Tampere University

Jon Leskinen

INVESTOR HERDING IN THE FINNISH STOCK MARKET

Faculty of Engineering and Natural Sciences
Master of Science Thesis
February 2020

ABSTRACT

Jon Leskinen: Investor herding in the Finnish stock market
Master of Science Thesis
Tampere University
Industrial Engineering and Management
February 2020

This thesis studies herding in the Finnish stock market between 1.7.2005 and 30.6.2009. Key questions are (1) was there herding in the Finnish market during that time, (2) is there a link between herding and returns and (3) is there a link between herding and volatility. Herding was measured using the herding measure developed by Lakonishok, Shleifer and Vishny (1992). The relationship between herding and returns and herding and volatility was studied using a mixed-effects regression model. The data used was provided by Euroclear Finland Oy.

During the study period considerable amount of herding was found. The average amount of herding during the whole study period was 10.10 %. In the pre-crisis period the average amount was 9.79 %; in the crisis period 10.40 %. In the pre-crisis period the buy herding measure and sell herding measure had almost identical values. During the crisis period a clear gap emerged, when the buy herding measure decreased to 7.23 % while the sell herding measure increased to 12.11 %. This implies that during a crisis sell-side herding is more common than buy-side herding.

It seems that household investors and institutional investors both herd less than the average investor in the Finnish market. When looking at the whole study period of 2005 - 2009, institutional investors herded on average less than household investors. However, in the crisis period of 2007 - 2009 institutions herded slightly more than household investors.

During the whole study period, on average, large capitalization stocks experienced more herding than small capitalization stocks. The lowest levels of herding were observed in health care and consumer staples industries. Highest levels of herding were observed in telecommunication services, materials and utilities.

A statistically significant link between past returns and herding was found. High past returns correlate with higher sell-side herding, while poor past returns correlate with higher buy-side herding. This implies that investors in the Finnish market followed a contrarian investment strategy, selling past winners and buying past losers.

Regarding herding and future returns some statistically significant correlations were found, but on the whole the results were inconclusive. When studying the link between herding and both past and future volatility, no conclusive results were obtained.

Keywords: herding, LSV, investor behaviour, behavioral finance, stock market

TIIVISTELMÄ

Jon Leskinen: Sijoittajien laumautuminen Suomen osakemarkkinoilla
Diplomityö
Tampereen yliopisto
Tuotantotalouden tutkinto-ohjelma
Helmikuu 2020

Tässä diplomityössä tutkittiin laumautumista Suomen osakemarkkinoilla aikavälillä 1.7.2005 - 30.6.2009. Keskeisiä kysymyksiä ovat (1) tapahtuiko Suomen markkinoilla laumautumista mainittuna aikana, (2) onko laumautumisen ja tuottojen välillä yhteys ja (3) onko laumautumisen ja volatiliteetin välillä yhteys. Laumautumista mitattiin mittarilla, jonka ovat kehittäneet Lakonishkok, Shleifer ja Vishny (1992) . Laumautumisen ja tuottojen sekä laumautumisen ja volatiliteetin välisen yhteyden tutkimiseen käytettiin lineaarista regressioanalyysiä. Käytetty data saatiin Euroclear Finland Oy:ltä.

Tutkitulla aikavälillä havaittiin merkittävää laumautumista. Koko tutkimusajalta laumautumisen keskiarvo oli 10.10 %. Ennen kriisiä olevalla periodilla (2005 - 2007) keskimääräinen laumautuminen oli 9.79 %; kriisiperiodilla (2007 - 2009) 10.40 %. Ennen kriisiä olevalla periodilla osto- ja myyntilaumautumisen arvot olivat lähes identtiset. Kriisiperiodin aikana ne eriytyivät, kun ostolaumautuminen laski 7.23 %:iin ja myyntilaumautuminen nousi 12.11 %:iin. Vaikuttaa siis siltä, että kriisin aikana myyntilaumautuminen on yleisempää kuin ostolaumautuminen.

Tulosten perusteella sekä yksityiset sijoittajat että institutionaaliset sijoittajat laumautuvat vähemmän kuin keskimääräinen Suomen markkinoilla oleva sijoittaja. Kun tarkastellaan koko 2005 - 2009 tutkimusperiodia, institutionaaliset sijoittajat laumautuivat vähemmän kuin yksityiset sijoittajat. Kuitenkin, kriisiperiodin aikana instituutiot laumautuivat hieman enemmän kuin yksityiset sijoittajat.

Koko tutkimusperiodilla suuren kapitalisaation osakkeet kokivat enemmän laumautumista kuin pienen kapitalisaation osakkeet. Toimialoista terveydenhuolto ja kuluttajahyödykkeet kokivat vähiten laumautumista. Eniten laumautumista havaittiin telekommunikaatiopalveluissa, materiaaleissa ja yleishyödyllisissä palveluissa.

Tilastollisesti merkitsevä linkki menneiden tuottojen ja laumautumisen välillä löydettiin. Korkeat menneet tuotot korreloivat korkeamman myyntilaumautumisen kanssa, kun matalat menneet tuotot korreloivat ostolaumautumisen kanssa. Tämä implikoi, että sijoittajat Suomen markkinalla sijoittavat vastavirtaan, myyden menneitä voittajia ja ostaen menneitä häviäjiä.

Laumautumisen ja tulevien tuottojen väliltä löydettiin joitain tilastollisesti merkitseviä yhteyksiä, mutta kokonaisuudessaan tulokset eivät ole vakuuttavia. Myöskään laumautumisen ja tulevan tai menneen volatiliteetin väliltä ei löydetty yksiselitteistä korrelaatiota.

Avainsanat: laumautuminen, LSV, sijoittajien käyttäytyminen, behavioraalinen rahoitus, osakemarkkinat

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -ohjelmalla.

PREFACE

The brain is a wonderful organ; it starts working the moment you get up in the morning and does not stop until you need to write the preface for your Master's thesis.

I'd like to thank professor Juho Kanniainen for his valuable support during this process. His classes inspired me to study finance in the first place. In addition I want to thank my family and friends for their love and support.

In Helsinki, 3rd February 2020

Jon Leskinen

CONTENTS

LIST OF SYMBOLS AND ABBREVIATIONS

BHM	Buy-side herding measure, which is derived from the LSV herding measure
$\mathbb{E}$	The symbol for expected value
LSV	The LSV herding measure first presented by Lakonishok, Shleifer and Vishny (1992)
p.p.	percentage point
SHM	Sell-side herding measure, which is derived from the LSV herding measure

1 INTRODUCTION

Herding is a well-known phenomenon where investors disregard their own information and instead imitate the observed decisions of others (Hwang and Salmon 2004). The purpose of this thesis is to empirically study herding in the Finnish stock market between 2005 and 2009. Special focus is put on the effect of the global financial crisis of 2007 - 2008 and the difference between institutional investors and household investors. On one hand, herding can be seen as irrational, since in classical economic theory all investment decisions should reflect the agents' rationally formed expectations, which are formed by using all available information in an efficient manner (Scharfstein, Stein et al. 1990). On the other hand, herding behavior may sometimes be rational to individual agents. These situations might include professional investors who worry about their reputation (Scharfstein, Stein et al. 1990) or situations where there is little fundamental information available, but many opportunities to observe the investment decisions of others (so-called information cascades) (Alevy, Haigh and List 2007; Hirshleifer and Hong Teoh 2003).

Herding has been observed to happen in a variety of markets, including real-estate (Philippas et al. 2013), bonds (Oehler and Chao 2000; Raddatz and Schmukler 2011) and stocks (e.g. Lakonishok, Shleifer and Vishny 1992; Nofsinger and Sias 1999; Wermers 1999). Herding is commonly thought to destabilize markets and increase volatility (Alevy, Haigh and List 2007; Philippas et al. 2013; Scharfstein, Stein et al. 1990), but some authors argue that herding can in fact speed up the price process and drive prices towards equilibrium values (Hirshleifer, Subrahmanyam and Titman 1994).

Multiple reasons for herding have been suggested in earlier literature. These include reputational risks for managers (Scharfstein, Stein et al. 1990), correlated private information (Froot, Scharfstein and Stein 1992; Hirshleifer, Subrahmanyam and Titman 1994), information cascades (Hirshleifer and Welch 1992) and preference for stocks with certain characteristics (Falkenstein 1996). These reasons are investigated more in-depth in chapter 2.

The topic of this thesis is relevant for multiple reasons. First, only a few studies (Grinblatt and Keloharju 2000; Grinblatt and Keloharju 2001; Kyrolainen and Perttunen 2003) have focused on the Finnish stock market. Secondly, the effect of the global financial crisis on herding gives an interesting perspective and it has never before been studied in the Finnish market. Third, this thesis also studies the relationship between herding and both past and future returns. This may give new insights about the reasons behind herding and the circumstances under which herding might occur. Fourth, the behaviour of household investors is rarely studied, perhaps due to data limitations.

This Master's thesis uses a rare data set from Euroclear Finland Oy (previously Finnish Central Security Depository) to empirically study herding. This data set includes investors' trades on daily and even intra-day levels. The data spans 21 years, from 1995 to 2016. This thesis focuses on the time period between 2005 and 2009. The data is presented more thoroughly in chapter 3. This thesis has three main themes: the amount of herding in the Finnish stock market between 2005 and 2009, the relationship between herding and returns and the relationship between herding and stock price volatility. There are 5 main questions and some sub-questions:

1. Was there herding in the Finnish stock market between 2005 and 2009, and if so, how much?
   - Does herding differ between the pre-crisis (2005 - 2007) and crisis (2007 - 2009) periods?
   - Is there a difference between buy-side and sell-side herding prevalence?
   - Is there a difference between the herding of institutional investors and household investors?
   - Does herding differ between industries?
2. Do past stock returns affect herding?
3. Does herding affect future stock returns?
4. Does past volatility affect herding?
5. Does herding affect future volatility?

These questions are addressed in chapters 4 and 5. Previous work in this field include e.g. Choe, Kho and Stulz (1999), Gębka and Wohar (2013), Goodfellow, Bohl and Gebka (2009), Hsieh (2013), Kim and S.-J. Wei (2002), Lakonishok, Shleifer and Vishny (1992), Nofsinger and Sias (1999), Wermers (1999) and Zheng, Li and Chiang (2017).

Choe, Kho and Stulz (1999) studied herding in the Korean market using the LSV herding measure and found herding amounts in excess of 20 % for some large Korean stocks. Gębka and Wohar (2013) studied herding in different industries and found that on an international level, basic materials, consumer services and oil and gas stocks show some evidence of herding. Unfortunately comparing the results of this thesis with those of Gębka and Wohar (2013) is difficult, since they utilize a different industry classification. Using data from Poland, Goodfellow, Bohl and Gebka (2009) measure herding using the CSAD measure developed by Chang, Cheng and Khorana (2000). They find that individual investors herd more during bear markets and less during bull markets. This is consistent with the findings of this thesis. Goodfellow, Bohl and Gebka (2009) find no evidence of institutional investor herding. In this thesis some evidence of insitutional investor herding is found. Hsieh (2013) found significant evidence of herding in Taiwan. Using the LSV measure, the average herding amount of institutional investors was 16.92

% while the herding amount of individual investors was 6.73 %. He also found that individual investors are contrarian investors, who buy stocks with negative past returns and sell stocks with positive past returns. These results are well in line with the findings of this thesis.

Studying investors in Korea using the LSV herding measure, Kim and S.-J. Wei (2002) found herding levels between 8.4 % and 13.2 % for individual foreign investors, but lower levels of herding between 4.8 % and 8.3 % for resident individual investors. Similar levels of herding are found in this study. Lakonishok, Shleifer and Vishny (1992) originally developed the herding measure later used by many others, including this thesis. They measured the herding of mutual fund money managers and found low herding amounts of 2.7 %. In this thesis the level of institutional investor herding is somewhat larger, but considering the differences in data, market size and time period, still well in line with the original findings of Lakonishok, Shleifer and Vishny (1992). Nofsinger and Sias (1999) study the herding of institutional and individual investors. They also investigate post-herding returns. They find that the stocks institutional investors flock to buy perform better than those they sell. They do not use the LSV herding measure and their herding period is one year. In this thesis no conclusive results about the relationship between herding and future returns were found.

Wermers (1999) studied mutual fund herding between 1975 and 1994. Using the LSV measure he found only low levels of herding, 3.4 %. He also found that mutual funds seem to use momentum investment strategies. His findings imply that momentum investment strategies are rational, since no evidence of subsequent return reversals were found. The levels of herding observed in this thesis are consistent with the findings of Wermers (1999). In this thesis it is found that investors in the Finnish market show contrarian tendencies and no evidence of momentum investment strategies are found. Zheng, Li and Chiang (2017) study herding in 9 Asian markets using the CSAD measure. They find that cross-industry herding occurs most commonly in the telecommunication and financial industries, and that herding is more pronounced in down markets. These results are well in line with the findings of this thesis.

This thesis also adds to the body of literature investigating whether herding destabilizes or stabilizes markets. If buy-side herding destabilizes markets, then there should be and observable price increase followed by a price decrease. If, on the other hand, herding stabilizes markets and drives prices towards their equilibrium values, then there should be no observable price decrease after the initial price increase (Wermers 1999). This thesis studies the stabilizing or destabilizing nature of herding through the relationship between herding and volatility.

The remainder of this thesis is structured as follows: Chapter 2 describes herding as a phenomenon, some possible reasons behind it and different methods used to measure herding. Chapter 3 introduces the data and methodology. Chapter 4 presents the results regarding herding in the Finnish stock market. Chapter 5 analyzes the relationship between herding, returns and volatility. Chapter 6 concludes.

2 HERDING THEORY

This chapter introduces herding as a phenomenon, takes a look at some possible reasons behind herding and presents different measures used to study herding.

2.1 Basic theory of herding

Herding is a well-known and widely researched phenomenon in financial markets. The term herding comes from the perceived habit of investors to "follow the herd" in their investment decisions. Different researches have defined herding in a multitude of ways. Hwang and Salmon (2004) say that "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". Bikhchandani and Sharma (2000) say that "herding results from an obvious intent by investors to copy the behaviour of other investors". Lakonishok, Shleifer and Vishny (1992) define herding by money managers as "buying (selling) the same stocks as other managers buy (sell) at the same time".

Even in these definitions we can observe an important difference: does herding happen only when investors knowingly decide to follow others, or can trades in the same stocks at the same time be defined as herding? Relating to this Bikhchandani and Sharma (2000) have defined "spurious herding". Spurious herding occurs when investors face similar decision problems and have the same information, so they make similar investment decisions. Here investors do not follow each others' trades, but still buy (or sell) at the same time. When empirically researching herding, the problem is that it is challenging, maybe even impossible, to separate spurious herding from "real" herding just by looking at the data. Changes in fundamentals may drive many investors to buy (or sell) the same stocks at the same time independent of each other. Then, in reality, no herding has occurred, since investors make their decisions independent of each others' investment decisions. But separating this from a situation where investors base their investment decisions on the previously observed decisions of other investors is close to impossible, since both situations look identical in most commonly used data types (e.g. transaction data or changes in portfolio composition).

Research on herding usually focuses either on herding by institutional investors or by individual investors (Merli and Roger 2013). Institutional investors include banks, mutual funds, pension funds and other institutions where professional managers are responsible for making the investment decisions. Individual investors are retail investors, non-professionals, who invest through brokerage firms or savings accounts. This division between institutional and individual investors is justifiable, since professional investors and non-professional investors may differ in their goals, incentives and information available to them (Lakonishok, Shleifer and Vishny 1992). There are also important differences in the availability of data concerning the groups: many institutions have to report their holdings on regular basis (e.g. quarterly 13F filings in the US), while there is usually little data available on the trades made by individual investors.

Many possible reasons for institutional herding have been identified. These include information cascades, investigative herding, reputational herding, fads and characteristic herding (Sias 2004). These will be looked at more closely in the next section. Some of these reasons are also applicable in the case of individual investors (such as informational cascades), while e.g. reputation is not usually a concern for individual investors.

Herding is an important phenomenon to investigate since it has been linked to a number on financial phenomena, including excess volatility, momentum, reversals in stock prices and investment recommendations (Graham 1999; Nofsinger and Sias 1999). It has been said that herding hinders the price discovery process, since investors who herd do not reveal their own private information but instead choose to ignore it (Philippas et al. 2013). It has also been theorized that herding by institutional investors may destabilize markets and drive asset prices away from their fundamental values (Alevy, Haigh and List 2007; Philippas et al. 2013; Scharfstein, Stein et al. 1990). There is also a contrary view that argues that herding by institutional investors might in fact speed up the price discovery process and drive asset prices towards their fundamental values faster (Hirshleifer, Subrahmanyam and Titman 1994).

The reasoning behind these ideas is as follows: if institutional investors follow each others' example in trades without any regard to fundamentals, asset prices can increase (or decrease) past their fundamental values. If markets are efficient, after this initial price increase there should be a price decrease as other investors now realize that the asset is overpriced and sell it. If this holds, then after buy-side herding has occurred, there should be an increase in asset prices which is followed by a subsequent decrease in prices. Thus herding would cause increased volatility, drive asset prices away from their fundamental values and create reversals in said asset prices.

On the other hand, institutional investors may actually drive prices toward their fundamental values and stabilize markets. If markets are efficient, but new information is not reflected in asset prices **instantly** but is instead incorporated over time, and if institutional investors are better informed than other investors, then herding by institutional investors may speed up the price discovery process and drive asset prices towards their fundamental values. If this is the case, then the price increase following a buy herd should not be followed by a subsequent price decrease.

This brings us back to the concept of "spurious herding". In the latter example institutional investors might all get the same signal about a change in fundamentals at the same time and because of that signal make the same investment decision. Because the trades happen at the same time and in the same direction, many researches would classify this as herding. This trading would also move prices towards the fundamental value (since the signal that started the trading was new information about fundamentals). Thus one might come to the conclusion that institutional herding stabilizes markets, while in fact there was no "real" herding, only a rational reaction to new information. The measuring of herding and the shortcomings of different measures are discussed in a later section.

2.2 Reasons for herding

As previously mentioned, many possible reasons for herding have been identified in literature. Some of the reasons apply to both institutional and individual investors, some to just one group. A common assumption in herding literature is that individual investors are poorly informed noise traders. It has been suggested that individual investors are influenced by investment fads or styles, that they irrationally extrapolate past growth rates into the future, that they buy attention grabbing stocks because of limited attention spans or that they sell past winners and hold on to past losers (also called the disposition effect) (Barber, Odean and Zhu 2009; Nofsinger and Sias 1999). Different possible reasons for herding are discussed more in-depth in this section.

Information cascades can occur when decisions are made sequentially and earlier decisions are publicly observable. An information cascade arises when individuals choose to ignore their private information and instead choose to repeat the actions of previous decision makers (Alevy, Haigh and List 2007; Bikhchandani, Hirshleifer and Welch 1992). This may create a situation where investors herd on an investment decision that is wrong for all of them.

Bikhchandani and Sharma (2000) give an example: let us suppose that the are 100 investors who each have their own assessments, possibly different, about the profitability of an investing opportunity. 20 of these investors believe that the investment is profitable while 80 investors believe it is not. Every investor knows only their own estimate of the profitability; they do not know the assessments of others' or known which opinion is more common. If all the investors got together and pooled their knowledge, they would decide that this investment opportunity does not seem profitable. In addition, these investors do not make their investment decision at the same time. Investing at the same time would create a situation where 20 optimistic individual invest and 80 more pessimistic ones do not invest. Instead, the decisions are made sequentially and investors can observe the decisions made by people before them. Now, if the first few investors are among the 20 optimistic ones and they decide to invest, they can influence some of the 80 pessimistic investors to invest too, which in turn leads to a snowballing effect where almost everyone invests. This kind of behaviour can be rational, since if an investor's information about the profitability of the investment opportunity is not certain, or there is little information available, following the decisions of others' can be rational because they might know more and their investment decision should reflect their knowledge (Alevy, Haigh and List 2007). So, if all the people before you have invested, you could assume they know something that you don't and that you should invest too.

This example illustrates three aspects of information cascades. (1) The decisions made by the early investors may determine what the majority decides. (2) The decision that the investors herd on may be wrong. (3) If the majority decision is wrong, then upon the arrival of new information or with experience the herd may reverse their decision and start a herd in the opposite direction. This kind of behaviour increases volatility in the market. (Bikhchandani and Sharma 2000)

Investigative herding means that investors focus on examining a piece of information they believe that others will also examine (Graham 1999). The investor would like to sell their investment at a profit, but this is possible only if other investors study the same signal and thus drive the price to the direction anticipated by the first investor. Froot, Scharfstein and Stein (1992) have developed a model describing this kind of behaviour. They give an example: let us assume that there are two variables, *a* and *b*, which have useful information about the value of an investment. Individual investors can learn information about *a* or *b* but not both. For all the information to be correctly included in the asset price, half the traders should study *a* and the other half *b*. Froot, Scharfstein and Stein (1992) argue that if traders have long investment horizons, this is what happens. On the other hand, if traders have short horizons the situation is different. If all traders choose to study *a*, there is no incentive for an individual trader to study *b*. Even though *b* should affect the asset price (and with a long horizon it eventually would), with a short horizon, if nobody else is studying *b*, information obtainable from studying *b* will not affect the asset price. Thus all short-term traders operate only on information obtainable from studying *a*.

This sort of informational inefficiency caused by short-term trading mean that all traders may focus on one source of information. Froot, Scharfstein and Stein (1992) argue further that "the informational spillovers can be so powerful that groups of traders may choose to focus on very poor quality data, or even on completely extraneous variables that bear no relation at all to fundamentals". This is a Keynesian beauty contest where traders do not try to define the fundamental value of an asset. Instead they try to guess what other traders think the price for the asset should be.

Hirshleifer, Subrahmanyam and Titman (1994) make another complementary point regarding investigative herding. In their model investors receive information relating to an asset at different times. Competent (or lucky) investors receive this information before others. This gives them the ability to buy (or sell) the asset before the information spreads further and is fully reflected in the asset price. The earlier they receive the relevant information, the more profit there is to be made, since the new information is not yet reflected in the asset price. Now, if many investors study the same asset, then the new information will spread quickly and the profits can be realized faster and with more certainty. On the other hand, if no other investors follow the same asset, then the new information would never be reflected in the asset price and there would be no profits available. Hirshleifer, Subrahmanyam and Titman (1994) argue that this leads to investors focusing on certain assets on the expense of others.

Reputational herding occurs when an investors decides to ignore their private information and instead copy the actions of previous investors, similar to what happens in information cascades. However, reputational herding models have an added layer of mimicking other investors (Graham 1999). These reputational herding models, such as the one developed by Scharfstein, Stein et al. (1990), take into account reputational benefits obtainable through herding. Because of this, reputational herding applies primarily to analysts or professional managers whose compensation or career development is linked to their perceived ability.

In the model developed by Scharfstein, Stein et al. (1990), and also used by Graham (1999), professional analysts can be either smart of dumb. Smart analysts receive informative signals about the value of an investment, while dumb analysts receive purely random signals. No analyst knows whether they are smart or dumb. All analysts can observe the investment decision made by others before them. This means that smart managers receive correlated signals, while dumb managers receive purely random signals. This means that if an analyst copies the behaviour of other managers, he implies to others (peers, employers) that he, too, has received a signal that is correlated with the signals of others, meaning he is more likely to be smart. In contrast, if an analyst takes a contrarian position, they are more likely to be dumb (all other things being equal). This creates a "sharing-the-blame" effect. If the analyst is a part of the majority that invests, and the investment is profitable, the analyst is perceived to be smart and they are rewarded. On the other hand, if the investment turns out to be unprofitable, the analyst can "share the blame" and claim that all the analysts who invested are smart, but this outcome was simply unpredictable. This way the analyst does not seem to be dumber than the others. (Scharfstein, Stein et al. 1990)

Style investing (also known as investment trends or fads) means that some investors group risky assets into different styles (categories) and move their funds between these styles (Barberis and Shleifer 2003). Assets in a single category usually share some common characteristics. These categories might be e.g. government bonds, small-cap stocks or real-estate. Barberis and Shleifer (2003) state that interest in style investing has risen over the years and many funds now follow a certain investment style, such as investing in growth stocks, value stocks or technology stocks.

An investment style is usually born when good fundamental news regarding securities within this style are released. This could mean e.g. that many technology companies release financial statements reporting much better profits and growth than anticipated. This would trigger the relocation of funds from other styles towards technology stocks. The problem is, funds will move to all stocks that are grouped as "technology", even though the fundamentals may differ completely. Style investing creates common factors in the *returns* of assets that have been grouped in the same style, even though there might be *no common factors in underlying cash flows*. This means that the returns of fundamentally different assets may converge just because they have been classified into the same style, while the returns of fundamentally similar assets may disperse just because they have been grouped into different styles. (Barberis and Shleifer 2003) This is problematic, since high past returns will direct funds towards a certain style, which will further raise the prices of assets belonging into that style, driving even more funds to the same style. According to Barberis and Shleifer (2003) the style will finally collapse, either because of arbitrage or because of bad new about fundamentals.

One possible explanation for individuals' tendency to focus on certain styles may be individual investors *limited attention span*. Barber, Odean and Zhu (2009) find that individual investors' purchases are concentrated on "attention-grabbing" stocks, i.e. stocks with abnormally high trading volumes in recent history. When deciding which stocks to buy, individual investors face a search problem: there are thousand of stocks to choose from. To simplify this search, individual investors may concentrate on stocks that have recently grabbed their attention. If a certain investment style is popular (e.g. technology stocks), these stocks will receive abnormally high trading volumes and may catch the eye of individual investors. This may increase the buys done by individual investors towards stocks that represent the aforementioned style.

Characteristic herding means that investors have a preference or aversion towards securities with certain kind of characteristics. Falkenstein (1996) studied the preferences of US open-end mutual funds. The most important findings are as follows: the funds prefer stocks with high volatility and high liquidity, but show an aversion towards low-price stocks and small firms. Funds also avoid companies with little information available about them. Gompers and Metrick (2001) have also studied the preferences of institutional investors. They found that large institutions prefer stocks with larger market capitalizations, more liquidity, higher book-to-market ratios and lower returns for the previous year. They also found that large institutions have almost doubled their share of the market between 1980 and 1996. A related model is the search cost model developed by Vayanos and Wang (2007). Here investors flock towards stocks with low search costs, meaning that investors trade stocks where buyers and sellers are quickly matched with each other. This drives investors to focus on more liquid stocks.

This kind of behaviour from institutional investors cannot always be classified as herding. If institutions have observed the investment decisions of other institutions and decided to follow their trades, that behaviour is herding and it would lead to institutions buying stocks with similar characteristics. It is also possible that institutions have decided to buy stocks with similar characteristics *independent* of each other. For example, buying small and illiquid stocks is difficult for institutional investors because making bids on large amounts of illiquid stocks would increase the stock price and result in the investor paying a price higher than the current market price. Thus large institutional investors may not be able to trade smaller stocks without affecting the stock price and incurring extra costs for themselves.

Gompers and Metrick (2001) also found that the institutional investors' preference for large and liquid stocks, combined with the rising total market share of institutions, has increased the demand for such stocks and thus resulted in price increases and high returns for large stocks relative to small stocks between 1980 and 1996.

Using a **momentum investment strategy** an investors buys and sells stocks based on their past returns, meaning that investors buy past winners and sell past losers (Bikhchandani and Sharma 2000). If markets are thought to be efficient, this kind of behaviour is not rational, since all available information is thought to be included in the prices. Thus, after a price increase, no further price increase is to be expected (until new positive information is published). According to Bikhchandani and Sharma (2000) this kind of behaviour can exaggerate price movements and increase volatility. On the other hand, if incorporating new information into prices takes some time, this kind of behaviour can be rational.

A lot of empirical research has been done about momentum investment strategies. For example, Grinblatt, Titman and Wermers (1995) found that mutual funds show a tendency to buy past winners. However, the tendency to sell past losers was less pronounced. They also found that funds who buy past winners realize higher returns. Froot, O'connell and Seasholes (2001) found that international portfolio flows are strongly influenced by past returns, which is consistent with international institutional investors using momentum investment strategies. They also found that these inflows of money have forecasting power for future returns, meaning that momentum investment strategies may create higher future returns.

Using a data set spanning from 1975 to 1994 Wermers (1999) found that mutual funds show evidence of momentum investing strategies. The findings also indicate that using momentum investment strategies is rational, since he found no evidence of subsequent return reversals. Similar to Grinblatt, Titman and Wermers (1995), he found that buying past winners is more common than selling past losers. Their findings support the view that momentum investment strategies may be rational and drive prices towards their fundamental values.

Using a more recent data set from 1994 to 2003, Brown, K. D. Wei and Wermers (2013) get an opposite result. They found that buy herding by institutional investors increases prices during the herding quarter, but the prices decrease during the following year. Similarly, stocks sold by institutional investors show a price decrease during the herding quarter but show a reversal in returns during the following year. They also found that funds are more likely to herd on the sell-side than on the buy-side, contrary to earlier findings. These findings suggest that momentum investment strategies may not be rational, since price increases following excess buying are followed by return reversals. Brown, K. D. Wei and Wermers (2013) do not dispute earlier, contrasting results, but instead suggest that mutual fund trading may have become destabilizing in more recent times.

Individual investors may also participate in momentum trading. Barber, Odean and Zhu (2009) present possible psychological reasons for this: the representativeness heuristic and the limited attention of individual investors. The representativeness heuristic means that investors expect a small sample and a short time-series of data to represent the underlying distribution. This will lead to investors over-weighting recent returns in forecasting future returns, directing investors towards buying recent winners. The limited attention span of investors means that when choosing which stocks to buy, investors will focus on attention grabbing stocks that have abnormally high recent trading volumes. As mentioned in the chapter describing style investing, focusing on attention grabbing stocks with high trading volumes (and high recent returns) will cause investors to buy these stocks. While the reason for purchasing such stocks isn't necessarily high past returns themselves, this psychological factor may direct individual investors towards buying such stocks nevertheless.

2.3 Measuring herding

Measuring herding is a difficult task. Herding as a phenomenon is well established and many possible reasons for herding have been presented. But when trying to empirically measure herding one faces multiple problems. One problem is that is it difficult, maybe even impossible, to tell apart "spurious" herding from "real" herding just by looking at the data (Bikhchandani and Sharma 2000). Even if an abnormally large amount of investors is trading on the same side of the market, this may not be evidence of herding. This trading might be caused by a change in fundamentals, to which all investors are reacting simultaneously and independently. On the other hand, these changes in fundamentals might justify a smaller change in the price of asset than actually happened. In this case, excess buying/selling because of herding may be partially to blame. To separate which part of trading is based on fundamentals and which part is because of herding is, in practice, impossible.

Another problem with measuring herding is related to deducing the reason behind herding. Even though many possible reasons for herding have been identified, telling these apart by looking at the data is challenging to say the least. Even if herding itself is successfully found and measured, the reason behind this herding can't reliably be deduced from the data types most commonly in use. For example, how would one tell apart herding caused by information cascades and herding caused by reputational concerns if the end results looks identical in the data?

One reason for these difficulties in measuring herding is the data available. Many articles (e.g. Brown, K. D. Wei and Wermers 2013; Grinblatt, Titman and Wermers 1995; Lakonishok, Shleifer and Vishny 1992; Wermers 1999) use quarterly reported data of mutual fund holdings. While this data does tell about the net purchases/sales done between quarters, it tells nothing about the reasoning behind the investment decisions. Indeed, Bikhchandani and Sharma (2000) state that "One cannot distinguish between different causes of herd behavior directly from the analysis of a data set on asset holdings and price changes since it is difficult, if not impossible, to discern the motive behind a trade that is not driven by “fundamentals””. However, they suggest it may be possible to try and filter out reactions to public information by allowing for changes in fundamentals. If after this filtering herding is still found, then the reason might e.g. information cascades, reputational herding or some other identified reason for herding. But which kind of herding? Because no data on the subjective reasoning of each individual investor (professional or non-professional) is available, differentiating between different kinds of herding is difficult (Bikhchandani and Sharma 2000).

As previously stated, many articles have used data about mutual fund holdings or about the holdings of other institutional investors. The herding of individual investors has been investigated less, at least partly because data on the trades and holdings of individual investors is not widely available. Some exceptions to this are Barber, Odean and Zhu (2009), Dorn, Huberman and Sengmueller (2008) and Grinblatt and Keloharju (2000).

Because data on the actual reasons behind investors’ investment decisions is not available, practically all empirical research focuses on finding abnormal clustering of trades with purely statistical methods. Thus there is no clear link between the theoretical discussion about the reasons behind herding and the empirical measurement of herding. Graham (1999) and Wermers (1999) can be mentioned as exceptions, since they link their empirical findings to momentum investment strategies and reputational herding models. Many studies do not try to differentiate "spurious" herding from "real" herding. This is at least partly because it is hard to say what constitutes as a change in fundamentals and in addition it is challenging to measure and quantify changes in fundamentals. Also, as previously stated, even if changes in fundamentals are observed, it is difficult to separate which part of e.g. price changes is justifiable with changes in fundamentals and which part is attributable to herding. (Bikhchandani and Sharma 2000)

2.4 Herding measures

In the market as a whole, all trades net to zero. For every stock sold there is a stock bought. Thus, if the data studied is about trades or changes in portfolio composition, market-wide herding cannot be observed. Instead herding can be measured within a *subgroup* of investors. The most common herding measure used with this type of data is the measure developed by Lakonishok, Shleifer and Vishny (1992) (LSV from here on after). Some measures also try to capture market-wide herding. Examples of these measures are the CSSD developed by Christie and Huang (1995) and the CSAD developed by Chang, Cheng and Khorana (2000).

2.4.1 The LSV herding measure

One of the most commonly used herding measures is the so-called LSV herding measure developed by Lakonishok, Shleifer and Vishny (1992). It attempts to capture how many investors more were buying (or selling) stocks compared to the expected amount of buyers (or sellers). The basic formulation for the measure is as follows:

$$H = \left| \frac{B}{B+S} - p \right| - \mathsf{AF}, \tag{2.1}$$

where

$$\mathsf{AF} = \mathbb{E}\left| \frac{B}{B+S} - p \right|. \tag{2.2}$$

B is the amount of net buyers of a specific stock within a specified time interval, while *S* is the amount of net sellers of the same stock withing the same interval. p is the proportion of buyers (of all active investors) of any stock within the same time interval and *AF* is the adjustment factor. This herding measure is computed for each combination of stock and time intervals and then averaged across different subgroups (Lakonishok, Shleifer and Vishny 1992).

In a given time interval (e.g. quarter, month or day) there should not necessarily be the exact same number of buyers and sellers of each stock. The amount of net buyers over all stock in a single time interval varies, thus a different *p* is calculated for each interval. The null hypothesis is that there is no herding. The adjustment factor *AF* is included, because even if there is no herding, the absolute value of *B/(B+S) - p* is greater than 0. *AF* is the expected value of this absolute value term and it is calculated using a binomial distribution. The actual process through which the LSV measure is calculated is presented next.

Let us go through the terms in the formulas using examples. Let's set our time period as week 26 in 2005 and the stock we are calculating the LSV measure for is Nokia. Now, *B* is the number of investors who are net buyers of Nokia during week 26 of 2005. *S* is the equivalent number of net sellers. *p* is the number of net buyers of **all stocks** during week 26 in 2005 divided with the number of active investors in week 26 in 2005. *p* gives us a reference level of sorts; has the Nokia stock seen more buying or selling than the average stock during that same time? *p* is calculated again for each week. Then *p* is subtracted from $\frac{B}{B+S}$. If Nokia has been bought in excess amounts, this difference is positive. If, on the other hand, Nokia has been sold more than the average stock, this difference is negative. Both may be evidence of herding, and because of this, the absolute value of this difference is considered.

The adjustment factor (AF) is included because some deviation from the average buying propensity *p* is always expected. The adjustment factor is the expected value of the previously presented absolute value term. The method for calculating the AF is presented well in Jones, Lee and Weis (1999) and summarized here.

Let us suppose that in a given time period (month, week, quarter...) investors bought and sold stocks in similar amounts, so that *p* = 0.5. Then let us suppose that the Nokia stock

Table 2.1. *Calculating the adjustment factor with 2 active investors. Taken from Jones, Lee and Weis (1999).*

# of buys	Probability	Value	Product
	$\frac{(n! * (p)exp(k) * (1-p)exp(n-k))}{([(n-k)! * k!])}$	$\left\lvert \frac{B}{(B+S)} - p \right\rvert$	
0	0.25	|0/2-0.50|	0.125
1	0.50	|1/2-0.50|	0.000
2	0.25	|2/2-0.50|	0.125
AF			**0.250**

had only two investors active in it during this time period. The adjustment factor for the Nokia stock is the expected value of $\frac{B}{B+S} - p$ and it's calculation is presented in table 2.1.

As previously stated, the adjustment factor decreases when the number of active investors goes up. To demonstrate this, let us now suppose that the are five active investors in the stock. This calculation is presented in 2.2.

Table 2.2. *Calculating the adjustment factor with 5 active investors. Taken from Jones, Lee and Weis (1999).*

# of buys	Probability	Value	Product
	$\frac{(n! * (p)exp(k) * (1-p)exp(n-k))}{([(n-k)! * k!])}$	$\left\lvert \frac{B}{(B+S)} - p \right\rvert$	
0	0.03125	|0/5-0.50|	0.016
1	0.15625	|1/5-0.50|	0.047
2	0.31250	|2/5-0.50|	0.031
3	0.31250	|3/5-0.50|	0.031
4	0.15625	|4/5-0.50|	0.047
5	0.03125	|5/5-0.50|	0.016
AF			**0.188**

As can be seen from comparing tables 2.1 and 2.2, when the number of active investors increases, the adjustment factor decreases. This was also noted by Frey, Herbst and Walter (2014) and Wylie (2005). It is noteworthy that, if Frey, Herbst and Walter (2014) are correct and the AF over corrects, and is more pronounced with smaller amounts of active traders, and small cap stocks have a smaller amount of active traders, it might be that LSV herding results for small cap stocks are systematically too low. It is important to keep this in mind when considering the low amount of active institutional investors, especially in small cap stocks.

With all these terms now calculated, we can solve the LSV herding measure for a single stock-week combination (Nokia during week 26 in 2005). This procedure is now repeated for all stock-week-combinations, meaning that 103 stocks and 4 years result in over 21 000 separate stock-week observations. After the LSV measure for each stock-week is calculated, further classification into buy herding measure (BHM) and sell herding measure (SHM) as in e.g. Wermers (1999) and Hsieh (2013) is made. They are conditional measures and defined as follows:

$$BHM = H|\frac{B}{B+S} > p \tag{2.3}$$

and

$$SHM = H|\frac{B}{B+S} < p. \tag{2.4}$$

This is done to simplify later regression model calculations. The regressions will model the relationship between the LSV measure and both past and future returns and volatilities. Here having only one herding measure without regard to its direction would be problematic: if, for example, high past returns lead to buy herding and low past returns lead to sell herding, this would not be accurately captured in the regression model since both low and high returns would result in a high herding measure.

The herding measure given by the LSV measure is quite easily interpreted. For example, in their original paper Lakonishok, Shleifer and Vishny (1992) get a mean herding measure of 2.7%. This means that if *p*, the average amount of buyers, was 0.5, then in the average stock 52.7 % of the subgroup of investors under investigation were changing there holdings in the same direction.

The mean herding measure of 2.7 % by Lakonishok, Shleifer and Vishny (1992) can be considered quite low and at most evidence of weak herding among money managers between 1985 and 1989. Similar low results have been gotten by e.g. Grinblatt, Titman and Wermers (1995) and Wermers (1999). A somewhat higher result of 3.3 % has been reported by Brown, K. D. Wei and Wermers (2013) when using data from 1994 to 2003. Significantly higher results have been reported by e.g. Kim and S.-J. Wei (2002) and Choe, Kho and Stulz (1999) regarding international investors, 13.2 % and over 20 %, respectively. Hsieh (2013) reports numbers of 16.92 % for institutional investors and 6.73 % for individual investors.

The LSV herding measure has also received some criticism. Bikhchandani and Sharma (2000) state that LSV uses only the number of investors to assess herding, with no regard to how much stock they buy or sell. In a situation where there is an equal amount of buyers and sellers, but buyers demand a higher amount of stocks, LSV would not find any herding while in reality there is buy-side herding. They also criticize that LSV cannot measure whether the exact same investors continue to herd over time. They also point that choosing the herding interval (e.g. quarter, month, week) is very important. If the average time between trades is a quarter or more, then choosing a quarter as the time interval if justified. If, on the other hand, trades are done more often (e.g. monthly), then a quarter is too long a time interval to capture herding. Since many studies have only quarterly data available, this might be an issue. For example, both Kim and S.-J. Wei (2002) and Choe, Kho and Stulz (1999) study herding during the Korean crisis in 1996 - 1997. Kim and S.-J. Wei (2002) use monthly intervals and get a herding measure of 6.23 % for foreign investors. Choe, Kho and Stulz (1999) use daily intervals and get a results of over 20 % for foreign investors. Wylie (2005) points out that the LSV measure expects short selling to be possible for all investors, which might not be the case. Wylie (2005) also says that the LSV ignores money inflows to mutual funds. This inflow of new money forces managers to make new buys, which might be seen as an increase in the herding measure. On the other hand, even if fund managers are forced to make new buys, they are not forced to buy the same stocks as other managers. Bikhchandani and Sharma (2000) also say that while the LSV is called a herding measure, it actually measures the *correlation* of trading patterns for a group of trades. And while herding leads to correlated trading, the reverse may not be true.

The mathematical formulation of the LSV measure has also been under some scrutiny. Frey, Herbst and Walter (2014) state that while the LSV measure is well suited to study whether there is herding (i.e. is the herding measure $\neq 0$), it is less well suited in measuring the actual amount of herding. They claim that when there is herding, the LSV measure is biased downwards. They show that the adjustment factor in LSV overcorrects and leads to an understatement of herding. They also show that the bias decreases when the number of trades increases. Thus the bias is more pronounced with smaller trade amounts. Wylie (2005) has used the LSV measure to study herding in the UK, and he also observes that when the number of active investors increases, the herding measure increases too. It is possible that at least some of this observed increase in the herding measure stems from the bias in the adjustment factor. To correct this bias, Frey, Herbst and Walter (2014) suggest using square values instead of absolute values. They have defined the new measure:

$$\text{FHW}^2_{j,t} = \left(\left(\frac{B}{B+S} - p \right)^2 - E\left[\left(\frac{B}{B+S} - p \right)^2 \right] \right) \frac{I_{j,t}}{I_{j,t}-1}, \tag{2.5}$$

where $I_{j,t}$ is the amount of active investors for stock *j*. For a time period of *t - 1; t* and universe of *J* stocks, the average *FHW* measure is

$$\text{FHW}_t = \sqrt{\frac{1}{J} \sum_{j=1}^{J} FHW^2_{j,t}}. \tag{2.6}$$

Frey, Herbst and Walter (2014) claim that their measure is superior to the LSV measure in terms of the mean square error for as little as only 20 observations and 5 managers trading a stock. They also point out that their measure is less well suited to testing the existence of herding. They suggest first using the LSV measure to test for the existence of herding, and if herding if found, then use their measure to test the level of herding. However, Bellando (2010) tested the FHW measure more and claims that it is unbiased only with certain assumptions made by Frey, Herbst and Walter (2014). He claims that as soon as the probability of no herding is not 0 or when there is some asymmetry, the FHW measure is biased *upward*. Thus it seems that the LSV measure is in some cases biased downwards, while the FHW measure is biased upwards. Merli and Roger (2013) comments that when using these measures it is impossible to know the true value of herding; however we know that the true value of herding is somewhere in between the results given by LSV and FHW.

2.4.2 The CSSD and CSAD herding measures

Both the CSSD (cross-sectional standard deviation) and the CSAD (cross-sectional absolute deviation) are measures which are based on the deviation of stock returns. The CSSD was introduced by Christie and Huang (1995) and the CSAD by Chang, Cheng and Khorana (2000). Both measures try to find herding by comparing the actual returns to what would be expected if investors were fully rational and there would be no herding.

Christie and Huang (1995) were the first to study herding through the dispersion of stock returns. They claim that during periods of markets stress and large price changes the dispersion of returns should increase if there is no herding. This is because different assets differ in their sensitivity to the market return. In the capital asset pricing model this sensitivity is described with β. For example, if the market return is 10 %, then a stock with a β of 2 should see a return of 20 %. Similarly, a different stock with a β of 0.5 should see a return of 5 %. Christie and Huang (1995) thus state that if assets are priced rationally, then during large changes in market returns the dispersion of individual asset returns would increase. Correspondingly, in the presence of herding, investors would reject these rational models and the returns for all individual assets would converge towards the market return. This would mean that the dispersion of returns decreases. Thus they conclude that a decrease in the dispersion of individual returns during periods of market stress and large changes in market return would be evidence of herding. In their research Christie and Huang (1995) focus on periods of market stress because they deduce that herd behavior is most likely to emerge during these periods of unusual market movements. Mathematically the CSSD herding measure is defined as

$$\text{CSSD}_t = \sqrt{\frac{\sum_{i=1}^{N} \left(r_{i,t} - \overline{r_t}\right)^2}{N-1}}. \tag{2.7}$$

Here *N* is the number of stocks being observed, r_i is the return of a individual stock during time *t* and $\overline{r_t}$ is the average return of all *N* stocks during time *t*.

The CSSD measure also has some drawbacks, some of which are mentioned by Christie and Huang (1995) themselves. While herding might lead to a small CSSD value, a small CSSD value is not necessarily evidence of herding. For example, if there is no new information about the observed stock during time *t*, the dispersion in returns would be low without any herding taking place. Also, in their research Christie and Huang (1995) compare especially the dispersion in periods of "extreme" markets returns to the average dispersion. They have defined "extreme" as being in the 1 % or 5 % tails of all observed trading days. This definition of extreme is arbitrary, as stated by the researchers themselves. Hwang and Salmon (2004) also point out that focusing on these periods of "extreme" returns may exclude some herding. For example, there might be large price changes in different sectors during the same time (e.g. financial institution stocks going down while technology stocks going up) which would balance each other out when looking at the market as a whole. Thus there may be significant reallocation of funds without large changes in the average market return. Focusing only on "extreme" days would exclude these examples of herding (Hwang and Salmon 2004). Also Richards (1999) criticizes the CSSD (and the soon to be presented CSAD) by stating that they only look for a certain kind of herding and only using data on returns. He states that there may be other types of herding present which the CSSD measure will not pick up. Thus, according to Richards (1999), the absence of evidence on herding gained with the CSSD should not be taken as evidence of there being no herding at all.

The CSAD is a related measure developed by Chang, Cheng and Khorana (2000). It is based on the same idea as CSSD and looks for differences in the dispersion of returns. Chang, Cheng and Khorana (2000) go further than Christie and Huang (1995) and suggest that the return dispersion does not only *increase* with the market returns, but it *increases linearly*. While a decrease in return dispersion with large average price movements is still considered evidence of herding, also a decrease in the rate of dispersion increase will also be considered evidence of herding. So even a decrease in the expected linear relation between return dispersion and the market return could be evidence of herding. Mathematically the CSAD is defined as

$$CSAD_t = \frac{\sum_{i=1}^{N} |r_{i,t} - \overline{r_t}|}{N}. \tag{2.8}$$

Here *N* is the number of stocks being observed, $r_{i,t}$ is the return of a individual stock during time *t* and $\overline{r_t}$ is the average return of all *N* stocks during time *t*.

Neither the CSSD or the CSAD are herding measures in themselves. Instead, they are used as a part of a regression. Christie and Huang (1995) use the following regression in their paper:

$$CSSD_t = \alpha + \beta_1 D_t^L + \beta_2 D_t^U + \epsilon_t, \tag{2.9}$$

where $D_t^L = 1$ if the market return on day *t* is the extreme **lower** tail on the distribution and 0 otherwise and

$D_t^U = 1$ if the market return on day *t* is the extreme **upper** tail on the distribution and 0 otherwise.

Chang, Cheng and Khorana (2000) utilize the following regressions:

$$CSAD_t^{UP} = \alpha + \gamma_1^{UP} \left| \overline{r_t^{UP}} \right| + \gamma_2^{UP} \left(\overline{r_t^{UP}} \right)^2 + \epsilon_t, \tag{2.10}$$

$$CSAD_t^{DOWN} = \alpha + \gamma_1^{DOWN} \left| \overline{r_t^{DOWN}} \right| + \gamma_2^{DOWN} \left(\overline{r_t^{DOWN}} \right)^2 + \epsilon_t, \tag{2.11}$$

where $\left|\overline{r_t^{UP}}\right|$ (or $\left|\overline{r_t^{DOWN}}\right|$) is the absolute value of equally weighted returns of all observed assets on days when the market is up (or down).

Using these measures Christie and Huang (1995) find no evidence of herding in US stocks during periods of market stress in 1925 to 1988. Using data for multiple markets, Chang, Cheng and Khorana (2000) find significant evidence of herding in South Korea and Taiwan, partial evidence for herding in Japan and no evidence for herding in the US and Hong Kong. The advantages of the CSSD or CSAD measures lie in the good availability of data and on the fact that they are based on rational pricing models. The drawbacks include the ability to find only certain kind of herding and the inability to pinpoint herding to certain investor groups or securities.

3 DATA AND METHODOLOGY

The empirical portion of this Master's Thesis consists of measuring herding using the LSV herding measure. This section discusses the research philosophy behind this thesis, and presents the data and methods used.

The research philosophy underlying this thesis is most closely linked to positivism. This thesis focuses on what can be measured and tries to produce credible and meaningful data. This is quite clear in the selection of the used method: using quantitative data which is studied through statistical methods. The results are compared with earlier research and existing theories about herding. As the researcher, I aim to minimize the effect of my own values in the results. It must be noted that this is not completely possible, since some choices always have to be made. For example, the choice of which herding measure to use and which time period to choose reflect my own values and beliefs, even though I have tried to base these choices in earlier research. Also, because this study takes place in the field of behavioral finance, the role of human interaction and psychology cannot be passed. This mean that the are also critical realist research features present. As mentioned in Saunders (2016), "critical realist research therefore focuses on providing an explanation for observable organisational events by looking for the underlying causes and mechanisms through which deep social structures shape everyday organisational life". When researching the reasons behind herding, some social aspects always come into play.

This thesis' approach to theory development has features of both the inductive and the abductive approach. Saunders (2016) states that "if your research starts by collecting data to explore a phenomenon and you generate or build theory (often in the form of a conceptual framework), then you are using an inductive approach". This is at least partly the case in this thesis. On the other hand, Saunders (2016) also states that "abduction begins with the observation of a 'surprising fact'; it then works out a plausible theory of how this could have occurred". In this thesis this is also true, and drawing a line between the use of these approaches is very difficult and maybe even unnecessary.

This thesis has a quantitative approach. This is a mono method quantitative study, since only one data collection technique is used (Saunders 2016). This study is also best characterised as descripto-explanatory, since it has two somewhat separate angles: first, has there been herding in the Finnish stock market and second, is there a link between returns (or volatility) and herding? This means there is a description of the situation (e.g. there has been herding in the Finnish stock market) and also maybe an explanation for this (e.g.

there is a correlation between recent returns and herding). The research strategy used mostly resembles an experiment strategy. As Saunders (2016) state, "the purpose of an experiment is to study the probability of a change in an independent variable causing a change in another, dependent variable". While the data used in this study has not been collected specifically for this use, a large part of this study focuses on the relationship between herding and stock returns, and herding and volatility. Time horizon wise, this study is a longitudinal one, since the data does not represent a "snapshot" of a certain time but instead allows us to study how herding has changed as time went on.

3.1 The data

The data set in use is provided by Euroclear Finland Oy (previously Finnish Central Security Depository). The data set contains daily level records of investor's trades and portfolios, including all Finnish households, Finnish institutions and some foreign institutions. The data is on a transaction level, so each purchase and sale is separated, even if they happen within the same day. The daily level records are duplicates of the official certificates of ownership and trades and thus reliable. This data spans 15 years, from 1995 to 2009. In addition to this, there is additional data spanning 6 years, from 2010 to 2016. This data is also on a daily level, but transactions are reported as net transactions, where the purchases and sales of the same stock are summarized.

This data reports the shareholdings of all Finnish investors, both retail and institutional, who have traded in Finland. The book entry system requires registration of holdings for all Finnish individuals and institutions. Foreign investors are partially relieved from this registration because they can choose to be registered under a nominee name. If they register under a nominee name, the foreign investors cannot be separated from each other. The same data from different time periods has been used in many articles, including Baltakys, Baltakienė et al. (2019), Baltakys, Kanniainen and Emmert-Streib (2018), Grinblatt and Keloharju (2000), Grinblatt and Keloharju (2001) and Siikanen et al. (2018).

The data set is very comprehensive and includes a lot of information relating to each trade. The data includes investor-specific ID's, the ISIN code for the traded stock, information on the transaction itself (buy/sell, volume, trading date, registration date, price) and sector codes identifying households and different types of institutions. For individual investors the data also includes date of birth, gender, postal code, language and tax payer status. Much of this data is not relevant in the scope of this thesis. Some dummy data with relevant columns is presented in table 3.1.

The data set is located in a SQL database. In addition to this transaction data some price data is also used. The price data is from Nasdaq and it includes daily closing prices for all stocks listed in the Helsinki stock exchange from 1995 to 2009 (or for a shorter period if the stock is listed or de-listed within this time period). The prices have been adjusted

Table 3.1. *Dummy data depicting relevant columns of the Euroclear data.*

owner id	sector code	holding type	isin	volume	transaction type	trading date
1001	100	1	FI0009502101	500	10	6.9.2004 10:30
1001	100	1	FI0009502103	350	20	11.10.2004 12:42
5001	200	1	FI0009501234	300	10	6.9.2004 10:24

for splits and dividends.

3.1.1 Arguments for data cuts chosen

The main goal of this thesis is to use the LSV herding measure to study whether there has been herding in Finnish stocks between 1.7.2005 and 30.6.2009. In addition to this main goal, analyses on the relation between herding and stock returns, and herding and volatility are conducted. Because of the large size of the data set in use, some subset of the data must be chosen for further analysis.

The time period of 1.7.2005 - 30.6.2009 was chosen mainly for two reasons. First, it includes a time period prior to the financial crisis of 2007 - 2008 and a time period during the crisis. This makes it possible to compare herding measures between non-crisis and crisis periods. The period from 1.7.2005 to 30.6.2007 is considered to be the pre-crisis period, while 1.7.2007 - 30.6.2009 is considered to be the crisis period. To clarify the difference between these two periods in the stock markets, figure 3.1 shows the development of the OMX Helsinki 25 stock index during these periods. Second, this time period fits well with some qualities of the data. The main table in the data set in use spans from 1995 to 2009, but it has some problems with data quality during the period of 2000 - mid 2007. For this period there is a different table with better quality data. Choosing these two periods (pre-crisis and crisis) allows for simpler extraction of data. All of the transaction data for the pre-crisis period (1.7.2005 - 30.6.2007) can be extracted from the data set containing good quality data for the period of 2000 - 30.6.2007. In the same manner, all of the transaction data for the crisis period (1.7.2007 - 30.6.2009) can be extracted from the larger data set, because for this time period the data has good quality.

Another time-related question is what time period is chosen for calculating the LSV measure. As mentioned in section 2, the choice herding interval may significantly affect the results (Bikhchandani and Sharma 2000; Choe, Kho and Stulz 1999; Kim and S.-J. Wei 2002). Many previous studies have used a quarter as the herding interval; this is mainly due to the fact that they have had only quarterly data available. The data set in use for this thesis would theoretically allow the use of a single day as the herding interval. But which interval length is most suited for finding herding?

If herding is defined as investors copying the actions of other investors, then investors have to be aware of the actions of other. In this regard a day might be too short. It seems

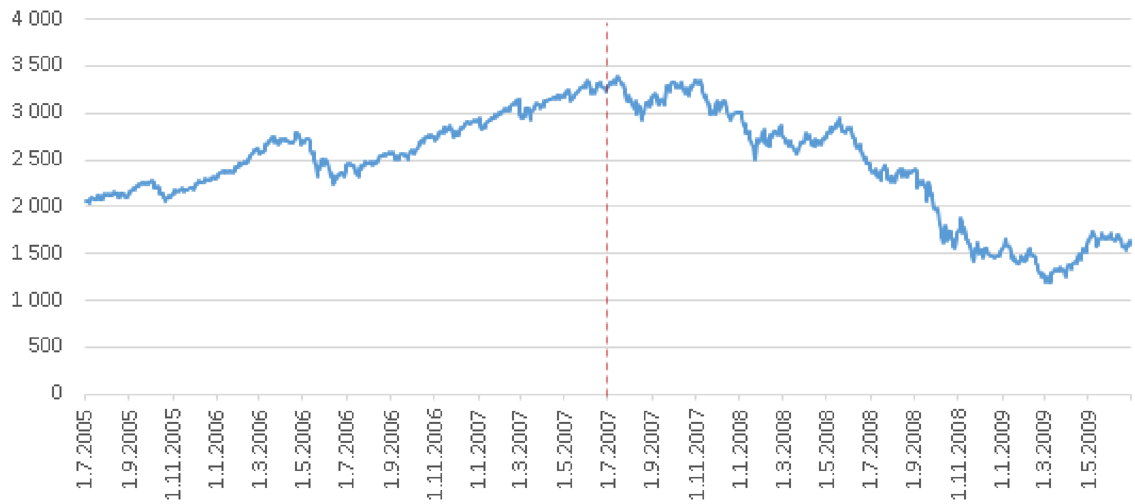

Figure 3.1. *The development of the OMX Helsinki 25 index from 1.7.2005 to 30.6.2009. The red line marks the transition from the pre-crisis period into the crisis period.*

unlikely that institutional investors would follow each others' trades on daily basis and then execute their own trades during the same day. The same may apply for households: it seems likely that the information about the trades of others (peers, institutions) takes some time to diffuse into the individual investor population. Households may also face some constraints regarding liquidity: they may not have cash on hand every day to make purchases. For these reasons a day seems to be too short of an interval to accurately measure herding. A quarter, on the other hand, may be too long. Because the LSV herding measure uses **net** buyers and sellers, investors who first buy and then sell (or vice versa) the same amount of stock during the herding interval are not registered by the measure. It is also not impossible to imagine a scenario where, during a single quarter, investors first herd to buy a certain stock and then herd to sell it. This would not be picked up by the LSV measure, even though significant herding might have actually occurred. Institutional investors may also re-balance their portfolios more than once a quarter, which also makes a quarter too long of a time interval to study herding.

For this study a week is chosen as the herding interval. It can be considered long enough so that information on the trades of others is diffused into the investor population, but short enough so that trades might have an actual correlation. Choosing a week as the herding interval also decreases the probability of missing intra-period herding. The average number of weekly active investors for the whole study period of 1.7.2005 - 30.6.2009 is 11 792. This is definitely enough to get reliable results. Shorter herding intervals are of more interest also because most data sets do not contain data at this level. This creates an excellent opportunity to study herding with shorter intervals and perhaps create new understanding about the optimal interval for studying herding. However, just to be sure some sensitivity analysis was conducted with different period lengths. Another suitable period considered for this study was a month. When the calculations were made with a month as the herding interval no significant differences in the amount of herding was found compared to the results obtained when using a week as the herding interval.

The choice of which stocks to focus on is also important. Some previous studies have chosen to focus on e.g. 16 largest stocks (Grinblatt and Keloharju 2000) to avoid skewed results possibly caused by the illiquidity of smaller stocks. In this thesis, however, no such limitations are used. The stocks to be studied are chosen as follows: first, all stocks that are traded in the Helsinki stock exchange on December 2006 are listed. This date is chosen because of some limitations of data; the Nasdaq website does not contain records older than this. From this pool of stocks some additional omissions have to be made for each of the periods (pre-crisis and crisis). For the pre-crisis period (1.7.2005 - 30.6.2007), we only include stocks that are listed for that whole period. If a stock is listed or delisted during that period, it is omitted. For the crisis period (1.7.2007 - 30.6.2009) we start with the pool of stocks used for the pre-crisis period and do similar omissions. This means a few more stocks are omitted from the crisis period because of their delisting during that period. Second, for the regression analysis on the relationship between the LSV herding measure and stock returns, data on the returns of one previous year and one future year is needed. Since the study period is 1.7.2005 - 30.6.2009, this means that for the regression analysis price data is needed from 1.7.2004 to 30.6.2010. Because of this additional limitation, some stocks are omitted from the study.

With these omissions made, we are left with 103 stocks for the pre-crisis period and 99 stocks for the crisis period. Even with some omissions we are still left with a good number of stocks representing companies of different sizes and industries. The 28 large cap stocks are listed in table 3.2 and the 37 mid cap stocks in table 3.3. 38 small cap stocks were included in the pre-crisis period from 1.7.2005 to 30.6.2007 but only 34 for the crisis period from 1.7.2007 - 30.6.2009. These stock are listed in table 3.4.

As previously stated, the LSV herding measure cannot identify market wide herding. A subgroup of investors must be chosen to properly measure herding. Bikhchandani and Sharma (2000) say that to examine herding, one should find a sufficiently homogeneous group of investors. If investors within this group act similarly and face similar decision problems, then herding is more likely to arise. This group cannot also be too large compared to the whole market. If the group represents e.g. 80% of the market, then it is likely that buyers and sellers are both represented in this group in equal amounts. Common subgroups in previous literature include mutual funds (Grinblatt, Titman and Wermers 1995; Lakonishok, Shleifer and Vishny 1992), foreign investors (Kim and S.-J. Wei 2002) and individual investors (i.e. households) (Grinblatt and Keloharju 2000). Because of the sector codes present in the data, in this thesis it is possible to compare the herding of different subgroups of investors. The investor subgroups chosen are households and institutional investors. For households only one sector code is used in the data, so the separation of household investors is quite straightforward. However, multiple sector codes are used for different types of institutional investors. The following 13 investor groups are classified as institutional investors and are used in this study: financial and insurance corporations, Bank of Finland, other monetary financial institutions, deposit money corporations, money market funds, financial institutions practicing financial intermediation, other financial intermediaries, insurance corporations, social security funds, employment

Table 3.2. *The 28 large cap stocks included in this study and their respective industries.*

Stock	**Industry**	**Market cap**
Amer Sports Oyj	Consumer Discretionary	Large cap
Elisa Oyj	Telecommunication Services	Large cap
Fortum Oyj	Utilities	Large cap
Huhtamäki Oyj	Materials	Large cap
Kemira Oyj	Materials	Large cap
Kesko Oyj A	Consumer Staples	Large cap
Kesko Oyj B	Consumer Staples	Large cap
Metso Oyj	Industrials	Large cap
M-real Oyj A	Materials	Large cap
M-real Oyj B	Materials	Large cap
Nokia Oyj	Information Technology	Large cap
Nokian Renkaat Oyj	Consumer Discretionary	Large cap
Nordea Bank AB FDR	Financials	Large cap
OKO A	Financials	Large cap
Outokumpu Oyj	Materials	Large cap
Rautaruukki Oyj K	Materials	Large cap
Sampo A	Financials	Large cap
SanomaWSOY	Consumer Discretionary	Large cap
Stockmann Oyj Abp A	Consumer Discretionary	Large cap
Stockmann Oyj Abp B	Consumer Discretionary	Large cap
Stora Enso A	Materials	Large cap
Stora Enso R	Materials	Large cap
TeliaSonera	Telecommunication Services	Large cap
TietoEnator	Information Technology	Large cap
UPM-Kymmene Oyj	Materials	Large cap
Uponor Oyj	Industrials	Large cap
Wärtsilä Oyj Abp B	Industrials	Large cap
YITOyj	Industrials	Large cap

pension schemes, other social security funds, non-profit institutions serving households and state churches. In this study investments made under a nominee registered holding are excluded. Only normal holdings are considered.

Of these groups, households are of special interest. Few studies have been made about the herding of households, since data on household trades is rarely available. Some notable exceptions to this are Grinblatt and Keloharju (2000) and Kyrolainen and Perttunen (2003), both using Finnish data. During the study period there were 1 374 active weekly household investors. Similarly there were 119 active weekly institutional investors. Though it must be noted that when looking at institutional investors and small cap stocks,

Table 3.3. *The 37 mid cap stocks included in this study and their respective industries.*

Stock	Industry	Market cap
Aldata Solution Oyj	Information Technology	Mid cap
Aspo Oyj	Industrials	Mid cap
Atria Yhtymä Oyj A	Consumer Staples	Mid cap
Basware Oyj	Information Technology	Mid cap
CapMan Oyj B	Financials	Mid cap
Citycon Oyj	Financials	Mid cap
Comptel Oyj	Information Technology	Mid cap
Cramo	Industrials	Mid cap
Elcoteq SE A	Information Technology	Mid cap
Elektrobit Group Oyj	Information Technology	Mid cap
Exel Oyj	Materials	Mid cap
Finnair Oyj	Industrials	Mid cap
Finnlines Oyj	Industrials	Mid cap
Fiskars Oyj Abp A	Consumer Discretionary	Mid cap
Fiskars Oyj Abp K	Consumer Discretionary	Mid cap
F-Secure Oyj	Information Technology	Mid cap
HK Ruokatalo Group A	Consumer Staples	Mid cap
Ilkka-Yhtymä 2	Consumer Discretionary	Mid cap
KCI Konecranes Oyj	Industrials	Mid cap
Kyro Oyj Abp	Industrials	Mid cap
Lassila & Tikanoja	Industrials	Mid cap
Lemminkäinen Oyj	Industrials	Mid cap
PKC Group Oyj	Industrials	Mid cap
Ponsse 1	Industrials	Mid cap
Pöyry Oyj	Industrials	Mid cap
Raisio Oyj Vaihto-osake	Consumer Staples	Mid cap
Ramirent Oyj	Industrials	Mid cap
Rapala VMC	Consumer Discretionary	Mid cap
Scanfil Oyj	Information Technology	Mid cap
Sponda Oyj	Financials	Mid cap
Talentum Oyj	Consumer Discretionary	Mid cap
Tamfelt Etu	Industrials	Mid cap
Technopolis Oyj	Financials	Mid cap
Teleste Oyj	Information Technology	Mid cap
Vacon Oyj	Industrials	Mid cap
Vaisala Oyj A	Information Technology	Mid cap
Ålandsbanken Abp B	Financials	Mid cap

Table 3.4. *The 38 small cap stocks included in this study and their respective industries. The 4 stocks market with an asterisk (*) are only present in the pre-crisis period and omitted from the crisis period.*

Stock	**Industry**	**Market cap**
AffectoGenimap Oyj	Information Technology	Small cap
Amanda Capital Oyj	Financials	Small cap
Aspocomp Group Oyj	Information Technology	Small cap
Benefon S	Information Technology	Small cap
Biohit Oyj B	Health Care	Small cap
Biotie Therapies Oyj	Health Care	Small cap
Birka Line Abp B*	Consumer Discretionary	Small cap
Cencorp Oyj	Information Technology	Small cap
Componenta Oyj	Industrials	Small cap
Done Solutions Oyj	Industrials	Small cap
Efore Oyj	Industrials	Small cap
eQ*	Financials	Small cap
Etteplan Oyj	Industrials	Small cap
Evox Rifa Group Oyj*	Information Technology	Small cap
Incap Oyj	Information Technology	Small cap
Lännen Tehtaat Oyj	Consumer Staples	Small cap
Marimekko Oyj	Consumer Discretionary	Small cap
Norvestia Oyj	Financials	Small cap
Okmetic Oyj	Information Technology	Small cap
Olvi Oyj A	Consumer Staples	Small cap
Proha Oyj	Information Technology	Small cap
QPR Software Oyj	Information Technology	Small cap
Raute Oyj A	Industrials	Small cap
Ruukki Group Oyj	Industrials	Small cap
Satama Interactive	Information Technology	Small cap
Solteq Oyj	Information Technology	Small cap
SSH Communications	Information Technology	Small cap
Stonesoft Oyj	Information Technology	Small cap
Stromsdal Oyj*	Materials	Small cap
Suominen Yhtymä Oyj	Consumer Staples	Small cap
SysOpen Digia Oyj	Information Technology	Small cap
Tecnomen Oyj	Information Technology	Small cap
Tekla Oyj	Information Technology	Small cap
Tieto-X Oyj	Information Technology	Small cap
Tiimari	Industrials	Small cap
TJ Group Oyj	Information Technology	Small cap
Tulikivi Oyj A	Industrials	Small cap
Turvatiimi Oyj	Industrials	Small cap

there were weeks with only 1 active investor. Indeed, with small cap stocks and institutional investors, the average amount of active investors was only 13.6 per week. Such low investors amounts create problems for the reliability of results regarding the herding of institutional investors. The results are presented in chapter 4 where their reliability is also examined more closely. On the other hand, the groups (households and institutions) do not represent too large a part of the market. Thus it is more likely that they act together as a homogeneous group and their herding may differ from the average herding in the market.

It has also been suggested that company size may affect herding propensity. Bikhchandani and Sharma (2000) point out that herding is usually lower in large capitalization stocks. Lakonishok, Shleifer and Vishny (1992) also find that herding is more pronounced in small capitalization stocks. Wermers (1999) comes to a similar conclusion. Different reasons for this have been suggested. Lakonishok, Shleifer and Vishny (1992) suggests that this may be because money managers "window-dress" their portfolios and sell off small, obscure and poorly performing stocks. This would be consistent with the reputational herding theory presented by Scharfstein, Stein et al. (1990). Vice versa, money managers might be inclined to buy well performing small stocks because their higher capitalization increases their liquidity and analyst coverage. More herding in smaller capitalization stocks is also consistent with the informational cascade theory (Bikhchandani and Sharma 2000). If there is less public information available on smaller stocks, then the investment decisions of others may be observed and followed more closely. In this thesis the stocks are grouped into 3 different size categories: large, mid and small capitalization. This grouping is based on the size cutoff points used by Nasdaq (large = market value of over 1 billion euros, mid = market value between 150 million and 1 billion euros, small = market value below 150 million euros) and stocks are classified based on their size in December 2006.

Another grouping of interest is the amount of herding by industry. The industry classification used is based on the Industry Classification Benchmark (ICB) used by Nasdaq and maintained by the FTSE group. The ICB system classifies companies based on which activity or business area generates the largest proportion of revenue for the company. A question of interest is whether industries vary in the amount of herding, or if herding around a specific industry changes within the observed time period. For example, the financial crisis was especially hard for many banks and financial institutions. Thus it might be that herding around the financial industry increases in the beginning of the crisis period.

3.2 Use of the LSV measure

The LSV herding measure was chosen for this thesis because it is (1) the most commonly used herding measure and (2) easily implemented with the type of data that is available. The widespread use of the LSV measure is evidence of at least some level of acceptance

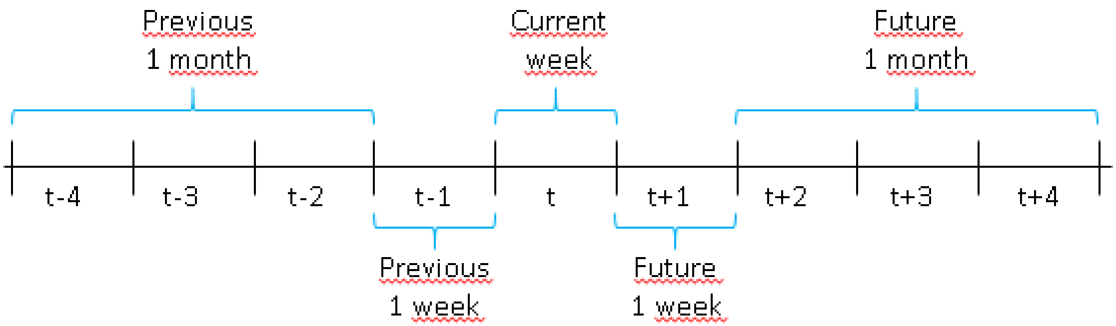

Figure 3.2. *A visualization of how different time periods used in the regression models do not overlap.*

by the scientific community. In addition, having multiple previous results available makes it possible to contrast and compare new results with the old results. While Frey, Herbst and Walter (2014) and Bellando (2010) have criticized the LSV measure and suggested improvements, these suggestions are not utilized in this thesis. Not a lot of research or commentary has been made on the work of either Frey, Herbst and Walter (2014) or Bellando (2010). Thus at the moment there is little evidence that their suggestions will actually improve the LSV measure. Even if e.g. the measure suggested by Frey, Herbst and Walter (2014) would yield more accurate results, comparing them to the results found in previous herding literature would be difficult. On the other hand, if the critique presented by both Frey, Herbst and Walter (2014) and Bellando (2010) is correct and the LSV measure is biased downwards, and significant herding is found nevertheless, this would only make the results more robust.

3.3 Regression models

In this thesis, a mixed effects model is used to study the relationship between herding and returns, and herding and volatility. Two different analyses are made for each group. First, the relationship between past returns (or volatility) and herding is studied. Here the BHM and SHM, which are based LSV herding measure, are the dependent variables and past returns act as the explanatory variables. Different periods for past returns are used. Past returns are calculated for the previous 1 year, 6 months, 3 months, 1 month, 1 week and current week periods. Second, this relationship is flipped, using the BHM and SHM as the explanatory variables and future returns as the dependent variable. Future returns are also calculated for different periods: 1 year, 6 months, 3 months, 1 month and 1 week into the future. The periods have been calculated so that no overlapping occurs. This is visualized in figure 3.2.

In both analyses the specific stock and year-week combination are used as random effects. When the relationship between past returns and herding is examined, the different periods of past returns are the fixed effects.

For example, the mathematical notation for studying the relationship between sell-side herding and past returns with all time periods included looks like this:

$$\begin{aligned}\text{SHM}_t =& \alpha + \beta_1 * R_{t-126,t-252} + \beta_2 * R_{t-64,t-125} + \beta_3 * R_{t-21,t-63} + \\ & \beta_4 * R_{t-6,t-20} + \beta_5 * R_{t-1,t-5} + \beta_6 * R_{t,t+4} + \text{isin}_i + \text{YearWeek}_n\end{aligned} \tag{3.1}$$

where the different β 's are fixed effect coefficients, $isin_i$ is a stock-specific random effect and $YearWeek_n$ is a time-period-specific random effect. Here $R_{j,k}$ is the return between days *j* and *k* and it is defined as

$$R_{j,k} = \frac{S_j - S_k}{S_k}, \tag{3.2}$$

where S_j is the stock price on day *j* and S_k is the stock price on day *k*. As mentioned, the relationship between herding and volatility is also studied in a similar manner.

When the relationship between herding and future returns is studied, the herding measure is used as the fixed effect. This means that a new model is created for each different time period. For example, when studying the relationship between sell-side herding and future 1 week returns, the regression model is as follows:

$$R_{t+4,t+9} = \alpha + \beta\text{SHM}_{t,t+4} + \text{isin}_i + \text{YearWeek}_n \tag{3.3}$$

3.4 Reliability and validity

In all studies questions about the reliability and validity of the findings are of great importance. Saunders (2016) summarizes reliability and validity as follows:

“Reliability refers to replication and consistency. If a researcher is able to replicate an earlier research design and achieve the same findings, then that research would be seen as being reliable. In essence, validity refers to the appropriateness of the measures used, accuracy of the analysis of the results and generalisability of the findings.”

The data and methods used in this study guarantee a fairly good reliability. It is my belief that any researcher armed with the same data, and using the methods outlined in this thesis, would achieve the same results. Validity, on the other hand, poses more challenges.

First there is the question about measurement validity, i.e. does the LSV herding measure really measure herding? As previously mentioned, the LSV herding measure has received some criticism despite it’s popularity. This critique is presented more thoroughly in chapter 2. Most notably, as Bikhchandani and Sharma (2000) point out, the LSV measures the correlation of trading patterns for a group of trades. While herding leads to correlated trading, all correlated trading may not be because of herding. The term "spurious herding" is very relevant here. Just because the same stocks are sold or bought at the same time, it does not necessarily mean that herding has occurred. It might be that some new information has entered the market and many investors react to it the same way independent of each other. To separate spurious herding from actual herding is nearly impossible with the data that is available.

Second, the external validity of the results. Can the results be generalised to other markets or time periods? One must be very careful when making generalisations. For example, if evidence of herding is found, that results is true only for the study period of 1.7.2005 to 30.6.2009. Based on the results of this thesis, one cannot say anything about the possible current amounts of herding in the Finnish stock market.

4 HERDING RESULTS

In this chapter the results regarding herding are presented. The results are presented and discussed in a few different sections: herding in time, herding by investor group and herding by industry.

4.1 Herding in time

This study is split into two different time periods: pre-crisis period from 1.7.2005 to 30.6.2007 and crisis period from 1.7.2007 to 30.6.2009. One of the central questions of this study is whether there is a difference in the amount of herding between these two time periods.

The average LSV herding measure for all stocks and investors during the pre-crisis period was 9.79 %. During the crisis period the average LSV herding rose to 10.40 %. The average amount of herding for all stocks, all investors and the whole study period was 10.10 %. All numbers can be found in table 4.1. This amount of herding can be considered to be quite high. For example, Lakonishok, Shleifer and Vishny (1992) found that the mean herding measure for money managers in the US between 1985 and 1989 was 2.7 %. Higher results have been reported by e.g. Kim and S.-J. Wei (2002) and Choe, Kho and Stulz (1999). Kim and S.-J. Wei (2002) reported herding levels between 8.4 % and 13.2 % for individual foreign investors in Korea, but only levels between 4.8 % and 8.3 % for resident individual investors. In their research the level of herding is generally lower during the crisis period than the pre-crisis period. For their pre-crisis period Choe, Kho and Stulz (1999) report herding levels of over 20 %. For their crisis period they report slightly lower numbers, but still very high compared to earlier studies. Studying data from Finland similar to what was now used, but from a different time period, Kyrolainen and Perttunen (2003) also find relatively high levels of herding. They report numbers of 16.95 % and 13.7 % for the herding of passive investors in winner and loser stocks, respectively. Hsieh (2013) reports relatively high numbers of 16.92 % for institutional investors and 6.73 % for individual investors in Taiwan between 2002 and 2003.

Contrasted with these earlier results, the herding levels measured in this study can be considered high, but not unprecedented. In fact, they are quite well in line with earlier studies focusing on periods of crisis and smaller stock markets. One difference is that in earlier studies herding usually decreased when moving from the pre-crisis period to the

Table 4.1. *LSV herding measures for all investors split by study period and stock capitalization. The standard deviation of the LSV measure also presented for each period and stock capitalization.*

	2005 - 2007	**2007 - 2009**	**2005 - 2009**
All stocks	9.79 %	10.40 %	10.10 %
Std. of herding	11.86 %	13.31 %	12.60 %
Large cap	13.69 %	12.83 %	13.26 %
Std. of herding	12.61 %	12.92 %	12.77 %
Mid cap	9.85 %	10.47 %	10.16 %
Std. of herding	11.33 %	13.46 %	12.45 %
Small cap	6.85 %	8.32 %	7.59 %
Std. of herding	10.91 %	13.13 %	12.03 %

crisis period. In this study the level of herding increased slightly (from 9.79 % to 10.40 %). Though it must be noted that the difference is small, and in large capitalization stocks the level of herding actually decreased (from 13.69 % to 12.83 %). Some contrary results are also presented in earlier research. Blasco, Corredor and Ferreruela (2012), who studied herding at the Spanish market, state that "herding intensity significantly increases in crisis or down market periods". Also Klein (2013) found that herding increases during times of crisis. While this increase from 9.79 % to 10.40 % reported in this study cannot be called "significant", other results pointing in a similar direction do exist. There is also the possibility that no two crisis are the same and thus the amount of herding in pre-crisis and crisis periods does not follow a certain pattern, but instead varies from crisis to crisis.

The results also show lower amounts of herding for smaller capitalization stocks, with large cap stocks having the highest herding and small cap stocks the lowest. This finding contradicts some theories behind the reasons of herding. Generally it can be thought that there is less information available on smaller capitalization stocks and that fewer analysts follow them. According to Bikhchandani and Sharma (2000), informational cascades are more likely when there is less public information available. Since there is less public information available on smaller stocks, this would mean that informational cascades are more likely to occur with small stocks. These results do not support this theory. On the other hand, according to Graham (1999) and Froot, Scharfstein and Stein (1992), investors may focus more on stocks which they believe also others are following. Since large stocks have more analysts and other investors following them, this may drive more investors to trade these stocks on the expense of smaller stocks. This theory would be in line with the findings of this study. Blasco, Corredor and Ferreruela (2012) also found that large capitalization stocks experience more herding that small capitalization stocks. They point out that Sias (2004) argues that the greater amount of information available on large capitalization stocks increases the chance of herding. Higher amounts of herding in large capitalization stocks could happen because uninformed investors prefer large companies over small ones because of their familiarity or because institutional investors invest mainly in large companies. Similar results were also obtained by Lin, Tsai and Sun (2009), who found that stocks with higher quality information available experience more herding, and larger companies commonly have better quality information available on them. Lin, Tsai and Sun (2009) also suggest that herding is caused by the "search cost effect", meaning that investors trade stocks which have low search costs, i.e. large capitalization stocks. The standard deviation of herding also increases when moving from the pre-crisis period into the crisis period.

The difference between buy- and sell-side herding is also studied. As explained in chapter 3, the buy herding measure (BHM) and sell herding measure (SHM) are conditional measures based on the LSV measure and they were defined in equations 2.3 and 2.4.

Table 4.2 presents the results. When looking at the whole study period from 2005 to 2009 and all investors, sell-side herding is about 2.4 % larger. However, there is a clear difference when looking at the pre-crisis and crisis periods separately. During the pre-crisis period there is almost no difference between buy- and sell-side herding (only 0.05 p.p.'s). During the crisis period a clear gap emerges: BHM decreases to 7.23 % and SHM increases to 12.11 %. This implies that during the crisis period sell-side herding is a lot more common than buy-side herding. This is in line with Brown, K. D. Wei and Wermers (2013), who found that funds are more likely to herd on the sell-side than on the buy-side. Herding more on the sell-side seems to hold true at least in times of crisis. Similar findings have been reported by Hsieh (2013), who found that sell-side herding is more common than buy-side herding. On the other hand, Hsieh (2013) reports that sell-side herding is even more common than buy-side herding during a bullish market than a bearish market; this contradicts the findings of this thesis.

Table 4.2. *Comparison between the average BHM's and the SHM's for all investors. The standard deviations for both SHM and BHM and the correlation between SHM and BHM are also presented.*

	2005-2007	**2007-2009**	**2005-2009**
SHM	9.59 %	12.11 %	10.86 %
Std. of SHM	2.02 %	3.59 %	3.18 %
BHM	9.64 %	7.23 %	8.42 %
Std. Of BHM	2.38 %	2.74 %	2.83 %
Corr. of SHM and BHM	-0.1942	-0.4787	-0.4788

This change in the difference between BHM and SHM can clearly be seen in figure 4.1. The figure shows the 5 week moving average of both BHM and SHM for all investors. While there is some fluctuation during the pre-crisis period, the difference between the

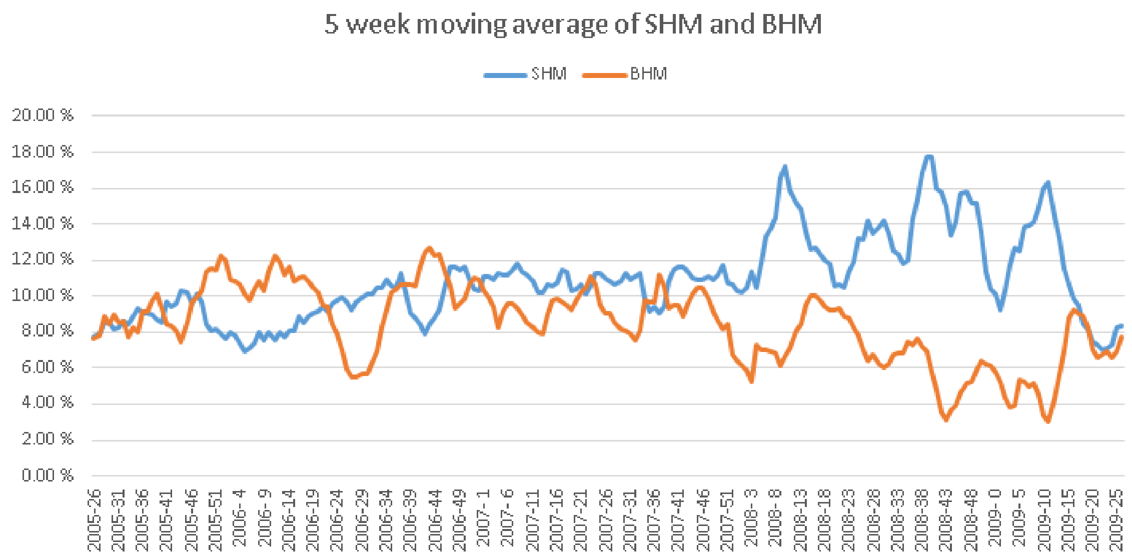

Figure 4.1. *The 5 week moving averages of BHM and SHM. The period changes from pre-crisis to crisis in week 26 in 2007.*

pre-crisis and crisis periods is quite clear in the graph. The crisis period was defined to begin from week 26 in 2007. The SHM and BHM are also negatively correlated with each other; on average, when SHM increases, BHM decreases (and vice versa). This implies there is rarely both sell-side and buy-side herding occurring at the same but instead investors focus on either selling or buying during one week. This negative correlation increases in strength when moving from the pre-crisis period into the crisis period. This implies that in market downturns this polarization between buying and selling is even more pronounced.

To study this statistically, the variance of the 5 week moving averages was calculated. For SHM, the pre-crisis period variance of the 5 week moving average is 0.01281. The crisis period variance for SHM is 0.02459. For SHM there is a clear increase in variance when moving from the pre-crisis period into the crisis period. For BHM, the pre-crisis period variance of the 5 week moving average is 0.01595. The crisis period variance for BHM is 0.01944. While the SHM variance increased by 0.01178, the BHM variance increased only 0.00350. For both the SHM and the BHM, the variance was larger in the crisis period, but the difference is much greater with SHM. This can also be seen in table 4.2, where the standard deviation of both SHM and BHM increase when moving from the pre-crisis period into the crisis period. In the crisis period the standard deviation of SHM is also larger than the standard deviation of BHM. This increase in herding volatility, especially with sell-side herding, implies that during some weeks in the crisis period a significant amount of investors rush to sell the same stocks at the same time. This might be because e.g. momentum investent strategies, where investors start to sell stocks that have seen recent losses. It is also possible that this simultaneous selling is caused by a change in fundamentals, if many companies report poor results at the same time.

4.2 Herding by investor group

In addition to studying herding by all market participants, two groups of investors are separated for closer inspection: household investors and institutional investors. An interesting questions is if there is a difference between the average level of herding, the herding of households and the herding of institutions, and how these results relate to findings in earlier research.

The results for household investor herding are presented in table 4.3. Interestingly, the herding levels of households are lower than those of the whole investor population, the average herding in all stocks being 6.8 % (compared with 10.10 % for all investors). The results follow the same pattern as before, with higher average herding in large capitalization stocks and lower herding in small capitalization stocks. The levels of herding are lower than the whole investor populations' also in each market capitalization group. For all market capitalizations the standard deviation of herding increases when moving from the pre-crisis period into the crisis period. On average, the amount of household investor herding increases when moving from the pre-crisis period into the crisis period. Similar results gave been reported by Hsieh (2013), who also found that individual investor herding increased under "high pressure" market conditions. Goodfellow, Bohl and Gebka (2009) also report that individual investors herd more during market downswings.

Table 4.3. *Results for average household investor LSV herding measure by time period and stock market capitalization. The standard deviation of the LSV herding measure is also presented for each period and market capitalization.*

	2005-2007	**2007-2009**	**2005-2009**
All sizes	6.5 %	7.1 %	6.8 %
Std. of herding	14.5 %	17.2 %	15.8 %
Large cap	8.9 %	8.4 %	8.7 %
Std. of herding	13.7 %	15.1 %	14.4 %
Mid cap	6.2 %	6.9 %	6.6 %
Std. of herding	14.7 %	17.4 %	16.0 %
Small Cap	5.1 %	6.0 %	5.5 %
Std. of herding	14.7 %	18.6 %	16.5 %

The results for institutional investors are presented in table 4.4. The results for institutional investors are lower still, with the average herding over all stocks being 5.7 %. Institutions having lower levels of herding is in line with many previous articles. Lakonishok, Shleifer and Vishny (1992) report herding of 2.7 % for mutual funds, Wermers (1999) report 3.4 % and Brown, K. D. Wei and Wermers (2013) 3.3 %. Considering that they all study US based funds and this paper studies the Finnish stock market, these somewhat higher levels of herding seem logical considering e.g. the difference in the amount of stock choices. Also as previously mentioned, average levels of herding seem to be higher in smaller, perhaps less developed, stock markets.

Herding by institutions seems to differ from the average herding in regards to different market capitalizations. In the pre-crisis period there is almost no difference in herding between mid cap and small cap stocks (3.7 % and 3.8 %). And during the crisis period all three groups have almost the same level of herding (7.4 %, 7.5 % and 7.1 %). Also when comparing to the whole investor population, the herding rose quite sharply (from 4.8 % to 7.4 %) between the pre-crisis and crisis periods. These results suggest that institutions herd slightly more on larger capitalization stocks than smaller capitalization stocks. This contradicts the findings of Hsieh (2013), who found that institutional investors herd more on small capitalization stocks.

When comparing institutional investors to households, the average total levels of herding are lower for institutions. However, in the crisis period the average level of herding for institutions was higher than that of households. It this result is true, it would imply that during the crisis institutional investors herded more than household investors. This contradicts some earlier research, e.g. Kim and S.-J. Wei (2002). On the other hand, Hsieh (2013) found that institutional investors herd more than individual investors.

Some reasons that could explain this are reputational concerns of professional investors, window-dressing of portfolios during a bear market or characteristic herding. One must still remember that the number of active weekly institutional investors was quite low for some weeks of the data. This lowers the reliability of these results. As with household investors, the standard deviation of herding increases for all market capitalizations when moving from the pre-crisis period into the crisis period. The standard deviation of herding is also larger for institutional investors than for household investors in all market capitalizations and time periods, excluding the pre-crisis period small capitalization stocks. This implies that institutional herding varies more from week to week than household investor herding. It might be that institutional investors trade less often and that when they trade, the trades happen at the same time. This might be because of e.g. monthly of quarterly reviews or portfolio re-balancing that always happens during specific intervals.

When looking at the groups of household investors and institutional investors one must keep in mind the possible limitations of the data used. For example, if a private investor does his investing through a limited company (e.g. for tax reasons), they would most likely be classified as a Finnish-owned private corporation instead of a household investor. Similarly it is possible that all relevant institutional investors are not included in the group, even though the grouping was made according to the best available information about the contents of each sector code in the data. Excluding investments made under a nominee registered holding may also exclude trades that may be, in reality, made by either households or institutional investors.

Table 4.4. *Results for average institutional investor LSV herding measure by time period and stock market capitalization. The standard deviation of the LSV herding measure is also presented for each period and market capitalization.*

	2005-2007	**2007-2009**	**2005-2009**
All sizes	4.8 %	7.4 %	5.7 %
Std. of herding	16.0 %	20.6 %	17.9 %
Large cap	6.2 %	7.4 %	6.7 %
Std. of herding	16.7 %	20.2 %	18.2 %
Mid cap	3.7 %	7.5 %	5.0 %
Std. of herding	16.0 %	20.9 %	18.0 %
Small Cap	3.8 %	7.1 %	5.1 %
Std. of herding	14.5 %	21.0 %	17.3 %

4.3 Herding by industry

Herding was also calculated between industries. An interesting question is whether there is a difference in herding between industries and does herding change within an industry when moving from the pre-crisis period into the crisis period. These results are presented in table 4.5.

Table 4.5. *Average LSV herding results for each industry. The standard deviation of herding is also presented for each industry and time period.*

Industry	**2005 - 2007**		**2007 - 2009**	
	avg. LSV	Std.	avg. LSV	Std.
Consumer Discretionary	9.85 %	12.82 %	8.56 %	13.16 %
Consumer Staples	7.36 %	10.29 %	9.64 %	12.18 %
Financials	9.88 %	11.36 %	11.17 %	13.34 %
Health Care	4.80 %	9.38 %	6.35 %	10.45 %
Industrials	9.11 %	11.23 %	9.40 %	12.66 %
Information Technology	9.35 %	11.33 %	11.15 %	14.17 %
Materials	11.71 %	12.26 %	11.88 %	12.97 %
Telecommunication Services	26.30 %	15.08 %	17.83 %	14.86 %
Utilities	14.28 %	10.77 %	13.49 %	11.91 %

There are clearly differences in herding between industries. Low levels of herding can be found in the health care and consumer staples industries. Higher levels of herding can be found in materials, utilities and most notably in telecommunication services. Here it must be pointed out that both the health care industry and the telecommunications industry contain only 2 companies. This means that the industry level herding reflects very strongly the herding of these few companies.

On average, the level of herding increased slightly when moving from the pre-crisis period to the crisis period. This is of course reflected also in the industry level herding numbers. The largest difference in herding between periods can be found in telecommunication services, where herding decreased from 26.3 % to 17.83 %. The largest increase in herding was in consumer staples, where herding rose from 7.36 % to 9.64 %. No clear pattern emerges from these results and mostly the changes in herding between periods is quite small (< 2 %). The standard deviation of herding increased in almost all industries when moving from the pre-crisis period into the crisis period with the only exception being telecommunication services.

In earlier research, Gębka and Wohar (2013) found that on an international level, basic materials, consumer services and oil and gas stocks show some evidence of herding. Comparison with these results is difficult, since Gębka and Wohar (2013) use a different industry classification. In their research, Zheng, Li and Chiang (2017) find evidence that herding is more prevalent in the financial and telecommunication industries and weaker in industrial and consumer services industries. These are quite well in line with the findings of this thesis.

Telecommunication services stands out from other industries because of its high amounts of herding. Some possible reasons for such high herding numbers can be sought from earlier literature. One explanation might be that telecommunication stocks were *in style* at that time. Elisa Oyj, one of the two telecommunications companies, reported growing sales and profits in 2006 and 2007, but decreasing sales and profits on 2008. The Elisa Oyj stock price reached its highest value in February 2007. The other telecommunication company, TeliaSonera, reported record earnings for the year 2006 and its stock price reached its highest point in April 2007. It might be that good news about fundamentals caught the attention of investors and created a buy-side herd, driving the stock prices up. When the crisis period came around the direction of the herd reversed and investors sold their shares, causing the stock price to decrease. It is noteworthy that e.g. TeliaSonera reported mainly increasing sales and EBITDA numbers through 2008 to 2010. This could be interpreted as evidence of the possible illogical results of herding (decreasing stock price while reporting increasing earnings). This sort of behaviour could be explained by style investing: the certain types of stocks are in style, they are bought (without regard to fundamentals) and when they are out of style, they are sold (again without regard to fundamentals). This behaviour is also consistent with using momentum investment strategies. For example the TeliaSonera stock price increased quite steadily from 2003 to 2007. Using a momentum investment strategy investors would have bought TeliaSonera because on recent positive returns. As soon as the stock price decreased in 2007 investors would have started to sell because of poor recent returns. Again, the trades would have been made without regard to fundamentals, basing decisions only on the recent changes in stock price.

5 HERDING, RETURNS AND VOLATILITY

One of the major questions of this thesis is whether there is a link between herding and stock returns. This is studied in two ways: first, do past returns affect herding in the present, and second, does present herding affect future returns? The link between volatility and herding is also studied.

5.1 Herding and past returns

The relationship between herding and past returns was studied using a mixed effects model. The specific stock and year-week combination are used as random effects, while the different periods of past returns are used as fixed effects. The periods for which past returns have been calculated are 1 year, 6 months, 3 months, 1 month, 1 week and the current week. To remove autocorrelation, the periods have been calculated so that they do not overlap. For example, if the "previous 1 week" period refers to week 52 of year 2005, then the "previous 1 month" period includes only weeks 51, 50 and 49. If week 52 had also been included, there would automatically be correlation between the returns of "previous 1 week" and "previous 1 month". The same logic applies for all time periods.

The analysis was made using Matlab. Different combinations of periods were tested and compared with each other until the one with most explanatory power was found. For the sell-side herding measure (SHM) the model with the best explanatory power included returns for the previous 1 year, previous 6 months, previous 1 month, previous 1 week and the current week. The results of the model for the SHM are in table 5.1. The volatility of returns for different periods and the economic signifance of each explaining variable is also included in the table. The economic significance is calculated as Estimate $*$ Volatility of returns. If the explaining variable changes one times it's standard deviation, then the LSV herding measure changes one times the economic significance.

As can be seen from table 5.1, there is a strong and statistically significant correlation between past returns and sell-side herding. High past returns result in higher sell-side herding. This would imply that investors in the Finnish market are contrarian investors, selling past winners and thus locking in profits. The economic significance of the results seems to increase slightly when moving closer to the herding week. To put the economic significance values in perspective, the standard deviation of sell-side herding (SHM) is 13.93 %.

Table 5.1. *The coefficient estimates, p-values and economic significances for the mixed effects model used for explaining SHM with different periods of past returns. The stock specific ISINs and the year-week combination were used as random effects. The regression model used is* $SHM_{t,t+4} = \alpha + \beta_1 * R_{t-126,t-252} + \beta_2 * R_{t-64,t-125} + \beta_3 * R_{t-6,t-20} + \beta_4 * R_{t-1,t-5} + \beta_5 * R_{t,t+4} + isin_i + YearWeek_n$.

Name	**Estimate**	**pValue**	**Volatility of returns**	**Economic significance**
Intercept	0.08945	1.93E-42		
Past returns over 1 Year	0.006645	9.06E-04	0.8609	0.57 %
Past returns over 6 months	0.005502	3.75E-05	1.2226	0.67 %
Past returns over 1 month	0.011468	1.01E-28	1.6799	1.93 %
Past returns over 1 week	0.003982	6.15E-13	2.951	1.17 %
Returns of the current week	0.006697	4.19E-45	2.9541	1.98 %

For buy-side herding (BHM), the model with most explanatory power included the returns for previous 6 months, previous 1 month, previous 1 week and the current week. This model differs from the sell-side model. The result for the BHM model are in table 5.2.

Table 5.2. *The coefficient estimates, p-values and economic significances for the mixed effects model used for explaining BHM with different periods of past returns. The stock specific ISINs and the year-week combination were used as random effects. The regression model used is* $BHM_{t,t+4} = \alpha + \beta_1 * R_{t-64,t-125} + \beta_2 * R_{t-6,t-20} + \beta_3 * R_{t-1,t-5} + \beta_4 * R_{t,t+4} + isin_i + YearWeek_n$.

Name	**Estimate**	**pValue**	**Volatility of returns**	**Economic significance**
Intercept	0.075383	6.68E-60		
Past returns over 6 months	-0.00324	5.81E-04	1.2226	-0.40 %
Past returns over 1 month	-0.00413	4.36E-07	1.6799	-0.69 %
Past returns over 1 week	-0.00156	3.08E-04	2.951	-0.46 %
Returns of the current week	-0.00288	7.28E-16	2.9541	-0.85 %

Again we can see a strong and statistically significant correlation between past returns and herding. All of the coefficient estimates are negative, meaning that high past returns result in lower buy-side herding. Or turned the other way, negative past returns result in higher buy-side herding, meaning that investors buy past losers.

To put the economic significance values into perspective, the standard deviation of buy-side herding (BHM) is 10.96 %.

Combining the results for both of the models, we can say that investors in the Finnish market show contrarian tendencies, buying past losers and selling past winners. Using data from Finland, Grinblatt and Keloharju (2000) found that household investors were contrarian investors. The findings of this thesis are similar, but even more general since they suggest that the **average** investor in the Finnish market follows a contrarian strategy. Based on the estimated coefficients, it seems that the tendency to buy past losers is weaker than the tendency to sell past winners. Hsieh (2013) also found that individual investors in the Taiwanese market show contrarian tendencies, buying past losers and selling past winners.

As mentioned in chapter 2 of this thesis, momentum investment strategies are sometimes presented as a reason for herding. Grinblatt, Titman and Wermers (1995) found that mutual funds in the US have a tendency to buy past winners; Wermers (1999) also found that mutual funds show evidence of momentum investment strategies. Since the LSV measure does not take trading volume into account, only the number of active traders, institutions play a smaller part in the findings of this thesis. More weight is given to household investors and other non-institutional investors, since they are greater in number. These non-institutional investors may also use momentum investment strategies. Barber, Odean and Zhu (2009) suggest that individuals may be prone to momentum investing because of the representativeness heuristic and the limited attention span of individual investors. The representativeness heuristic means that investor take a small sample recent data and expect it to represent the underlying distribution. This means that investors over-weight recent returns in forecasting future returns. While (Barber, Odean and Zhu 2009) suggest this would lead to momentum investing, one could argue also the contrary.

If an investor sees that a stock price has risen to its all-time high they might think that this implies it cannot go any higher. This is something that Tversky and Kahneman (1974) called "misconceptions of chance". This means that people expect a sequence generated by a random process to represent the characteristics of that process, even when the sequence observed is short. If we think that the stock price process is random, and also know that the long term annual yield of stocks is about 7 %, then we might think that after abnormally high returns of, say 14 %, the stock price is likely to decrease so that the annual return converges towards 7 %. Of course in reality this is not the case, since the returns of different stocks and the average returns in different years may vary widely.

The limited attention span mentioned by Barber, Odean and Zhu (2009) means that when choosing which stocks to trade, people may focus on the most attention grabbing ones. Stocks in your portfolio that have risen a lot in the near past and are strongly in the green are sure to grab your attention, as are stocks that have decreased in value recently (they might also seem cheap when compared with their pre-decrease price).

One possible reason for the tendency to sell past winners is *loss aversion*. Loss aversion means that "losses loom larger than corresponding gains" (Tversky and Kahneman 1991). For example, the pain one experiences when losing 100 euros is larger than the joy one feels when gaining 100 euros. This implies that investors are more willing to sell stocks that have already yielded a profit to lock those profits and thus avert a possible future loss.

The tendency to sell past winners instead of losers might also be explained by prospect theory (Kahneman and Tversky 2013). Kahneman and Tversky (2013) report that people are willing to take large risks when trying to avert losses. When given a choice between an 80 % chance to win 4000 or 3000 for sure, most people choose 3000 for sure over the gamble, even though the expected value of the gamble is larger than 3000. However, when these gains are reversed to losses, the preferences change. When given the choice between an 80 % chance to lose 4000 (and 20 % to lose nothing) or a sure loss of 3000, most people choose the gamble (Kahneman and Tversky 2013). It is probable that the same applies to stocks. Selling a stock at a loss is equal to choosing a sure loss of a certain amount. If people choose not to sell and instead keep holding the stock, there is a chance that the stock price may increase and decrease their loss (or even turn the loss into a profit). Even if this is not likely, it seems people are willing to take the risk and not sell their stocks at a loss.

5.2 Herding and future returns

The second major question of this chapter is whether herding affects future returns. This was studied in the same manner as the relationship between past returns and herding, using a mixed effects model. Here the specific stock and year-week combination are again used as random effects. Now the fixed effect is the amount of herding, which is used the explain future profits. The periods for which profits have been obtained are the current week (the week for which the herding was calculated), 1 week, 1 month, 3 months, 6 months and 1 year into the future.

For sell-side herding (SHM) the only statistically significant results were obtained for the periods of current week, 1 week into the future and 1 month into the future. The results can be seen in table 5.3.

Table 5.3. *The coefficient estimates, p-values and economic significances for the mixed effects models used for explaining future returns using SHM. The regression model used is $R_{j,k} = \alpha + \beta_1 * SHM_{t,t+4} + isin_i + YearWeek_n$, where $R_{j,k}$ has different values for each j and k.*

Explained variable	**Estimate**	**pValue**	**Std of SHM**	**Economic significance**
Return over future 1 year	-0.2525	2.50E-01	0.1393	-3.52 %
Return over future 6 months	0.4905	1.80E-01	0.1393	6.83 %
Return over future 3 months	-0.0700	7.85E-01	0.1393	-0.97 %
Return over future 1 month	-0.6008	4.87E-09	0.1393	-8.37 %
Return over future 1 week	-0.5607	2.42E-03	0.1393	-7.81 %
Return over current week	2.0514	3.01E-30	0.1393	28.58 %

Here we can see that high sell-side herding correlates with higher profits in the current week, but low profits in the following week and following month. It must be noted that the relationship between herding and current week returns is not clear. For example, if high profits have been obtained in the beginning of the week (from Monday to Thursday) and on Friday investors flock to sell the stock, it implies that high returns cause higher sell-side herding (as already observed) and not the other way around. However, it is also possible that the sell-side herding has occurred in the beginning of the week and after that the stock enjoys larger returns. Thus the results regarding the current week are not reliable.

The economic significance on these results is quite large. The results imply that if the SHM of a certain week increases one times it's standard deviation, then the returns of the following week decrease by 7.81 p.p.'s. A causation of this magnitude seems unlikely in real life.

For buy-side herding (BHM) the only statistically significant results were obtained for the period of current week. Interestingly no statistically significant relationship was found between buy-side herding and the returns of the following week or the following month, like was found in sell-side herding. The results can be seen in table 5.4.

Table 5.4. *The coefficient estimates, p-values and economic significances for the mixed effects models used for explaining future returns using BHM. The regression model used is* $R_{j,k} = \alpha + \beta_1 * BHM_{t,t+4} + isin_i + YearWeek_n$*, where* $R_{j,k}$ *has different values for each j and k.*

Explained variable	Estimate	pValue	Std of BHM	Economic significance
Return over future 1 year	-0.0970	6.91E-01	0.1096	-1.06 %
Return over future 6 months	-0.1309	6.03E-01	0.1096	-1.44 %
Return over future 3 months	0.0765	3.25E-01	0.1096	0.84 %
Return over future 1 month	-0.0544	7.02E-01	0.1096	-0.60 %
Return over future 1 week	-0.0326	8.91E-01	0.1096	-0.36 %
Return over current week	-1.9662	2.29E-15	0.1096	-21.55 %

Here the estimated coefficient for the current week is negative, implying that higher buy-side herding correlates with lower current week returns. As with sell-side herding, here the causation between current week returns and herding is not clear.

When considering the problems with causation relating to current week returns and herding, the fact that different periods were statistically significant when comparing sell-side herding with buy-side herding, and that both sell-side and buy-side herding seem to correlate with lower future returns, on the whole no strong conclusions can be drawn. Based on these results, there does not seem to be a clear relationship between herding and future returns. This is to be expected, since if a relationship had been found, investors could observe current trading patterns and deduce future stock price changes.

An interesting questions of course is whether herding seems to be rational or irrational. If high buy-side herding had correlated with higher future returns, that would have implied that herding is rational. Correspondingly, if high sell-side herding had correlated with low future returns, it would have implied that sell-side herding is rational. Since the results are inconclusive, nothing new can be said about the possible rationality of herding.

5.3 Herding and volatility

The relationship between herding and volatility was also studied using a similar mixed effects model as before. The volatilities were calculated as follows: first, data on intra-day trades was obtained, containing the exact time of the trade and the trade price. Then the day was divided into 5-minute intervals, and a mean price was counted for each interval as the arithmetic mean of all trades completed within this interval. This means that for each trading day there was a maximum of 102 5-minute intervals. For many stocks and days there were less intervals, since there sometimes were no trades in all possible intervals. Next, returns between these 5-minute intervals were calculated, resulting in a maximum of 101 return observations for each trading day. These returns were then used in calculating the daily volatilities. These daily volatilities were then used in calculating the average volatility for different time periods: previous 1 year, previous 6 months, previous 3 months, previous 1 month, previous 1 week, current week, future 1 week, future 1 month, future 3 months, future 6 months and future 1 year.

As with previous models, past volatilities were used in trying to explain herding. Here the specific stock and year-week combination were again used as random effects. For SHM, the model with the best explaining power included only the volatility over the past 1 year as the explaining variable. The results are presented in table 5.5.

Table 5.5. *The coefficient estimates, p-values and economic significances for the mixed effects model used for explaining SHM using past volatilities. The regression model used is* $SHM_{t,t+4} = \alpha + \beta_1 * \sigma_{t-126,t-252} + isin_i + YearWeek_n$.

Name	Estimate	pValue	Std of volatility	Economic significance
Intercept	0.09738	1.97E-56		
Volatility over past 1 year	-0.00289	8.14E-04	1.3591	-0.39 %

This result implies that an increase in volatility during the past year decreases the current SHM. While this result is statistically significant, it not economically significant. If volatility increases one times it's standard deviation, the SHM would decrease 0.39 p.p.'s. Compared to the standard deviation of SHM (13.93 %) this result is small.

When modelling buy-side herding (BHM), the best explanation power was with a model which had the previous 1 week as the fixed effect. This result is presented in table 5.6.

The result implies that higher volatility during the previous 1 week decreases the BHM in the current week. This result, as in the case of SHM, is statistically significant but

Table 5.6. *The coefficient estimates, p-values and economic significances for the mixed effects model used for explaining BHM using past volatilities. The regression model used is* $BHM_{t,t+4} = \alpha + \beta_1 * \sigma_{t-1,t-5} + isin_i + YearWeek_n$.

Name	Estimate	pValue	Std of volatility	Economic significance
Intercept	0.08023	2.25E-76		
Volatility over past 1 week	-0.00491	2.71E-02	1.3767	-0.68 %

not economically so. If the volatility of the previous 1 week increases by one time it's standard deviation then BHM decreases by 0.68 p.p.'s. Again, when compared to the standard deviaton of BHM (10.96 %) the result is small.

Based on these results it would seems that higher past volatility would cause lower SHM and lower BHM. This would mean that high past volatility would cause investors to herd less in all directions. Still, when considering that only one period had any statistically significant explaining power in both cases, it seems that past volatility has no effect on herding or the effect is very small.

To study this further, the relationship between past volatilities and the LSV herding measure values was also tested. Here the results are similar to the results regarding SHM; the only period with statistical significance is the previous 1 year. The results are presented in table 5.7.

Table 5.7. *The coefficient estimates, p-values and economic significances for the mixed effects model used for explaining LSV using past volatilities. The regression model used is* $LSV_{t,t+4} = \alpha + \beta_1 * \sigma_{t-126,t-252} + isin_i + YearWeek_n$.

Name	Estimate	pValue	Std of volatility	Economic significance
Intercept	0.10103	1.83E-121		
Volatility over past 1 year	-0.00303	1.15E-05	1.3591	-0.41 %

Again, as with SHM, the result is statistically significant but not economically so. If volatility over the previous year increases by one times it's standard deviation, LSV decreases by 0.41 %. Compared to the standard deviation of LSV (12.60 %), this is not economically very significant.

When trying to explain future volatility with herding, no statistically significant results were obtained. The sell-side herding (SHM) results can be seen in table 5.8 and the buy-side herding (BHM) in table 5.9. The relationship between future volatility and LSV was also tested, but no statistically significant results obtained there either. These results are not presented, since they add no value on top of the presented SHM and BHM results.

Looking at the sell-side results, we can see that current sell-side herding has no statistically significant correlation with future volatility in any of the tested periods. The coefficient estimates are mostly positive, implying that higher sell-side herding would increase

Table 5.8. *The coefficient estimates, p-values and economic significances for the mixed effects models used for explaining future volatility using SHM. The regression model used is $\sigma_{j,k} = \alpha + \beta_1 * SHM_{t,t+4} + isin_i + YearWeek_n$, where $\sigma_{j,k}$ has different values for each j and k (days).*

Period	**Estimate**	**pValue**	**Std of SHM**	**Economic significance**
Future volatility over 1 year	-0.3401	1.15E-01	0.1393	-4.74 %
Future volatility over 6 months	0.2165	2.22E-01	0.1393	3.02 %
Future volatility over 3 months	0.2221	1.64E-01	0.1393	3.09 %
Future volatility over 1 month	0.0173	8.71E-01	0.1393	0.24 %
Future volatility over 1 week	0.0919	3.90E-01	0.1393	1.28 %
Volatility over the current week	0.0231	8.35E-01	0.1393	0.32 %

future volatility. But considering that the results are not statistically significant, and their economic significance is quite low, no such claim can really be made.

Table 5.9. *The coefficient estimates, p-values and economic significances for the mixed effects models used for explaining future volatility using BHM. The regression model used is $\sigma_{j,k} = \alpha + \beta_1 * BHM_{t,t+4} + isin_i + YearWeek_n$, where $\sigma_{j,k}$ has different values for each j and k (days).*

Period	**Estimate**	**pValue**	**Std of BHM**	**Economic significance**
Future volatility over 1 year	-0.1749	3.73E-01	0.1096	-1.92 %
Future volatility over 6 months	-0.0075	9.56E-01	0.1096	-0.08 %
Future volatility over 3 months	-0.1367	1.46E-01	0.1096	-1.50 %
Future volatility over 1 month	0.0236	8.37E-01	0.1096	0.26 %
Future volatility over 1 week	-0.0752	5.25E-01	0.1096	-0.82 %
Volatility over the current week	-0.1996	7.47E-02	0.1096	-2.19 %

For buy-side herding, similar results are obtained. No statistically significant correlation between present buy-side herding and future volatility was found. Here the coefficient estimates are mostly negative, which would imply that higher buy-side herding causes lower future volatility. But again, since the results are not statistically significant and their economic significance is quite small, one cannot really back up such a claim.

When looking at the results about herding and future volatility as a whole, it seems that there is no statistically significant link between herding and future volatility. As previously mentioned, some previous authors (Alevy, Haigh and List 2007; Philippas et al. 2013; Scharfstein, Stein et al. 1990) suggest that herding destabilizes markets and increases volatility. These new results do not either confirm or contradict these claims.

In his well known study, Wermers (1999) found abnormal return differences between stocks most heavily bought and sold by institutional investors. He reports that the stocks most heavily bought experienced abnormally high returns during the next quarter and there seemed to be no consequent return reversal. This suggests that herding by institutional investors is rational and speeds up the price discovery process. In his paper Wermers (1999) mentions that "Of course, the limitations of our quarterly holdings data set prevent us from making conclusive statements about whether herding destabilizes daily or weekly stock prices". Now, using weekly data, no statistically significant link between buy-side herding and future volatility was found. This does not support the claim that herding would be rational, although they do not imply that herding would be irrational either. While the coefficient estimates do imply that sell-side herding increases future volatility and buy-side herding decreases future volatility, because the results are not statistically significant, no strong arguments can be made in any direction. In addition, Wermers (1999) studied only institutional investors, while the data used in this study includes all kinds of investors.

6 CONCLUSION

This Master's Thesis examined herding in the Finnish stock market, the relationship between herding and returns, and the relationship between herding and volatility. This chapter summarizes the findings, evaluates the research and suggests futher research topics.

6.1 Summary of results

The data used is provided by Euroclear Finland Oy (previously Finnish Central Security Depository. The records are duplicates of the official certificates of ownership and trades, meaning that the data is extremely reliable. The data contains daily level records of investor's trades. The data also includes sector codes so that the separation of household investors and institutional investors is possible. The frequency of observations and the ability to separate investor types from the data make this truly an unique set of data. Using this data, this thesis set out to answer 5 main questions: (1) was there herding in the Finnish stock market between 2005 and 2009, and if so, how much? (2) Do past stock returns affect herding? (3) Does herding affect future stock returns? (4) Does past volatility affect herding? (5) Does herding affect future volatility?

Between 2005 and 2009 there was considerable herding in the Finnish stock market. The average LSV herding measure for the whole period and all investors was 10.10 %. This can be considered to be quite high, but not unprecedented. Similar results have been reported by e.g. Choe, Kho and Stulz (1999), Kim and S.-J. Wei (2002) and Kyrolainen and Perttunen (2003). The average herding amount in the pre-crisis period was 9.79 % and it rose to 10.40 % during the crisis period. While some researches have reported lower amounts of herding during the crisis period (e.g. Choe, Kho and Stulz (1999)), results with higher amounts of herding during the crisis period have also been reported (e.g. Blasco, Corredor and Ferreruela (2012)).

It also seems that between 2005 and 2009 large capitalization stocks experienced more herding than smaller capitalization stocks. Some earlier empirical work (e.g. Lakonishok, Shleifer and Vishny (1992)) and theories (e.g. Bikhchandani and Sharma (2000)) suggest that smaller capitalization stocks should and, in fact, do experience more herding. On the other hand some theories (e.g. Froot, Scharfstein and Stein (1992) and Graham (1999)) and empirical results (Blasco, Corredor and Ferreruela 2012) point in the other direction. In this thesis the data quality is exceptional and there is a good amount of observations for all stock-week combinations, even for small capitalization stocks (minimum amount of active traders in a week for small capitalization stocks was 389 and the average amount of active traders in a week was 1671), which means the results are reliable. Still, no clear contradiction exists, because the earlier empirical work with contrarian results is from a different market and a different time period. It is possible that both markets and time periods differ in this regard.

A difference between buy- and sell-side herding was also found. During the pre-crisis period the amounts of buy- and sell-side herding were practically equal: 9.59 % for sell-side herding and 9.64 % for buy-side. However, during the crisis period a clear gap emerged: sell-side herding rose to 12.11 % and buy-side herding decreased to 7.23 %. This implies that during times of crisis sell-side herding is more common than buy-side herding. This is in line with Brown, K. D. Wei and Wermers (2013), who found that funds are more likely to herd on the sell-side than on the buy-side. The correlation between SHM and BHM was also studied and it was found to be negative: -0.1942 in the pre-crisis period and -0.4787 in the crisis period. This implies that when sell-side herding increases, buy-side herding decreases. This relationship becomes stronger in the crisis period. This implies that on a weekly level, on average, investors herd either on the sell-side or the buy-side, but usually not both at the same time. This might be because of e.g. momentum-investment strategies, where investors sell stocks that have seen recent losses (and buy recent winners). This is interesting because it is perfectly possible that during any one week, some stocks perform well and some stocks perform poorly. Thus, if an investor uses a momentum investment strategy, they should buy the well-performing stocks and sell the poorly performing ones. This means there should not necessarily be a negative correlation between SHM and BHM. This negative correlation could be seen as evidence of the irrationality of herding. Christie and Huang (1995) suggest that in the presence of irrational herding, the returns of all stocks converge because investors do not make a difference between different stocks, but treat all stocks equally despite their fundamentals. The results of this thesis support this claim. The negative correlation between SHM and BHM suggests that in any given week, investors flock to either sell or buy, but not both at the same time. This seems irrational, since it is unlikely that it is rational to only sell stocks during one week and only buy stocks during another. This kind of behaviour implies that when herding, investors do not make a difference between different stocks but instead just either sell or buy the average stock, following the herd.

There was also a difference in herding volatility between the pre-crisis and crisis periods. In the pre-crisis period, the average volatility of the LSV herding measure was 11.86 %. In the crisis period this rose to 13.31 %. The difference is more clearly seen when comparing SHM and BHM. In the pre-crisis period, the standard deviation of SHM was 2.02 % and it rose to 3.59 % in the crisis period. For BHM, the standard deviation increased from 2.38 % to 2.74 % when moving from the pre-crisis period into the crisis period. This implies that in times of crisis the amount of herding varies more wildly from week to week. A visual representation of this is presented in figure 4.1. This is especially true for sell-side herding. This might mean that in times of crisis investors are more sensitive to negative news and negative market sentiment than in non-crisis periods and are more quick to sell than to buy stocks.

A difference between the behaviour of household investors and institutional investors was also observed. The average herding amount for household investors during the whole study period was 6.8 %. In the pre-crisis period the average amount of herding was 6.5 %, increasing to 7.1 % in the crisis period. For institutional investors, the average herding amount for the whole study period was 5.7 %. For institutions, the average pre-crisis herding was 4.8 % while the crisis period herding was 7.4 %. Here an interesting observation can be made: in the pre-crisis period institutions herded less than household investors. During the crisis period institutions herded more than household investors (7.4 % for institutions vs. 7.1 % for households). In earlier research institutions have had low levels of herding: Lakonishok, Shleifer and Vishny (1992) reporting 2.7 %, Wermers (1999) 3.4 % and Brown, K. D. Wei and Wermers (2013) 3.3 % for US based funds. The pre-crisis herding for institutions (4.8 %) is quite well in line with these earlier findings. The increase in institutional herding from the pre-crisis period to the crisis period is significant. In some previous studies (Choe, Kho and Stulz 1999; Kim and S.-J. Wei 2002) the amounts of herding have been lower during times of crisis, but e.g. Blasco, Corredor and Ferreruela (2012) found that herding increases during times of crisis. Based on the results of this thesis, it seems that especially institutions in the Finnish stock market herd more in times of crisis.

One point of view also studied was herding by industry. All stocks were divided into 9 different industries and the average amounts of herding by each industry were compared. The results are presented in table 4.5. The lowest levels of herding were found in consumer staples and health care industries. The highest levels of herding were found in utilities and most notably telecommunication services, where pre-crisis herding levels were up to 26.30 %. It must be noted that both the telecommunications industry and the health care industry contain only 2 companies. This makes the results regarding them less reliable than the results regarding other industries. On average, the level of herding increased slightly when moving from the pre-crisis period into the crisis period. This is of course also reflected in industry-specific herding amounts. The largest increase in herding between periods was in consumer staples, where herding increased from 7.36 % to 9.64 %. The largest decrease in herding was in telecommunication services, where herding decreased from 26.30 % to 17.83 %. Gębka and Wohar (2013) also studied herding in different industries on an international level. They found that basic materials, consumer services and oil and gas stocks show most evidence of herding. Comparing with these results is unfortunately difficult, since Gębka and Wohar (2013) uses a different industry classification. Zheng, Li and Chiang (2017) also studied herding in different industries. They found that cross-industry herding occurs most commonly in the telecommunication and financial industries. They also found that herding is more pronounced in bear markets. These findings are well in line with the results of this thesis.

A link between past returns and herding was found. Based on a mixed effects model with the specific stock and year-week combination as random variables it seems that high past returns correlate with higher sell-side herding while negative past returns correlate with buy-side herding. This suggests that the average investor in the Finnish market uses a contrarian strategy, buying past losers and selling past winners. In their research, Grinblatt, Titman and Wermers (1995) and Wermers (1999) found that mutual funds show a tendency to use momentum investment strategies. The majority of investors in the Finnish stock market are non-institutional investors, which might explain this difference in findings. In their research, Grinblatt and Keloharju (2000) found that household investors show contrarian tendencies. Hsieh (2013) also found that household investors in Taiwan show contrarian tendencies. Barber, Odean and Zhu (2009) have suggested that individual investors might be more prone to momentum investment strategies than institutional investors. These findings do not support this argument. Some possible psychological reasons for this contrarian behaviour might be "misconceptions of chance" or loss aversion (Tversky and Kahneman 1974; Tversky and Kahneman 1991). The tendency to sell past winners instead of losers might also be explained by prospect theory (Kahneman and Tversky 2013).

When studying the relationship between herding and future returns, no conclusive results were obtained. While some statistically significant periods were found, some of the results contradict each other. The economic significance calculated does not seem believable either. There were some results that suggest that high sell-side herding is correlated with lower profits during the following week and following month. Considering that this relationship was found only for some time periods, and only with sell-side herding, no strong conclusion can be made. Based on these findings, nothing conclusive can be said about the rationality of herding on the basis of future returns.

For the relationship between past volatility and herding, statistically significant results were found only for one period for both SHM and BHM. Neither of these results is economically significant. On the whole no clear link between past volatility and herding was found. When trying to explain future volatility with herding, no statistically significant results were obtained either. Thus nothing can be said about the possible stabilizing (or destablizing) effect of herding. As previously stated, many authors (Alevy, Haigh and List 2007; Philippas et al. 2013; Scharfstein, Stein et al. 1990) argue that herding destabilizes markets while some (Hirshleifer, Subrahmanyam and Titman 1994) argue that herding might in fact stabilize markets. Sadly, this thesis does not add anything to this discussion.

6.2 Evaluation of the research

A good number of studies have been made about herding during recent years. Yet, I believe there are still aspects worth researching, and I hope that this thesis has done its part in increasing knowledge about herding. Still, even this thesis has some shortcomings.

The data that is used in this thesis is of exceptional quality and very reliable. This gives a strong base for conducting good research. For most weeks and the whole investor population there are enough observations to make reliable deductions. Unfortunately, with especially small capitalization stocks and institutional investors, some weeks contain only a few observations. This lowers the reliability of those results. The LSV herding measure chosen for this study is widely used and researched. Still, as mentioned in chapter 2, the measure has received some critique. Because of the flaws that may be present in the LSV herding measure, one must be wary in the interpretation of results. The biggest question is quite fundamental: does the LSV herding measure really measure herding? But on the other hand, what kind of a measure would accurately capture only real herding, but no spurious herding? To my knowledge, such a measure does not exist at the moment. Thus in my opinion, the LSV herding measure was the best one for this thesis.

The results of this thesis regarding the amount of herding are strong and compatible with earlier research. Perhaps the least reliable results are related to institutional investors and small capitalization stocks because of the lower amount of observations. Similarly the increase in herding that has observed when moving from the pre-crisis period into the crisis period has been observed also in some previous research. The results regarding the relationship between past returns and herding are strong and give an interesting opportunity to speculate about the possible reasons behind the observed relationship.

Unfortunately, the results about the relationship between herding and future returns are inconclusive. This is also quite logical, since if a relationship had been found, it would have implied that one can forecast future returns based on past herding. Similarly, when studying the relationship between herding and both past and future volatility, no clear links were observed. This is a shame, since the rationality of herding is an interesting topic. If a statistically significant link had been found, this would have perhaps told us something about the possible rationality of herding. But in light of the observed results this thesis does not add anything new to that discussion.

As already discussed on chapter 3, the reliability of this study is, in my opinion, good. The data is of good quality and methods have been outlined accurately. The possible issues are more related to validity. The LSV herding measure may not accurately capture herding. Also, the results may not be generalisable to other markets or time periods. At most, the results perhaps reflect the situation in a smaller market during the study period of 2005 - 2009.

6.3 Further research suggestions

There are still many markets and time periods where herding has not been researched. For example, while searching for earlier research about herding, I did not come across any studies that were made about other Scandinavian markets. Studying herding in countries similar to Finland might be interesting and could maybe reveal a link between market size and herding. Herding in different market conditions is also an interesting angle. For example, after 1.1.2009 the OMX Helsinki 25 index has seen a steady increase in value. It could be interesting to see whether herding has varied within a single, long bull market.

In this study separating real herding from spurious herding was given little thought. In future research one could try to control for spurious herding by using e.g. investor reports, press releases and other news. It might be possible to separate simultaneous trades caused by a change in fundamentals from real herding by controlling for new information releases. This could perhaps be done even with the herding measured currently available.

While the LSV herding measure is widely used, it has received critique and some upgrades have been suggested e.g. by Frey, Herbst and Walter (2014) and Bellando (2010). As previously suggested by Merli and Roger (2013), one could apply both the original LSV measure and the modified version suggested by Frey, Herbst and Walter (2014) to the same data set and compare the results. This might reveal something new about the possible shortcomings of the LSV measure and thus be a basis for re-evaluation of earlier research conducted by using the original LSV herding measure.

In herding literature, the empirical research on the amount of herding and the theoretical research about the reasons behind herding are not really connected. Many possible reasons for herding have been presented, and herding has been measured in many markets and different time periods. The link that ties these together is still somewhat missing. If one could bridge this gap between the theoretical behavioural psychology research about the reasons behind herding, and the empirical herding measurement research, they would surely have a huge impact on herding research as a whole. Unfortunately, it might be impossible to gather data that clearly links the thoughts and motives of individual investors to the trades they make.

REFERENCES

Alevy, J. E., Haigh, M. S. and List, J. A. (2007). Information cascades: Evidence from a field experiment with financial market professionals. *The Journal of Finance* 62.1, 151–180.

Baltakys, K., Baltakienė, M., Kärkkäinen, H. and Kanniainen, J. (2019). Neighbors matter: Geographical distance and trade timing in the stock market. *Finance Research Letters* 31.

Baltakys, K., Kanniainen, J. and Emmert-Streib, F. (2018). Multilayer aggregation with statistical validation: Application to investor networks. *Scientific reports* 8.1, 8198.

Barber, B. M., Odean, T. and Zhu, N. (2009). Systematic noise. *Journal of Financial Markets* 12.4, 547–569.

Barberis, N. and Shleifer, A. (2003). Style investing. *Journal of financial Economics* 68.2, 161–199.

Bellando, R. (2010). Measuring herding intensity: a hard task.

Bikhchandani, S., Hirshleifer, D. and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy* 100.5, 992–1026.

Bikhchandani, S. and Sharma, S. (2000). Herd behavior in financial markets. *IMF Staff papers* 47.3, 279–310.

Blasco, N., Corredor, P. and Ferreruela, S. (2012). Does herding affect volatility? Implications for the Spanish stock market. *Quantitative Finance* 12.2, 311–327.

Brown, N. C., Wei, K. D. and Wermers, R. (2013). Analyst recommendations, mutual fund herding, and overreaction in stock prices. *Management Science* 60.1, 1–20.

Chang, E. C., Cheng, J. W. and Khorana, A. (2000). An examination of herd behavior in equity markets: An international perspective. *Journal of Banking & Finance* 24.10, 1651–1679.

Choe, H., Kho, B.-C. and Stulz, R. M. (1999). Do foreign investors destabilize stock markets? The Korean experience in 1997. *Journal of Financial Economics* 54.2, 227–264. ISSN: 0304-405X. DOI: `https://doi.org/10.1016/S0304-405X(99)00037-9`. URL: `http://www.sciencedirect.com/science/article/pii/S0304405X99000379`.

Christie, W. G. and Huang, R. D. (1995). Following the pied piper: Do individual returns herd around the market?: *Financial Analysts Journal* 51.4, 31–37.

Dorn, D., Huberman, G. and Sengmueller, P. (2008). Correlated trading and returns. *The Journal of Finance* 63.2, 885–920.

Falkenstein, E. G. (1996). Preferences for stock characteristics as revealed by mutual fund portfolio holdings. *The Journal of Finance* 51.1, 111–135.

Frey, S., Herbst, P. and Walter, A. (2014). Measuring mutual fund herding–a structural approach. *Journal of International Financial Markets, Institutions and Money* 32, 219–239.

Froot, K. A., O'connell, P. G. and Seasholes, M. S. (2001). The portfolio flows of international investors. *Journal of financial Economics* 59.2, 151–193.

Froot, K. A., Scharfstein, D. S. and Stein, J. C. (1992). Herd on the street: Informational inefficiencies in a market with short-term speculation. *The Journal of Finance* 47.4, 1461–1484.

Gębka, B. and Wohar, M. E. (2013). International herding: Does it differ across sectors?: *Journal of International Financial Markets, Institutions and Money* 23, 55–84.

Gompers, P. A. and Metrick, A. (2001). Institutional investors and equity prices. *The quarterly journal of Economics* 116.1, 229–259.

Goodfellow, C., Bohl, M. T. and Gebka, B. (2009). Together we invest? Individual and institutional investors' trading behaviour in Poland. *International Review of Financial Analysis* 18.4, 212–221.

Graham, J. R. (1999). Herding among investment newsletters: Theory and evidence. *The Journal of Finance* 54.1, 237–268.

Grinblatt, M. and Keloharju, M. (2000). The investment behavior and performance of various investor types: a study of Finland's unique data set. *Journal of financial economics* 55.1, 43–67.

– (2001). What makes investors trade?: *The Journal of Finance* 56.2, 589–616.

Grinblatt, M., Titman, S. and Wermers, R. (1995). Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior. *The American economic review*, 1088–1105.

Hirshleifer, D. and Hong Teoh, S. (2003). Herd behaviour and cascading in capital markets: A review and synthesis. *European Financial Management* 9.1, 25–66.

Hirshleifer, D., Subrahmanyam, A. and Titman, S. (1994). Security analysis and trading patterns when some investors receive information before others. *The Journal of Finance* 49.5, 1665–1698.

Hirshleifer, D. and Welch, I. (1992). A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy* 100.5, 992–1026.

Hsieh, S.-F. (2013). Individual and institutional herding and the impact on stock returns: Evidence from Taiwan stock market. *International Review of Financial Analysis* 29, 175–188.

Hwang, S. and Salmon, M. (2004). Market stress and herding. *Journal of Empirical Finance* 11.4, 585–616.

Jones, S. L., Lee, D. and Weis, E. (1999). Herding and feedback trading by different types of institutions and the effects on stock prices. *Unpublished working paper, Indiana University, Indianapolis Kennesaw State University, Kennesaw GA, and Merrill Lynch and Company, New York*.

Kahneman, D. and Tversky, A. (2013). Prospect theory: An analysis of decision under risk. *Handbook of the fundamentals of financial decision making: Part I*. World Scientific, 99–127.

Kim, W. and Wei, S.-J. (2002). Foreign portfolio investors before and during a crisis. *Journal of international economics* 56.1, 77–96.

Klein, A. C. (2013). Time-variations in herding behavior: Evidence from a Markov switching SUR model. *Journal of International Financial Markets, Institutions and Money* 26, 291–304.

Kyrolainen, P. J. and Perttunen, J. (2003). Investors' activity and trading behavior. *EFA 2002 Berlin Meetings Presented Paper, EFMA Helinski Meetings*.

Lakonishok, J., Shleifer, A. and Vishny, R. W. (1992). The impact of institutional trading on stock prices. *Journal of financial economics* 32.1, 23–43.

Lin, W., Tsai, S.-C. and Sun, D. (2009). What Causes Herding: Information Cascade or Search Cost?:

Merli, M. and Roger, T. (2013). What drives the herding behavior of individual investors?: *Finance* 34.3, 67–104.

Nofsinger, J. R. and Sias, R. W. (1999). Herding and feedback trading by institutional and individual investors. *The Journal of finance* 54.6, 2263–2295.

Oehler, A. and Chao, G. G.-C. (2000). Institutional herding in bond markets.

Philippas, N., Economou, F., Babalos, V. and Kostakis, A. (2013). Herding behavior in REITs: Novel tests and the role of financial crisis. *International Review of Financial Analysis* 29, 166–174.

Raddatz, C. and Schmukler, S. L. (2011). *Deconstructing herding: evidence from pension fund Investment behavior*. The World Bank.

Richards, A. J. (1999). Idiosyncratic Risk An Empirical Analysis, with Implications for the Risk of Relative-Value Trading Strategies.

Saunders, M. (2016). *Research methods for business students*. eng. Seventh edition. New York: Pearson Education. ISBN: 9781292016627.

Scharfstein, D. S., Stein, J. C. et al. (1990). Herd behavior and investment. *American Economic Review* 80.3, 465–479.

Sias, R. W. (2004). Institutional herding. *The Review of Financial Studies* 17.1, 165–206.

Siikanen, M., Baltakys, K., Kanniainen, J., Vatrapu, R., Mukkamala, R. and Hussain, A. (2018). Facebook drives behavior of passive households in stock markets. *Finance Research Letters* 27, 208–213.

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science* 185.4157, 1124–1131.

– (1991). Loss aversion in riskless choice: A reference-dependent model. *The quarterly journal of economics* 106.4, 1039–1061.

Vayanos, D. and Wang, T. (2007). Search and endogenous concentration of liquidity in asset markets. *Journal of Economic Theory* 136.1, 66–104.

Wermers, R. (1999). Mutual fund herding and the impact on stock prices. *the Journal of Finance* 54.2, 581–622.

Wylie, S. (2005). Fund manager herding: a test of the accuracy of empirical results using UK data. *The Journal of Business* 78.1, 381–403.

Zheng, D., Li, H. and Chiang, T. C. (2017). Herding within industries: Evidence from Asian stock markets. *International Review of Economics & Finance* 51, 487–509.