ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics

Herding toward Market and Momentum Returns:

A Consideration on herding toward market's predictive power

[International Bachelor Economics and Business economics]

Bachelor Thesis

Student Name/Number	SANG OOK NAM/432520
Supervisor	Esad Smajlbegovic
Second assessor	Mao, Q

Date of Final Version: 04-07-2017

Summary

This paper tries to come up with the answer to the main question that whether there is a significant relationship between herding toward market, one type of herding, and momentum returns. Regarding this, the paper gives the question an answer based on the Hong and Stein (1999), Hwang and Salmon (2004), and the thesis' assumption that herding toward market reduces the market-wide private information diffusion. The proposed answer is that higher herding toward market during the condition period will yield higher future momentum returns after holding period, but it is not empirically supported by the data provided although the sign of herding is consistently positive as implied. The statistically insignificant relationship may stem from an implausible assumption or the measurement method.

Key Words: Momentum - Herding toward market - Relationship between herding and momentum - Bounded rational herding - Herding not due to information of fundamentals

Table of contents

1. Introduction

The simple trading rule based on buying previous winner stocks and selling previous loser stocks, referred to momentum strategy, has yielded significant excess returns not well justified by existing asset pricing models or systematic risk. Since Jagadeesh and Titman (1993) figured out that relative strength rules which selects stocks based on past 3 to 12 months' price movements realize abnormal returns over next 3 to 12 month periods, numerous further researches have been done to examine the existence of momentum profits. Jagadeesh and Titman (1993) says, over the 1965 to 1989 period, a momentum portfolio based on past 6-month returns earns 12.01% excess annualized returns if it is held for 6 months. In addition, Jagadeesh and Titman (1993) find that the momentum strategy which constructs the portfolio based on past 12 months' price movements and lasts for 3 months afterward is most successful. However, the momentum returns have a long-run reversal within 36 months, which means that the part of short-run momentum returns disappear within a long-run period after the formation, e.g. 36-months (Jagadeesh and Titman (1993)). Afterwards, several researches have been published for substantiation of momentum, explanation of momentum, and prediction for momentum returns. Among them, a group of researches on momentum is about the relationship with herding.

Herding can be defined in various way. One way to define herding is given as the statement that investors herd around a subject, which cannot be justified for reasons (Hwang and Salmon (2004)). The subject can be a strategy, an informed traders' behavior, or an index. The herding without rational reasons cannot be directly observable since the research needs to deliberately discriminate the herding without reasons from the herding with reasons. Regarding this, Hoitash and Krishnan (2008) find that return patterns generated by herding is consistent with momentum patterns.

This paper tries to find whether there is a relationship between herding toward market due to bounded rationality and momentum returns in aggregate level. The herding toward market means the investor in the stock market focuses so much on market-wide information (e.g. the direction of market or the market index returns) that the irrational investor causes the values of equities to deviate from fundamentals (Hwang and Salmon (2004)). The one thing that this paper is different from other papers is that the thesis will focus on the relationship between market-wide herding and momentum portfolio returns in market level over time while other papers' implications are based on a cross-sectional analysis. This is because if the market-wide herding measure has a significant relationship with momentum returns over time, the results can be used for explaining momentum returns or managing a risk of momentum portfolio. In that sense, herding toward market is appropriate in that herding toward market can be measured to give a degree of market-wide herding by Hwang and Salmon (2017).

Therefore, the research question is

What is the relationship between market toward herding and momentum returns?

Before the question is addressed, the previous momentum researches are reviewed in section 2.1. Section 2.2 answers the main question in the theoretical framework, and Section 2.3 generates the (alternative) hypothesis that posits the positive relationship based on Hwang and Salmon (2004) and Hong and Stein (1999) with the assumption of the thesis. For the formal empirical test of hypothesis, the methodology of measuring herding toward market by Hwang and Salmon (2017) is summarized in section 3.1, and the values of measures through actual data are presented in section 3.2. The results from models with different control variables are represented in section 4 and discussed in section 5. The results based on the period of 1989 Dec to 2016 Dec. indicate that the coefficients are consistently positive but insignificant. Section 6 concludes the thesis.

2. Theoretical Framework

2.1 Overall literature review on momentum researches

Although data from Jagadeesh and Titman (1993) was confined to the United States NYSE and AMEX stocks during the 1965 to 1989 period, various follow-up studies have confirmed the existence and effectiveness of momentum strategy. Momentum is a globally pervasive phenomenon. Rouwenhorst (1998) uses a sample of 12 European countries over the period of 1978 to 1995 to find the global existence of momentum return patterns. The diversified relative strength portfolio over 12 countries yielded returns ranging from 0.51% to 1.45% per month, which is consistent with the finding from Jagadeesh and Titman (1993). The momentum profit was not derived from only certain countries. Instead, momentum strategy was effective in all countries' stock markets except Sweden (0.36%, not significant) and earned around returns ranging from 0.64% to 1.32% per month return in 11 countries. Momentum was effective across all size deciles even though it was more effective in smaller size deciles. Moreover, Rouwenhorst (1998) give a significant implication that momentum profits in European countries are correlated with those in the United States.

Regarding this, Jagadeesh and Titman (2001) examines the momentum effectiveness once more using the United States data over the period of 1990's. Still, it earned 1.39% return per month, and the profits are derived from winner portfolio and loser portfolio equally. Griffin, Ji, and Martin (2003) documents momentum profits from 40 countries and confirms that momentum profits are pervasive across all continents except Asia (0.32, not significant). Furthermore, Griffin, Ji, and Martin (2003) find that there are low interregional or intraregional correlations in momentum profits over 40 countries' stock markets during the period of 1926 to 2000 for the U.S. and the period of 1975 to 1995 for the other countries except Egypt, as opposed to the statement that there are high correlations across momentum profits according to Rouwenhorst (1998).

Asness and Moskowitz (2013) extends the scope of momentum existence from stock markets to other asset classes. Asness and Moskowitz (2013) documents that momentum strategy works in all asset classes as well as across all countries except Japan. In addition, momentum profits have co-movements with each other within asset classes and across countries. This evidence is consistent with Rouwenhorst (1998) while it contradicts Griffin et al (2003). Finally, according to Asness and Moskowitz (2013), momentum has negative correlation with value performance within asset classes and across countries, which makes momentum strategy more profitable by combining it with value strategy.

Despite the existence of momentum, the source of profits is not fully discovered yet. Initially, Jagadeesh and Titman (1993) have examined whether the profits of momentum are due to common risk factors. Regarding this, Jagadeesh and Titman (1993) denies attribution of momentum profits to market risk or systematic risk. Instead, the profitability is consistent with the market underreaction to the firm specific information(Jagadeesh and Titman (1993)). Also, according to Rouwenhorst (1998), common momentum returns in many countries are not explained by conventional risk factors, and the correlations among European countries and the United States momentum profits suggests that the international common factor may drive the profit. Regarding this, Asness and Moskowitz (2013) proposes the funding liquidity as a partial source of momentum. However, the strategy of combining momentum and value strategy eliminates the exposure to funding liquidity and yields higher Sharpe ratio at the same time, which makes the phenomenon harder to be explained.

Instead, Barberis et al (1998) posits that there is only one type of investors who have conservatism bias to suggest the consistent underreaction, momentum patterns, based on recent good news. Rather, Hong and Stein (1999) assumes heterogenous investors, "news watchers" and "momentum traders". News watchers use only the privately-owned part of information about future fundamentals to forecast the performance of stocks, and those private information is diffused to other newswatchers gradually. This leads to an initial underreaction. In contrast, momentum traders only condition on past price movements, and they accelerate the consistent price movement. According to Hong and Stein (1999), those two different participants holding their own strategy make eventual overreaction to any news and cause the following price reversals.

Lastly, Chordia and Shivakumar (2002) gives a risk explanation view that the momentum returns can be attributed to lagged macroeconomic variables such as dividend yield, default spread, yield on three-month T-bills, and term structure spread using the United States data. That is, Chordia and Shivakumar (2002) says momentum returns are explained by *"difference in conditionally expected returns explained by high macroeconomics risk"*. Regarding this, however, Griffin et al (2003) examines this model in international level and states that *"momentum portfolio still performs well significantly even after model consideration".* Therefore, Griffin et al (2003) refutes the argument of Chordia and Shivakumar (2002). In addition, Griffin et al (2003) documents that momentum strategy generally works well in all different economic states classified by GDP growth or stock market performance.

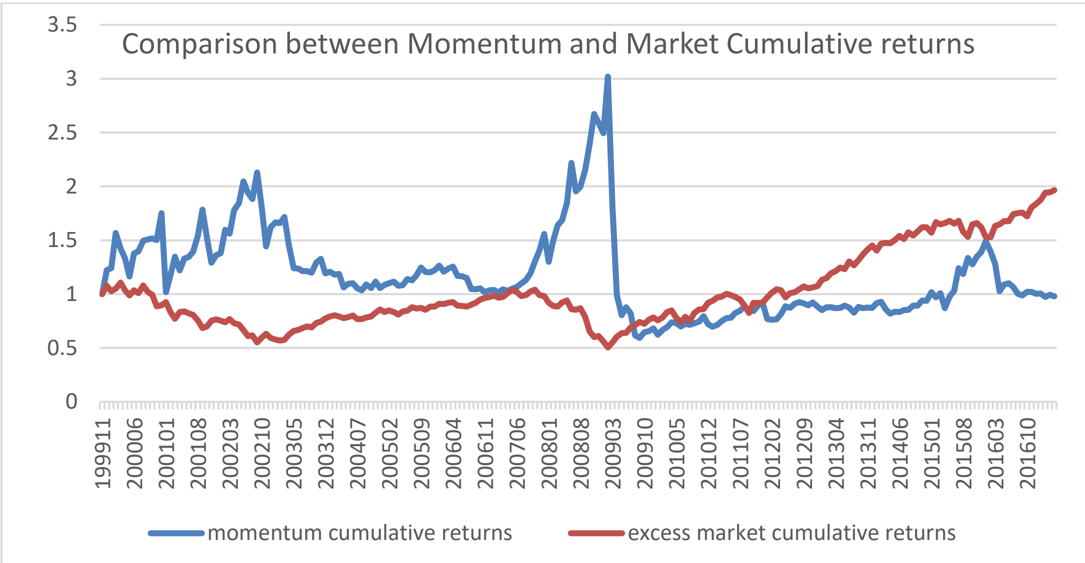
Momentum strategy cannot be fully dissipated by existing risk models, but yields stable abnormal returns. However, despite huge alphas, there is the major drawback of momentum strategy, large crashes as indicated in figure 1. During the period of the end of 1999 to the February of 2009, the momentum strategy earns almost 200% profits while the market earns -38.89% profits. Clearly, momentum on average outperforms market portfolio for a long time. However, those 9 years cumulative profits suddenly disappear altogether within next two months.[1]

Therefore, much efforts have been put on managing or predicting the momentum crashes as well as momentum returns themselves. Barroso and Santa-Clara (2015) claims that risk-managed portfolio could be obtained through using autocorrelation of realized variance of daily returns. By scaling momentum returns by a parameter (targeted volatility/average realized volatility during previous 6 months), the strategy not only yields higher Sharpe ratio, but also reduces momentum crash shocks. This strategy makes excess kurtosis smaller and skewness less negative, which means the outlier in the left tail, equivalent to disastrous negative return from momentum crashes, is reduced to a smaller size.

---

[1] Momentum returns and excess market returns can be found at Kenneth French's Website, 10 Portfolios Formed on Momentum and Fama/French 3 Factors, respectively: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html. Momentum portfolio is constructed by buying winner stocks and selling loser stocks. Firstly, stocks are sorted by past performance during the period of t -12 to t-2. Winner stocks includes stocks in the top decile, and loser stocks consist of stocks in the bottom decile. Holding period of a momentum portfolio is 1-month, and the portfolio is consisting of NYSE, AMEX, and NASDAQ stocks.

Figure 1



Daniel and Moskowitz (2016) states that momentum crashes do not occur immediately after market crashes. Rather, Daniel and Moskowitz (2016) predicts that crashes occur in times of contemporaneous positive market returns following bear markets and high ex ante volatility. Daniel and Moskowitz (2016) suggests that this could be partly caused by dynamic beta changes. That is, when a momentum portfolio is constructed in a bear market, the winner stocks tend to be low beta stocks while the loser stocks are more likely to be high beta stocks. Then, the momentum portfolio will have a negative market exposure in total. When market suddenly rebounds, momentum crash occurs. It is supported by the empirical evidence that 14 out of worst 15 momentum returns occurs with contemporaneous market rebound following bear market. Daniel and Moskowitz (2016) also founds that the loser stocks betas during the period of market rebound after bear markets are much larger than betas during the period of bear markets. As evidence, Daniel and Moskowitz (2016) shows that Winer-Minus-Loser(WML) portfolio's beta estimate in stable bear market is -0.742 while it is -1.796 during the rebound. Therefore, in a bear market, Momentum portfolio has similar payoffs with shorting a call option(Daniel and Moskowitz (2016)). Gains in stable bear market are frequent but small while loses in crashes are infrequent but extremely high.

Although Daniel and Moskowitz (2016) partly explains and predicts momentum crashes in the bear market state, this reasoning cannot apply to the bull market state. To elaborate on this, when market collapsed during the market being bullish, Momentum portfolio does not show any option-like payoffs, or momentum crash does not occur immediately.

Many literatures have addressed the relationship between herding and stock return patterns. The herding can be a source of explaining and predicting individual firms' momentum if the stock with high herding measure will yield a short-run momentum pattern and a long-run reversal. Regarding this, Nofsinger & Sias (1999) finds that the degree of herding defined as investors commonly buying and selling the same stock has a positive correlation with stock returns. As evidence, the stocks investors herd into yield the return of 18.38% while the stocks investors leave exhibit -13.12% over the herding period (Nofsinger & Sias (1999)). However, it is not entirely consistent with the momentum patterns in that there is no reversal afterwards. Nofsinger & Sias (1999) mention that it cannot clarify the causal direction between herding and stock returns. This is because the degree of herding is also affected by past returns patterns, called positive feedback trading (Nofsinger & Sias (1999)). These findings are consistent with the Wermers (1999). Wermers (1999) also found that the stocks the institutional investor herd to by purchases yield higher returns compared to stocks they sell, but there is no subsequent reversal. According to Wermers (1999), it means that the herding is on a rational basis and correct the price into the equilibrium price.

However, the above analyses do not discriminate the correlated behaviors due to rational basis (e.g. an information on fundamentals) from the groundless behaviors due to sentiments or bounded rationality. There may exist a relationship between herding and momentum if the herding is restricted to common behaviors due to bounded rationality. Hoitash and Krishnan (2008) find the positive cross-sectional relationship between a momentum returns and a lagged herding measure called "SPEC". SPEC is derived from the error term of regressions where the autocorrelation coefficient of trading volumes is regressed on variables indicating the rational motives to yield herding. It is because the error term is a not-explained part by rational motives and means the "speculative intensity or herding" (Hoitash and Krishnan (2008)). Hoitash and Krishnan (2008) also documented that the stock returns of higher lagged SPEC firms will exhibit larger reversal after, and the documented phenomenon of short-run momentum and long-run reversal is consistent with the momentum patterns.

2.2 Herding toward market and momentum returns: Derivation of a hypothesis

This paper tries to answer the main question that whether there is a relationship between momentum returns and herding toward market in aggregate model. According to Hwang and Salmon (2004), "herding arises when investors decide to imitate the observed decisions of others or movements in the market rather than follow their own beliefs and information", and there are numerous cases where people follow what others around us do or focus on others rather than reflect what the information itself is. Perhaps, the most famous story about herding would be "beauty contest" by Keynes (1936). Keynes (1936) states that *"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* (1936)*).*

The phenomenon that people herd toward something is not consistent with the assumption that classical finance theory takes on. In theory, asset pricing models frequently assume that investors make choices based on their future belief or information about the firm. This will make stock prices reflect all the information. It is called "fundamental trading". Even if the investors have some biased judgements, stock prices will converge to the equilibrium price in aggregate as long as the mean of judgement is equal to the equilibrium price (Miller (1977). However, when people herd to something, they no longer add their belief or interpretation to the price. Instead, they just follow what the first mover decided.

The implication of herding is that the herding behavior will cause the price to deviate from its fundamental value. Imagine the Banerjee (1992) 's example that there are two restaurants to be selected by 100 people. 99 of 100 will receive the signal that the restaurant B is better than the restaurant A while only the first chooser consider the A better. What would be a result if people herd? According to Banerjee (1992), A would be chosen over B because the second chooser will choose the A following the first person as opposed to his/her individual judgement. The third one will surely follow A since previous two people already have chosen A. Then, subsequently all individuals will choose A. This is not a desired decision in terms of social welfare (Banerjee (1992)). Likewise, suppression of information due to herding will cause the price not to reflect fundamental value (Hwang and Salmon (2004))

However, there are numerous types of herding in stock markets. Among them, this paper will focus on herding toward market. Herding toward market means that *"the investors herd around the consensus of all market participants as reflected in the market index" (Hwang and Salmon* (2004)*)*. This type of herding is plausible. Sometimes, we are focused on New York Stock Exchange Composite, S&P 500, or DOW-JONES INDUSTRIALS 30 STOCK indices and if one's stock wins the market, he is worried that my stock might be overrated while if the stock loses, he is irritated that his stock is treated unfairly. People do not add or look for their individual stock's information and belief about future fundamental. They will be based on biased information which mostly relates to market consensus. Therefore, the investor will argue that his stocks should yield returns close to market portfolio returns in times of market-wide herding toward market. Herding toward market affects private information diffusion in the market.

Higher herding toward market constrain the future private information to be reflected in the price. In addition, the one of the implications by Hong and Stein (1999) is that slower future private information diffusion will positively affect the momentum returns via underreaction caused by “newswatchers” and overreaction due to “momentum traders”, which is supported in Hong and Stein (2000). Therefore, if the two implications from herding toward market by Hwang and Salmon (2004) and Hong and Stein (1999) are combined, the conclusion can be reached that herding toward market affects the privates’ information diffusion, which determines the newswatchers’ and momentum traders’ trading patterns so that herding toward market eventually determines momentum returns. To understand the implication how the information diffusion affects momentum returns by Hong and Stein (1999), the three assumptions by Hong and Stein (1999) and their interaction must be understood.

In the market consisting of only two types of investors, the first assumption is that there is newswatcher. A newswatcher only predict a stock price only based on a small fraction of future information he has. The restriction is that he only has a small fraction of future information. Then, if a future event information has reached to investors, they cannot determine prices which reflect all the future information (Hong and Stein (1999)). The second restriction is that a newswatcher never conditions on past price changes. They do not reflect past price information and predict the impact on price the other players might cause by exploiting past price changes.

The second assumption by Hong and Stein (1999) is that momentum traders are the other type of investors and only condition on past price changes. According to Hong and Stein (1999), the restriction is that a momentum trader never conditions on fundamentals consisting of the private information. The second restriction is that the momentum trader should employ simple forecasting analysis such as univariate auto regressive model. As a result, Hong and Stein (1999) states that they only follow the price trend by determining the amount of stocks to be invested. Lastly, the final assumption is that each future information a newswatcher hold is diffused gradually so that the complete future information will be finally revealed to market.

Hong and Stein (1999) posits a stock market of a firm is only full of newswatchers at first. The price by only newswatchers cannot reflect all the available future information because they predict the price using a small part of future information, and the future information is not fully revealed yet but spreading to other news watchers slowly. It leads to a phenomenon where a consistent underreaction exists in the short-term relative to the equilibrium price adjustment in fully efficient market. However, momentum traders enter because a consistent underreaction attracts momentum traders who condition on past price movements (Hong and Stein (1999)). Subsequent momentum trader’s trend chasing behaviors will accelerate the underreacted price to the equilibrium price. Therefore, momentum traders in early stages will gain a profit. In other words, differences between the equilibrium price and the underreacted price will be earned by momentum traders (Hong and Stein (1999)).

However, momentum traders will eventually make the overreaction of price to future information (Hong and Stein (1999)). This is because they do not care about information itself, so they can not presume the exact equilibrium prices. Therefore, the consistently "accelerated price" by momentum traders will eventually exceed the equilibrium price (Hong and Stein (1999)). Then, Hong and Stein (1999) concludes that the price reversal should exist to arrive at the equilibrium price. After all, momentum traders who enter the market later will lose.

Until now, this paper summarizes how the not fully informed future information causes the underreaction and the overreaction of the stock price. Now, the implication that the slower information diffusion yields higher momentum by Hong and Stein (1999) is explained below. According to (1), the price by newswatchers is determined as

$$P_t = D_t + \{(z-1)e_{t+1} + (z-2)e_{t+2} + \ + e_{t+z-1}\}/z - \Theta S_t^2 \qquad (1)$$

where e_j is future information about a firm at time j and D stands for a liquidation value at the specified time (Hong and Stein (1999)). Hong and Stein (1999) assumes that only up to t+z-1 information will be publicized at time t. However, as stated above, the individual news watchers have only a fraction of information. For example, this can be expressed as e_{t+z-1}/z according to Hong and Stein (1999). This fraction of information is diffused very slowly, which means an additional fraction of information flows into others only after a period. Therefore, at t+1, an individual will have $2*e_{t+z-1}/z$. The full information for t+z-1 will be complete at time t+z-1. Here, z in Hong and Stein (1999) means how slowly the information is spreading. If z = 1, then perfect information is given, but no future information is signaled in advance. On the contrary, information is given in small fraction and diffused very slowly although more distance information is given if z is large.

The slower the speed of information diffusion is (the larger the z is), the higher autocorrelation the time series of a stock price will exhibit. This is because the price equations will share more common terms as in (1) (Hong and Stein (1999)). Then, momentum traders follow the price trend by putting larger orders because of the momentum traders' behavior assumption, which results in higher acceleration until the larger overreaction. This will be very profitable for momentum traders.

---

[2] According to Hong and Stein (1999) , "Θ means a function of risk aversion and variances of future information and S_t means available stocks for news watchers". The exact sentences and explanations can be found in pages from 2149 to 2151 in "Hong, H., & Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. The Journal of finance, 54(6), 2143-2184".

Proposed mechanism how the herding toward market explains momentum returns

ΔHerding toward market → Δ pace of information diffusions → Newswatchers prices the stock differently → Δthe degree of short run autocorrelation → ΔMomentum traders' orders -> more pronounced momentum

Diagram 1

In conclusion, this paper can answer the main question by combining Hwang and Salmon (2004) and Hong and Stein (1999). The herding toward market will positively explain the momentum returns. The herding toward market makes investors firstly focused on only the part of information related to the market, which delay information diffusion. Following Hong and Stein (1999), the price by newswatchers is more underreacted relative to the states of less herding. More autocorrelated price patterns will attract more momentum traders' orders, which accelerates the momentum.

2.3 The formal hypothesis

However, the answer is not complete, but in need of empirical supports if the answer wants to have an explanatory power because the thesis assumes an assumption that herding toward market will sufficiently restrict the information flows of individual firms, and the Hong and Stein (1999)'s implication will be valid in aggregate level. Perhaps, the assumption of identifying the herding toward market consensus with less private information diffusion can be inappropriate. The logical deduction of the conclusion from two arguments with the assumption can be false, so it needs to be tested even though the two basis arguments by Hong and Stein (1999) and Hwang and Salmon (2004) are recognized as a good explanation or being documented well. Therefore, this paper generates the hypothesis and aims to test it for the empirical supports. **The hypothesis is "herding toward market positively predicts momentum returns."**

$$\mathbf{R_{wml,t} = a + bHTM_{t-2} + Control\ Variables\ where\ b > 0\ is\ proposed}$$

The catch here is that the herding affects the future momentum returns. It is because it may need some time for new price patterns to be established by newswatchers even if the herding toward market is assumed to immediately changes information diffusion speed (refer to the diagram1). On top of this, momentum traders condition on past price changes (Hong and Stein (1999)). Momentum traders will change their behavior even after the price pattern changes. In sum, contemporaneous change in herding will yield the future momentum returns considering the theoretical model where information diffusion goes to momentum returns through underreaction by newswatchers and overreaction by momentum traders.

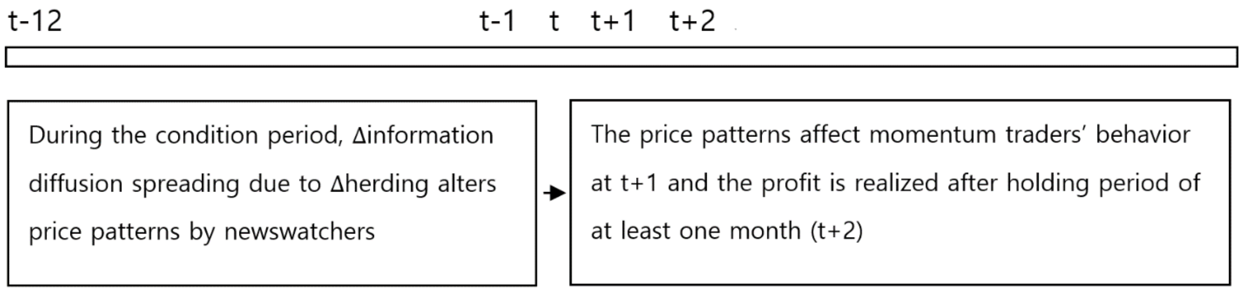


Diagram 2

For example, see the diagram 2 where the reduced information diffusion due to an increase in herding during the momentum condition period causes a series of prices to be more correlated (Hong and Stein (1999)). Early momentum traders notice this, so they will put larger purchase orders from the next period. Subsequently, future momentum traders will also carry the underreaction more until the higher overshooting. Finally, those early momentum traders will realize the higher momentum returns after the holding period of at least 1 month.

3. Methodology

3.1 Methodological framework - How to measure the herding toward market

Hwang and Salmon (2017) invents a herding measure using cross-sectional dispersion of individual sensitivities to market portfolio and *considers the change of the dispersion over time as the change in the degree of herding over time.* If the herding toward market increases, the dispersion of sensitivities will decrease. This is because investors only look for a market related information of holding stocks and insist their holding stocks should yield returns like market returns in times of herding toward market according to Hwang and Salmon (2004). The dispersion will be enlarged in times of adverse herding for the opposite reason.

The method is superior to the method by Lakonishok, Shleifer, and Vishny (1992) or the method of individual returns dispersion by Christie and Huang (1995) according to Hwang and Salmon (2001). First, the most famous method by Lakonishok, Shleifer, and Vishny (1992) does not exactly measure the herding toward market although numerous following researchers such as Grinblatt. Titman, and Wermers (1995) and Zhou and Lai (2009) replicated this method. Lakonishok, Shleifer, and Vishny (1992) assumes half of investors buy a stock while the other half of investors sell the stock in the state of no herding and checks that whether there is herding among institutional investors by calculating how disproportionately money managers buy/sell each stock every quarter across 769 tax-exempted equity portfolios (LSV method). In sum, the LSV method measures how similar portfolio selection strategies are across institutional investors. The method does not accurately measure herding toward market although it is useful in measuring common behaviors (e.g. another type of herding) among institutional investors. Also, there is a fundamental problem in using LSV method in that this method can be used in firm level analysis, but not in aggregate level. That is, the LSV method can detect which firms that institutional investors herd to, but it cannot detect the period when market-wide herding is high. In addition, Hwang and Salmon (2001) states that the method has two drawbacks: lack of data and lack of representativeness. To replicate the LSV method necessitates huge amount of records regarding transaction and its position, which are not easily obtained. Also, analysis is focused on the institutional sector since private investors' records are more difficult to obtain. The results from insufficient and biased composition cannot be said to be a good proxy for measuring market-wide herding (Hwang and Salmon (2001)).

According to Hwang and Salmon (2001), the alternative method for herding toward market is to measure the dispersion of returns given by Christie and Huang (1995). It is also inferior in terms of accuracy although the intuition is similar to Hwang and Salmon (2017). It is because the method of Christie and Huang (1995) cannot discriminate the degree of herding toward market from the rational but correlated behaviors due to the change in information about the risk factors (Hwang and Salmon (2001)). It is important to discriminate the irrational herding from rationally correlated behaviors because rationally correlated behaviors will not restrict the information diffusion speed. For a better understanding, assume that the return generating process is under Capital Asset Pricing Model(CAPM). Then, according to Hwang and Salmon (2001)

$$R_{it} = a_{it} + B^{\text{Herding-biased}}{}_{imt} R_{mt} + e_{it} \quad (2)$$

Cross-sectional Variance of returns and market betas of individual assets will be as follows.

$$Var(R_{it}) = Var(a_{it}) + Var(B^{\text{Herding-biased}}{}_{imt})R^2{}_{mt} + Var(e_{it}) \quad (3)$$

As seen in the equation, the herding measure as a dispersion of returns can vary due to change of the market portfolio returns whilst there is no change in herding attitude. It is rational that the variance of individual returns changes because a systematic factor returns change.

Also, if multiple factor models are assumed, the interpretation of a cross-sectional dispersion of returns can be more problematic. It is because the change in the return dispersion can stem from various sources other than herding. It can be due to change in factor mimicking portfolio returns or even covariance between systematic risk factors. According to Hwang and Salmon (2001),

$$R_{it} = a_{it} + B^{\text{Herding-biased}}{}_{imt} R_{mt} + \Sigma B^{\text{Herding-biased}}{}_{ikt} F_{kt} + e_{it} \quad (4)$$

$$Var(R_{it}) = Var(a_{it}) + Var(B^{Herding\text{-}biased}{}_{imt})R^2{}_{mt} + \Sigma\, Var(B^{Herding\text{-}biased}{}_{ikt})\, F^2{}_{kt}\, Var(e_{it}) + \text{Covariance terms} + Var(e_{it})\ ^{3} \quad (5)$$

In sum, A herding measure by Christie and Huang (1995) as a return dispersion cannot discriminate the variation due to herding from other variations while the measure as a beta dispersion by Hwang and Salmon (2017) can. Therefore, the measure in Hwang and Salmon (2017) is most appropriate for the research considering accuracy of measurement, data availability, and market participants representativeness.

Remaining is how to define a dispersion of factor sensitivities of individual assets. *Hwang and Salmon* (2017) *defines the herding measure as the variance of standardized betas.* It is because using cross sectional variance of standardized betas has three advantages. Firstly, it is easier to replicate. Secondly, standardization considers relative significance of individual betas. It is important because there is a possibility that some insignificantly large betas can drive the measurement values (Hwang and Salmon (2001)).

Lastly, Hwang and Salmon (2017) put emphasis on standardized beta usage in that it eliminates the undesirable change of herding measure because of the heteroskedasticity characteristics of cross sectional variance of estimation errors. According to Hwang and Salmon (2017), if the measure is the variance of betas,

$$\text{Variance of estimated betas} = \frac{1}{n-1}\Sigma\, \{B\wedge^{Herding\text{-}biased}{}_{it} - \frac{1}{n}\Sigma\, (B\wedge^{Herding\text{-}biased}{}_{it})\}\wedge 2 \quad (6)$$

Since the estimates can be divided into true value and corresponding estimation error, which are independent of each other,

$$B\wedge^{Herding\text{-}biased}{}_{it} = B^{Herding\text{-}biased}{}_{it} + (x'x)^{-1}x'e_{\text{appropriate element}}$$

Then,

$$\text{Variance of estimated betas} = \frac{1}{n-1}\Sigma\, \{B^{Herding\text{-}biased}{}_{it} - \frac{1}{n}\Sigma\, (B\wedge^{Herding\text{-}biased}{}_{it})\}\wedge 2 + \frac{1}{n-1}\Sigma\ \text{square of estimation error} \quad (7)$$

Variance of estimated betas = Variance of truly biased betas + ∈ (cross sectional estimation error variance)

The equation (7) means that the herding toward market estimate by Hwang and Salmon (2017) measures the real degree of herding but there is an error ∈. The ∈ is a cross-sectional variance of estimation errors of individual constituent betas. The problem is that the values of the measure will be more likely to be driven by the error ∈ over time if the variance of error ∈ is

---

[3] Hwang and Salmon (2001) ignored covariance terms because in asset pricing models, those risks are independent and systematic. However, risk factors exhibit significant correlations in many empirical literatures. In addition, even in my data, risk factors have significant correlation, see table 1 in the section 3.2

getting larger over time. This can happen because there is no guarantee that errors of individual constituent betas are time series homoscedastic. Then, the derived composite $\in$ can be also time series heteroskedastic. Therfore, the homoscedastic behavior of the error would be better for minimizing the effect of the error $\in$ (Hwang and Salmon (2017)). This can be done by the standardization of betas. Then Hwang and Salmon (2017) states that the herding measure as a composition of individuals betas is now free from the concerns due to heteroskedasticity because all the standardized betas will exhibit same distribution (t- distribution and homoscedasticity). The method eliminates the concern regarding variance of estimation error $\in$.[4]

However, the paper will use both the measures as the dispersion of betas and the dispersion of standardized betas. It is because this paper fears that the standardization may distort the measurement values but improve the method at the same time. The intuition of the measure is that the dispersion of market sensitivities captures the herding phenomenon because herding toward market will result in returns herding around market returns, so the corresponding betas will also herd around 1. However, the standardization might deviate from the intuition since the standardization makes the method determined not only by the betas, but also by the standard errors of betas. Although Hwang and Salmon (2001) states that the standardization improves the measurement because it considers the relative significances of betas, the concern from Hwang and Salmon (2001) with insignificantly large betas can be addressed by using the results from full samples only. Rather, the counter example also exists that, for example, the beta of a firm is around 1, but it can be extremely highlighted by standardization because of the minimal standard error. Then, the small beta around 1 can also drive the measurement value. Meanwhile, However, the standardization still makes the estimation error of the measure homoscedastic over time. In sum, the standardization might distort the intuition where the measure is derived but reduces the unwanted characteristic of the measure. Therefore, although the Hwang and Salmon (2017) put the importance on the usage of standardization, the paper will report the empirical analysis based on the measure as a dispersion of betas in appendix D.

3.2 Data and Implementation

This paper includes all the United States monthly stock returns with primary listings on the NYSE, American Stock Exchange and NASDAQ from Center for Research in Security Prices (CRSP) over the period of January 1985 to December 2016. The estimation window is 60 months. Since the first 60 months are estimated for December 1989, this paper has 325 monthly herding

[4] Detailed explanation and measures Chi-square distribution is given in Overconfidence, Sentiment, and Beta herding by Hwang and Salmon (2017), page 13-14

measures from December 1989 to December 2016. Penny stocks whose price are below $5 are trimmed. For risk factor proxies, monthly data of market portfolio less risk free, size mimicking portfolio (Small minus Big), and value mimicking portfolio (High minus Low) values are from Kenneth French's library. Finally, momentum returns (Winner Minus Loser portfolio returns) and risk-free returns are also from Kenneth French's library.

In theory, risk factors are independent and systematic. Therefore, if only market betas are needed, CAPM model will suffice. However, the market returns in the data has a significant correlation with some other risk mimicking portfolio values such as Size.

Table 1

Pairwise correlation among returns of risk mimicking portfolios

Pair-wise correlation	Market	Size	Value
Market	1.0000		
Size	0.2869*	1.0000	
Value	-0.0271	-0.1440*	1.0000

* significant at 5%

Pairwise correlation among returns of risk mimicking portfolios and Momentum portfolio

Pair-wise correlation	Market	Size	Value	Momentum
Market	1.0000			
Size	0.2869*	1.0000		
Value	-0.0271	-0.1440*	1.0000	
Momentum	-0.2765*	-0.1123*	-0.5990*	1.0000

* significant at 5%

Therefore, even if only herding toward the market is wanted, regressions should include other risk factors as control variables. Otherwise, Hwang and Salmon (2017) states that the herding measure will be biased because the constituent individual betas are biased due to omitted risk factors. The regression method is as follows. Betas of all individual equities at t are estimated using rolling estimation with estimation window of 60 previous months (including month t) as in Hwang and Salmon (2017). In this thesis, Carhart four-factor model as market, size, value, and momentum will be used following Hwang and Salmon (2017). In addition, standard error of each beta is calculated based on Newey-West standard error.[5] The catch here is that I will only use estimated beta with full observations. In this way, the threshold below which the results are discarded does have to be determined. Also, since only 60 months are used for estimation, there is a suspicion that the results from number of observation below 60 can be reliable. Still, there are sufficient results at each time point even after this strict restriction because this

[5] When estimating autocorrelation heteroskedasticity robust standard error, Stata necessitates maximum lag of residuals to be calculated. In this regression, maximum lag is 1.

sample uses all the stocks in the United States. There are 1,540,000 valid results with full number of observation over 325 periods, which indicates that there are on average 4738 data in each t. Following Hwang and Salmon (2017), the extreme top 1% and bottom 1% of betas/standardized betas in each time point are trimmed.

3.3 Transformation and Interpretation of Herding Results

The measures of herding toward market from estimated betas in CAPM is also provided in the appendix A for comparison although the major measures of herding toward market are from estimated betas in Carhart four factor model following Hwang and Salmon (2017).[6] There is one difference from the Hwang and Salmon (2017) that this paper will subtract the dispersion from a constant. The reason is that Hwang and Salmon (2017) states that the higher the dispersion is, the lower the degree of the herding would be, and this paper is worried that the inverse relationship might give a confusion for the interpretation. Therefore, to give a clear interpretation and avoid any distortion on the measures at the same time, this paper implements a linear transformation: cross-sectional variance of standardized beta V into 20 – V(H).[7]

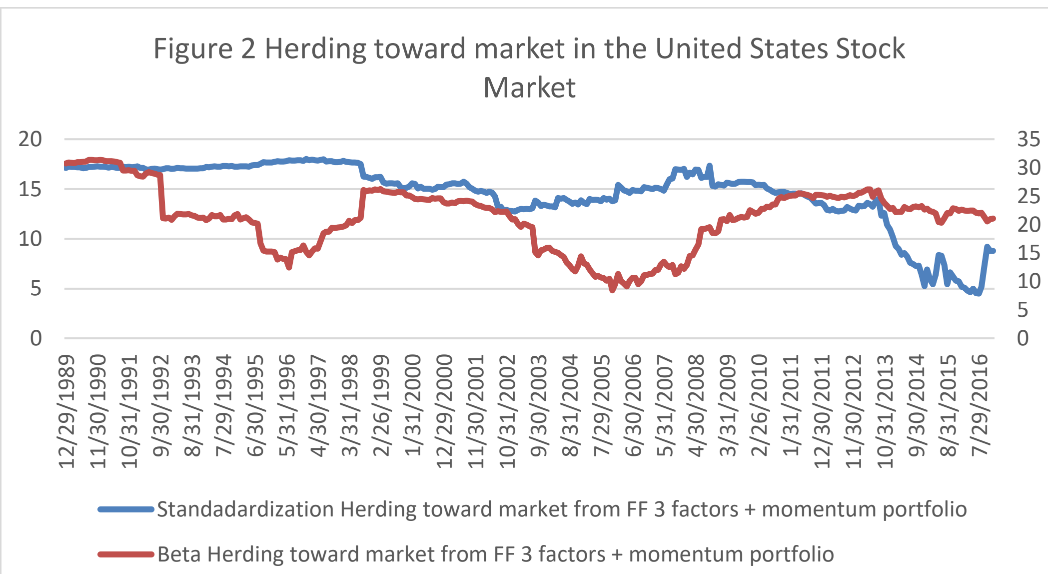


Figure 1 transforms the value of the variance of standardized betas linearly for a clear interpretation: V, standardized→ 20 – V (H) and V, beta -> 40 – 50V (H). Left(Right) axis represents values for Standardized Herding toward market (Beta Herding toward market from FF 3 factors and Momentum portfolio).

---

[6] However, note that the standard errors of beta estimations in CAPM are calculated by White standard errors, not by Newey-west standard error.

[7] This paper implements a linear transformation cross-sectional variance of betas V into 40 – 50V (H)

By the transformation, the measure has an easier interpretation. The first characteristic of the measures is that the two plot graphs are different especially after the financial crisis. Although the both of measurement values drops during the financial crisis, the value of beta herding increases after the crisis while standardized herding measure shows decreasing trends. The plot graphs are generally consistent with the Hwang and Salmon (2017)[8]. According to Hwang and Salmon (2017), *"the herding measure is independent of business cycle although it can be matched with some crisis events".* Rather, Hwang and Salmon (2017) proposes the one source of herding toward market is market sentiment. Sentiment explanatory power on herding measure is significant in regressions, but it does not fully explain the phenomenon (low R squared) according to Hwang and Salmon (2017). One characteristic of beta herding toward market measure is that it is very volatile compared to the standardization herding toward market. It may be due to heteroskedastic estimation error. On the other hand, the standardization herding toward market values suddenly exhibits a steeper decreasing trend after 2014 relative to beta herding toward market values. It seems that the steeper decreasing trend of standardization herding measure is because of some minimal standard errors.

The last characteristic of the herding toward market measure (H) is that it may not be stationary. Recall the equation (7).

Estimated dispersion = Truly biased dispersion + cross sectional estimation error variance

One should note that both herding measures try to capture the herding phenomenon in betas. In other words,

$$\text{Truly biased dispersion} = \text{Herding Phenomenon (HP)} * \text{True dispersion} \tag{8}$$

The thing is that there is no guarantee that the variance of true betas in the fully efficient market is constant over time. As time goes, the whole firms in United States stock market can become safer. For example, the firms which had very high risks (betas) could have gone bankrupt, or to avoid the risk, they could reduce the risk by a business diversification. In this case, the dispersion will be reduced. On the other hand, the other way is also possible. There can be newly born firms which have very high risks or reverse risks. On that case, the dispersion will be increased. The point of these examples is that there is a high probability that the variance of true betas/standardized betas can be volatile. Then, the variance of truly biased betas can also change due to the variations in the variance of true dispersion as well as in herding behaviors. In conclusion, the measure can be not stationary. This reasoning can be

[8] Note that the values of Hwang and Salmon (2017) are not transformed. Therefore, the directions of changes in the Hwang and Salmon (2017)'s figure are opposite to the directions of changes in this paper's figure.

tested by a stationary diagnostic test, called Dickey- Fuller test, and the test also fail to reject the argument that our measure is non-stationary. The test statistic for the standardization measure(beta measure) is -0.563(-1.924) and the corresponding p-value is 0.8792(0.3209).

Thus, the variable should be transformed into a stationary variable to be used in regressions and have a valid interpretation. Generally, the transformation takes the difference, and this paper puts an additional assumption that the variance of true betas in the fully efficient market will be constant for a short time(e.g. within at least 12 months). The important thing is to decide which lag should be chosen for the subtraction. It will depend on momentum traders' behavior. According to Hong and Stein (1999)'s model, the momentum traders will condition on the last month's price change while Jagadeesh and Titman (1993) states that traders condition on previous 3 to 12-month price changes. In this paper, the change in herding behaviors during the previous 11-months matters since the momentum returns data this paper uses are from momentum portfolios conditioning on previous 11-months returns[9]. Therefore 11-month lags will be chosen to take a difference of the original variable. In sum, the value of the new variable is a value of the difference between a current and a 11 month ago values(H). That is, the new variable stands for how overall herding toward the market has changed during a conditioning period. Finally, the differenced herding variable(Deherding) is stationary. The test statistic for the standardization measure(the beta measure) is -3.298(-3.372) and the corresponding p-value is 0.0150(0.0120).

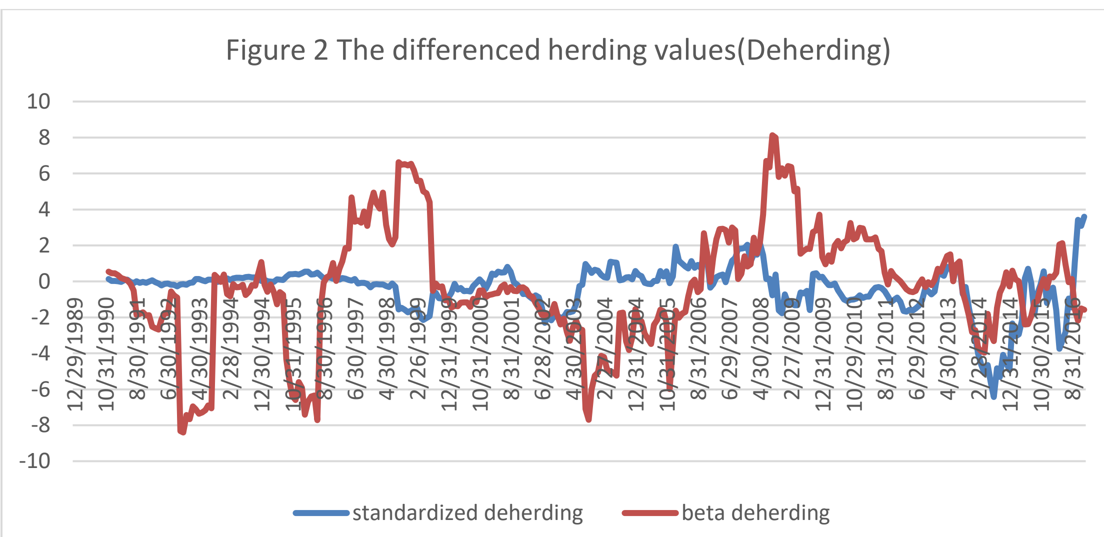


Figure 2 The transformed herding measure is differenced for being stationary. The both new variables Deherding is stationary.

[9] Remember that the momentum returns are from Kenneth French's Website, see foot note 1.

4. Empirical Analysis

Table 2

momentum returns and herding toward market

This table presents the coefficient for following regression specification: $R_{wml,t} = a + bHTM_{t-2} + \in$ where $R_{wml,t}$ is the monthly momentum returns during the period from 1989 12 to 2016 12. HTM_{t-2} stands for changes in herding toward market measure during past 11 months and are lagged by two months. The t-statistic will be given in parentheses. *, **, and *** means the coefficients are significant at 10%, 5%, and 1%, respectively. Momentum returns are constructed from 11 months condition periods (skipping the one month before the construction) and 1 month holding period. The standard error of estimates is calculated from Newey-west error with lag 12. Breusch-Godfrey p value is 0.0337.

Coefficient	Regression
Standardized Deherding (HTM_{t-2})	.4091669 (1.15)

First, this paper will examine the relationship between herding toward market and momentum returns without control variables. When only herding changes and the momentum returns are used without control variables, although there is a positive coefficient in standardized herding toward market as implied by Hong and Stein (1999), it is not significant. However, the regression should include other control variables recognized to explain/predict the momentum returns because of concern for the biasness of the coefficient. The coefficient must be downward biased If the herding phenomenon has any positive correlation with those omitted variables predicting the momentum negatively. On the contrary, the insignificant coefficients can be biased upward if the herding toward market contains some positive explanatory powers of other variables.

Chordia and Shivakumar (2002) uses lagged dividend yield, term spread, default spread, and yield on three-month T-bills to predict momentum returns in the United States. Also, as mentioned above, Barroso and Santa-Clara (2015) uses lagged monthly realized momentum volatility to protect the momentum portfolio from crashes. Finally, Baltzer, Jank, and Smajlbegovic (2014) focuses on the relationship between momentum returns and momentum trading with control variables such as market sentiment, market volatility, momentum volatility, real GDP growth, and market return. Regarding control variables, market sentiment and real GDP growth are relevant to check whether the herding toward market yield a distinct effect on momentum returns because Hwang and Salmon (2017) mentions that the herding toward market is independent of market states, but the source of herding might be sentiment. Also, momentum volatility and momentum returns have a significant relationship in Barroso and Santa-Clara (2015) in that it can reduces the effect of crashes. On the contrary, Chordia and Shivakumar (2002)'s explanatory variables are not often significant (see Griffin et al (2003)).

Therefore, this paper includes the control variables Baltzer, Jank and Smajlbegovic (2014) uses in table 3.[10]

The herding variable does not still predict the momentum returns significantly even when the widely used control variables are used although the signs of the coefficient are consistently positive as predicted. However, there are two things to be noted. As mentioned above, the herding's explanatory power does not depend on business cycle (GDP growth) as the coefficient does not change because they are independent of each other. However, the insignificant herding coefficient becomes more insignificant when the sentiment is included as a control variable. Therefore, the sentiment is related to herding as mentioned by Hwang and Salmon (2017).

The herding measure's predictive power is positive, which is consistent with cross-sectional analyses such as Nofsinger & Sias (1999), Wermers (1999), and Hoitash and Krishnan (2008) but insignificant in all regressions with different control variables and different herding measure (see table 3, Appendix C, and Appendix D). In other words, this paper fails to reject the null hypothesis that herding cannot positively predict the future momentum returns. Therefore, the herding toward market the paper proposes as a source of explaining momentum is not empirically verified. In addition, the herding toward market cannot add a value for predicting momentum returns.

Table 3

This table presents the coefficient for following regression specification: $R_{wml,t} = a + bHTM_{t-2} + \text{Control Variables}_t + \epsilon$ where $R_{wml,t}$ is the monthly momentum returns during the period from 1989 12 to 2015 09 (due to data availability). HTM_{t-2} stands for changes in herding toward market measure during past 11 months condition period and are lagged by two months. Control Variables include monthly momentum volatility, monthly market volatility, 12-month market returns, Real GDP growth based on 2009 dollar, and Market Sentiment. Momentum returns are constructed from 11 months condition periods (skipping the one month before the construction) and 1 month holding period. The t statistics for herding variable will be given in parentheses. Detailed control variable description is given in appendix B. The standard error of estimates is calculated from Newey-west error, lag(8).

Variable	(1)	(2)	(3)	(4)	(5)
Herding	.33762805 (1.01)	.36636172 (1.03)	.3563677 (0.99)	.34887544 (0.99)	.17051861 (0.44)
Mom Vol	-.0888704***	-.12750036***	-.13913901***	-.14043372***	-.13667336***
Mkt Vol		.04887313***	.04425247***	.04292211***	.04444422***
12 Month MR			-5.0571833	-4.5391702	-1.904918

[10] Instead, the results with macroeconomic variables by Chordia and Shivakumar (2002) will be given in Appendix C.

Real GDP growth				-.10991627	-.17659983
Market Sentiment					1.6446388**
CON	2.4348142***	1.7910545***	2.4900098***	2.7774667**	2.4170235*

*, **, and *** means the coefficients are significant at 10%, 5%, and 1%, respectively.

Lastly, the paper implements the out of sample forecast. Firstly, based on 29/12/1989 to 30/10/2009, the regression will predict the momentum value for 30/11/2009. Next, 30/11/2009 data is included, and based on 29/12/1989 to 30/11/2009, the next regression will predict the momentum value for 29/12/2012. This rolling forecast with the expansionary estimation window will last until the last data. Therefore, out of sample analysis is based on 85 forecast values. The out of sample forecast R squared will indicate how well the herding measure can predict in times of recent periods. The result is given as -0.01510[11], which means it does not predict at all. Even, it predicts worse than the sub sample mean of momentum returns predicts (where R squared is 0).

5. Discussion

The proposed hypothesis to main question is motivated by the Hwang and Salmon (2004)'s concept and the Hong and Stein (1999)'s implication. The hypothesis is a logical deduction from the Hwang and Salmon (2004)'s concept and the Hong and Stein (1999)'s implication with the assumption that focus on [herding toward] the signal from market portfolio will restrain market-wide private information flows. That is, The hypothesis stems from two scientific arguments and one premises. However, the empirical analysis fails to give a supporting evidence even though one empirical analysis is not sufficient to deny the argument. According to D-N models of explanation, the falsity of deductive arguments stem from the falsity of premises or a scientific model.

The Hong and Stein (1999)'s implication that slower information diffusion will yield more accelerated momentum patterns is supported by Hong and Stein (2000). Hong and Stein (2000) considers a degree of analysis coverage as the proxy for the information diffusion speed. The momentum portfolio based on low analyst coverage stocks (Hong and Stein (2000) refers to *"Sub 1"*) yields higher cumulative returns compared to high analyst coverage stocks. Moreover, the cumulative momentum returns in Sub 1 keeps accumulating after one year while the momentum returns for high analyst coverage stocks (*"Sub 3"* according to Hong and Stein

[11] Out of Sample Forecast R squared is calculated from 1 – Residuals Sum of Squares

(2000)) stop accumulating after 10 months although the return differences between Sub 1 and Sub 3 become less significant over time (Hong and Stein (2000)).

The potential problem may stem from the assumption this paper poses that herding toward market will restrain market-wide private information flows. The assumption is composed of two arguments:

(1) The herding toward market will restrain private information flows of all the individual firm.

(2) The Hong and Stein (1999)'s implication based on a firm level will be valid on a market level.

However, there might be a case where the herding toward market does not affect the stocks that momentum traders hold. This is possible because most of momentum traders are institutional investors, and they are more informed and rational traders. Or, the implication by the Hong and Stein (1999)'s implication may not be valid in an aggregate level. It is well-known fact that the results from a cross sectional individual analysis may not be used for the inference for a market level, which is called "individualistic fallacy or ecological fallacy (e.g. Selvin (1958)) on the other way".

An alternative reason of the insignificant relationship may be a problem in the research method although there is no theoretical implausibility. There exists the concern that the measure of herding toward market might not capture the herding phenomenon correctly. According to Hwang and Salmon (2004), the definition of herding toward market is the state where investors is obsessed with how other investors evaluate the market. Hwang and Salmon (2017) composes a previous intuition and argue that the overconfidence on aggregate market signal and market sentiment are the source of herding toward market. Accordingly, all stocks returns will converge to market returns in times of herding, and herded returns will result in betas herding around 1. However, there can be other sources that cause the returns to herd around market returns. That is, the measure values can vary due to changes in other source but not because of the herding phenomenon. If the momentum causes the herding to vary (reverse causality), then the measurement problem gets worse as implied by Nofsinger & Sias (1999).

In contrast, the momentum returns from Kenneth French's library may not reflect the full momentum forces. It is because the holding period is 1-month in momentum data Kenneth French's library while Jagadeesh and Titman (1993) states that the momentum strategy is powerful if the momentum portfolio lasts for 3 - 12 months after the portfolio construction. The data makes the empirical analysis only possible to see the relationship between the change in herding during the conditioning period and the holding returns after 2 months, which is not desirable approach in terms of Jagadeesh and Titman (1993). To see the exact relationship as implied Jagadeesh and Titman (1993), the momentum returns should be the returns from portfolios of at least 3-month holding period. Also, since the 1 month holding return is possible, it is not possible to see that whether there is a reversal in the long run.

6. Conclusion

In conclusion, the thesis proposes an answer to the research question that the herding toward market will positively affect the future momentum returns. The answer is tested with the measure of Hwang and Salmon (2017) and the empirical data during the period of December 1989 to December 2016. Although the signs of the coefficients are consistently positive across the regressions with control variables, they are not significant. In addition, the herding toward market does not have so much predictive values.

However, the reason of the positive but insignificant relationship is inconclusive. It may due to the falsity of the assumption. The herding toward market is distinct from the herding toward a strategy (e.g. positive feedback trading) used in other papers, and the herding toward market may have no direct relationship with momentum while the Herding toward positive feedback trading by institutional investors has a positive correlation with stock returns. However, there is also a possibility that the research method does not correctly reflect the phenomenon, and/or the data provided may only contain the partial force of momentum.

A single empirical analysis does not disconfirm the arguments. Perhaps, the follow-up researches can test the relationship(hypothesis) with a different data or method.

References

Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. *Value and momentum everywhere.* The Journal of Finance, 68(3), 2013.

Baltzer, M., Jank, S., & Smajlbegovic, E. *Who Trades on Momentum?* Bundesbank Discussion Paper No. 42/2014, 2014.

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

Barberis, N., Andrei, S., & Robert, V. *A model of investor sentiment.* Journal of financial economics, 49(3), 1998.

Barroso, P., & Santa-Clara, P. *Momentum has its moments.* Journal of Financial Economics, 116(1), 2015.

Chordia, T., & Shivakumar, L. *Momentum, business cycle, and time-varying expected returns.* The Journal of Finance, 57(2), 2002.

Christie, W. G., & Huang, R. D. *Following the pied piper: Do individual returns herd around the market?* Vol. 51(4). Financial Analysts Journal, 1995.

Daniel, K., & Moskowitz, T. J. *Momentum crashes.* Journal of Financial Economics, 122(2), 2016.

Griffin, J. M., Ji, X., & Martin, J. S. *Momentum investing and business cycle risk: Evidence from pole to pole.* The Journal of Finance, 58(6), 2003.

Grinblatt, M., Titman, S., & Wermers, R. *Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior.* The American economic review, 1995.

Hoitash, R., & Krishnan, M. *Herding, momentum and investor over-reaction.* Review of Quantitative Finance and Accounting, 30(1), 2008.

Hong, H., & Stein, J. C. *A unified theory of underreaction, momentum trading, and overreaction in asset markets.* The Journal of finance, 54(6), 1999.

Hong, H., Lim, T., & Stein, J. C . *Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies.* The Journal of Finance, 55(1), 2000.

Hwang, S., & Salmon, M. H. *A new measure of herding and empirical evidence.* 2001.

Hwang, S., & Salmon, M. *Market stress and herding.* Journal of Empirical Finance, 11(4), 2004.

—. *Overconfidence Sentiment and Beta Herding.* Available at SSRN: https://ssrn.com/abstract=299919 or http://dx.doi.org/10.2139/ssrn.299919, 2017.

Jegadeesh, N., & Titman, S. *Profitability of momentum strategies: An evaluation of alternative explanations.* The Journal of finance, 56(2), 2001.

—. *Returns to buying winners and selling losers: Implications for stock market efficiency.* The Journal of finance, 48(1), 1993.

—. *Returns to buying winners and selling losers: Implications for stock market efficiency.* The Journal of finance 48(1), 1993.

Keynes, J. M. *The general theory of employment, investment, and money.* London and New York, 1936.

Lakonishok, J., Shleifer, A., & Vishny, R. W. *The impact of institutional trading on stock prices.* Journal of financial economics, 32(1), 1992.

Miller, E. M. *Risk, uncertainty, and divergence of opinion.* The Journal of finance, 32(4), 1977.

Nofsinger, J. R., & Sias, R. W. *Herding and feedback trading by institutional and individual investors.* The Journal of finance, 54(6), 1999.

Rouwenhorst, K. G. *International momentum strategies.* The Journal of Finance, 53(1), 1998.

Selvin, H. C. *Durkheim's suicide and problems of empirical research.* American journal of sociology, 63(6), 1958.

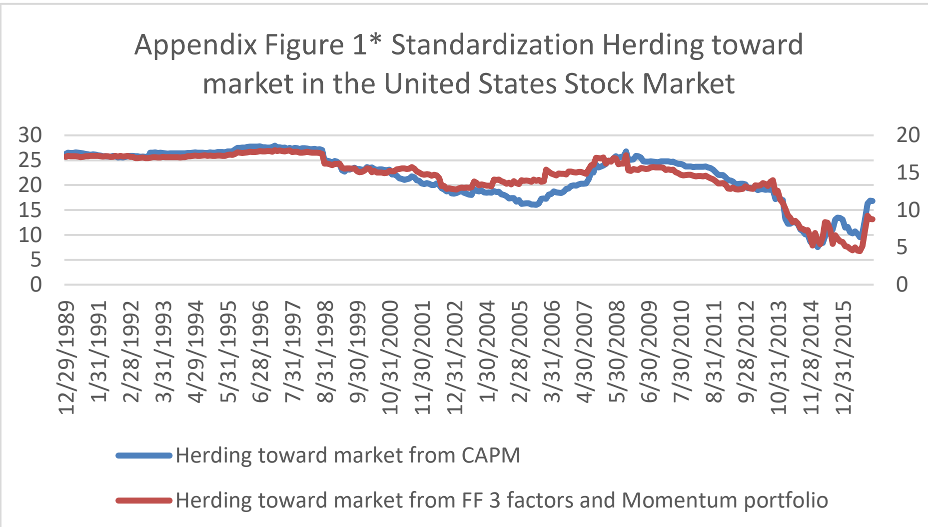
Wermers, R. *Mutual fund herding and the impact on stock prices.* the Journal of Finance, 54(2), 1999.

Zhou, R. T., & Lai, R. N. *Herding and information based trading.* Journal of Empirical Finance, 16(3), 2009.

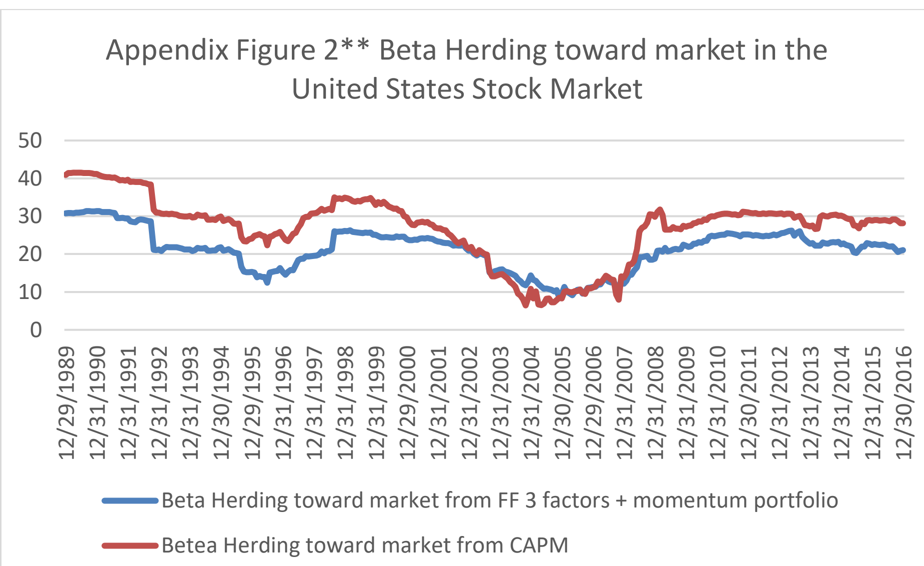
Appendix

Appendix A

Figure 2* Standardized Herding toward market in the United States Stock Market.



Appendix A Figure 1*. Left(Right) axis represents values for Standardization herding toward market from CAPM (Standardization herding toward market from FF 3 factors and Momentum portfolio). Standardization herding toward market from CAPM is derived from the transformation of the variance of betas (V*) into 30 – V*(H*). Note that the estimation error of betas in CAPM is from White standard error, not Newey-west HAC standard error.



Appendix A Figure 1**. Beta Herding toward market from CAPM is derived from the transformation of the variance of betas (V*) into 50 – 50V**(H**).

Appendix B

The description of control variables

Market sentiment index is the index from Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance*, *61*(4), 1645-1680. and Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *The Journal of Economic Perspectives*, *21*(2), 129-151. Data can be found on http://people.stern.nyu.edu/jwurgler/

Monthly momentum volatility and Monthly market volatility is manually constructed with the daily returns data from Kenneth French's library, http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html. and defined as $\sum_{y=0}^{20} R_{end-j}^{2}$. "R" means daily returns, and "end" means the last date for each month, following Barroso, P., & Santa-Clara, P. (2015). Momentum has its moments. Journal of Financial Economics, 116(1), 111-120.

Real GDP growth data of the United States is from https://fred.stlouisfed.org/series/GDPC1

12-month market return data is manually constructed by the addition of previous 11 months returns with the current month return

Yield on three-month T bill data is from Center for Research in Security Prices (CRSP)

Default Spread data is defined as the difference of rates between 3-month LIBOR and US 3-month T bill from https://fred.stlouisfed.org/series/TEDRATE

Term Spread data is defined as 10-year treasury rate minus 1-year treasury rate from https://research.stlouisfed.org/pdl/183?ob=vs&od=asc&filter%5B0%5D=&filter%5B1%5D=

Dividend Yield data is based on S&P 500 composite from http://www.econ.yale.edu/~shiller/data.htm

Appendix C

Table 3* Analysis with the control variable by Chordia and Shivakumar (2002)

This table presents the coefficient for following regression specification: $R_{wml,t} = a + bHTM_{t-2} + \text{Control Variables}_{t-1} + \in$ where $R_{wml,t}$ is the monthly momentum returns during the period from 1989 12 to 2014 12 (due to data availability). HTM_{t-2} stands for changes in herding toward market measure during past 11 months condition period and are lagged by two months. Control Variables include term spread, default spread dividend yield on S&P 500 firms, and Yield on 3-month T-bill. The control variables are lagged by one month. Momentum returns are constructed from 11 months condition periods (skipping the one month before the construction) and 1 month holding period. The t statistics for herding variable will be given in parentheses. Detailed control variable description is given in appendix B. The standard error of estimates is calculated from Newey-west error, lag(8).

Variable	(1)	(2)	(3)	(4)	(5)
Herding	.4091669 (1.19)	.35666929 (1.05)	.27473638 (0.80)	.40660313 (1.09)	.11645587 (0.26)
TERM		-.34508722	-.12779346	.5634644	1.636958*
DEF			.08913761	.13184006**	.14254482***
DIV				-268.10479*	-329.06951*
YLD					.74911974
CON	1.0252015**	1.556914**	.28365894	4.2770122	1.383942

*, **, and *** means the coefficients are significant at 10%, 5%, and 1%, respectively.

Appendix D

Herding measure – beta herding toward market

Table 2*

momentum returns and herding toward market

This table presents the coefficient for following regression specification: $R_{wml,t} = a + bHTM_{t-2} + \in$ where $R_{wml,t}$ is the monthly momentum returns during the period from 1989 12 to 2016 12. HTM_{t-2} stands for changes in herding toward market measure during past 11 months and are lagged by two months. The t-statistic will be given in parentheses. *, **, and *** means the coefficients are significant at 10%, 5%, and 1%, respectively. Momentum returns are constructed from 11 months condition periods (skipping the one month before the construction) and 1 month holding period. The standard error of estimates is calculated from Newey-west error with lag 12. Breusch-Godfrey p value is 0.0337.

Coefficient	Regression
Beta Deherding (HTM_{t-2})	-.1116231 (-0.58)

Table 3*

This table presents the coefficient for following regression specification: $R_{wml,t} = a + bHTM_{t-2} + \text{Control Variables}_t + \in$ where $R_{wml,t}$ is the monthly momentum returns during the period from 1989 12 to 2015 09 (due to data availability). HTM_{t-2} stands for changes in herding toward market measure during past 11 months condition period and are lagged by two months. Control Variables include monthly momentum volatility, monthly market volatility, 12-month market returns, Real GDP growth based on 2009 dollar, and Market Sentiment. Momentum returns are constructed from 11 months condition periods (skipping the one month before the construction) and 1 month holding period. The t statistics for herding variable will be given in parentheses. Detailed control variable description is given in appendix B. The standard error of estimates is calculated from Newey-west error, lag(8).

Variable	(1)	(2)	(3)	(4)	(5)
Herding (Beta Deherding)	.19780431 (1.16)	.06167668 (0.45)	.07800287 (0.57)	.07145306 (0.54)	.1109891 (0.90)
Mom Vol	-.0957089***	-.12886722***	-.14125171***	-.14241006***	-.13887585***
Mkt Vol		.04746143***	.04235802***	.04115328***	.04234889***
12 Month MR			-5.2604942	-4.7249038	-1.9824631
Real GDP growth				-.11101059	-.17101086
Market Sentiment					1.7694902**
CON	2.4695043***	1.7193888***	2.4681471***	2.7524792**	2.4538609*

*, **, and *** means the coefficients are significant at 10%, 5%, and 1%, respectively.