n)p_{+-}(x) - (N+n)p_{-+}(x)\right]$$

$$= \left[(1-x)p_{+-}(x) - (1+x)p_{-+}(x)\right] \tag{3.3.2}$$

In order to grasp the very idea of contagion $p_{+-} > p_{-+}$ when the prevailing disposition of the population is already optimistic ($x > 0$) and *vice versa*[6]. Moreover, Lux assumed that the relative change in the probability to switch from pessimism to

[5] *Ibid.*, p. 884.
[6] *Ibid.*

optimism increases linearly with changes in x, $\frac{dp_{+-}}{p_{+-}} = adx$. Symmetric behavior in both directions leads to $\frac{dp_{-+}}{p_{-+}} = -adx$.

Then, we have

$$p_{+-}(x) = ve^{ax}, \quad p_{-+}(x) = ve^{-ax} \tag{3.3.3}$$

Here, a gives a measure of herd behavior and v is a variable for the speed of change. When $x = 0 \Rightarrow p_{+-} = p_{-+} = v > 0$. This means changes of attitudes occur due to personal circumstances not comprised in the model.

With this specification of transition rates the time development of the mean value of the index x becomes:

$$\frac{dx}{dt} = \left[(1-x)ve^{ax} - (1+x)ve^{-ax}\right]$$

from the definition of hyperbolic sine and cosine, it is possible to derive the following identities $e^{x} = Cosh(x) + Sinh(x)$, $e^{-x} = Cosh(x) - Sinh(x)$.

Thus, we can write[7]

$$\frac{dx}{dt} = 2v[Sinh(ax) - xCosh(ax)] = 2v[Tanh(ax) - x]Cosh(ax). \tag{3.3.4}$$

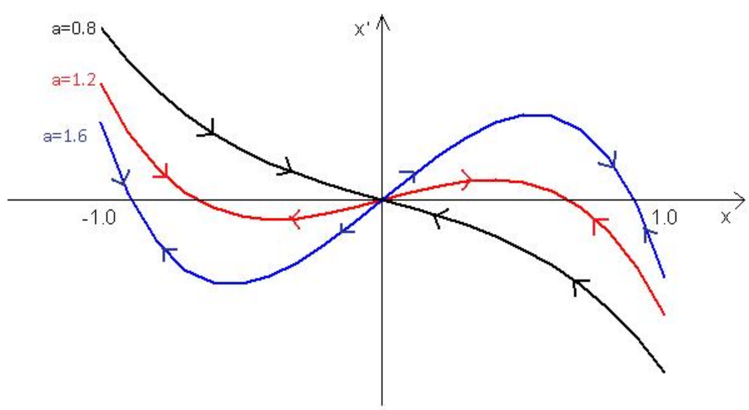

Figure 3.3.1: Pure contagion dynamics

Source: Lux, T. (1995). p. 886.

Proposition 6. *(i) For $a \leq 1$, (3.3.4) possesses a unique stable point at $x = 0$. (ii) For $a > 1$, the equilibrium $x = 0$ is not stable and there exists two additional stable equilibrium, say $x_+ > 0$, $x_- < 0$ exist ($x_+ = -x_-$).*[8]

[7] *Ibid.*, p. 885.
[8] *Ibid.*

The results are proved by considering the equilibrium condition $Tanh(ax) = x$. Proposition 6 says that if the herd effect is relatively weak ($a \leq 1$), then all defection into one direction will die out in the course of events and the system will return to a state where there is an equal number of optimistic and pessimistic individuals ($x = 0$). For $a > 1$, on the other hand, small deviations from the balanced state are sufficient to make a majority of traders bullish or bearish through mutual infection. Hence, the equilibrium $x = 0$ is unstable[9].

[9] *Ibid.*

Part II

Empirical Analysis of Herd Behavior in Financial Markets

Chapter 4

Herding Measures

4.1 Introduction

Several measures already exist to investigate herd behavior in financial markets although most empirical study do not test a particular theoretical model directly. Excluding the paper by Cipriani and Guarino (2014) that introduces a theoretical model of herd behavior and then tests it empirically, in general there still exists a gap between the theoretical and empirical literature.

Empirical investigations of herd behavior in financial markets have branched into two paths. The first track focuses on co-movement behavior based on the measurement of dynamic correlations[1]. Contagion is defined by Forbes and Rigobon (2002) as significant increases in cross-market co-movements, while interdependence is considered any continued market correlation at high levels [2], the existence of contagion must involve evidence of a dynamic increment in correlations. Contagion, as opposed to 'interdependence', conveys the idea that there are breaks in the international transmission mechanism due to financial panics, herd behavior or switches of expectations across instantaneous equilibrium. Studying the international transmission of shocks from the Hong Kong stock market crisis in October 1997, Corsetti et al. (2005) found that the strong result of "no contagion, only interdependence" obtained by previous contributions is quite dubious for a number of countries, and they found evidence of "some contagion, some interdependence". Chiang et al. (2007) identified two different phases of the Asian crises. The first phase entails a process of increasing volatility in stock returns due to contagion spreading from the earlier crisis-hit countries to other coun-

[1]Chiang, T. C. and Zheng, D. (2010). An empirical analysis of herd behavior in global stock markets. *Journal of Banking and Finance*, 34(8), pp. 1911-1921.

[2]Forbes, K.J. and Rigobon, R. (2002). No contagion, only interdependence: Measuring stock market co-movements. *Journal of Finance*, 57(5), pp. 2223-24.

tries. In this phase, investor trading activities are governed mainly by local information. However, in the second phase, as the crisis grew in public awareness, the correlations between stock returns and their volatility are consistently higher, as evidenced by herd behavior. Boyer et al. (2006) found that, in emerging stock markets, there is greater co-movement during high volatility periods, signifying that crises stretch through the holdings of international investors are mainly due to contagion rather than changes in fundamentals.

In the second path of empirical work on herd behavior we can distinguish between measures of herding by individuals or small-group of investors and market-wide herding. The former analyse the tendency of individuals or group of investors such as fund managers and financial analysts to imitate each other and trade the same asset at the same time. In this case, it is necessary to have detailed records of investors' trading activities. On the other hand, measures of market-wide herding aim to analyse the collective behavior of all participants towards the market views and therefore buying or selling a particular asset at the same time. Chang et al. (2000) found considerable evidence of herd behavior in South Korea and Taiwan. However, there was no evidence of herd behavior in the United States, Hong Kong and Japan. A study by Demirer and Kutan (2006), which examined data from individual firms and at the sector level, found persistent herd behavior in Chinese markets and suggested that the dispersions of equity returns are significantly higher during periods of large changes in the aggregate market price.

In the following we report some empirical measures of both herding by individuals or small group of investors, and market-wide herding.

4.2 Measures of Herding by Individuals or Small Group of Investors

4.2.1 Lakonishok, Shleifer and Vishny measure

Lakonishok et al. (LSV) (1992) used the data on holdings of 769 tax-exempt funds, and they concluded that money managers in their sample do not exhibit significant herd behavior. Their criterion is based on trades conducted by an homogenous group of fund managers over a period of time. Let $B(i,t)$ ($S(i,t)$) be the number of fund managers who buy (sell) the asset i in quarter t and $H(i,t)$ be the measure of herd behavior in stock i of the quarter t. LSV define $H(i,t)$ as follow:

$$H(i,t) = |\, p(i,t) - p(t)\,| - AF(i,t) \tag{4.2.1}$$

where $p(i,t) = \frac{B(i,t)}{B(i,t)+S(i,t)}$, and $p(t)$ is the average of $p(i,t)$ over all assets i that were traded by at least one fund manager. $AF(i,t)$ is the adjustment factor and it accounts for the fact that under the null hypothesis of no herd behavior $|\, p(i,t) - p(t)\,|$ is greater than zero. $AF(i,t)$ is defined as follows:

$$AF(i,t) = E[|\, p(i,t) - p(t)\,|]$$

note that the expectation is calculated under the null hypothesis that $B(i,t)$ follows a binomial distribution with parameter $p(t)$.

Under the null hypothesis of no herd behavior the probability of a randomly chosen fund managers being net buyer of stock i is $p(t)$. If $B(i,t) + S(i,t)$ is large then under the null hypothesis $p(i,t)$ will tend to $p(t)$ and $AF(i,t)$ will be close to zero. The adjustment factors is included in the herding measure to take account for the bias in $|\, p(i,t) - p(t)\,|$ for stock-quarters which are traded by a small number of participants. If $H(i,t)$ is significantly different from zero, this result should be interpreted as evidence of herd behavior[3].

The main drawbacks of this criterion are that it does not take account for the quantity of stock investors buy or sell, moreover it is not possible to identify intertemporal trading patterns[4].

4.2.2 Portfolio Change Measure

Wermers (1995) developed a new measure of herd behavior, called *portfolio-change measure* (PCM) of correlated trading, which captures both the direction and the intensity of trading by investors. The cross correlation PCM of lag τ between portfolio I and J is defined as follows:

$$\hat{\rho}^{I,J}_{t,\tau} \equiv \frac{\left(\frac{1}{N_t}\right) \sum_{n=1}^{N_t} (\Delta \widetilde{w}^{I}_{n,t})(\Delta \widetilde{w}^{J}_{n,t-\tau})}{\hat{\sigma}^{I,J}(\tau)} \tag{4.2.2}$$

where $\Delta \widetilde{w}^{I}_{n,t}$ and $\Delta \widetilde{w}^{J}_{n,t-\tau}$ are respectively, the change in portfolio I's weight of n during the period, quarter, $[t-1,\, t]$, and the change in portfolio J's weight of n during the period, quarter, $[t-\tau-1,\, t-\tau]$, N_t is the number of stocks in the

[3] Lakonishok, J., Shleifer, A. and Vishny, R. W. (1992). The impact of institutional trading on stock prices. *Journal of Financial Economics*, 32, pp. 29-30.

[4] Bikhchandani, S. and Sharma, S. (2000). *art.cit.*, p. 18.

intersection of the set of tradable securities in portfolio I during period $[t-1,\, t]$, and the set of tradable securities in portfolio I during period $[t-\tau-1,\, t-\tau]$, and $\hat{\sigma}^{I,J}(\tau) = \frac{1}{T}\sum_t \left\{ \frac{1}{N_t} \left[\sum_n \left(\Delta\widetilde{w}^I_{n,t}\right)^2 \sum_n \left(\Delta\widetilde{w}^J_{n,t-\tau}\right)^2 \right]^{\frac{1}{2}} \right\}$ is the time series average of the product of the cross-sectional standard-deviations.

Using the PCM measure, Wermers found evidence of a significant level of herd behavior by mutual funds. He randomly splits his sample of mutual funds into two groups and then used the PCM measure of correlated trading to compare the revisions of the net asset value weighted portfolios of the two groups.

4.3 Market-wide herding measures

4.3.1 Cross-sectional Standard Deviation

Christie and Huang (1995) suggested that under the traditional definition of herd behavior, an intuitive measure of its market impact is dispersion. To measure the return dispersion, they proposed the *cross-sectional standard deviation* (CSSD) method, which is expressed as:

$$CSSD_t = \sqrt{\frac{\sum_{i=1}^{N}(R_{i,t} - R_{m,t})^2}{(N-1)}} \tag{4.3.1}$$

where N is the number of firms in the portfolio, $R_{i,t}$ is the observed stock return of industry i at time t, and $R_{m,t}$ is the cross-sectional average stock of N returns in the portfolio at time t[5].

Dispersions quantify the average proximity of individual returns to the mean. They are bounded from below at zero when all returns move in perfect union with the market. As individual returns begin to vary from the market return, the level of dispersion increases.

During periods of extreme market movements, investors are more likely to suppress their own information, and to mimic collective actions in the market. Individual stock returns under these conditions tend to cluster around the overall market return then, dispersions tend to decrease.

Differently, the rational asset pricing models predict that, during periods of market stress, large changes in the market return would translate into an increase in dispersion, because individual assets differ in their sensitivity to the market return. In other words,

[5]Christie, W. G., and Huang, R. D. (1995). Following the pied piper: Do individual returns herd around the market?. *Financial Analysts Journal*, pp. 31-37.

dispersion in factor sensitivities will repel individual returns away from the market. Thus, herd behavior and rational asset pricing models offer conflicting predictions for the behavior of dispersions during periods of market stress[6].

Christie and Huang empirically examined whether equity return dispersions are significantly lower than average during periods of extreme market movements. They estimated the following empirical specification:

$$CSSD_t = \alpha + \beta_1 D_t^L + \beta_2 D_t^U + \epsilon_t \tag{4.3.2}$$

where $D_t^L = 1$, if the market return on day t lies in the extreme lower tail of the distribution, and equal to zero otherwise; and $D_t^U = 1$, if the market return on day t lies in the extreme upper tail of the distribution, and equal to zero otherwise.

The dummy variables are designed to capture differences in investor behavior in extreme up or down versus relatively normal markets. The presence of negative and statistically significant β_1 and β_2 coefficients should be interpreted as evidence of herd behavior. Christie and Huang used one or five percent of the observations in the upper and lower tail of the market return distribution to define extreme price movement days[7].

Christie and Huang empirical evidence showed that dispersions increase significantly during periods of large absolute price changes supporting the predictions of rational asset pricing models, and suggesting that herd behavior is not an important factor in determining equity returns during periods of market stress[8].

4.3.2 Cross-sectional Absolute Deviation

Chang et al. (2000) proposed an alternative measure of dispersion. The return dispersion is measured by the *cross-sectional absolute deviation* (CSAD):

$$CSAD_t = \frac{1}{N} \sum_{i=1}^{N} |R_{i,t} - R_{m,t}| \tag{4.3.3}$$

where N is the number of firms in the portfolio, $R_{i,t}$ is the observed return on asset i at time t, and $R_{m,t}$ is the return on the market portfolio at time t.

The$CSAD$ is not a measure of herd behavior, instead the relationship between $CSAD_t$ and $R_{m,t}$ is used to detect herd behavior[9]. In order to detect herd behavior, Chang et al. proposed the following specification:

[6] *Ibid.*, p. 32.
[7] *Ibid.*, p. 33.
[8] *Ibid.*, pp. 36-37.
[9] Chang, E. C., Cheng, J.W., Khorana, A. (2000). An examination of herd behavior in equity markets: An international perspective. *Journal of Banking and Finance*, 24, p. 1654.

$$CSAD_t = \alpha + \gamma_1 |R_{m,t}| + \gamma_2 R_{m,t}^2 + \epsilon_t \tag{4.3.4}$$

Chang et al. noted that rational asset pricing models predict that equity return dispersions are an increasing and linear function of the market return. If market participants tend to follow aggregate market behavior and ignore their private information during periods of large average price movements, then the relation between dispersion and market return will no longer be linear and increasing. Instead, the relation could become non-linearly increasing or even decreasing. For this reason, a non-linear market return, $R_{m,t}^2$, is included in the econometric model, and a significantly negative coefficient γ_2 in the empirical test would be consistent with the occurrence of herd behavior.

To allow for the possibility that the degree of herd behavior may be asymmetric in the up-versus the down-market, Chang et al. run the following empirical specification[10]

$$CSAD_t^{UP} = \alpha + \gamma_1^{UP} |R_{m,t}^{UP}| + \gamma_2^{UP} (R_{m,t}^{UP})^2 + \epsilon_t$$

$$CSAD_t^{DOWN} = \alpha + \gamma_1^{DOWN} |R_{m,t}^{DOWN}| + \gamma_2^{DOWN} (R_{m,t}^{DOWN})^2 + \epsilon_t$$

Chang et al. empirical analysis showed that during periods of extreme price movements, equity return dispersions for the US, Hong Kong and Japan actually tend to increase rather than diminuish, hence providing evidence against the presence of any herd behavior. However, for South Korea and Taiwan they found evidence of herd behavior during both extreme up and down price movement days. The differences in return dispersions across the developed and emerging markets may partly be the result of incomplete information disclosure in the emerging markets[11].

4.3.3 Beta Herding

Hwang and Salmon (2004) developed a new approach to measuring market-wide herding based on observing deviations from the equilibrium expressed by CAPM prices. By conditioning on the observed movements in fundamentals it is possible to separate adjustment to fundamentals news from herd behavior due to market sentiment and hence extract the latent herding component in observed asset prices.

Beta herding measures the behavior of traders who follow the performance of the market, or other macroeconomic factors, and hence buy or sell individual assets at the same time disregarding the underlying risk-return relationship.

[10] *Ibid.*, p. 1655-57.
[11] *Ibid.*, p. 1677.

When investors herd toward the performance of the market portfolio, the CAPM betas for individual assets will be biased and the cross-sectional dispersion of the individual betas smaller than it would be in equilibrium.

Consider the following CAPM in equilibrium

$$E_t[r_{i,t}] = \beta_{i,m,t} E[r_{m,t}],$$

where $r_{i,t}$ is the excess return on asset i, $r_{m,t}$ is the excess return on the market portfolio and $\beta_{i,m,t}$ is the systematic risk measure.

The conventional CAPM assumes that the betas does not change over time. Although, there is considerable empirical evidence that shows the betas are not constant, even if this evidence does not suggest that the betas change over time in equilibrium. A significant proportion of the betas time-variation reflects changes in investor sentiment and equilibrium betas generally vary very slowly as firms evolve.

When herd behavior arises, market participants' beliefs shift so as to follow the performance of the overall market more than they should in equilibrium, they disregard the equilibrium relationship (β) and will try to harmonise the return on individual assets with that of the market[12].

When there is herd behavior towards the market portfolio and the equilibrium relationship no longer holds; both the beta and the expectation of the asset i's return will be biased. Hwang and Salmon assumes that $E[r_{m,t}]$ is set by a common market-wide view and the investor first forms a view of the market as a whole and then considers the value of the individual asset. Then, in the presence of herd behavior towards the market portfolio, the following relationship is assumed to hold:

$$\frac{E_t^b[r_{i,t}]}{E_t[r_{m,t}]} = \beta_{i,m,t}^b = \beta_{i,m,t} - h_{m,t}(\beta_{i,m,t} - 1)$$

where $E_t^b[r_{i,t}]$ is the market's biased short run conditional expectation on the excess returns of asset i, $\beta_{i,m,t}^b$ is the biased beta of asset i at time t, and $h_{m,t}$ is a latent herding parameter that changes over time. Note that $h_{m,t}$ process must be stationary around zero for the CAPM equilibrium to exist.

If $h_{m,t} = 0 \Rightarrow \beta_{i,m,t}^b = \beta_{i,m,t}$ there is no herd behavior, on the contrary if $h_{m,t} = 1 \Rightarrow \beta_{i,m,t}^b = 1$ there is perfect herd behavior toward the market portfolio, in general when $0 < h_{m,t} < 1$ some degrees of herd behavior exists.

[12]Hwang, S. and Salmon, M. (2004). Market stress and herding. *Journal of Empirical Finance*, 11, pp. 590-91.

Consider the case of $\beta_{i,m,t} > 1$ and thus $E_t[r_{i,t}] > E_t[r_{m,t}]$, the equity is herded toward the market portfolio so that $E_t[r_{i,t}] > E_t^b[r_{i,t}] > E_t[r_{m,t}]$. Therefore, the equity looks less risky than it should. The opposite, if $\beta_{i,m,t} < 1$[13].

As, $\beta_{i,m,t}$ and $h_{m,t}$ are not observable, Hwang and Salmon proposed to use the cross-sectional standard deviation of the individual asset betas.

$$Std_c(\beta^b_{i,m,t}) = \sqrt{E_c[(\beta^b_{i,m,t} - E_c[\beta^b_{i,m,t}])^2]}$$

$$= \sqrt{E_c[(\beta_{i,m,t} - h_{m,t}(\beta_{i,m,t} - 1) - 1)^2]} = \sqrt{E_c[(\beta_{i,m,t} - 1)^2]}(1 - h_{m,t})$$

$$Std_c(\beta^b_{i,m,t}) = Std_c(\beta_{i,m,t})(1 - h_{m,t}) \tag{4.3.5}$$

where $E_c[.]$ and $Std_c(.)$ represents the cross-sectional expectation and standard deviation, respectively, and $E_c[\beta^b_{i,m,t}] = 1$.

As $Std_c(\beta_{i,m,t})$ can not change significantly within a short time interval, any significant changes in $Std_c(\beta^b_{i,m,t})$ over a short time period will be attributed to changes in $h_{m,t}$[14].

To extract $h_{m,t}$ from $Std_c(\beta^b_{i,m,t})$, Hwang and Salmon took logarithms of equation (4.3.5)

$$log[Std_c(\beta^b_{i,m,t})] = log[Std_c(\beta_{i,m,t})] + log(1 - h_{m,t})$$

As $Std_c(\beta_{i,m,t})$ can vary slowly, it is assumed:

$$log[Std_c(\beta_{i,m,t})] = \mu_m + \nu_{m,t} \tag{4.3.6}$$

where $\mu_m = E[log[Std_c(\beta_{i,m,t})]]$ and $\nu_{m,t} \sim iid(0, \sigma^2_{m,v})$. The model is formalized in the following state-space form

$$log[Std_c(\beta^b_{i,m,t})] = \mu_m + H_{m,t} + \nu_{m,t} \tag{4.3.7}$$

$$H_{m,t} = \phi H_{m,t-1} + \eta_{m,t}$$

where $H_{m,t} = log(1-h_{m,t})$, and it follows an autoregressive process, $\eta_{m,t} \sim iid(0, \sigma^2_{m,\eta})$. The authors estimated the model using the Kalman filter.

[13] *Ibid.*
[14] *Ibid.* pp. 589-592.

A significant $\sigma^2_{m,\eta}$ can be interpreted as the existence of herd behavior and a significant ϕ as a support for this particular autoregressive structure. Another restriction is that the process of herd behavior should be stationary, hence $|\phi| < 1$ is required.

The $Std_c(\beta^b_{i,m,t})$ is expected to change over time in response to herd behavior in the market. However, it is necessary to check if the herd behavior extracted from $Std_c(\beta^b_{i,m,t})$ is robust in the presence of market variables or variables reflecting macroeconomic fundamentals. If $H_{m,t}$ becomes insignificant after having included these variables in the model, then changes in $Std_c(\beta^b_{i,m,t})$ are due to changing in fundamentals instead of herd behavior.

Some alternative models:

1. This model includes market volatility ($\sigma_{m,t}$) and returns ($r_{m,t}$) as independent variables in the measurement equation,

$$log[Std_c(\beta^b_{i,m,t})] = \mu_m + H_{m,t} + c_{m1} ln\sigma_{m,t} + c_{m2} r_{m,t} + \nu_{m,t} \tag{4.3.8}$$

$$H_{m,t} = \phi H_{m,t-1} + \eta_{m,t}$$

2. This model adds SMB and HML factors of Fama and French (1993) as further independent variables in (4.3.8).

$$log[Std_c(\beta^b_{i,m,t})] = \mu_m + H_{m,t} + c_{m1} ln\sigma_{m,t} + c_{m2} r_{m,t} + c_{m3} SMB_t + c_{m4} HML_t + \nu_{m,t} \tag{4.3.9}$$

$$H_{m,t} = \phi H_{m,t-1} + \eta_{m,t}$$

3. The most general model is given by adding macroeconomic variables, such as the dividend price ratio (DP_t), the relative treasury bill rate (RTB_t), the term spread (TS_t), and the default spread (DS_t).

$$log[Std_c(\beta^b_{i,m,t})] = \mu_m + H_{m,t} + c_{m1} ln\sigma_{m,t} + c_{m2} r_{m,t} +$$

$$+ c_{m5} DP_t + c_{m6} RTB_t + c_{m7} TS_t + c_{m8} DS_t + \nu_{m,t} \tag{4.3.10}$$

$$H_{m,t} = \phi H_{m,t-1} + \eta_{m,t}$$

The measurement of herd behavior towards other factors can be investigated using the standard linear factor model.

$$r_{i,t} = \alpha^b_{i,t} + \sum_{k=1}^{K} \beta^b_{i,k,t} f_{k,t} + \epsilon_{i,t}, \;\; i = 1, ..., N \, and \, t = 1, ..., T$$

where $\alpha^b_{i,t}$ is an intercept that changes over time, $\beta^b_{i,k,t}$ are the biased coefficients under herd behavior on factor k at time t, $f_{k,t}$ is the realized value of factor k at time t, and $\epsilon_{i,t}$ is mean zero with variance σ^2_ϵ.[15]

Herd behavior towards factor k at time t, $h_{k,t}$, can then be captured by

$$\beta^b_{i,k,t} = \beta_{i,k,t} - h_{k,t}(\beta_{i,k,t} - E_c[\beta_{i,k,t}])$$

Thus, with the same assumption as before, the model can be written:

$$log[Std_c(\beta^b_{i,k,t})] = \mu_k + H_{k,t} + \nu_{k,t} \tag{4.3.11}$$

$$H_{k,t} = \phi H_{k,t-1} + \eta_{k,t}$$

where $\mu_k = E[log[Std_c(\beta_{i,k,t})]]$,$\eta_{m,t} \sim iid(0, \sigma^2_{m,\eta})$, $\nu_{k,t} \sim iid(0, \sigma^2_{k,v})$, and $H_{k,t} = log(1 - h_{k,t})$.[16]

Hwang and Salmon (2004) applied their approach to the US, and South Korean stock market and found that herd behavior toward the market portfolio showed significant movements and persistence even when the level of market volatility and returns was taken into account. Macro factors were found to offer almost no help in explaining these herding patterns. They also found evidence of herd behavior towards the market portfolio both when the market was rising and when it was falling. In addition, they examined herding relationships between different countries and found that herd behavior towards the market portfolio between US and South Korea was not significantly correlated. Finally, they have shown that in the US market there were relatively few periods in the entire sample that herd behavior, while present in the market, was a major concern.

4.3.4 Quantile Regression

Saastamoinen (2008) investigated herd behavior in the Finland stock market. He applied the methodology of Chang et al. (2000) to daily closing price quotes for the large-capital companies in the OMXH.

The main contribution of this paper is the application of quantile regression (QR) to the analysis of stock returns dispersion. He argued that there are several reasons to use quantile regression in attempts to detect herd behavior in equity markets. First, financial data usually are not normally distributed. Second, since the market stress

[15] *Ibid.* pp. 592-593.
[16] *Ibid.*

models are prevalent in the empirical literature of herd behavior in financial markets, QR is a versatile tool in analyzing extreme quantiles of return distribution. Third, Christie and Huang noted that the method based on CSSD is sensitive to outliers, while QR is not sensitive to outliers[17].

Saastamoinen found statistically significant negative coefficient up to the first quartile (25%) of stock returns distribution. After this, the coefficient remained negative but statistically insignificant almost up to the median. A possible explanation for this could be that stock sell-offs disturb the information set of investors, who then cut their losses or consolidate earnings by selling when other investors sell too. One cause for this finding could be the observation period, which coincides with an economic expansion and a bull market. Overall, this might constitute the evidence of herd behavior. However, this is just speculation because there is no control device that would implicate herd behavior as the causal factor[18].

4.4 Cipriani and Guarino (2014)

Cipriani and Guarino (2014) develop a theoretical model of herd behavior in financial markets that can be estimated with financial market transaction data. This methodology permits to measure the quantitative importance of herd behavior, to identify when it happens, and to assess the informational inefficiency that it generates.

Their theoretical model is built on the work of Avery and Zemsky (1998) (see section 2.5), which shows that herd behavior occurs if there is multidimensional uncertainty and more precisely when there is *event uncertainty* in addition to *value uncertainty*.

Cipriani and Guarino estimated the model using data for a NYSE stock (Ashland Inc.) in 1995 and identified the periods in each trading day when informed traders herd. They found that herd behavior was present in the market and fairly pervasive on some trading days. Moreover, herd behavior generated important informational inefficiencies.

In the following the theoretical model proposed by Cipriani and Guarino (2014) and its estimation method will be discussed.

4.4.1 The Theoretical Model

An asset is traded by a sequence of traders who interact with a market maker. Trading days are indexed by $d = 1, 2, 3, \ldots$. Time within each day is discrete and indexed by

[17] Saastamoinen, J. (2008). Quantile regression analysis of dispersion of stock returns - evidence of herding?. Discussion Papers, No. 57. University of Joensuu, Economics, p. 10.

[18] *Ibid.*, p. 14.

$t = 1, 2, 3, \ldots$.

The fundamental asset value in day d is denoted by V_d. The asset value does not change during the day but it can change from one day to the next. At the beginning of the day, the asset value remains the same as in the previous day ($V_d = v_{d-1}$) with probability $1-\alpha$, and it changes with probability α. In each day d, the value of the asset in the previous day $d-1$, v_{d-1} , is known to all market participants. When the asset value changes from one day to the other, an information event occurred and the asset value will decrease to $v_d^L = v_{d-1} - \lambda^L$ with probability $1-\delta$ in case of bad informational event, and increase to $v_d^H = v_{d-1} + \lambda^H$ with probability δ in case of good informational event, where $\lambda^L > 0$ and $\lambda^H > 0$[19]. Informational events are independently distributed over the days of trading. Finally, $(1-\delta)\lambda^L = \delta\lambda^H$ is assumed to guarantee that the closing price is a martingale[20].

The asset is traded in a specialist market. A market maker, who interacts with a sequence of traders, set the asset price. At any time $t = 1, 2, 3, \ldots$ during the day a trader is randomly chosen and it can choose among to buy, sell, or not to trade (the trader's action space is $\mathcal{A} = \{buy,\ sell,\ no\, trade\}$). Each trade consists of the exchange of one unit of the asset for cash. The symbol X_t^d indicates the action of the trader at time t in day d, and H_t^d is the history of trades and prices until time $t-1$ of day d.

At any time t of day d, when the market maker posts the asset prices, he must take account for the possibility of trading with investors who have some private information on the asset value. He will set different prices for buying and for selling, thus there will be a bid-ask spread[21]. The ask price (the price at which a trader can buy) at time t is denoted by a_t^d and b_t^d is the bid price (the price at which a trader can sell) at time t in day d. Due to (unmodeled) potential competition, the market maker makes zero expected profits when he sets the ask and bid prices equal to the expected value of the asset conditional on the information available at time t and on the chosen action

[19] Cipriani, M., and Guarino, A. (2014). Estimating a structural model of herd behavior in financial markets. *American Economic Review*, 104(1), p. 228.

[20] An adapted process V is a martingale if

$$E(V_d \mid \mathcal{F}_{d-1}) = V_{d-1}$$

where V_{d-1} is known in day d and $\mathcal{F}_{d-1}$ is the information available at time $d-1$. As $E(V_d \mid \mathcal{F}_{d-1}) = (v_{d-1} + \lambda^H)\delta + (v_{d-1} - \lambda^L)(1-\delta)$, if V is a martingale it will be verified:

$$(v_{d-1} + \lambda^H)\delta + (v_{d-1} - \lambda^L)(1-\delta) = v_{d-1}$$

the above equation is true if and only if $(1-\delta)\lambda^L = \delta\lambda^H$. Thus, this condition guarantees that the closing price is a martingale.

[21] Glosten, L. R., and Milgrom, P.R. (1985). Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1), pp. 71-100.

$$a_t^d = E(V_d \,|\, h_t^d,\, X_t^d = buy,\, a_t^d,\, b_t^d)$$

$$b_t^d = E(V_d \,|\, h_t^d,\, X_t^d = sell,\, a_t^d,\, b_t^d)$$

There are a countable number of traders. Traders act in an exogenous sequential order. Each trader is chosen to take an action only once, at time t of day d. Traders can either be informed or noise. The former have private information on the asset value, while the latter trade for unmodeled reasons. In no-event days, all traders in the market are noise traders. In information-event days, at any time t an informed trader is chosen to trade with probability μ and a noise trader with probability $1-\mu$, with $0 < \mu < 1$[22].

Noise traders buy with probability $\frac{\epsilon}{2}$, sell with probability $\frac{\epsilon}{2}$, and do not trade with probability $1-\epsilon$ (with $0 < \epsilon < 1$). Informed traders receive a private signal on the new asset value and observe the previous history of trades and prices, and the current prices. The private signal S_t^d has the following value-contingent densities

$$g^H(s_t^d \,|\, v_d^H) = 1 + \tau(2s_t^d - 1)$$

$$g^L(s_t^d \,|\, v_d^L) = 1 - \tau(2s_t^d - 1)$$

with $\tau \in (0, \infty)$.

Given the value of the asset, the signals S_t^d are $i.i.d.$ The signals satisfy the monotone likelihood ratio property. At each time t, the likelihood ratio after receiving the signal, $\frac{Pr(V_d=v_d^H \,|\, h_t^d, s_t^d)}{Pr(V_d=v_d^L \,|\, h_t^d, s_t^d)} = \frac{g^H(s_t^d \,|\, v_d^H)}{g^L(s_t^d \,|\, v_d^L)} \frac{Pr(V_d=v_d^H \,|\, h_t^d)}{Pr(V_d=v_d^L \,|\, h_t^d)}$, is higher than that before receiving the signal if $s_t^d > 0.5$ ("good signal") and lower if $s_t^d < 0.5$ ("bad signal"). The parameter τ measures the informativeness of the signals. When $\tau \to 0$, the densities are uniform, and the signals are completely uninformative. As τ increases, the signals become more and more informative. Given that signal structure, informed traders are heterogenous, since they receive signal realizations with different degrees of informativeness about the asset fundamental value. In addition to capturing heterogeneity of information in the market, a linear density function for the signal makes it possible to compute the traders' strategies and the market maker's posted prices analytically. An informed trader's payoff function is defined as

[22] Cipriani, M., and Guarino, A. (2014). *art.cit.*, pp. 230-31.

$$U(v_d,\, X_t^d,\, a_t^d,\, b_t^d) = \begin{cases} v_d - a_t^d & if\ X_t^d = buy, \\ 0 & if\ X_t^d = no\,trade, \\ b_t^d - v_d & if\ X_t^d = sell. \end{cases}$$

An informed trader finds optimal to buy whenever $E(V_d \,|\, h_t^d,\, s_t^d) > a_t^d$ and to sell whenever $E(V_d \,|\, h_t^d,\, s_t^d) < b_t^d$. He chooses not to trade when $b_t^d < E(V_d \,|\, h_t^d,\, s_t^d) < a_t^d$. Otherwise, he is indifferent between buying and not trading, or selling and not trading.

Moreover, the trading decision of an informed trader can be simply characterized by two threshold, σ_t^d and β_t^d. An informed trader will buy for any signal realization greater than β_t^d and sell for any signal realization smaller than σ_t^d[23].

Definition 7. An informed trader engages in herd buying at time t of day d if

(i) he buys upon receiving a bad signal,

$$E(V_d \,|\, h_t^d,\, s_t^d) > a_t^d\, for\, s_t^d < 0.5$$

(ii) the price of the asset is higher than the price at time 1,

$$p_t^d = E(V_d \,|\, h_t^d) > p_1^d = v_{d-1}$$

Similarly, an informed trader engages in herd selling at time t of day d if

(i) he sells upon receiving a good signal,

$$E(V_d \,|\, h_t^d,\, s_t^d) < b_t^d\, for\, s_t^d > 0.5$$

(ii) the price of the asset is lower than the price at time 1, [24],

$$p_t^d = E(V_d \,|\, h_t^d) < p_1^d = v_{d-1}$$

In other words, a trader herds when he trades against his own information in order to conform to the information contained in the history of trades. At any given time t, it is possible to detect whether an informed trader herds for a positive measure of signals by simply comparing the two thresholds σ_t^d and β_t^d to 0.5.

Definition 8. There is herd behavior at time t of day d when there is a positive measure of signal realizations for which an informed trader either herd buys or herd sells, which is, when

[23] *Ibid.*

[24] *Ibid.*, p. 232

$$\sigma_t^d > 0.5 \, or \, \beta_t^d < 0.5.$$

Figure 4.4.1 and 4.4.2 show an example of herd buy and herd sell in a day with a good information event[25].

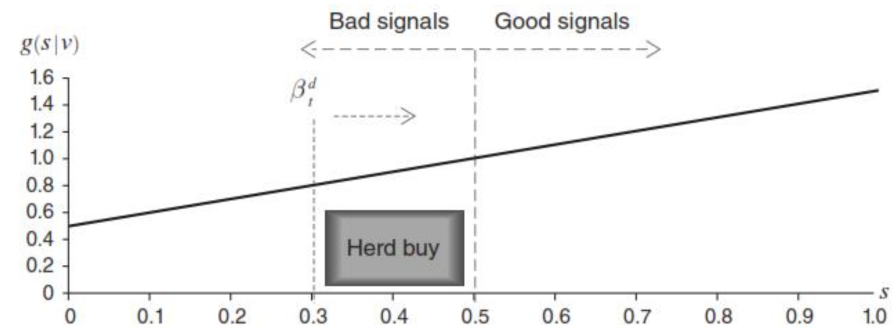

Figure 4.4.1: Herd buy

Source: Cipriani and Guarino. (2014), p. 234

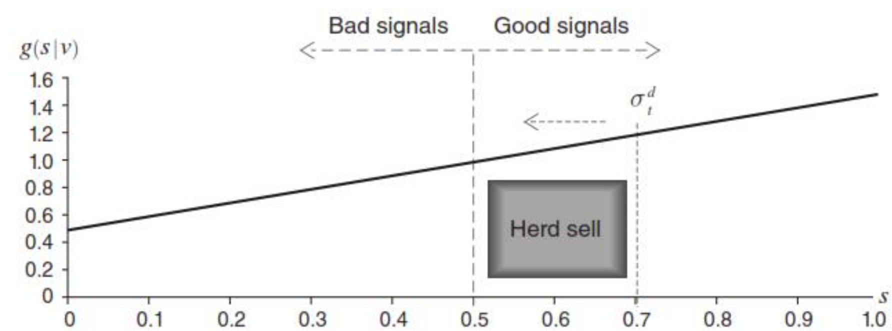

Figure 4.4.2: Herd sell

Source: Cipriani and Guarino. (2014), p. 234

The reason why herd behavior occurs is that prices move “too slowly” as buy and sell orders arrive in the market. Suppose that, at the beginning of an information event day, there is a sequence of buy orders. Informed traders, knowing that there has been an information event, attach a certain probability to the fact that these orders come from informed traders with good signals. However, the market maker attaches a lower probability to this event, as he takes account for the possibility that there was no information event and that all the buys came from noise traders. Therefore, after

[25] *Ibid.*, p. 233

a sequence of buys, he will update the prices upwards, but by less than the movement in traders' expectations. Because traders and the market maker interpret the history of trades differently, the expectation of a trader with a bad signal may be higher than the ask price, in which case he herd buys.

As in Avery and Zemsky (1998) in this model herd behavior occurs because of uncertainty on whether an information event has occurred.

The probability of herd behavior depends on the parameter values. For instance, when α (the probability of an information event) is close to zero, the market maker has a strong prior that information events do not occur. He barely updates the prices as trades arrive in the market, and herd behavior arises as soon as there is an imbalance in the order flow. On the contrary, if α is close to 1, the market maker and the informed traders update their beliefs in a similar manner and herd behavior rarely happens[26].

4.4.2 Estimation of the Theoretical Model

To estimate the herd behavior model presented above, it is necessary to specify its likelihood function. Cipriani and Guarino write the likelihood function for the history of trades only, disregarding bid and ask prices, as in their theoretical model there is no public information and for this reason, there is a one-to-one mapping from trades to prices, and adding prices to the likelihood function would be redundant. Remember also that information events are assumed to be independent and, before the market opens, market participants have learned the realization of the previous day's asset value. Because of this, the probability of the sequence of trades in a day depends only on the value of the asset that day[27]. Therefore, the likelihood of a history of trades over multiple days can be written as

$$\mathcal{L}(\Phi; \{h^d\}_{d=1}^{D}) = Pr(\{h^d\}_{d=1}^{D} \,|\, \Phi) = \prod_{d=1}^{D} Pr(h^d | \, \Phi)$$

where h^d is the history of trades at the end of a trading day d, and $\Phi \equiv \{\alpha, \, \delta, \, \mu, \, \tau, \, \epsilon\}$ is the vector of parameters.

As the sequence of trades, and not just the number of trades, conveys information, Cipriani and Guarino focus on the probability of a history of trades in a single day. Therefore, it is necessary to compute the probability of a history of trades recursively

[26] *Ibid.*, p. 233-35.
[27] *Ibid.*

$$Pr(h^d|\,\Phi) = \prod_{s=1}^{t} Pr(x_s^d|\,h_s^d,\,\Phi)$$

where $Pr(x_t^d|\,h_t^d,\,\Phi)$ depends on the measure of informed traders who buy, sell, or do not trade after a given history of trades h_t^d. Using the law of total probability, at each time t, $Pr(x_t^d|\,h_t^d,\,\Phi)$ is computed in the following way:

$$Pr(x_t^d|\,h_t^d,\,\Phi) = Pr(x_t^d|\,h_t^d,\,v_d^H,\,\Phi)Pr(v_d^H|\,h_t^d,\,\Phi)$$

$$+Pr(x_t^d|\,h_t^d,\,v_d^L,\,\Phi)Pr(v_d^L|\,h_t^d,\,\Phi)$$

$$+Pr(x_t^d|\,h_t^d,\,v_{d-1},\,\Phi)Pr(v_{d-1}|\,h_t^d,\,\Phi)$$

Consider first the probability of an action conditional on a good-event day in order to show how to compute these probabilities. For the sake of exposition, let us focus on the case in which the action is a buy order. As illustrated above, at each time t, in equilibrium there is a signal threshold β_t^d such that an informed trader buys for any signal realization greater than β_t^d,

$$E(V_d|\,h_t^d,\,\beta_t^d) = a_t^d = E(V_d|\,h_t^d,\,X_t^d = buy,\,a_t^d,\,b_t^d)$$

which can be written as

$$v_{d-1} + \lambda^H Pr(v_d^H|\,h_t^d,\,\beta_t^d) - \lambda^L Pr(v_d^L|\,h_t^d,\,\beta_t^d)$$

$$= v_{d-1} + \lambda^H Pr(v_d^H|\,h_t^d,\,buy_t^d) - \lambda^L Pr(v_d^L|\,h_t^d,\,buy_t^d)$$

or, after some manipulation, as[28]

$$Pr(v_d^H|\,h_t^d,\,\beta_t^d) - Pr(v_d^L|\,h_t^d,\,\beta_t^d)$$

$$= \frac{\delta}{1-\delta}\left(Pr(v_d^H|\,h_t^d,\,buy_t^d) - Pr(v_d^L|\,h_t^d,\,buy_t^d)\right) \tag{4.4.1}$$

The probabilities in this equation can easily be expressed as a function of the traders' and market maker's beliefs at time t - 1 and of the parameters.

$Pr(v_d^H|\,h_t^d,\,\beta_t^d)$

[28]Remember the assumption $(1-\delta)\lambda^L = \delta\lambda^H$

$$= \frac{g^H(\beta_t^d|\, v_d^H)Pr(v_d^H|\, h_t^d,\, V_d \neq v_{d-1})}{g^H(\beta_t^d|\, v_d^H)Pr(v_d^H|\, h_t^d,\, V_d \neq v_{d-1}) + g^L(\beta_t^d|\, v_d^L)Pr(v_d^H|\, h_t^d,\, V_d \neq v_{d-1})}$$

and $Pr(v_d^H|\, h_t^d,\, buy_t^d)$

$$= \frac{Pr(buy_t^d|\, v_d^H,\, h_t^d)Pr(v_d^H|\, h_t^d)}{Pr(buy_t^d|\, v_d^H,\, h_t^d)Pr(v_d^H|\, h_t^d) + Pr(buy_t^d|\, v_{d-1},\, h_t^d)Pr(v_{d-1}|\, h_t^d) + Pr(buy_t^d|\, v_d^L,\, h_t^d)Pr(v_d^L|\, h_t^d)}$$

By substituting these expressions into (4.4.1) β_t^d can be computed. It is important to note first that the above expressions themselves contain the probabilities of a buy order by an informed trader in a good, bad, and no-event day; all these probabilities obviously depend on the threshold β_t^d itself (as illustrated below). That is, the threshold is a fixed point. Second, at time $t = 1$, the prior beliefs of the traders and of the market maker are a function of the parameters only. Therefore, it is possible to compute β_1^d as the solution to equation (4.4.1), and from β_1^d, the probability of a buy order at time 1. After observing x_1^d, the market maker's and traders' beliefs are updated, the same procedure for time 2 is repeated, and β_2^d and the probability $Pr(buy_2^d|\, x_1^d,\, v_d^H)$ are computed. This procedure is repeated recursively for each time t, always conditioning on the previous history of trades. Note that in order to maximize the likelihood function, the thresholds β_t^d (and the analogous threshold σ_t^d) have to be computed for each trading time in each day of trading, for each set of parameter values[29].

Once solved for β_t^d, the probability of a buy order in a good-event day can be computed. Let us focus on the case in which $\tau \in [0,\, 1)$; that is let us concentrate on the case of bounded beliefs. In this case,

$$Pr(buy_t^d|\, h_t^d,\, v_d^H) = \mu \int_{\beta_t^d}^1 (1 + \tau(2s_t^d - 1))ds_t^d + (1-\mu)\left(\tfrac{\epsilon}{2}\right)$$

$$= \left(\tau(1 - \beta_t^{d^2}) + (1-\tau)(1-\beta_t^d)\right)\mu + (1-\mu)\left(\frac{\epsilon}{2}\right)$$

By following an analogous procedure, it is possible to compute σ_1^d and the probability of a sell order in a good-event day,

$$Pr(sell_t^d|\, h_t^d,\, v_d^H) = \left((1-\tau)\sigma_t^d + \tau\sigma_t^{d^2}\right)\mu + (1-\mu)\left(\frac{\epsilon}{2}\right)$$

The probability of a no-trade is just the complement to the probabilities of a buy and of a sell.

[29] Cipriani, M., and Guarino, A. (2014). *art.cit.*, p. 236-37.

The analysis for the case of a bad information event ($V_d = v_d^L$) follows the same steps. The case of a no-event day ($V_d = v_{d-1}$) is easier, since the probability of a buy or of a sell is $\frac{\epsilon}{2}$ and that of a no-trade is $1 - \epsilon$. Also the case of unbounded beliefs, where $\tau \geq 1$, can be dealt with in a similar manner. The only changes are the extremes of integration when computing the probability of a trade. Finally, to compute $Pr(x_t^d |\, h_t^d,\, \Phi)$, the conditional probabilities of V_d given the history until time t is needed, which is $Pr(V_d = v |\, h_t^d,\, \Phi)$ for $v = v_d^L,\, v_{d-1},\, v_d^H$. These can also be computed recursively by using Bayes's rule.

Then, Cipriani and Guarino estimated the parameters through maximum likelihood, using both a direct search method (Nelder-Mead simplex) and the Genetic Algorithm[30].

[30] *Ibid.*, p. 238-39.

Chapter 5

An Estimate for the Italian Stock Exchange

5.1 Methodology

We use the approach proposed by Hwang and Salmon (2004) (see section 4.3.3) to detect herd behavior in the Italian stock market in the period from January 1998 to December 2012. We have decided to adopt this measure because of its empirical and theoretical properties in the sense that this measure automatically conditions on fundamentals and can also measure herd behavior towards other factors. In addition, the influence of time series volatility is accounted automatically by this measure.

Firstly, we calculate the OLS estimate of betas using daily data over monthly intervals in the Fama and French three factors model (FF model)

$$r_{it_d} = \alpha_{it} + \beta_{imt}(r_{mt_d} - r_{ft_d}) + \beta_{iSt}SMB_{t_d} + \beta_{iHt}HML_{t_d} + \epsilon_{it_d}$$

where t_d denotes day d in month t. The estimated betas are then used to create a monthly time series of the cross section standard deviation of the betas

$$\widehat{Std_c(\hat{\beta^b}_{ikt})} = \sqrt{\frac{\sum_{i=1}^{N_t}\left(\hat{\beta^b}_{ikt} - \overline{\hat{\beta}^b_{ikt}}\right)^2}{N_t}},$$

where $\hat{\beta^b}_{ikt}$ is the OLS estimate of the biased betas relative to the factor k with $k = m$, S, H, $\overline{\hat{\beta}^b_{ikt}} = \frac{1}{N_t}\sum_{i=1}^{N_t}\hat{\beta^b}_{ikt}$ and N_t is the number of equities in month t.

Then, we estimate the following state space models by using the Kalman Filter

$$\begin{cases} log[Std_c(\beta^b_{i,m,t})] = \mu_m + H_{m,t} + \nu_{m,t} \\ H_{m,t} = \phi H_{m,t-1} + \eta_{m,t} \end{cases} \tag{1}$$

$$\begin{cases} log[Std_c(\beta^b_{i,m,t})] = \mu_m + H_{m,t} + c_{m1} ln\sigma_{m,t} + c_{m2} r_{m,t} + \nu_{m,t} \\ H_{m,t} = \phi H_{m,t-1} + \eta_{m,t} \end{cases} \tag{2}$$

$$\begin{cases} log[Std_c(\beta^b_{i,m,t})] = \mu_m + H_{m,t} + c_{m1} ln\sigma_{m,t} + c_{m2} r_{m,t} + c_{m3} SMB_t + c_{m4} HML_t + \nu_{m,t} \\ H_{m,t} = \phi H_{m,t-1} + \eta_{m,t} \end{cases} \tag{3}$$

where $ln\sigma_{m,t}$ denotes the log-market volatility calculated using squared daily returns as in Schwert (1989) and $r_{m,t}$ is the market excess return. The unknown parameters are estimated by maximum likelihood[1].

5.2 Data and Descriptive Statistics

We use daily data from 1 January 1998 to 31 December 2012 to investigate herd behavior in the Italian Stock Exchange. We have calculated the herding measure considering 262 stocks and we have used the FTSE MIB index and the Bank of Italy BOT to calculate the excess market return .

As Hwang and Salmon (2004), we have decided to use Fama and French's SMB and HML factors. SMB and HML stand for "small [cap] minus big" and "high [book/price] minus low", these factors measure the historic excess returns of small caps and high book-value-to-price ratio stocks over the market as a whole. The SMB and HML factors are calculated using the method described in Fama and French (1993). At the end of June every year, all stocks are ranked on size. The market capitalization (ME) median is used to split the stocks into two groups, small and big (S and B). The stocks are also split into three book-to-market (BE-ME) equity groups, bottom 30% (Low), middle 40% (Medium) and top 30% (High) of the ranked values of BE-ME. As Fama and French, we do not use negative-BE firms when calculating the breakpoints for BE-ME or when forming the size-BE/ME portfolios. Six portfolios are constructed (SL, SM, SH, BL, BM and BH) from the intersection of the two ME and the three BE-ME groups. Daily value-weighted returns on the six portfolios are calculated. Then SMB and HML are obtained as follows

[1]All the calculations and estimates are made using R. In the estimate of the state space model we use package dlm.

$$SMB = \frac{1}{3}(SL + SM + SH) - \frac{1}{3}(BL + BM + BH),$$

$$HML = \frac{1}{2}(SH + BH) - \frac{1}{2}(SL + BL).$$

Table 5.1 reports some statistical properties of the excess market return and the SMB and HML factor returns. As can be seen, the three distributions are leptokurtic and positively skewed (except the market excess return).

	Market Excess Return	SMB	HML
Mean	-0.00022	0.00009	0.00076
Standard Deviation	0.01565	0.01233	0.02610
Skewness	-0.06751	15.86206	50.98766
Excess Kurtosis	4.31075	634.7973	2976.297
	Correlation Matrix		
	Market Excess Return	SMB	HML
Market Excess Return	1		
SMB	-0.53830	1	
HML	0.06857	0.50131	1

Table 5.1: Statistical Properties of the Excess Market Returns and Fama-French's SMB and HML Factor Returns

5.3 Empirical Results

5.3.1 Properties of the Cross-Sectional Standard Deviations of Betas

Table 5.2 illustrates some statistical properties of the estimated cross-sectional standard deviation of the betas.

As can be seen, the $\widehat{Std_c(\hat{\beta}^b{}_{ikt})}$ are positively skewed, and leptokurtic. In addition, the Jarque Bera statistics for normality indicates that the cross-sectional standard deviation of betas and their logarithm are not Gaussian, except the Log-cross sectional standard deviation of betas on HML factor.

5.3.2 Herd Behavior towards the Market Portfolio

We first investigate herd behavior towards the market portfolio in the period from January 1998 to December 2012. The results of Model 1 are given in the first column

Cross-Sectional Standard Deviations of OLS Betas			
	Market Returns	SMB	HML
Mean	0.9064	1.3195	0.9264
Standard Deviations	0.4158	0.6147	0.2478
Skewness	4.3576	4.6733	1.1747
Excess Kurtosis	31.3046	34.2908	2.4770
Jarque Bera Statistics (p-value)	7919.52 (<2.2e-16)	9474.123 (<2.2e-16)	87.4143 (<2.2e-16)
Log-Cross-Sectional Standard Deviations of OLS Betas			
	Market Returns	SMB	HML
Mean	-0.1638	0.2130	-0.1094
Standard Deviations	0.3369	0.3299	0.2553
Skewness	1.1388	1.3355	0.1500
Excess Kurtosis	2.9330	3.6858	0.6009
Jarque Bera Statistics (p-value)	103.4253 (<2.2e-16)	155.3972 (<2.2e-16)	3.3831 (0.1842)

Table 5.2: Properties of the Cross-Sectional Standard Deviation of Betas

of Table 5.3, the results of Model 2 and 3 in the second and third column. We can see immediately that H_{mt} is highly persistent with $\hat{\phi}_m$ large and significant and the proportions of signal are also of a similar order of magnitude indicating that herd behavior explains around 12% of the total variability in $Std_c(\beta^b_{imt})$. Moreover, the estimates of $\sigma_{m\eta}$ (the standard deviation of η_{mt}) are significant and thus we can conclude that there is herd behavior towards the market portfolio.

Model 2 takes the level of market volatility and returns into account. Note that the significance of the two market variable coefficients should be interpreted as adjustment in the mean level (μ_m) of $log(Std_c(\beta^b_{imt}))$ in the measurement equation and not as herd behavior, this permits us to examine the degree of herd behavior given the state of the market. The sign of the two coefficients is interesting; in fact $log(Std_c(\beta^b_{imt}))$ decreases as market volatility rises but increases with the level of market returns. So when the market becomes riskier and is falling, $Std_c(\beta^b_{imt})$ decreases, while it increases when the market becomes less risky and rises. Using the definition of herd behavior of Hwang and Salmon (2004), which is a reduction in $Std_c(\beta^b_{imt})$ due to the H_{mt} process, these results suggest that herd behavior is significant and exists independently from and given the particular state of the market.

Model 3 includes the SMB and HML factors as explanatory variables with results very similar to those of Models 1 and 2, and the estimated coefficients on SMB and HML are found not to be significant. The Akaike Information Criterion[2] (AIC) selects Model 2.

[2] Akaike derived an asymptotically unbiased estimator of expected Kullback-Leibler information as follows:

Further, the Jarque Bera normality test illustrates that the residuals of the Kalman filter are not normal in the three models analysed.

Herd Behavior Towards the Market Portfolio

Period: January 1998 - December 2012

	Model 1		Model 2		Model 3	
μ	-0.125		-0.747	**	-1.452	
ϕ	0.983	**	0.992	**	0.998	**
σ_{mv}	0.287	**	0.254	**	0.264	**
$\sigma_{m\eta}$	0.041	**	0.036	**	0.037	**
$log-Vm$	-		-0.115	**	-0.158	**
$r_m - r_f$	-		26.234	**	0.108	
SMB	-		-		-0.025	
HML	-		-		0.059	
$\frac{\sigma_{m\eta}}{sd\left(Log\left(\widehat{Std_c(\hat{\beta}^b{}_{imt})}\right)\right)}$	0.122		0.107		0.110	
Correlation with the Market Index	0.647	**	0.709	**	0.718	**
AIC	-237.224		-275.026		-255.721	
Jarque Bera Test for Kalman Filter residuals (p-value)	89.082 (<2.2e-16)		580.695 (<2.2e-16)		332.057 (<2.2e-16)	

** represents significance at 5% level

Table 5.3: Herd Behavior Towards the Market Portfolio (Jan.1998-Dec.2012)

Figure 5.3.1 shows the evolution of the FTSE MIB and of the herding measure $h_{mt} = (1 - e^{H_{mt}})$ calculated with the betas of the FF model using Model 1, 2 and 3. First, we can note that the difference between Models 1 and 2 does not seem to be large. In addition, we can see that the largest value of h_{mt} is far less than one (bounded above roughly by 0.4 and below roughly by -0.8) which indicates that there was never an extreme degree of herd behavior towards the market portfolio during our sample period. We should also note that the confidence intervals only indicate two periods in Model 1 and three periods in Model 2 where herd behavior is significantly different from zero with a 95% confidence interval. In Model 1 these periods are from January 1998 to June 1999 and then from May 2010 to the end of the sample, in Model 2 from January

$$AIC = -2log(\mathcal{L}(\hat{\theta})\,|\,x) + 2K$$

where K is the number of parameters. (Anderson, D.R. (2008). Model based inference in the life sciences - A primer on evidence. Springer, p. 60)

We have not used the second order variant of AIC, $AICc = AIC + \frac{2K(K+1)}{n-K-1}$, as the sample size ($n$) is large and there is not a big difference between them.

1998 to January 2000, between August 2003 and January 2008 and then from February 2010 to the end of the sample period. At the beginning of 2009 herd activity began to decrease and finally adverse herd behavior towards market portfolio took over in June 2009. Differently, in Model 3 h_{mt} is always negative and significantly different from zero with a 95% confidence interval. The first high level or peak in herd behavior can be found around the end of 1998 and the beginning of 1999. herd behavior decreases from 1998 to January 2000, then it is almost steady in the period 2000 - 2004 and during 2007 it starts to decrease again.

We can also observe that the correlation between h_{mt} and the FTSE MIB is around 0.65. This means that herd behavior is likely to increase in bull markets and decrease during periods of bear markets.

Then, we investigate herd behavior towards the market portfolio dividing the sample period into two sub-periods, from January 1998 to December 2005 and from January 2006 to December 2012. The results of Model 1 are presented in the first two columns of Table 5.4, the results of Model 2 and 3 are reported from column 3 to 6. As can be noted, H_{mt} is highly persistent with $\hat{\phi}_m$ large and significant in each sub-period. The proportions of signal in Model 2 and 3 are of a similar order of magnitude and higher in the first period than in the second, indicating that herd behavior explains around 14% of the total variability in $Std_c(\beta^b_{imt})$ in the first period and around 12% in the second interval. Moreover, the estimates of $\sigma_{m\eta}$ (the standard deviation of η_{mt}) are significant at 10% level in the period from January 1998 to December 2005 and at 5% significance level in the second sub-period and thus we can conclude that there is herd behavior towards the market portfolio. Differently, the estimates of $\sigma_{m\eta}$ are not significant in the period from January 2006 to December 2012 in Model 1. Another interesting result is that the correlation of h_{mt} with the FTSE MIB is positive in the second interval, while it is not significantly different from zero in the first period.

In addition, Model 2 is again selected by the AIC value in the first sub-period, while Model 3 is chosen for the second sub-period.

Furthermore, the Jarque Bera normality test shows that the residuals of the Kalman filter are not normal in the three models analysed, with the exception of Model 1 during the first sub-period.

Figure 5.3.2 displays the evolution of the herd behavior measure h_{mt} calculated with the betas of the FF model using Models 1, 2 and 3 in the two sub-periods. We can first note that the difference between the three Models does not seem to be large. In the first period the confidence intervals indicate that herd behavior is never significantly different from zero with a 95% confidence interval, except from February 1998 to January 1999 in

Herd Behavior Towards the Market Portfolio

(I Period: Jan.1998 - Dec.2005, II Period: Jan.2006 - Dec.2012)

	Model 1				Model 2				Model 3			
	I period		II period		I period		II period		I period		II period	
μ	-0.244	**	-0.029		-0.933	**	-0.686	*	-0.914	**	-0.952	**
ϕ	0.908	**	0.969	**	0.920	**	0.988	**	0.923	**	0.987	**
σ_{mv}	0.242	**	0.326	**	0.199	**	0.302	**	0.199	**	0.290	**
$\sigma_{m\eta}$	0.046	*	0.054		0.039	*	0.045	**	0.039	*	0.048	**
$log - Vm$	-		-		-0.113	**	-0.117	**	-0.110	**	-0.159	**
$r_m - r_f$	-		-		30.304	**	20.307		30.974	**	3.151	
SMB	-		-		-		-		10.889		-47.746	**
HML	-		-		-		-		-1.858		15.680	**
$\frac{\sigma_{m\eta}}{sd\left(Log\left(\widehat{Std_c(\hat{\beta}^b{}_{ikt})}\right)\right)}$	0.172		0.141		0.146		0.118		0.146		0.125	
Correlation with the Market Index	-0.017		0.750	**	-0.125		0.613	**	-0.115		0.590	**
AIC	-155.971		-83.604		-188.074		-92.081		-184.779		-93.843	
Jarque Bera Test for Kalman Filter residuals (p-value)	2.868 (0.238)		49.219 (2.05e-11)		9.382 (0.009)		225.612 (<2.2e-16)		9.266 (9.73e-03)		108.191 (<2.2e-16)	

** represents significance at 5% level, * represents significance at 10% level

Table 5.4: Herd Behavior towards the Market Portfolio (Jan.1998 - Dec.2005 and Jan.2006 - Dec.2012)

Model 1. Differently, in the second period we can see that herd behavior is significantly different from zero between March 2007 and March 2009 and from April 2012 to the end of the sample period in Model 1. In Model 2 and 3 herd behavior is significantly different from zero only at the beginning and at the end of the sample period. As can be seen from the graph, adverse herd behavior towards market portfolio characterizes the period from January 2010 to the end of the sample period.

5.3.3 Herd Behavior towards Size Factors

We have also carried out the same analysis in order to investigate herd behavior towards SMB factor. Table 5.5 reports the maximum likelihood estimates for the state space model parameters. As can be seen, the standard deviations of the herding error ($\sigma_{S\eta}$) are significantly non-zero for all of the models, suggesting that there was herd behavior towards SMB in the Italian stock market. Moreover, we can see immediately that H_{St} is highly persistent with $\hat{\phi}_{St}$ high and significant, herd behavior towards SMB factor explains nearly 10% of the total variability in $Std_c(\beta^b_{iSt})$. In addition, as in the case of herd behavior towards the market portfolio, we find that market volatility and the market return level are significant with negative and positive signs, respectively, while the coefficients on SMB and HML are not significant in Model 3. Again, the AIC criterion selects Model 2 as the best model.

Furthermore, the Jarque Bera normality test does not allow us to not reject the null hypothesis of normality of the Kalman filter residuals in the three models analysed.

Then, we plot herd behavior towards SMB, $h_{St} = (1 - e^{H_{St}})$, in Figure 5.3.3. Note that the herd behavior movements towards SMB obtained from Model 1 and 2 are not so different from those implied by Model 3. We can see that the largest value of h_{St} is far less than one (bounded above roughly by 0.4 and below roughly by -0.6) which indicates that there was never an extreme degree of herd behavior towards SMB factor during our sample period. In addition, using a 95% confidence level, we can identify only few interesting periods with h_{St} significantly different from zero. In many cases, these periods are coincident with those of h_{mt} in Figure 5.3.2. These are January 1998 - June 2001, January 2010 - December 2012, and in Model 2 and 3 January 2003 - June 2006. Thus, when there is herd behavior towards the market portfolio we are also likely to observe herd behavior towards SMB and vice versa. In fact, as herd behavior towards the market portofolio, h_{St} is positive correlated with the FTSE MIB.

Now, we investigate herd behavior towards the SMB factor dividing the sample period into two sub-periods, from January 1998 to December 2005 and from January 2006 to December 2012. The results of Model 1 are illustrated in the first two columns

Herd Behavior Towards the SMB Factor

Period: January 1998 - December 2012

	Model 1		Model 2		Model 3	
μ	0.243		-0.186		-0.299	
ϕ	0.990	**	0.991	**	0.992	**
σ_{Sv}	0.276	**	0.259	**	0.264	**
$\sigma_{S\eta}$	0.031	**	0.032	**	0.032	**
$log - Vm$	-		-0.077	**	-0.099	**
$r_m - r_f$	-		17.298	**	0.113	
SMB	-		-		-0.053	
HML	-		-		0.010	
$\frac{\sigma_{S\eta}}{sd\left(Log\left(\widehat{Std_c(\hat{\beta}^b{}_{iSt})}\right)\right)}$	0.094		0.097		0.097	
Correlation with the Market Index	0.679	**	0.731	**	0.745	**
AIC	-255.819		-275.026		-261.763	
Jarque Bera Test for Kalman Filter residuals (p-value)	186.089 (<2.2e-16)		447.105 (<2.2e-16)		341.966 (<2.2e-16)	

** represents significance at 5% level

Table 5.5: Herd Behavior Towards the SMB Factor (Jan.1998-Dec.2012)

of Table 5.6, the results of Model 2 and 3 are reported from columns 3 to 6. As can be noted, H_{St} is highly persistent with $\hat{\phi}_S$ large and significant in each sub-period and $\sigma_{S\eta}$ significantly different from zero, except in Model 1 in the period January 2006 - December 2012. The proportions of signal in Model 2 and 3 are of a similar order of magnitude and higher in the first period than in the second, indicating that herd behavior explains around 13% of the total variability in $Std_c(\beta^b_{iSt})$ in the first period and around 12% in the second interval.

The Jarque Bera normality test suggests that the residuals of the Kalman filter are not normal in the three models analysed.

Figure 5.3.4 shows the evolution of the herding measure h_{St} calculated with the betas of the FF model using Models 1, 2 and 3 in the two sub-periods. We can first note that the difference between the three Models does not seem to be large. In the first period the confidence intervals indicate that herd behavior is significantly different from zero with a 95% confidence interval from June 1998 to May 2000 in Model 1 and in the period September 1998 - January 2000 in Model 2 and 3. Differently, in the second period we can see that herd behavior is never significantly different from zero in Model 1 and between January 2006 and July 2007 and from January 2011 to the end

Herd Behavior Towards the SMB Factor

(I Period: Jan.1998 - Dec.2005, II Period: Jan.2006 - Dec.2012)

	Model 1				Model 2				Model 3			
	I period		II period		I period		II period		I period		II period	
μ	0.104		0.356	**	-0.231		-0.242		-0.246		-0.429	
ϕ	0.969	**	0.802	**	0.947	**	0.981	**	0.945	**	0.982	**
σ_{Sv}	0.228	**	0.300	**	0.215	**	0.299	**	0.214	**	0.293	**
$\sigma_{S\eta}$	0.029	*	0.106		0.033	**	0.038	**	0.034	**	0.041	**
$log - Vm$	-		-		-0.054	*	-0.107	**	-0.058	*	-0.139	**
$r_m - r_f$	-		-		17.660	**	15.879		17.526	**	5.126	
SMB	-		-		-		-		-3.169		-31.756	
HML	-		-		-		-		-9.036		8.392	
$\frac{\sigma_{S\eta}}{sd\left(Log\left(\widehat{Std_c(\hat{\beta^b}_{iSt})}\right)\right)}$	0.113		0.302		0.128		0.108		0.132		0.117	
Correlation with the Market Index	0.529	**	0.680	**	0.533	**	0.620	**	0.519	**	0.593	**
AIC	-169.423		-90.702		-175.718		-96.367		-172.362		-94.891	
Jarque Bera Test for Kalman Filter residuals	33.762		97.030		32.799		249.054		32.492		153.320	
(p-value)	(4.66e-08)		(<2.2e-16)		(7.55e-08)		(<2.2e-16)		(8.8e-08)		(<2.2e-16)	

** represents significance at 5% level, * represents significance at 10% level

Table 5.6: Herd Behavior towards the SMB Factor (Jan.1998 - Dec.2005 and Jan.2006 - Dec.2012)

of the sample period in Model 2 and 3. The period from January 2010 to the end of the sample period is characterized by adverse herd behavior towards SMB factor.

Further, herd behavior towards the SMB factor is correlated with the FTSE MIB with a level around 0.52 in the first period and around 0.60 in the second sub-period.

5.3.4 Herd Behavior towards Value Factors

We have also investigated herd behavior towards HML factor (H_{Ht}). Table 5.7 reveals that there is significant herd behavior in the Italian stock market towards HML. As opposed to the previous two sets of results, $Std_c(\beta^b_{iHt})$ is now not explained by the level of market returns. In addition, H_{Ht} is highly persistent with a proportion of signal around 12% in Model 1 and around 15% in Model 2 and 3.

Moreover, the Jarque Bera normality test do not reject the null hypothesis of normality only for Model 1. Again, Model 2 is selected by the AIC value.

Herd Behavior Towards the HML Factor

Period: January 1998 - December 2012

	Model 1		Model 2		Model 3	
μ	-0.120	*	-0.503	**	-0.468	**
ϕ	0.968	**	0.958	**	0.958	**
σ_{Hv}	0.225	**	0.216	**	0.214	**
$\sigma_{H\eta}$	0.032	**	0.039	**	0.040	**
$log-Vm$	-		-0.068	**	-0.062	**
$r_m - r_f$	-		2.501		5.554	
SMB	-		-		10.369	
HML	-		-		-3.827	
$\frac{\sigma_{H\eta}}{sd\left(Log\left(Std_c(\widehat{\hat{\beta}^b}_{iHt})\right)\right)}$	0.125		0.153		0.157	
Correlation with the Market Index	0.389	**	0.579	**	0.563	**
AIC	-327.970		-333.513		-330.841	
Jarque Bera Test for Kalman Filter residuals (p-value)	4.286 (0.117)		9.562 (0.008)		10.015 (0.007)	

** represents significance at 5% level, * represents significance at 10% level

Table 5.7: Herd Behavior Towards the HML Factor (Jan.1998-Dec.2012)

Figure 5.3.5 reports the evolution of the herding measure towards the HML factor $h_{Ht} = (1 - e^{H_{Ht}})$ in Model 1, 2 and 3. We can see that there was never an extreme degree of herd behavior towards HML factor. In fact, using a 95% confidence level, we can identify only a few interesting periods with h_{Ht} significantly different from zero. These are January 1998 - June 2000, and in Model 2 and 3 also from June 2012 to the end of the sample period. In addition, the correlation of h_{Ht} with the market index is positive, but less pronounced.

Then, we investigate herd behavior towards the HML factor dividing the sample period into two sub-periods, from January 1998 to December 2005 and from January 2006 to December 2012. We can easily note that H_{Ht} is highly persistent with $\hat{\phi}_H$ large and significant and $\sigma_{H\eta}$ significantly different from zero in the first period, while $\hat{\phi}_H$ is not significant in the second sub-period and $\sigma_{H\eta}$ is significantly different from zero only in Model 1. These results suggest that there is evidence of herd behavior towards HML factors only in the first period analysed.

In addition, the Jarque Bera normality test indicates that the residuals of the Kalman filter are normal only in the second sub-period analysed. Further, the AIC value selects Model 1 for the first interval and Model 2 for the second one.

Figure 5.3.6. shows the evolution of herd behavior towards the HML factor when we divide the sample period into two sub-periods. As can be seen, the plot relative to the second period confirms the evidence of no herd behavior.

5.3.5 Relationship between Herd Behavior towards different factors

In this section we investigate the relationships between herding patterns towards market portfolio, SMB and HML. Table 5.9 and 5.10 present the correlation matrices of herd behavior towards different factors obtained from models selected by the AIC criterion. As revealed by the data in Table 5.9, h_{mt} is positive correlated to some degree with both h_{St} and h_{Ht} . An interesting observation is that h_{mt} and h_{St} show almost perfect positive correlation. These results suggest that herd behavior towards the market portfolio is likely to be accompanied by either herd behavior towards SMB or herd behavior towards HML.

When we divide the sample period into two sub-periods we can immediately note that from 2006 to 2012 h_{Ht} is not correlated with herd behavior towards the other factors (in fact there was no evidence of herd behavior towards the HML factor in that period), while in the first sub-period it shows a correlation of nearly 0.90 with h_{St}. These results suggest that herd behavior towards HML, when it arises, is likely to be accompanied by herd behavior towards SMB or herd behavior towards the market portfolio.

Herd Behavior Towards the HML Factor

(I Period: Jan.1998 - Dec.2005, II Period: Jan.2006 - Dec.2012)

	Model 1				Model 2				Model 3			
	I period		II period		I period		II period		I period		II period	
μ	-0.181	*	-0.059	**	-0.497	**	-0.424	**	-0.506	**	-0.423	**
ϕ	0.971	**	0.082		0.955	**	0.104		0.953	**	0.107	
σ_{Hv}	0.230	**	0.001		0.226	**	0.153		0.221	**	0.153	
$\sigma_{H\eta}$	0.034	**	0.226	**	0.039	**	0.154		0.040	**	0.154	
$log-Vm$	-		-		-0.053		-0.065	**	-0.057		-0.065	**
$r_m - r_f$	-		-		-3.333		0.014		-1.779		0.016	
SMB	-		-		-		-		3.981		0.008	
HML	-		-		-		-		-19.326	*	0.018	
$\frac{\sigma_{H\eta}}{sd\left(Log\left(Std_c(\widehat{\hat{\beta}^b}_{iHt})\right)\right)}$	0.125		0.991		0.144		0.675		0.148		0.675	
Correlation with the Market Index	0.683	**	-0.144		0.717	**	-0.192	*	0.696	**	-0.193	*
AIC	-165.989		-157.903		-164.220		-164.501		-163.342		-156.508	
Jarque Bera Test for Kalman Filter residuals	8.878		0.115		14.471		1.441		14.341		1.434	
(p-value)	(0.012)		(0.944)		(7.20e-04)		(0.486)		(7.69e-04)		(0.488)	

** represents significance at 5% level, * represents significance at 10% level

Table 5.8: Herd Behavior towards the HML Factor (Jan.1998 - Dec.2005 and Jan.2006 - Dec.2012)

	Herding towards Market Portfolio		Herding towards SMB factor		Herding towards HML factor
Herding towards Market Portfolio	1				
Herding towards SMB factor	0.931	**	1		
Herding towards HML factor	0.453	**	0.599	**	1

** represents significance at 5% level

Table 5.9: Correlation between herd behavior towards different factors

I period: January 1998 - December 2005

	Herding towards Market Portfolio		Herding towards SMB factor		Herding towards HML factor
Herding towards Market Portfolio	1				
Herding towards SMB factor	0.426	**	1		
Herding towards HML factor	0.308	**	0.887	**	1

II period: January 2006 - December 2012

	Herding towards Market Portfolio		Herding towards SMB factor		Herding towards HML factor
Herding towards Market Portfolio	1				
Herding towards SMB factor	0.985	**	1		
Herding towards HML factor	0.120		0.085		1

** represents significance at 5% level

Table 5.10: Correlation between herd behavior towards different factors (I period: 1998 - 2005, II period: 2006 - 2012)

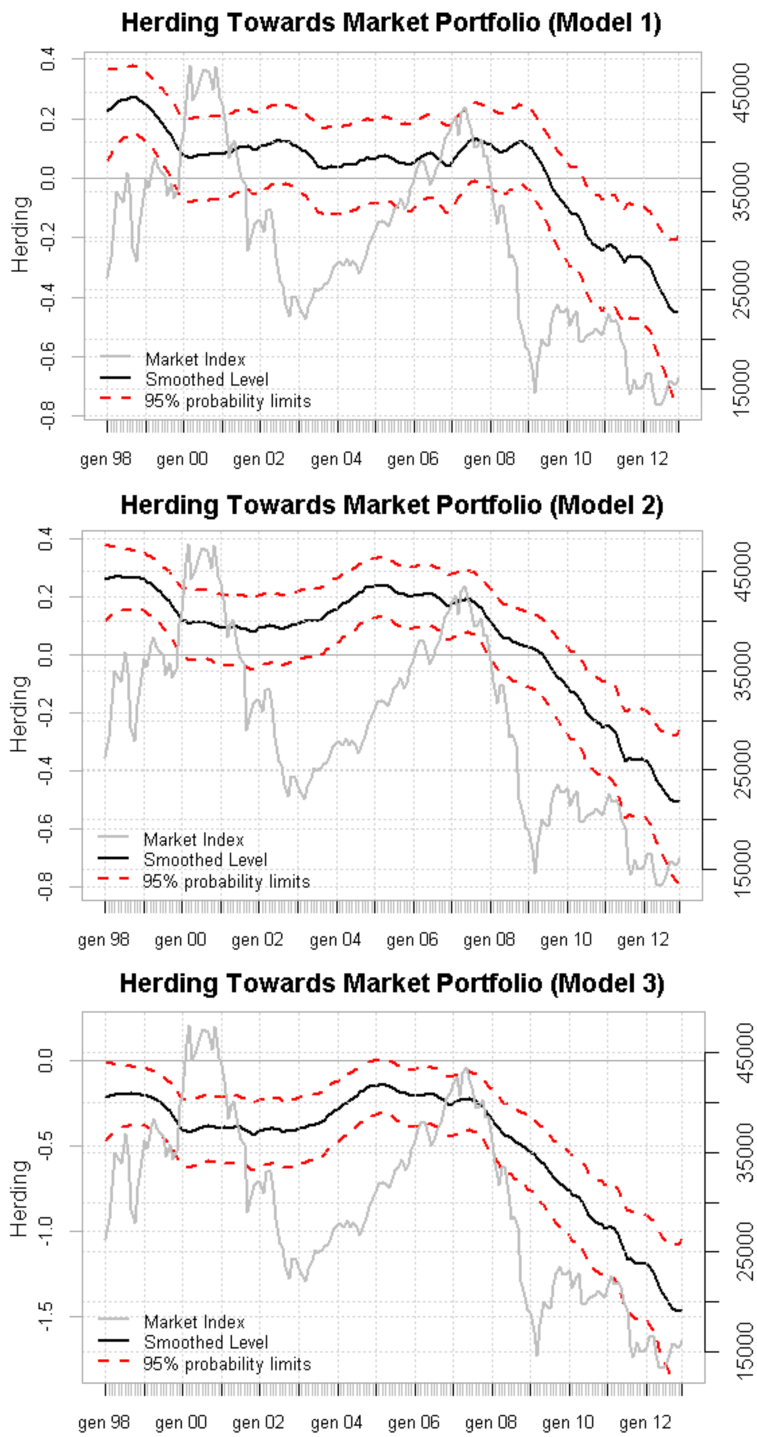

Figure 5.3.1: Herd Behavior Towards the Market Portfolio (Jan.1998 - Dec.2012)

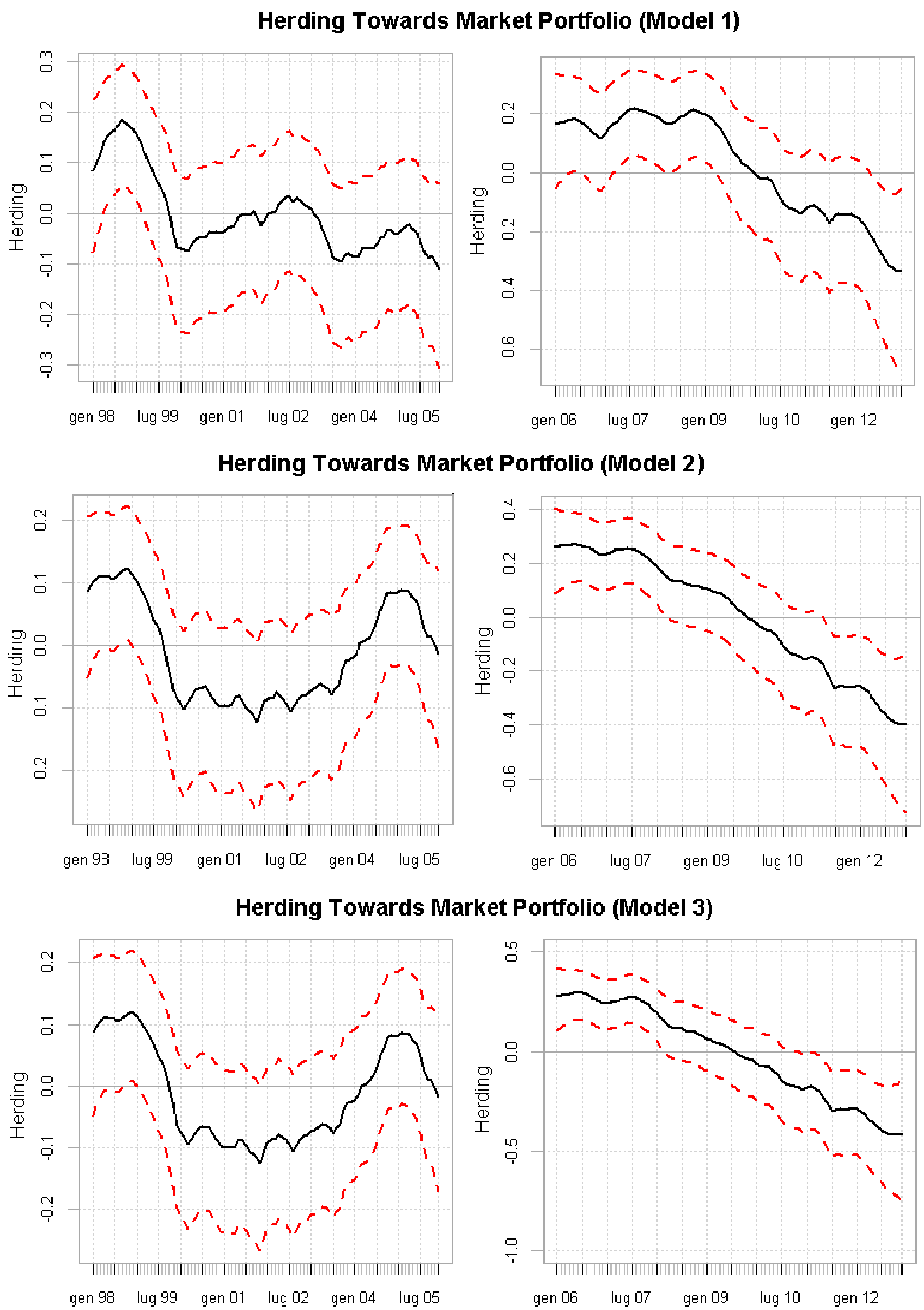

Figure 5.3.2: Herd Behavior towards the Market Portfolio (Jan.1998 - Dec.2005 and Jan.2006 - Dec.2012)

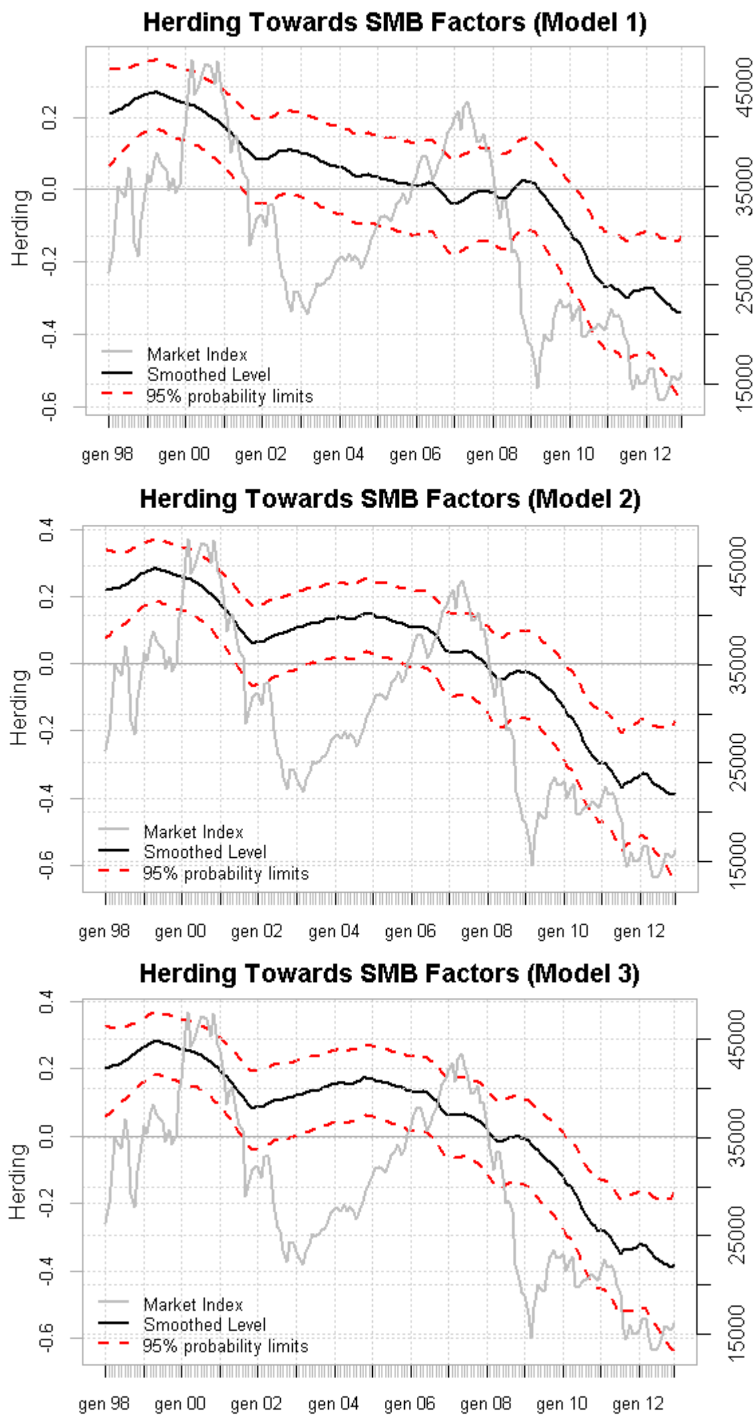

Figure 5.3.3: Herd Behavior Towards the SMB Factor (Jan.1998-Dec.2012)

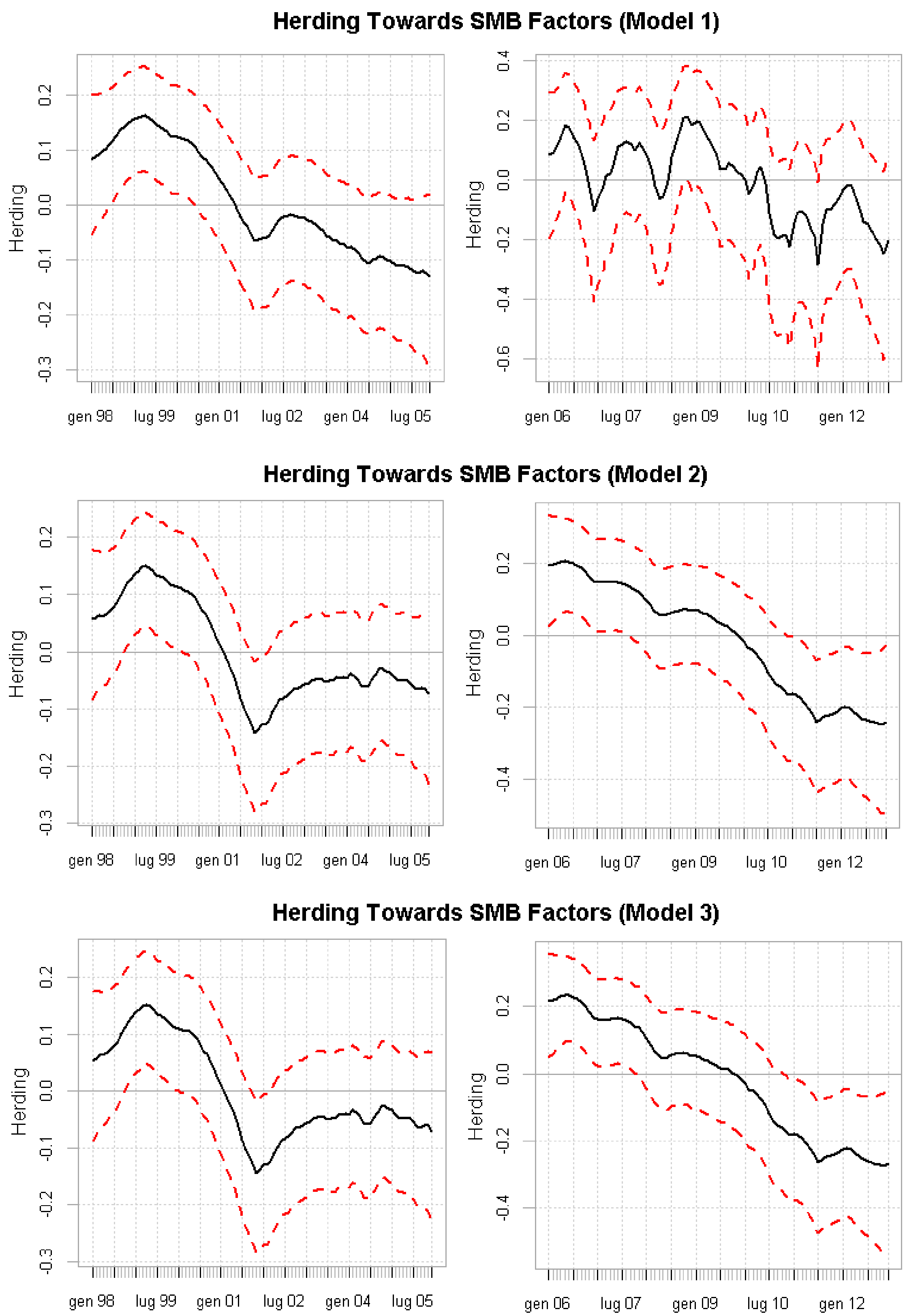

Figure 5.3.4: Herd Behavior towards the SMB Factor (Jan.1998 - Dec.2005 and Jan.2006 - Dec.2012)

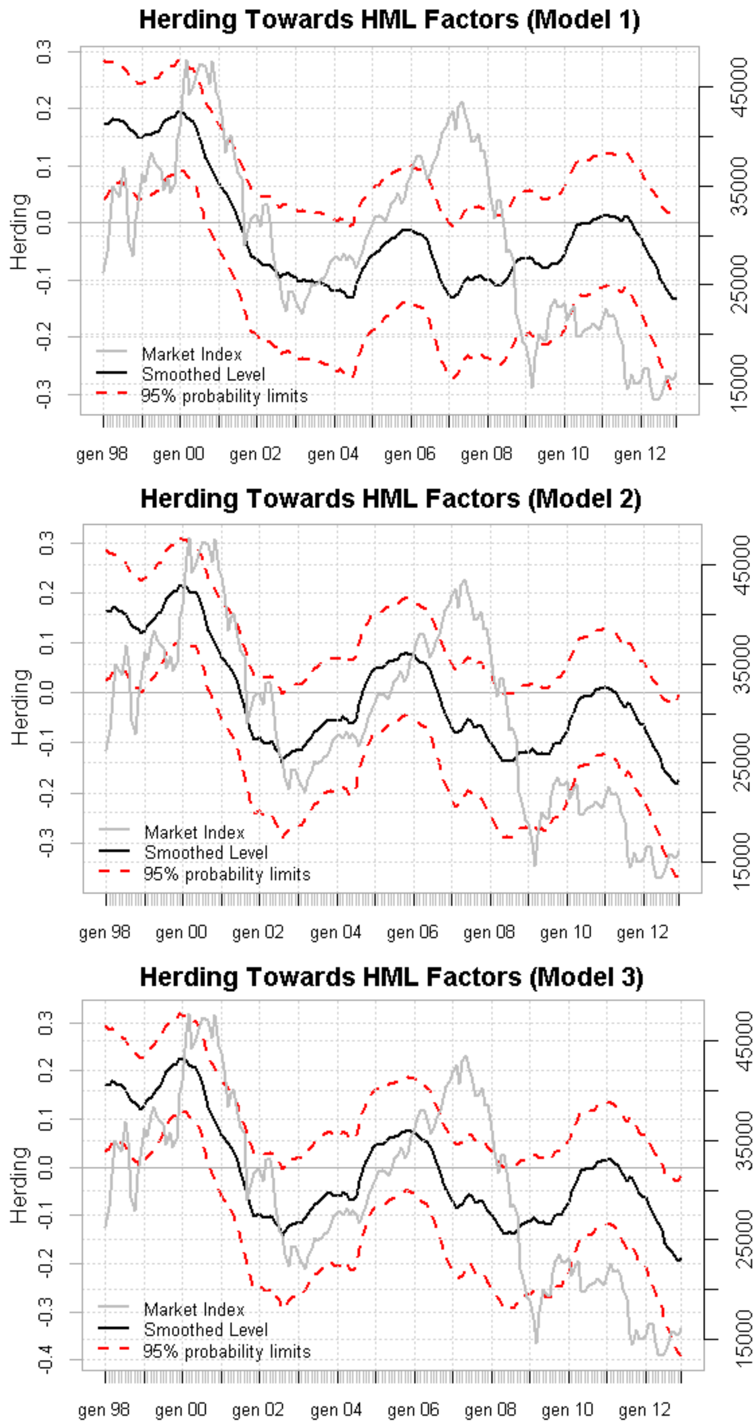

Figure 5.3.5: Herd Behavior Towards the HML Factor (Jan.1998-Dec.2012)

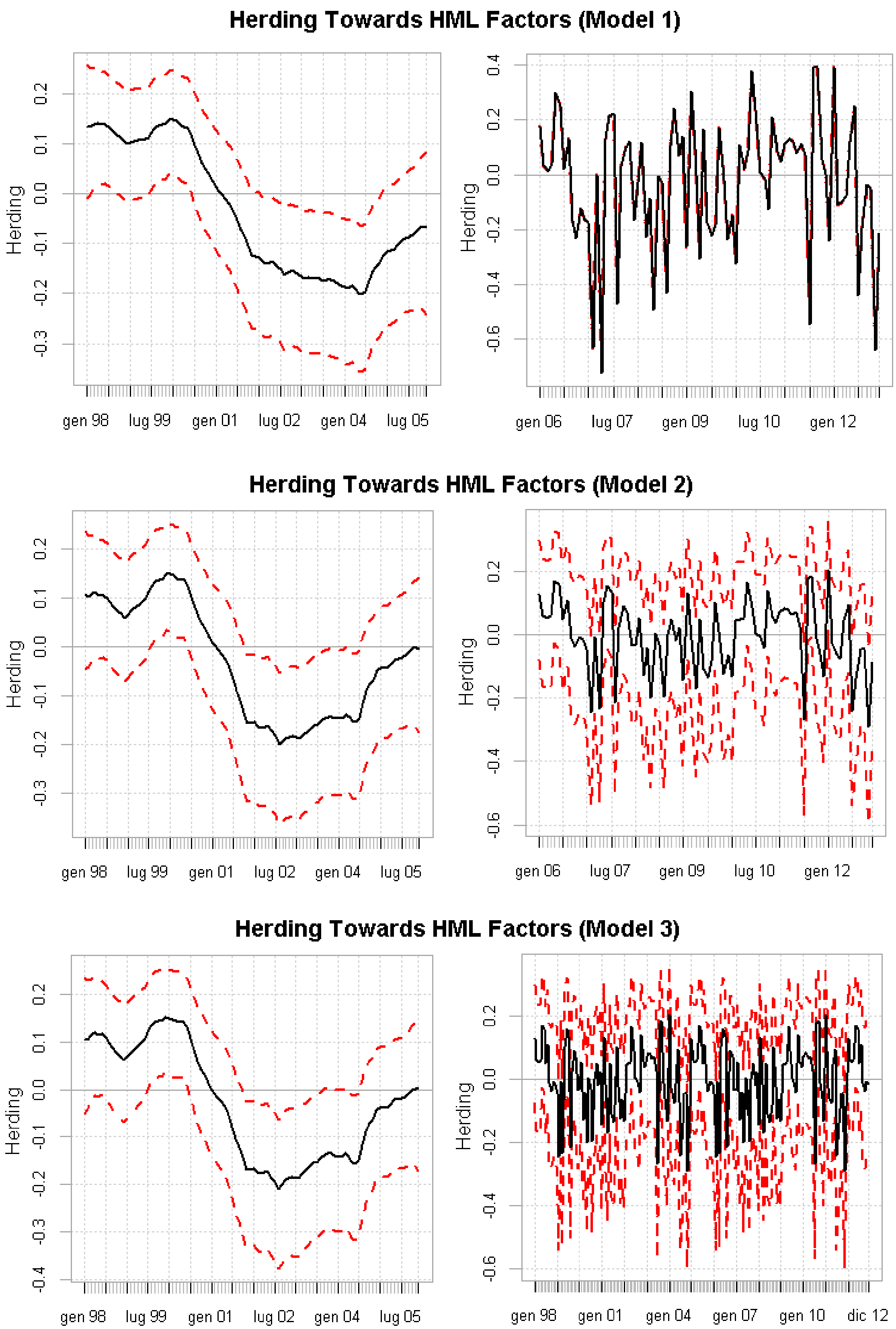

Figure 5.3.6: Herd Behavior towards the HML Factor (Jan.1998 - Dec.2005 and Jan.2006 - Dec.2012)

Chapter 6

Conclusion

In the recent years there has been many theoretical and empirical contributions into the causes and effects of herd behavior in financial markets.

The theoretical papers generally offer models that provide an explanation of why agents can rationally choose to imitate the actions of others. We discussed the papers by Bikhchandani et al. (1992) and Banerjee (1992). They showed that the herd behavior of consumers and investors is rational under the condition that other decision makers have more or more accurate information. In addition, they demonstrated that herd behavior can lead to an inefficient equilibrium. Zhiyong et al. (2010) showed that herd behavior does not occur when informed traders and market makers have the same ambiguity aversion (there is ambiguity when the outcome of an act is uncertain and the probability distribution of the possible events is unknown). Then, Avery and Zemsky (1998) proved that herd behavior arises only in the presence of two dimensions of uncertainty. Other authors, like Kirman (1993) and Lux (1995), dealt with the determinant of the behavior of traders who do not have access to information about fundamental values.

The empirical papers generally investigate whether too many agents appear to take the same action. The thesis has analysed some of the available empirical measures of herding by individuals or small group of investors and measures of market wide herding. Within the former, we analysed the Lakonishok et al. (1992) criterion which is based on trades conducted by a subset of market participants over a period of time, and the *portfolio-change measure* (PCM) proposed by Wermers (1995), this measure captures both the direction and intensity of trading by investors. Christie and Huang (1995) discussed a method for measuring market-wide herding which investigates the magnitude of *cross-sectional standard deviation* (CSSD) of individual stock returns during large price changes. Later, Chang et al. (2000) proposed an alternative measure of dispersion based on the *cross-sectional absolute deviation* (CSAD). Recently, a paper by Cipriani and Guarino (2014) contributed to the reduction of the gap between the theoretical and empirical literature that exists in this area. Indeed, they proposed an empirical test of a theoretical model of herd behavior.

In our empirical analysis we applied the approach to measuring and testing herd behavior proposed by Hwang and Salmon (2004) to investigate herd behavior in the Italian Stock Exchange for the period January 1998 - December 2012. We decided to adopt this measure because of its empirical and theoretical properties in the sense that this measure automatically conditions on fundamentals and can also measure herd behavior towards other factors, such as the market excess return, SMB and HML factors. In addition, the influence of time series volatility is accounted automatically by this measure.

We found that herd behavior towards the market portfolio was significant and persistent independently from and given the particular state of the market as expressed in market volatility and return. We have also found that herd behavior towards the market portfolio was positive correlated with the FTSE MIB in the period from January 2006 to December 2012, which means that herd behavior is likely to increase when the market is rising. Perhaps more importantly, given that herd behavior can lead to significant mispricing, it is interesting to note that in the Italian stock market there were only three periods in the sample when herd behavior was statistically significant and it was never greater than the 40% of its maximum potential value. We have also examined herd behavior towards SMB and HML factors and found evidence of significant periods of herd behavior towards SMB factors and evidence of herd behavior towards HML only in the period January 1998 - December 2005. Further, we have found that herd behavior towards different factors shows a positive and high correlation coefficient, which means that for example herd behavior towards the market portfolio is likely to be accompanied by either herd behavior towards SMB or herd behavior towards HML.

Limitations of the current study are that we have estimated the monthly time series of the betas with daily return over monthly intervals, a period that is too short to reduce the influence of unusual good or bad events of the company on the betas. Further, in our analysis we have found that the majority of the Kalman filter residuals were not normally distributed. However, it is important to highlight that even when the error terms are not normally distributed if we restrict the attention to linear estimator the Kalman filter minimizes the mean square error of the estimate[1].

Finally, as a future direction for research, it would be worthwhile to use a robust

[1] Hamilton, J. D. (1994). State-space models. In R. F. Engle and D. L. McFadden (Eds.), Handbook of econometrics (Vol. IV, Chap. 50). Elsevier Science B.V., p. 3053.

regression approach to calculate the betas and to use a non-Gaussian and nonlinear state space model to investigate herd behavior in financial markets.

Chapter 7

Appendix A

7.1 Kalman Filter

Filtering aims to update the knowledge about the system as each observation y_t comes in. On the other hand, smoothing allow to base the estimation of quantities of interest on the entire sample $y_1, ..., y_n$.

Consider the linear Gaussian state space model

$$y_t = Z_t\alpha_t + \epsilon_t \; \epsilon_t \sim N(0,\, H_t), \tag{7.1.1}$$

$$\alpha_{t+1} = T_t\alpha_t + R_t\eta_t \; \eta_t \sim N(0,\, Q_t),\, \alpha_1 \sim N(a_1,\, P_1), \tag{7.1.2}$$

where y_t is a $p \times 1$ vector of observations and α_t is an unobserved $m \times 1$ vector. Equation 7.1.1 is called the observation equation, whereas equation 7.1.2 is called the state equation. The matrices Z_t, T_t, R_t, H_t and Q_t are assumed to be known and the error terms ϵ_t and η_t are assumed to be serially independent and independent of each other at all time points. Moreover, the initial state vector α_1 is assumed to be independent of $\epsilon_1, ..., \epsilon_n$ and $\eta_1, ..., \eta_n$[1].

The objective of the Kalman filter is to obtain the conditional distribution of α_{t+1} given Y_t for $t = 1, ..., n$ where $Y_t = \{y_1, ..., y_n\}$. Since all distributions are assumed to be normal, the conditional distribution of α_{t+1} given Y_t will be normal.

It is assumed that α_t given Y_{t-1} is $N(a_t, P_t)$. Now a_{t+1} and P_{t+1} are calculated recursively from a_t and P_t.

$$a_{t+1} = E[T_t\alpha_t + R_t\eta_t \,|\, Y_t] = T_tE[\alpha_t \,|\, Y_t], \tag{7.1.3}$$

[1]Durbin, J. and Koopman, S.J. (2001). Time series analysis by state space methods. Oxford University Press, p. 38.

$$P_{t+1} = Var(T_t\alpha_t + R_t\eta_t \,|\, Y_t) = T_t Var(\alpha_t \,|\, Y_t) T_t' + R_t Q_t R_t', \tag{7.1.4}$$

Define the one-step forecast error of y_t given Y_{t-1}, $v_t = y_t - E[y_t \,|\, Y_{t-1}] = y_t - E[Z_t\alpha_t + \epsilon_t \,|\, Y_{t-1}] = y_t - Z_t a_t$.

$$E[\alpha_t \,|\, Y_t] = E[\alpha_t \,|\, Y_{t-1},\, v_t] = E[\alpha_t \,|\, Y_{t-1}] + Cov(\alpha_t,\, v_t)[Var(v_t)]^{-1} v_t =$$

$$= a_t + M_t F_t^{-1} v_t, \tag{7.1.5}$$

where

$$M_t = Cov(\alpha_t,\, v_t) = E[(\alpha_t - a_t)(Z_t\alpha_t + \epsilon_t - Z_t a_t)' \,|\, Y_{t-1}]$$

$$= E[(\alpha_t - a_t)(\alpha_t - a_t)' Z_t' \,|\, Y_{t-1}] = P_t Z_t'$$

and

$$F_t = Var(v_t) = Var(Z_t\alpha_t + \epsilon_t - Z_t a_t) = Z_t P_t Z_t' + H_t$$

F_t it is assumed non singular. Substituting (7.1.5) in (7.1.3) gives

$$a_{t+1} = T_t(a_t + M_t F_t^{-1} v_t) = T_t a_t + K_t v_t, \tag{7.1.6}$$

with $K_t = T_t M_t F_t^{-1} = T_t P_t Z_t' F_t^{-1}$.

$$Var(\alpha_t \,|\, Y_t) = Var(\alpha_t \,|\, Y_{t-1},\, v_t) = Var(\alpha_t \,|\, Y_{t-1}) - Cov(\alpha_t,\, v_t)[Var(v_t)]^{-1} Cov(\alpha_t,\, v_t)' =$$

$$= P_t - M_t F_t^{-1} M_t' = P_t - P_t Z_t' F_t^{-1} Z_t P_t. \tag{7.1.7}$$

substituting in (7.1.4) gives

$$P_{t+1} = T_t(P_t - P_t Z_t' F_t^{-1} Z_t P_t) T_t' + R_t Q_t R_t' = T_t P_t L_t' + R_t Q_t R_t', \tag{7.1.8}$$

with $L_t = T_t - K_t Z_t$.

The recursion (7.1.8) and (7.1.6) constitute the Kalman filter for the model[2].

When the observations are univariate, the standardized innovations, defined by $\bar{v}_t = v_t/\sqrt{F_t}$, are a Gaussian white noise. Hence, if the model is correct, the sequence $\bar{v}_1, ..., \bar{v}_t$ computed from the data should look like a sample of size t from a standard

[2] *Ibid.*, pp. 64-67.

normal distribution[3]. However, it is important to note that when the error terms are not normally distributed and we restrict the attention to estimates which are linear in the y_t's the Kalman filter minimizes the mean square error of the estimate of each component of α_{t+1}[4].

It was assumed that the matrices Z_t, T_t, R_t, H_t, Q_t and a_1, P_1 were known, however in practical works the model usually depends on unknown parameters. These parameters can be estimated by maximum likelihood. For the linear Gaussian model the likelihood function can be calculated by a routine application of the Kalman filter[5].

Here, the log-likelihood function

$$log\mathcal{L}(y) = -\frac{np}{2}log2\pi - \frac{1}{2}\sum_{t=1}^{n}\left(log|F_t| + v_t'F_t^{-1}v_t\right).$$

7.1.1 State smoothing

Smoothing enables us to estimate α_t given the entire sample $y_1, ..., y_n$. The vector y is used to denote the stacked vector $(y_1', ..., y_n')'$, note that y is fixed when Y_{t-1} and $v_t, ..., v_n$ are fixed. Therefore we have

$$\hat{\alpha}_t = E[\alpha_t \,|\, y] = E[\alpha_t \,|\, Y_{t-1}, v_t, ..., v_n]$$

$$= a_t + \sum_{j=t}^{n} Cov(\alpha_t,\, v_j)F_j^{-1}v_j, \tag{7.1.9}$$

for $t = 1, ..., n$, with $Cov(\alpha_t,\, v_j) = Cov(\alpha_t,\, v_j') = E[\alpha_t(Z_j(\alpha_j - a_j) + \epsilon_j)'] = E[\alpha_t(\alpha_j - a_j)']Z_j',\ j = t, ..., n.$

Moreover,

$$E[\alpha_t(\alpha_t - a_t)'] = E\left\{E[\alpha_t(\alpha_t - a_t)' \,|\, y]\right\} = P_t$$

$$E[\alpha_t(\alpha_{t+1} - a_{t+1})'] = E\left\{E[\alpha_t(T_t\alpha_t + R_t\eta_t - T_ta_t - K_tv_t)' \,|\, y]\right\} =$$

$$= E\left\{E[\alpha_t(T_t\alpha_t + R_t\eta_t - T_ta_t - K_tZ_t(\alpha_t - a_t) - K_t\epsilon_t)' \,|\, y]\right\} =$$

$$= E\left\{E[\alpha_t(L_t(\alpha_t - a_t) + R_t\eta_t - K_t\epsilon_t)' \,|\, y]\right\} = P_tL_t'$$

[3]Petris, G., Petrone S., and Campagnoli, P. (2007). Dynamic linear model with R. Springer.
[4]Hamilton, J. D. (1994). *op.cit.*, p. 3053.
[5]Durbin, J. and Koopman, S.J. (2001). *op.cit.*, p. 138.

$$\vdots$$

$$E[\alpha_t(\alpha_n - a_n)'] = P_t L_t' ... L_{n-1}'.$$

Substituting in (7.1.9) gives

$$\hat{\alpha}_n = a_n + P_n Z_n' F_n^{-1} v_n,$$

$$\hat{\alpha}_{n-1} = a_{n-1} + P_{n-1} Z_{n-1}' F_{n-1}^{-1} v_{n-1} + P_{n-1} L_n' Z_n' F_n^{-1} v_n,$$

$$\hat{\alpha}_t = a_t + P_t Z_t' F_t^{-1} v_t + P_t L_t' Z_{t+1}' F_{t+1}^{-1} v_{t+1} + ... + P_t L_t' ... L_{n-1}' Z_n' F_n^{-1} v_n,$$

The smoothed state vector can be expressed as follows

$$\hat{\alpha}_t = a_t + P_t r_{t-1} \tag{7.1.10}$$

where $r_{t-1} = Z_t' F_t^{-1} v_t + L_t' Z_{t+1}' F_{t+1}^{-1} v_{t+1} + ... + L_t' L_{t+1}' ... L_{n-1}' Z_n' F_n^{-1} v_n$[6].

[6]Durbin, J. and Koopman, S.J. (2001). pp. 70-71.

Bibliography

[1] Anderson, D.R. (2008). Model based inference in the life sciences - A primer on evidence. Springer.

[2] Avery, C. and Zemsky, P. (1998). Multidimensional uncertainty and herd behavior in financial markets. *The American Economic Review*, 88(4), 724-48.

[3] Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics,* 107(3), 797-817.

[4] Bernoulli, D. Originally published in 1738; translated by Dr. Louise Sommer. (1954). Exposition of a new theory on the measurement of risk. *Econometrica,* 22, 22-36.

[5] Bikhchandani, S., Hirshleifer, D. and Welch, I. (1992). A theory of fads, fashion, custom, and cultural changes as informational cascades. *Journal of Political Economy*, 100(5), 992-1026.

[6] Bikhchandani, S. and Sharma, S. (2000). Herd behavior in financial markets: A review. IMF Working Paper.

[7] Boyer, B., Kumagai, T. and Yuan, K. (2006). How do crises spread? Evidence from accessible and inaccessible stock indices. *Journal of Finance*, 61, 957-1003.

[8] Chang, E. C., Cheng, J.W., Khorana, A. (2000). An examination of herd behavior in equity markets: An international perspective. *Journal of Banking and Finance*, 24, 1651-1679.

[9] Chiang, T. C., Jeon, B. N., and Li, H. (2007). Dynamic correlation analysis of financial contagion: Evidence from Asian markets. *Journal of International Money and Finance,* 26(7), 1206-1228.

[10] Chiang, T. C. and Zheng, D. (2010). An empirical analysis of herd behavior in global stock markets. *Journal of Banking and Finance*, 34(8), 1911-1921.

[11] Christie, W. G., and Huang, R. D. (1995). Following the pied piper: Do individual returns herd around the market?. *Financial Analysts Journal*, 31-37.

[12] Cipriani, M., and Guarino, A. (2014). Estimating a structural model of herd behavior in financial markets. *American Economic Review*, 104(1), 224-251.

[13] Corsetti, G., Pericoli, M., and Sbracia, M. (2005). Some contagion, some interdependence: more pitfalls in tests of financial contagion. *Journal of International Money and Finance*, 24, 1177-1199.

[14] DeBondt, W. F. M. and Thaler, R. (1985). Does the stock market overreact?. *Journal of Finance*, 40, 793-808.

[15] Demirer, R. and Kutan, A.M. (2006). Does herding behavior exist in Chinese stock markets? *Journal of International Financial Markets, Institutions and Money*, 16(2), 123-142.

[16] Durbin, J. and Koopman, S.J. (2001). Time series analysis by state space methods. Oxford University Press.

[17] Fama, E.F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2), 383-417.

[18] Fama, E.F. and French K.R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33, pp. 3-56.

[19] Forbes, K.J. and Rigobon, R. (2002). No contagion, only interdependence: Measuring stock market comovements. *Journal of Finance*, 57(5), 2223-61.

[20] Glosten, L. R., and Milgrom, P.R. (1985). Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1), 71-100.

[21] Hamilton, J. D. (1994). State-space models. In R. F. Engle and D. L. McFadden (Eds.), Handbook of econometrics (Vol. IV, Chap. 50). Elsevier Science B.V.

[22] Hwang, S. and Salmon, M. (2004). Market stress and herding. *Journal of Empirical Finance*, 11, 585-616.

[23] Kahneman, D. (2011). Thinking, fast and slow. Penguin Books.

[24] Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263-91.

[25] Keynes, J. M. (1964). The general theory of employment, interest and money. New York: Harcourt, Brace and World.

[26] Kirman, A. (1993). Ants, rationality, and recruitment. *The Quarterly Journal of Economics*, 137-56.

[27] Lakonishok, J., Shleifer, A. and Vishny, R. W. (1992). The impact of institutional trading on stock prices. *Journal of Financial Economics*, 32, 23-43.

[28] Lux, T. (1995). Herd behavior, bubbles and crashes. *The Economic Journal*, 105, 881-96.

[29] Neumann, J. and Morgenstern, O. (1953). Theory of games and economic behavior. Princeton, NJ. Princeton University Press.

[30] Petris, G., Petrone S., and Campagnoli, P. (2007). Dynamic linear model with R. Springer.

[31] Saastamoinen, J. (2008). Quantile regression analysis of dispersion of stock returns - evidence of herding?. Discussion Papers, No. 57. University of Joensuu, Economics.

[32] Schwert, G.W. (1989). Why does stock market volatility change over time?. *Journal of Finance*, 44, 1115-1153.

[33] Shiller, R. J. (2005). Irrational exuberance (2nd ed.). Broadway Books.

[34] Shiller, R. J. (1995). Conversation, Information and Herd Behavior. *Rhetoric and Economic Behavior*, 85(2), 181-185.

[35] Shiller, R. J. (1984). Stock prices and social dynamics. *Brookings Papers on Economic Activity*, 457- 98.

[36] Thaler, R. H. and Sunstein, C. R. (2008). NUDGE: Improving decisions about health, wealth, and happiness. Yale University Press.

[37] Varian, H. R. (1992). Microeconomic analysis (3rd ed.). Norton & Company.

[38] Wermers, R. (1995). Herding, trade reversals, and cascading by institutional investors. mimeo, University of Colorado, Boulder.

[39] West, K. D. (1988). Bubbles, fads, and stock price volatility tests: a partial evaluation. *Journal of Finance*, 43, 639-60.

[40] Wilkinson, N. and Klaes, M. (2012). An introduction to behavioral economics (2nd ed.). Palgrave Macmillan.

[41] Zhiyong Dong, Qingyang Gu, and Xu Han. (2010). Ambiguity aversion and rational herd behaviour. *Applied Financial Economics*, 20, 331-43.