Aalto University's School of Business
Department of Economics
Fall 2016

A Basic Theory of Rational Herd Behavior and Informational Cascades

Does it apply to Financial Markets?

Bachelor's thesis in Economics
Aalto University's School of Business 04.01.2017
Author: Nicky Tuominen (484710)
Opponent: Krista Kuuttiniemi
Instructors: Mikko Mustonen & Pauli Murto

Abstract

This paper pursues to motivate and intrigue towards a research of theories of rational herd behavior. The basic theories of rational herd behavior and informational cascades are introduced by using examples and reviewing the most recognized research studies of rational herd behavior by Banerjee (1992), Bikhchandani, Hirshleifer & Welch (1992&1998) and Chamley (2004). It is shown that herding can be modeled with rational individuals using the Bayes' theorem as their decision rule. This model is tested through empirical research, which gives us somewhat controversial evidence of the existence of informational cascades and, though almost all of them find herding to be present, they do not always second the assumption of rational individuals. Towards the end of the paper an implementation of the herd behavior theory to the financial markets is introduced and it is showed, that through the impact of herd behavior and with some extra assumptions to the theory there lies a state where herding occurs in the financial markets giving birth to a price bubbles. It is shown that there is a lot of robust research done among the herd behavior but more evidence to support some conclusions is still needed.

Table of Contents

1. INTRODUCTION 4

2. BASIC THEORY OF HERD BEHAVIOR 5

2.1. INTUITIVE MODEL OF HERD BEHAVIOR 6

2.2. RATIONAL INDIVIDUAL: BAYESIAN DECISION THEOREM 8

2.2.1. Example of decision making 8

2.3. APPLICATION: COST OF ADOPTION 10

3. INFORMATIONAL CASCADES AND IMPACTS OF HERD BEHAVIOR 11

3.1 INFORMATIONAL CASCADES AND REVERSE CASCADES 13

3.1.1. Fragility 14

3.2. IMPACTS 15

3.3. EMPIRICAL EVIDENCE IN LABORATORY 16

4. HERD BEHAVIOR IN FINANCIAL MARKETS 18

4.1. EMPIRICAL STUDIES 18

4.2. MOTIVATION 19

4.3. RATIONAL HERDING IN FINANCIAL MARKETS 20

4.3.1. Multidimensional Uncertainty (Avery & Zemsky, 1998) 21

5. CONCLUSION 22

REFERENCES 25

1. Introduction

It is commonly known that animals and humans are influenced by others surrounding them, others more than others. In human case, it is mostly because we do not live in a void. In this study I assume (mostly in common daily actions) that individuals are acting rationally and do their decisions comparing their personal preferences/knowledge against the actions of preceding individuals i.e. to the public information. When an individual acts due to the public information we say that he is herding.

Chamley (2004) introduces the example of umbrellas. Basically, it goes like this: You see someone in the street carrying an umbrella and you rationally think that the other carries it because he thinks there's a high chance of rain today. So you also take your umbrella with you. Other people see now two people carrying an umbrella and they infer the same. This leads to an outcome where most of the people are carrying an umbrella. Now, here comes the interesting part: What if the first few who started carrying an umbrella weren't carrying it because of the high probability of rain? What if they just wanted to move their umbrellas to some other location. Then the people would have started herding and ended up in a socially inefficient equilibrium.

One of the effects considered to occur from herd behavior is the stagnation of information accumulation or the deceleration of social learning. In the umbrella case the information aggregation stopped when people decided to act on a basis of what they see others doing instead of e.g. checking the weather forecast. This may lead to an efficient or inefficient outcome in society depending on which action the first ones to decide choose to pick.

Bikhchcandani, Hirshleifer & Welch (1998) and Chamley (2004) presented a case called High Sales Promote High Sales. In 1995 two management gurus secretly purchased 50 000 copies of their own book. The shops they bought it from were shops that were monitored for the bestseller list of the New York Times. The book made it to the bestseller list and continued to sell enough to keep itself in the list. In this case people were assumed to observe the sales, but not the payoffs, of the purchase i.e. is the book a good read. Even if the manipulation would have been known it would not have had any effect since then the people could have deduced the payoff not to be affected by the authors' purchase. This is a good example of people's herding and the opportunities it may offer when we get a better understanding of the phenomenon and learn how to use it.

The purpose of this thesis is to get familiar with a phenomenon called the herd behavior from an economic point of view and to give prerequisites to do further research on the subject. The thesis is based on existing literature and does not contain own empirical research. I go through the most general studies considering the herd behavior and empirical research concerning the phenomenon. I come to a conclusion that the mathematical theory still needs more evidence to support it. In the last section of the thesis we apply the theory to financial markets and present the required modifications.

The theories in this thesis are mostly based on the general models of herd behavior introduced by Banerjee (1992) in his study "A simple model of herd behavior", Bikhchcandani, Hirshleifer & Welch's (1992) paper of "A theory of fads, fashion, custom and cultural change as informational cascade" and Chamley's (2004) "Rational Herds".

The paper is structured as follows: The first section is introduction. In the second section I will introduce the basic theory of herd behavior, Bayesian framework, behind the presumption of rational behavior of individuals and other presumptions behind the theory. In chapter three, I go through the impacts of herd behavior, what is an informational cascade, are informational cascades fragile and what they are affected by. In chapter four I apply the theory to financial markets and study what other factors need to be taken into account when applying it. The fifth section concludes our results.

2. Basic theory of Herd Behavior

Herd Behavior means a state where the individuals are acting according to the actions of others rather than using their *private signal* (private information) to act. Meaning that the individual chooses to act differently when he knows the history of other individuals' actions than when he doesn't know them. When he changes his way to act as a consequence of the history we call him herding.

As most of the theories in economics make a presumption of rational individuals or consumers this one is no different. The basic assumption in the theory of herd behavior is individual's rationalism that is modeled by Bayes' theorem. The other important assumption, given the Bayes' framework, is that the decisions made are made in sequence and that individuals are aware of the decisions made by predecessors so that there is a sequence of time periods.

Theory of herd behavior presumes that when making a decision, an individual will act rationally given all the information he is capable of acquiring. In this theory, the information individual uses to make his decisions is divided in two. The first one is the individual's own signal which reflects his own preferences. This is called a private signal. The other is the history of actions made by predecessors who had faced the same decision problem. This history of known acts is called *public information*. Within these two types of information the individual encounters a decision problem which he solves using a Bayesian decision theorem. He weighs the public information against his private signal.

The public information is formed from the actions taken by the individuals who were following their private signal when making a decision. When the pool of public information grows, it is called *social learning* (Chamley, 2004). It is assumed that all individuals know the public information. Normally, it is assumed that private signals can't be seen but the actions taken are visible to others i.e. to posterior decision makers.

Accumulation of the public information continues as long as the individual makes a decision based on his private signal. When the individual disregards his private signal and chooses to act according to the public information, he no longer contributes information to the public information and we say that he is herding. The accumulation of public information stops and no information is conveyed to the future individuals. If everyone ends up disregarding their private signal, herding, and acting according to the public information, we're in the state of an informational cascade, which will be explained in the section three.

I will now start presenting the basics of the theory of herd behavior. I will not go through all the models presented of the subject since it would not serve the purpose of this thesis. The theories used here are a mix from different sources. To get more acquainted with the theories and applications I recommend to read Chamley's (2004) book of Rational Herds. In this section I focus on the basic theorem and assumptions behind the theory. First, we will go through the intuitive model of herd behavior.

2.1. Intuitive model of Herd Behavior

The most important thing to understand is the intuition of herd behavior model. It makes the herd behavior more comprehensible. In this section the model presented is basically the same that Banerjee (1992) presented but with few influences from Chamley (2004) and Bikhchandani, Hirschleifer and Welch (1992). To understand the model we will make a few extra assumptions:

Let's say that there are two restaurants A and B and there lies an optimum choice between the restaurants that is restaurant A. Everybody knows there's an optimal choice but no one knows which one of the two it is.

There are an infinite amount of time periods and individuals. Each individual makes a decision in a sequence and the order in which he decides is predetermined and known to all. The restaurants are located side by side and the individuals can see how many people are in both restaurants the moment they arrive i.e. all the actions made are publicly known to all and are used in the decision process of the forthcoming individual. Each individual has a private signal (s) which either tells to go to the restaurant A (s=*a*) or to the restaurant B (s=*b*). All the private signals have an equal precision and all the actions have the same weight and affect the public information when the individual's decision is made by his private signal.

Let's state that between the individual's there are 51% with signal a and the rest have signal b. So in ex ante sense it's more likely that most of the individuals will end up choosing the restaurant A. Now, let's assume the first individual arrives and sees two empty restaurants. He has a private signal *a,* so he chooses the restaurant A. Now the second individual has signal *a* and comes in front of the restaurant and sees that in restaurant A there is already one customer (In this case he follows his private signal so the predecessor's decision does not make a difference) and chooses restaurant A. The third individual has a signal *b* and comes to the restaurants and sees two people sitting in the restaurant A. According to his signal he would choose the restaurant B but he weighs his private signal in against to the history of actions made (2xA) and decides to go to a restaurant A. In this situation we call the third individual herding. Now the following individual will see 3 people in the restaurant A and despite of what his signal is he will choose restaurant A since the history of actions outweighs his private signal and this applies to all the subsequent individuals.

In the case above we ended up in a state where all individuals after the first two started herding and chose the restaurant A and thus chose the optimal decision. In the view of society we ended up in the right decision but the chances that the first two individuals would had signal *b* and chosen the restaurant B would have started a herding to the opposite direction and every individual after that would have ended up choosing the non-optimal decision (Restaurant B).

Bikhchandani, Hirschleifer and Welch (1992) have introduced a concept where all the forthcoming individuals will end up herding. They call this an informational cascade.

2.2. Rational Individual: Bayesian decision theorem

We assume that the individuals making the decision are rational. This presumption is stated in the theory of herd behavior by Banerjee (1992). This doesn't mean that everyone should have the same preferences or the same precision on their signal. It mostly contributes to the individual's decision making process. The theory used to model this is Bayes' theorem or Bayes' rule.

Bayes' theorem in general is a very remarkable theorem in mathematics. It can be said that Bayes' theorem is to the theory of probability what Pythagorean theorem is to geometry. I present the central idea of Bayes' rule in this chapter, as it gives the best understanding of the decision making process of individuals.

Bayes' rule is, in probability theory, an application which takes into account the probability of a before event X (prior event) and conditions it to another known/noticed event's Y probabilities. It gives us a general answer of conditional probability when these two events (X,Y) are not independent. The basic theorem is represented as below:

$$P(X|Y) = \frac{P(X) \times P(Y|X)}{P(Y)} \qquad (1)$$

$$= \frac{P(X) \times P(Y|X)}{P(X) \times P(Y|X) + P(Y) \times (1 - P(X))} \qquad (2)$$

The theorem states the probability of event X happening when event Y has already happened. The function (1) represents the general Bayes' theorem model and the function (2) is the application of Bayes' theorem used in the herd behavior theory. This theorem is used to illustrate the decision model of an individual.

2.2.1. Example of decision making

Let us first assume that the first individual to decide has no information about the history of earlier decisions. Let there be two possible actions taken $A = \{0,1\}$. The X_i denotes the event that an individual chooses the action $A = 1$ by his private signal. Let Y denote the action A=1 preferred by the public information. Every time the individual chooses to act according to his private signal the information accumulates to public information and updates the probability(=precision) of Y (A=1) to the forthcoming individuals.

Let us denote that $\mathrm{P(X_i)}$ is the probability of an i:th individual choosing action A=1 i.e. the probability denotes the individual's private signal's precision to action A=1. P(Y) is the probability of public information guiding to the action A=1 i.e. it is the accumulation of history of the actions taken by predecessors. $\mathrm{P(X_i|Y)}$ is therefore the probability of i:th individual choosing to take the action A=1 taking into account his own private signal given that there is public information he uses to weigh his decision.

In this example, the initial private signal and public information's precisions are given. This example could have been done by any other precisions though the outcome would change. The function used here is the Bayes' theorem function (2) that was presented in the previous section (2.1.).

In this example, I make an assumption that the first individual does not acknowledge the public information and therefore he makes his decision purely on his private signal which tells him to take action A=1 in a precision (probability) of $\mathrm{P(X_1)} = 0{,}6$. Let us denote that the public information precision to act A=1 is $P(Y|X) = 0{,}48$ i.e. according to public information the choice would be to take action A=0. If the first individual's private signal is known to the forthcoming individuals then the updated probability of P(Y) is:

$$(2) \qquad P(X_1|Y) = \frac{0{,}6 \times 0{,}48}{0{,}6 * 0{,}48 + 0{,}4 * 0{,}52} = \sim 0{,}58$$

As we can see the public information turned in the favor of A=1 from the point where it was more likely to end up in A=0. Now let us assume that the forthcoming individuals, who will decide in sequence, acknowledge the history of actions and use the public information against their private signal. The public information has been updated to have the precision $P(Y|X) = 0{,}58$. In this case we assume the preceding individual's private signal precision was observed by forthcoming individuals. Let us say the second individual's precision to act A=1 is $\mathrm{P(X_2)} = 0{,}45$. So when only considering his private signal he would choose action A=0. Now he makes the/his decision through the Bayes' theorem and chooses to act on the grounds which action is more likely to be the right one:

$$(2) \qquad P(X_2|Y) = \frac{0,45\times0,58}{0,45*0,58+0,55*0,42} = \sim0,53$$

Bayes' theorem shows that according to this information it's more favorable to choose the action A=1 and, therefore, the second individual chooses A=1, even though, his private signal tells otherwise. When the individual acts according to public information the public information aggregation stops, since it only aggregates when individuals make decisions using their private signal, and the public information stays unaltered $\mathrm{P}(Y|X) = 0,58$. We call the second individual herding since he disregarded his private signal.

This example presented how the public information affects the decision made by individuals in the herd behavior model. This outcome doesn't mean that all the following individuals would end up herding. If an individual has a low enough precision in his private signal to act A=1 then he would most likely choose to act according to his private signal, contribute to the public information and change the favor of the public information.

This example and Bayes' theorem are presented here just to clarify what lies behind the herd behavior theory. The Bayes' theorem is what makes the herd behavior theory possible and it is the most important assumption in this theory. In the thesis I will not get in to the details of every outcome but mostly present the intuition of the theory from here on. To get there, I will present one of the simplest models of herd behavior.

2.3. Application: Cost of Adoption

BHW (1992), Avery & Zemsky (1998) and Chamley (2004) presented an applied model of herd behavior which includes a term: *cost of adaption*, denoted as *C*. The term affects directly the individual's decision making. The cost of adaption falls for individual's cost when an individual decides to take an action that the cost is set on. The cost can be monetary, time, or some other factor. For example, when we are dealing with a decision problem of buying a car we have the option of not buying the car which costs us nothing or to buy the car paying its price (monetary). In this case the cost of adaption would be the price.

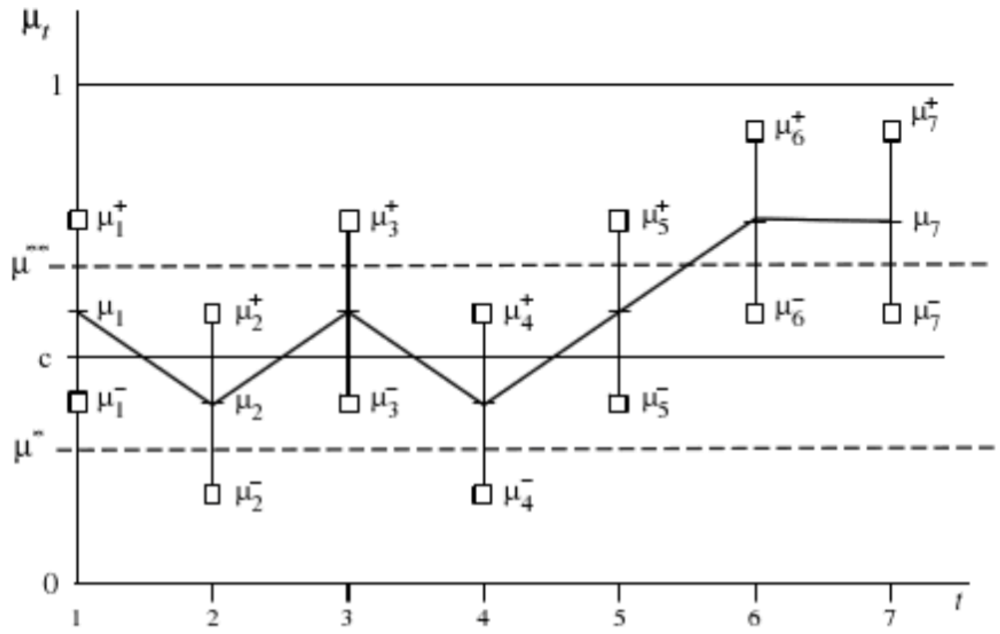
I briefly introduce the application through the model originally represented by BHW (1992). The cost of adoption is fixed. This assumption will be relaxed later when we get acquainted with the Avery & Zemsky's (1998) research in chapter 4, Herd Behavior in Financial Markets.

In the model there is a *value of adoption*, V and it is $\{0,1\}$. Each individual gets an independent, imperfect signal (=private signal) from V which is either the good (x=1) or bad news (x=0), $x \in \{0,1\}$ denotes the private signal. When making the decision, an individual uses it against the public information, in this model stated as history of actions H_t (t=period).

The individual now compares an expected value of adaption to the cost of adaption. The expected value is defined via the private signal's belief of what the value of adaption is, P(x=V), and the history of actions of predecessors. If $E[V|x, H_t] > C$ then the individual chooses to adapt and when the history i.e. the public information is relatively high the herding begins. In this model this can be denoted as $E[V|x = 1, H_t] > E[V|x = 0, H_t] > C$. In this outcome we would actually be in a state of an informational cascade.

3. Informational Cascades and Impacts of Herd Behavior

At this point we are familiar with the phenomenon of herd behavior. We have learned how we can end up in a state where rational individuals are herding and are in a belief that their choice is the optimal and efficient one even though it might not always be so. Bikchcandani, Hirshleifer & Welch (1992) (later referred as BHW) introduced the concept of informational cascade. Like stated above, in the informational cascade all future individuals will end up herding, in all periods, to the same decision and thus disregard their private signals. In a model which Chamley (2004) introduced as binary model that is quite similar to the intuitive model of herd behavior it is proven that eventually informational cascade will occur when the number of periods is finite.



Picture above is from Chamley's (2004) book and it represents the convergence over time to informational cascade in a binary model. In this picture the t=period and μ+ is an individual with a private signal to take action 1 and μ- is an individual whose private signal tells him to take action 0. μ denotes the public information. In period 6 an informational cascade has started and individuals starting from period 6 are herding.

To understand the concept of informational cascade Chamley (2004) makes an important difference between informational cascades and herd behavior. He states that when we're in the state of informational cascade all future individuals will herd but not vice versa. Chamley (2004) introduces the informational cascade as a long term constant state that doesn't change over time. In the researches of Bikchcandani, Hirshleifer & Welch (1992&1998), who first introduced the concept, it is stated that informational cascades are fragile and can be reversed by certain factors in which we take a closer look later in this chapter. Let us clarify the difference between the herd behavior and informational cascade through an example presented below:

Let us consider a state of world such as in the intuitive model of herd behavior (2.1.) with a difference that all individuals' private signals are heterogeneous and vary in between [0,1]. This given when the first individual who ends up herding, and therefore, does not contribute to the public information *accumulation, might end up herding with a relatively low public information signal. Let us assume that the public signal outcome is the same as introduced in the 2.1.1. Example of decision making chapter P(Y) = 0,58. Now if the forthcoming individual has a private signal of P(X) = 0,2 he does not end up herding but chooses to act in contrary of the previous individual who ended up herding (This can be proven by applying the signal information to the Bayes' theorem function (2) that was introduced earlier).*

In the example above it comes clear that, since the latter individual chooses not to herd but the first one was herding, herding can happen even though it would not lead to an informational cascade. In the example, I relaxed the assumption of homogenous signals, which makes the model closer to the real world model. It can be seen, that people with strong signals can change the trajectory of the outcome and the public information. Usually these individuals are better informed than an average individual and are recognized among the other individuals so that their decision has a greater impact on the public information than the average individual's decision, done via private signal. These individuals are referred to as *opinion leaders* (Chamley 2004, BHW 1992 & 1998 & Vives 1997).

If the assumption hadn't been relaxed in the example, then we would've been in a state of informational cascade after the first individual who disregarded his signal and decided to herd. Chamley (2004) states that informational cascades are merely a theoretical model and they do not occur in practice because of the fact that the individuals' signals are not homogenous. Chamley (2004) also states that even though informational cascades do not comply well when implemented in real world, they are interesting in the representation of social learning models. If we treat informational cascade as a short-term phenomenon they do occur, but usually only a brief of time until one individual chooses not to herd and deviates from the informational cascade.

3.1 Informational cascades and reverse cascades

We have established the existence of informational cascade and the fact that after it has started, all the forthcoming individuals will end up selecting the same decision that the preceding individual had decided. Chapter 2 introduced a model of restaurants and made the statement that in the world there lies an optimal and non-optimal decision choice and individuals choose between these without knowing which is which. In the example, we ended up in an outcome where all the forthcoming individuals end up herding and choose the optimal choice. This outcome is called an informational cascade. But what if the informational cascade occurs and favors an unfavorable outcome? BHW (1992) named this outcome to a reverse cascade.

Optimal and non-optimal choice can be interpreted as efficient and inefficient choices for society's welfare. So when we end up in a reverse cascade we are actually dealing with a welfare problem. The individuals are minimizing the welfare function by choosing to act according to a reverse cascade. Although the concept of informational cascades is mostly theoretical and does not very likely occur as a long term effect it does occur as a short term effect and thus a reverse cascade as well as active herding of majority to the non-optimal decision contains a problem that affects negatively on the whole society. Considering this problem, what can alter the direction of cascades?

3.1.1. Fragility

BHW (1992 &1998) introduce in their papers the fragility of an informational cascade. This fragility holds a set of factors that can alter the direction of a reverse cascade to an informational cascade and vice versa. These factors are discussed as *shocks*. BHW (1998) names three factors:

1) Better informed individuals (opinion leaders)
2) The release of public information
3) Shifts in the value of underlying value/cost of adaption versus rejection

Better informed individuals (opinion leaders)

Individuals that are widely recognized or respected among other individuals are called opinion leaders. These can be people who are extensively thought to have better information regarding the decision in hand e.g. when buying a car a car mechanic may be considered as an opinion leader or when buying stocks a successful investor like Warren Buffet could be considered as an opinion leader.

In an intuitive model of herd behavior, every individual's action weighted the same when aggregating to the public information. The opinion leaders have a greater impact on the public information and a more precise private signal (Chamley 2004, BHW 1992, Watts & Dodds 2007). Due to their precise private signal they are more likely to act according to their private signal and due to their strong influence on others their actions have a better chance of influencing the forthcoming individuals' actions, especially when there are more than one opinion leader making their decision in sequence. The opinion leaders therefore have a chance of changing the direction of a cascade.

The release of public information

Up to this point, we have discussed the arising of informational cascades through the actions of other individuals but haven't considered the impact of e.g. actions of government or other widely recognized institutions. The disclosure of public information, such as new research results, can be directly aggregated to the public information (that individuals use by their private signal) when they have a great reach within the audience (BHW 1992). If there is a large number of public disclosures that favors the optimal decision then by the law of large numbers the society should settle to the correct cascade. The public disclosures are a tool to influence a possible informational cascade and accumulate the public information when in a state of an informational cascade the social learning has stopped.

Shifts in the value of underlying value/cost of adaption versus rejection

As introduced in chapter 2 in the applied model of herd behavior the term, cost of adoption was added. In the model, it was given as a fixed cost. Considering the cost to be fixed we can think of a scenario where we keep the cost fixed but change it once when an informational cascade is occurring. Let us say that the cost of adaption rises so that from an ex ante situation $E[V|x = 0, H_t] > C_{t-1}$ the C alters so that $E[V|x = 0, H_t] < C_t$. Now there is a chance that social learning starts again since, at least to one individual, the action x=0 is again possible according to his private signal and the cascade may be changed to a reverse cascade depending on the influence of the deciding individual and the forthcoming individuals' private signals. If the value of adoption decreases, the effect would be the same.

3.2. Impacts

Informational cascades and herding have strong effects on social learning which makes them worthwhile for research. Social learning is a concept used throughout the research studies I've introduced in this thesis. It is learning that is achieved by an individual through the public information. Every time an individual chooses to act by his private signal he accumulates his action's information to the public information. This is called social learning since now all the other individuals have access to this information and the pool of public information has grown from the ex ante position. BHW (1998) uses the concept of information externalities for this phenomenon.

Chapter 3.1. introduced the welfare question between the optimal and non-optimal choice and stated that the optimal decision is, in a sense of efficiency and welfare, the welfare maximizing choice. The herd behavior affects the welfare in other ways also. Since information asymmetry inflicts inefficiency (moral hazards, mispricing etc.) and subsequent welfare losses, the cascades and widely active herding have a potential of affecting negatively on welfare. Even if all individuals would end up choosing the optimal decision, and entering the informational cascade, in a sense of conveying information, we are in an inefficient state. When an individual herds, it is said he inflicts a negative *externality* (BHW 1998) as his action convey no information.

BHW (1998) states that in their model of informational cascades, where actions of predecessors are only observable to the other individuals, the existence of the informational cascades actually has a positive impact on directing the cascade to becoming an informational cascade when the initial distribution of signals favors the truth than when no actions or signals are observable. They give an example where, when 51% of the distribution of signals are in favor of truth and all the signals are equally weighted, the likeliness of ending up in an informational cascade is 51,3%. Therefore the existence of public information and its accumulation should raise the chances of the outcome to end up in an initially more favorable outcome.

The informational cascades are often based on limited amount of information and are therefore fragile to shocks (represented in section 3.1.). Due to this, early decision makers have a great influence on the birth of an informational cascade, and therefore to the public information's aggregation.

3.3. Empirical evidence in laboratory

Banerjee's (1992) theory of herd behavior and Bikhchcandani, Hirshleifer & Welch's (1992) paper of information cascade, with assumption that individuals make their decisions based on Bayesian decision theorem, was put in a laboratory test by Anderson and Holt (1997). BHW (1992) introduced the idea of active herd behavior sometimes leading to informational cascades. The very idea of the occurrence of these informational cascades, in which the information aggregation and learning stops when entered, was tested with 72 human subjects with the amount of 56 periods by Anderson & Holt (1997).

Anderson and Holt's (1997) results concluded that, via Bayesian decision theory, most of the times, information cascades do form in a laboratory. This result favours the basic theory of informational cascades. The test had the initial assumption that the subjects will act rationally and use the Bayesian decision theorem to make the decision without it being taught to them.

The rationality of individuals (Bayesian decision theorem) was put to test by Huck and Oechssler (2000) in their research. The results were contradictory to Anderson and Holt's results. The conclusion was that the heuristic "follow your own signal" seemed to be a more precise description of the individual decision theorem than Bayesian.

In Huck and Oechssler 's (2000) laboratory test, the test subjects were students of the microeconomics' -course and had been taught about the Bayesian decision theory and were familiar with the Bayesian updating. They were given the Bayesian framework beforehand and the test affected their grading on the course.

The rational behavior of individuals in economics is assumed to be the base of the individuals behavior given that the individuals are not aware of this model. So when Huck and Oechssler (2000) taught the model of desired action model to the test subjects it almost surely affected the decisions the subjects made. There are possibilities that some of them did not learn the model correctly and answered through using the Bayesian model incorrectly even though their intuition may have guided them to the right answer.

Huck and Oechssler (2000) gave the data of decisions made by the predecessors to the subject. The rationality of the subject included that a rational individual recognizes the starting point of a cascade and knows which signals/actions to account for when using the public information. Meaning that when the cascade starts, the information aggregation stops and thus the decisions made after the cascade started are not significant to the decision maker and he should ignore them.

In many cases when we get information of history of actions made, we get them in charts e.g. how many has already bought the product but not in what order the bought and non-bought decisions were made (Chamley 2004, BHW 1998). Therefore, in these situations, even though the decision maker can infer the start of a cascade, he most likely won't be able to tell at which point it started and therefore know which actions to account in when using public information to make his own decision. Therefore further tests should also be ran through these conditions to see how this affects the behavior.

4. Herd Behavior in Financial Markets

4.1. Empirical Studies

Chiang, Zheng (2010) ran an empirical test for the global stock markets. They used 18 countries' daily data of industry and market price indices from the time period of 1988-2009. In their survey, they divided the countries into three categories: Latin America Markets, Advanced Markets and Asian Markets. The study finds evidence that herding is present in global markets and in particular within the Advanced markets and Asian Markets. They find evidence of herding to be present in both up and down markets and that a crisis triggers herding in the crisis country of origin and inflicts a contagion which spreads to the neighboring countries.

Zhou & Lai (2009) ran a regional herd behavior test on Hong Kong's Stock Exchange market with bid/ask quotes and trade records from 01/2003 – 12/2004. They compared the differences of herding in different industries and got evidence that in Hong Kong's main industries (financial sector and the property and construction sector) the herding is greater than in other industries. Zhou & Lai (2009) claim it also gives credence that investors are more likely to follow others when making trading decisions in order to play it safe. Bikhchandani & Sharma (2000) call this Reputation-Based Herding, which in brief means that when herding, an investor gives cover to himself and to the preceding investor by taking the same decision (They enhance the common opinion that their choice was the right one). The Hong Kong's market is said to be a transparent market, which should be one of the factors preventing possibility of herding due to the claim of informationally efficient financial markets. This argument is covered later in this chapter.

Although the surveys represented here find evidence of the herding behavior in the stock markets they do not claim that the outcomes are based solely on the assumption of rational individuals (Chiang & Zheng 2010). In addition, the cause and the conditions of herd behavior seem to differ between the regions and countries. In India the herd behavior is present during the upswings in market trends and the herding behavior is independent of the level of trading volume while in China the herd behavior is detected during the fall of market when the trading volume is high (Lao & Singh, 2011). Zhou & Lai (2009) find evidence that in Hong Kong's Stock Exchange market frequency of herding on the sell-side is higher than on the buy-side of the market. The surveys do not seem to find a common initiative factor of herding.

4.2. Motivation

Why are we interested in herd behavior in financial markets? The answer lies in the sum of the herd behavior impacts that are represented in Chapter 3. Herding and informational cascades create an information asymmetry due to the slowdown or stagnation of the social learning (Chamley, 2004), which can be interpreted as inefficiency in society. When there's information asymmetry, there's a chance of mispricing and when there's mispricing there's a chance of price bubbles and therefore a chance of occurance of the local or global financial crisis when these bubbles pop. The herding can therefore potentially create a mispricing in financial markets.

The herd behavior's existence challenges the validity of the informationally efficient financial markets. As in an efficient market all the information is being conveyed to the prices, the herd behavior should be nonexistent in financial markets. The Banerjee's (1992) and BHW's (1992) basic model assumes that the investment opportunity is available to all individuals for the same price and therefore it does not fit directly to e.g. stock markets (Bikhchandani & Sharma ,2000).

When the price adjusts to the accumulated information it should negate the directional impact of the public information's accumulation (Avery & Zemsky, 1998). For an example when an individual settles on a buy decision of a stock, the private signal used to buy, is conveyed directly to the public information and thus enhances the precision of public information to buy. At the same time the price of the stock raises due to the increased demand of the stock an therefore the cost of adaption is now higher than it was at the beginning and if the effects on both are equal then they completely negate each other. So for herding, the private expectation and prices must diverge.

As we can see, given the empirical studies presented above, there is evidence that herding do happen in financial markets though the cause of it isn't clear. There's indeed herding to be acknowledged but is there a general theory behind it that is economically resistant and that includes the assumption of rational individual or is the herding happening through irrational instinctive individuals? To offer answers to this question I will present few studies which have studied the occurance of herd behavior in financial markets.

4.3. Rational herding in financial markets

There are restrictions to the herd behavior in financial markets. A financial asset bought by one market player has to be sold by another market player, therefore all individuals can't be part of the buying or selling herd (Bikhchandani & Sharma, 2000).

The possibility of herding is characterized in the context of simple, and informationally efficient financial market. In Avery & Zemsky's (1998) and Park & Hamid's (2011) papers there is a market maker who organizes the trade. The market maker does not have private information and in every period, he posts a bid and an ask price. All the public information available is reflected to the prices. So he updates the prices once in a period.

The markets also have two kind of traders: noise traders and informed traders (Avery & Zemsky, 1998 and Par & Hamid, 2011). The noise traders trade randomly and have no information to use when they trade. The informed traders are rational traders who, in these models, are assumed to be risk neutral. So in these models there's irrational behavior included.

In the literature of herd behavior in financial markets there are two types of behavior recognized: herding and contrarianism (Park & Hamid, 2011 and Avery & Zemsky, 1998). Herding occurs like defined above when an individual chooses to act according to the public information rather than his private signal which would have advised otherwise. Contrarianism occurs when an individual decides to act against the public information.

Bikhchandani & Sharma (2000) name two kind of herding, spurious and intentional. Spurious is driven by the noise traders and other factors that aren't rational. Intentional herding can lead to excess volatility and systematic risk (Bikhchandani & Sharma, 2000). The prerequisites of rational herd behavior differ between studies. Bikhchandani & Sharma (2000) name three causes of rational herd behavior: imperfect information (information asymmetry between market maker and investors), concern of reputation and compensation structure.

During herding, prices can move substantially and it can induce lower liquidity and higher price volatility (Park & Hamid, 2011). In financial markets the liquidity is one factor to be considered since in financial markets the buyers are usually also the sellers.

As can be seen, when implementing the herd behavior in financial markets there are many things to consider and many aspects that are needed to be applied to the herd behavior theory to also make it valid in financial markets. I now present shortly the model introduced by Avery & Zemsky, 1998.

4.3.1. Multidimensional Uncertainty (Avery & Zemsky, 1998)

In chapter 2, the application of the herd behavior model which included cost and value of adoption was presented. In this model, the cost of adoption was a fixed cost. In Avery & Zemsky's (1998) study the model of BHW (1992) is used as a premise of their own model. They use the model with an alteration where cost of adoption isn't fixed, instead, it is a price which is updated according to publicly available information in every period. In an informationally efficient market, this is to be interpreted so that every buy and sell decision should be reflected in the price.

If the uncertainty is only about the value of the underlying investment, the stock market price is informationally efficient and herd behavior will not occur (Avery & Zemsky, 1998). Due to this assumption Avery & Zemsky (1998) introduced models with more than one dimension of uncertainty and demonstrated that with two dimensions of uncertainty, herding can happen in financial markets, although it does not cause mispricing but with three dimensions, the herding can cause mispricing and, in certain circumstances, price bubbles.

First dimension of uncertainty is *value uncertainty*. The individuals decide using their private signals and public information. In informationally efficient market the market maker is able to use public information when pricing, and therefore the prices convey public information directly and therefore the individuals are always left to decide with their private signal compared to the price and herding does not occur (Bikhchandani & Sharma 2000 and Avery & Zemsky, 1998). Therefore the uncertainty only lies within the individual's expected value of adoption and the individuals act only based on the information asymmetries between themselves and the market maker. Avery & Zemsky (1998) say that traders' signals are monotonic and therefore their private signal vary with the price, the assumption is that if the price goes up and the individual decides to buy then he would've also bought in an ex ante situation.

The second dimension of the uncertainty is *event uncertainty*. It makes the herd behavior possible but does not distort the prices since the herd behavior is assumed to be transparent to all market participants. With only value uncertainty, there's only one interpretation of the history of actions. The second dimension makes the signals non-monotonic and therefore enables the herd behavior. In the event of uncertainty there lies a possibility that there's public information that is not known to all, and therefore the market maker may not always have all the public information at his use.

The third dimension of the uncertainty is *composition uncertainty*. The third dimension considers the transparency of herd behavior. It proposes that there lies uncertainty within the knowledge of the composition in market participants. Are the participants well-informed or badly informed and how many of them exists in the market, is not common knowledge.

With these three dimensions of uncertainty Avery & Zemsky (1998) proposes that the price bubbles can arise even though it is very unlikely. This model is one of the many models demonstrated to illustrate the rational herd behavior in financial markets. There is no one model that is generally approved.

Avery & Zemsky (1998) argue that with informationally efficient asset prices herd behavior is not possible unless signals are non-monotonic. With unidimensional or monotonic signal structures herding is impossible. Park & Hamid's (2011) results conclude that herding may take place in more general fashion and the signals do not need to be non-monotonic for prices to move extremely.

5. Conclusion

The thesis gives a basic understanding of the herd behavior phenomenon, explains it and its impacts. The thesis is written to give a general understanding of the topic and to motivate the reader for further research. It's also written to familiriaze the author on the phenomenon of herd behavior, keeping an eye on a master's thesis.

The herd behavior theory needs more empirical research in establishing its position as a general economic theory. The rationality and the use of Bayes' theorem in decision making process is hard to prove to be an intrinsic part of an individual's decision making process empirically. In laboratory tests (Anderson & Holt, 1997 and Huck & Oechssler, 2000), the application of the theory has been controversial and no unanimous proof of the use of Bayes' theorem being accurate when modeling behavior of the individual, has been provided.

The economic theories of herd behavior and the assumption of rational individuals give the herd behavior a more manageable form. The surveys cover the theoretical framework of herd behavior but the evidence to prove it isn't yet robust enough. If the theory is developed further and establishes its position as a general economic theory it would have a great impact in many fields. When understood, and proved even further, the theory could be used to fix the inefficient outcomes and information asymmetries, and therefore it could enhance the welfare in society. It would also consolidate certain business strategies and could be used for the advantage of the seller like in the case of High Sales Promote High Sales.

The theory now(Banerjee, 1992) has a lot of presumptions. It assumes that the individuals arrive in a sequence and that in each time period only one individual is able to decide. There's also an assumption of the actions of the individuals to be visible to all other individuals and therefore the assumption of public information being common knowledge.

Considering the public information, it is worth for further research to analyze how wide these influence networks are and how smaller networks of public information affect each other. In one example Watts & Dodds (2007) introduce local and global networks in their research. Another question to be considered is if the individuals arrive at the same time, en masse, or if there is an arbitrary amount of individuals arriving in different periods.

In the financial markets, rational herd behavior models are provided (e.g. Avery & Zemsky, 1998 and Park & Hamid, 2011) and in them it's proved theoretically that the rational herd behavior is present and it can affect the prices when certain assumptions have been made. Even though there is a theoretical framework for the herd behavior in financial markets, it still lacks the evidence from empirical research.

The empirical surveys in financial markets ran through the statistical methods using the historical data of trades (Lao & Singh, 2011, Zhou & Lai 2009 and Chiang & Zheng, 2010) prove the existence of herding but they do not prove that it would be caused by rational individuals. Park & Hamid (2011) present in their model that there are noise traders, alongside the rational informed traders, who trade randomly and therefore their model does not rest on the rationality assumption solely. Zhou & Lai 2009 argues to have found evidence on informational cascade (BHW 1992) on their empirical research. This statement is intriguing but in the light, that informational cascade is mostly theoretical concept (Chamley, 2004), it needs to be examined further.

Even though herd behavior has been under active research for a few decades now, there's still a lot to be researched for. Therefore, for the interest in the topic, it bears a lot of possible subjects to research e.g. the influence of social media on public information, the impact of herding on populistic parleys in politics and financial markets in general.

References

Anderson, L.R., Holt, C.A. (1997). Information Cascades in the Laboratory American Economic Review, 87 (5), pp. 847-862

Avery, C., and Zemsky, P. (1998). "Multidimensional uncertainty and herd behavior in financial markets." American economic review: 724-748.

Banerjee, A. (1992) A simple model of herd behavior. Quart. J. Econ. 107, 797–817

Bikhchandani, S., Hirshleifer, D., Welch, I.

(1992) A theory of fads, fashion, custom, and cultural change as informational cascade. J. Polit. Economy 100, 992–1026

(1998), "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades", The Journal of Economic Perspectives, 12, 151–170.

Bikhchandani, S. and Sharma, S. (2000). Herd behavior in financial markets. IMF Staff Papers, 47, 279–310.

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

Chamley, C.P., Rational herds (2004). Economic models of social learning. Cambridge University Press, 1-39&58-91.

Huck, S. and Oechssler, J. (2000). "Informational cascades in the laboratory: Do they occur for the right reasons?." Journal of Economic Psychology, 21(6), 661-671.

Lao, P., & Singh, H. (2011). Herding behaviour in the Chinese and Indian stock markets. Journal of Asian Economics, 22(6), 495-506.

Park, A. and Hamid, S. (2011). "Herding and contrarian behavior in financial markets." Econometrica 79.4: 973-1026.

Vives, X. (1997). Learning from others: a welfare analysis. Games and Economic Behavior, 20(2), 177-200.

Watts, D. J., & Dodds, P. S. (2007). Influentials, networks, and public opinion formation. Journal of consumer research, 34(4), 441-458.

Zhou, R. T., & Lai, R. N. (2009). Herding and information based trading. Journal of Empirical Finance, 16(3), 388-393.