Aalto University
School of Business

Word of mouth in social learning: The effects of word of mouth advice in the smartphone market

Economics

Master's thesis

Mikael Head

2013

Department of Economics
Aalto University
School of Business

Abstract

Objectives

The objective of this thesis is to examine word of mouth advice and its relationship with product sales and market shares in the context of the smartphone market. The thesis aims to determine the key properties of valuable word of mouth advice from a consumer's perspective and seeks to identify the effects of sources and transmission methods on the valuation of word of mouth advice. Furthermore, the thesis aims to clarify the market wide effects of positive word of mouth on product market shares through empirical analysis of the effects of word of mouth advice on operation system market shares in the smartphone market.

Framework, data and methods

The methodology of the thesis is based on building a framework for empirical analysis by reviewing relevant literature on social learning and word of mouth advice and subsequently examining the individual level effects, as well as the market wide effects, of word of mouth advice though empirical analysis of smartphone market related data. The data utilized for the empirical analysis section of the thesis consists of two datasets including survey questionnaire responses and sales figures of smartphone handsets in various markets. The empirical methods utilized in analyzing the datasets include the Heckman selection model as well as the ordinary least squares method, the fixed effects estimation method and the random effects estimation method.

Findings

The main findings of this thesis are that, in terms of the determinants of valuable advice, the effectiveness of word of mouth is highly correlated with the strength of the social tie between the advice giver and the receiver of the advice. A closer social tie implies a higher rating for the advice received. Other factors contributing positively towards a high probability of a high valuation for advice are active search for advice and the receiver's familiarity with the subject of advice. Also, a variety of socio-economic factors such as gender and the place of residence of the respondent were found to result in in a higher probability for favorable ratings for advice.

In addition to this, in terms of the market wide effects of positive word of mouth, the thesis finds strong correlation between high shares positive word of mouth and high market shares for smartphone operation systems. The findings however experience large variations across markets and more research will be required to completely uncover the nature of the relationship between word of mouth and product sales. While the type of herd behavior implied by social learning theories certainly seems possible as a real world market outcome, more research in the field is needed to determine the actual strength of the phenomenon and the factors contributing to its origins.

Keywords

Social learning, herd behavior, word of mouth, advice, social ties, information technology

Table of Contents

List of Tables

1. Introduction

The purpose of this chapter is to give an introduction to this thesis and to offer an overview of the key aspects of the thesis, its research questions and its findings. First, the motivation for the thesis and its background are discussed to bring context to the analytical questions of the thesis. Second, the objectives and methodology of the thesis are outlined to illustrate the methods of empirical analysis utilized in deriving the empirical findings of the thesis, and to show how the thesis is related to the relevant bodies of academic literature. Third, the empirical research findings, and the main contributions of the thesis summarized. Fourth, and finally, the structure of the rest of the thesis is outlined for guidance.

1.1 Background and motivation

The importance of word of mouth generated advice in shaping human decisions has been recognized throughout time, but formal analysis of the widely observed phenomenon of word of mouth has been few and far between. As noted in the work by evolutionary biologists, such as Hoppitt & Laland (2012), the integral propensity of learning from the behavior of others has been in advantageous in the development of the human species, and still today remains as an influential force in guiding human decision making. The role of traditional advice, whether actively sought or passively received, has always been recognized as influential, but the actual effects of this kind of word of mouth have traditionally been seen as vaguely defined and, little work has been done in terms of academic literature on attempting to quantify the market wide effects of word of mouth advice.

The theoretical base of word of mouth advice as an influence is firmly rooted in theories of social learning developed though the course of the last two decades. The recent advances in social learning theories have provided new context to word of mouth and its role as a tool for aggregating information. Herding models, such as those of Banerjee (1992) and Bikhchandani, Hirshleifer & Welch (1992), started a stream of literature that has brought forward dozens of models for social learning illustrating the different possibilities for word of mouth to facilitate information aggregation and asymptotic learning, i.e. convergence of the learning outcomes. Especially the pioneering works of Ellison & Fudenberg (1995) & Banerjee & Fudenberg (2004) has been instrumental in broadening our understanding of how word of mouth can facilitate social learning.

The theoretical models of social learning have also been tested a laboratory environment and mainly supportive evidence in their favor has been found. While some studies such as Anderson & Holt (1997) have found validating support for the theoretical social learning models such as that of Bikhchandani et al. (1992), there has also been additional evidence on the power of advice facilitated social learning been brought forward by recent experimental studies such as Çelen, Kariv & Schotter (2010). Based on their experiment on word of mouth advice facilitated social learning Çelen et al. found that advice to be beneficial for the subjects payoff, not only for the subject receiving advice, but also for the subject forced to advice.

The importance of social ties in the transmission of word of mouth advice has for long been recognized as in important piece of the puzzle and the conventional wisdom on the matter has been that closer social ties lead to more influential advice. This view though has come under critique due to recent research findings. In their study on the factors influencing college course choices Steffes & Burgee (2009) find impersonal and anonymous advice received via an online platform to be more influential in guiding decision making than peer advice from a well-known source with a closer social tie, contradicting the conventional view on the matter.

While on a theoretical level word of mouth advice has been successfully incorporated to the numerous theoretical models for social learning and the effects of advice on learning outcomes have been tested in laboratory experiments, apart from some few recent contributions such as Chevalier & Mayzlin (2006) and Duan, Gu & Whinston (2008), there has not been much empirical analysis of its implications on real world product markets brought forward. The scarcity of empirical research on the determinants of the effectiveness of word of mouth both on an individual level as well as at a market level urges more research to be done in the field to deepen our knowledge on the effects of word of mouth on the market behavior of economic agents.

1.2 Objectives and methodology

The objectives of this thesis are to deepen our understanding on how advice and word of mouth are treated in the context of a product purchase decision and how they affect the market outcomes through consumer behavior. The thesis aims to provide answers to such questions as where does the most valuable word of mouth advice come from in terms of sources and channels of transmission, and is actively sought word of mouth different from passively received in terms of effectiveness? Additionally, the thesis aims to discover how word of mouth advice affects the relative market shares of smartphone operation systems, and whether word of mouth advice can plausibly create significant herd behavior in a product market context.

The methodology of this thesis is to establish a theoretical framework for studying the effects of word of mouth via an extensive literature review of previous theoretical and empirical literature, and to utilize empirical analysis in this framework to estimate the effects of word of mouth, both at an individual consumer level as well as at a market level. In the empirical part of the thesis the effectiveness of word of mouth at the individual consumer level is examined with the help of a Heckman selection model first presented by Heckman (1979). In the panel data estimations of the empirical part of the thesis the ordinary least squares method, the fixed effects method as well as the random effects method are used to estimate the effects of word of mouth on operation system market shares in the smartphone market.

1.3 Findings and contribution

As the areas of interest in the empirical part of this thesis are divided between the individual level and market level effects of word of mouth, so can the findings of thesis can be divided between the two. In terms of sources for word of mouth advice, this thesis finds the ratings given for advice received to be highly correlated with the strength of the social tie between the advisor and the receiver of advice. This result reached when differentiating between the sources of word of mouth is supported by the estimation results for transmission channel effects. Both actively seeking advice and being a technological forerunner are also found to have a favorable effect on the probability giving high ratings for advice received on the purchase of a high technology device in the form of a smartphone. Interestingly, consumers living developed countries also had a considerably higher probability to rate advice highly compared to consumers living in developing countries, likely reflecting cultural differences and the of inherent product quality effects.

In terms of market wide effects of word of mouth, the results of the empirical analysis of this thesis implicate word of mouth to be strongly correlated with market shares and positive word of mouth having a positive effect on operations system market shares in many of the smartphone markets analyzed. These results are however far from conclusive as there is large variation in estimation results between different operation systems and markets analyzed. Additional limitations to the generalizability of these results are also brought by data limitations hindering the exact identification of forces affecting product market shares in the smartphone operation system market. As the market is highly dynamic and in constant rapid development, there is a significant possibility for individual product specific time variant effects hindering the econometric identification of causality between word of mouth and market shares.

The contribution of this thesis to the body of knowledge on social learning and word of mouth is to identify some of the factors related to the effectiveness of word of mouth from an individual agent's perspective and to attempt to quantify the market dynamic effects of positive word of mouth in a real world product market setting. As neither of these subjects has been previously extensively researched from an economics perspective, the purpose of this thesis is to broaden our knowledge of these subjects and phenomenon, and to hopefully in its part help to facilitate further future research in the field.

1.4 Structure of the thesis

The structure of the rest of the thesis is as follows. Chapter 2 outlines an overview of social learning and the different economic theories and models used to model it. Chapter 3 will discuss the important earlier research on social learning and word of mouth. Chapter 4 gives a description of the smartphone market and describes the sources of data used in the empirical analysis sections of the thesis. Chapter 5 lays out the key empirical research questions of the thesis and describes the empirical framework and econometric models used for empirical analysis. Chapter 6 presents the empirical results of the thesis. Chapter 7 concludes by discussing the findings of the thesis, its main limitations and its key implications for future research.

2. Theories of social learning

Imitation and conformation to social norms have always been an integral part of human behavior. As already put into words by Machiavelli (1988) in the early sixteenth century, “For men almost always follow in the footsteps of others, imitation being a leading principle of human behavior.” (p. 19) This tendency inherent in all of us to conforming to norms and imitating others is also not limited only to human beings. For example Morgan, Rendell, Ehn, Hoppitt & Laland (2012) among others argue that the bases of human social learning are deeply rooted in our evolutionary development as a species. Galef & Laland (2005) have brought forward empirical evidence and theoretical models concerning social learning among animals. Learning from others has thus always been an invaluable tool for various species for acquiring and aggregating information. Aggregation of information among group members has traditionally been seen as an efficient way to make the most out of the scarce observational and cognitive resources possessed by individuals.

The exact definition for social learning varies between scholars, but it can be broadly defined as studying how agents learn by observing the behavior of others in an asymmetric information environment, and how the aggregation of information among agents affects the equilibrium outcome. While not always defined as such, the concept of social learning has nevertheless been around for centuries now. Indeed, according to Chamley (2003) the earliest formal analysis of information aggregation and the efficiency of group behavior date back to the 18th century and the pioneering work on jury behavior and election design by Marie Jean Antoine Nicolas de Caritat, marquis de Condorcet. De Caritat (1785) presented the first formal illustration of analysis and modeling of the information aggregation process focal in all social learning theory, and modern research on the subject owes much to the much to its methodology.

As economies and societies fundamentally consist of individual agents, economic and social outcomes are essentially a product of the decisions made by these agents under varying degrees of social interaction. It is therefore important to consider the effects of the information aggregation process between the agents to learning outcomes and their efficiency. Although the methods of social learning have undoubtedly benefitted the evolution of various species, humans in particular, the equilibrium outcomes produced by social learning are however not always strictly efficient or desirable.

In recent decades, a wide variety of models for social learning have been introduced by a number of different economists. Each of these models typically studies a certain aspect of social learning under some distinct behavioral- and institutional assumptions that strongly affect agent behavior, the information aggregation process and the equilibrium outcome. Although the concept of social learning is common to many social sciences, it can be seen as a vital part in determining the behavior of economic agents, and therefore as an integral component of all economic theory. Most theories for social learning in economics are usually set in a microeconomic environment and describe the behavior of agents in a microeconomic setting. The implications of the theories are however not limited to microeconomics only, as they can play a significant part in describing agent behavior and the information aggregation process in the current micro founded models popular in contemporary macroeconomics.

In economics the focal point of social learning theories has been on the process of information aggregation and how it contributes to the possibility of asymptotic learning outcomes. Asymptotic learning implies that the social learning system will converge to an asymptotic equilibrium with all agents choosing one and the same action or alternative. The possibility for asymptotic learning is important as it can have large implications on efficiency and welfare. If the system experiences asymptotic learning and eventually converges to an asymptotic equilibrium, the equilibrium might or might not be an efficient one. This can, at least in theory, result in the system converging to an inefficient equilibrium, such as an inferior technology being adopted over the societally more efficient one. Asymptotic learning is also a key aspect in economic theories explaining observable phenomenon such as herd behavior, a subject which I turn to next.

2.1 Herd behavior

Herd behavior is a well-known and an often observed phenomenon where agents decide to choose the same alternative from a set of multiple available alternatives, in other words herd towards one of the options. As the name suggests, the term herd behavior pertains to the tendency of animals to gather together into flocks and herds. In human behavior, herding is also commonly observed. Herd behavior has so far been already used to explain various observable events such as how peaceful street demonstrations can sometimes turn into violent riots, the spreading of fashion trends and why irrational financial bubbles are frequently observed, as noted by Bikhchandani et al. (1992) among others. Although herd behavior is such a commonly observed phenomenon in real life, no definitive explanation for this observed tendency has yet been given. Many possible explanations have been voiced in literature by various economists, but no point of view can be seen as completely dominating others. In fact, it is a generally held belief that the reasons for herding vary from one situation to another, and that there are usually multiple factors simultaneously affecting the outcome.

2.1.1 Informational cascades

From a social learning perspective the contributions of economics to the body of knowledge on the reasons and implications of herd behavior and its related concepts are usually seen to have begun with the nearly simultaneously published seminal works by Banerjee (1992) and Bikhchandani et. Al. (1992). The two papers are widely seen as having given birth to the whole field of modern economic analysis of social learning and they form the foundation on which almost all subsequent work in the economics of social learning has been built upon. Both studies utilize somewhat similar models to study the effects of information aggregation on asymptotic learning and herd behavior. The key drivers behind herding suggested by both studies are informational cascades. Formally informational cascades are defined by Bikhchandani et al. (1992) as, “An informational cascade occurs when it is optimal for an individual, having observed the actions of those ahead of him, to follow the behavior of the preceding individual without regard to his own information.” (p.992)

The idea behind the informational cascades theory is roughly speaking that, in a social learning setting where each agent possesses at least some amount of private information, it is rational for decision makers to consider the behavior of other agents before making one’s own decision The reason for this being that the decisions of others should at least in theory reflect their private information that other agents are not privy to. Including the publicly available information into the decision problem will on the other hand reduce the information value of each agent’s private information that the other agents receive via observing their decisions. This of course steers the agents to utilize more of the public information, and hence become less responsive to their own private information, further reducing its informational value.

The likely result of this all as argued by Bikhchandi et al. (1992) is that, because all of the agents are trying to benefit from the private information of other agents, everyone will end up choosing the same choice as all the other agents, even in a situation where their own private information would advise them to do otherwise. The agents thus choose to ignore their own private information in favor information received from other agents creating an informational cascade. The natural implication of such a cascade is asymptotic learning and herding towards one of the possible choices. The subject of which types of circumstances are required for informational cascades to appear and cause significant herding among agents is an actively discussed topic among researchers. Asymptotic learning outcomes described in social learning literature have been very sensitive to model construction and to different model specific parameters. It also needs to be remembered that that, while the presence of an informational cascade implies herding to result, herd behavior is not necessarily the caused by the presence of a cascade.

2.1.2 Externalities

Apart from informational cascades just discussed, other reasons for herding voiced in literature have range from internal needs for conformity and the punishment of deviants to positive payoff externalities and network effects. A general unifying theme to the non-social learning explanations for herd behavior is that, while social learning models center on studying the effects of purely informational externalities on the equilibrium outcome, the other avenues of research usually include payoff related externalities as well as other type of externalities for agents' actions to be included in the decision rules applied by agents.

Perhaps the most famous argument in economics literature for attributing herd behavior to be driven by other factors besides informational externalities is given by Becker (1991), who introduces a demand externality model for consumer demand for restaurants. In the model, an individual's demand for the product of restaurant services is positively dependent on the aggregate demand for the product. As one can immediately suspect, the model produces strong herding behavior, as a small amount of initial demand for one of the products will strengthen the future demand of that product. In the long run Becker's model ultimately converges to an asymptotic equilibrium, with all consumers choosing the same restaurant. For illustrations of payoff externalities driven herding one can also consider the obvious benefits of all drivers heading in the same direction driving on the same side of the road and the positive network effects related to e.g. the diffusion of telephones or the internet, which hard to question.

It is of course at least in theory also possible for herd behavior to arise from the fact that similar agents facing a similar choice situation are bound to make similar choices. For example, if there is one product in a market that is distinctly superior to all the other products, it would only be natural that all agents would then decide to purchase that superior product. Nonetheless, it is more than plausible to assume at least some degree of variation between the agents' preferences and hence their private information: This in turn implies that there is at least some degree of information- or payoff related externalities being at play in the presence of considerable herding of behavior.

Multiple factors including payoff externalities can be seen as having considerable explanatory power in explaining herd behavior in various types of real world settings. Nevertheless, from the viewpoint of social learning literature they are not seen as relevant as informational cascades in the settings of interest to social learning theories. As the focus of this thesis is on social learning, and more specifically word of mouth facilitated social learning, the rest of this thesis the focus will mainly be on models with purely informational externalities and the other aspects affecting learning outcomes are only discussed when needed.

2.1.3 Communication in learning

The majority of social learning models in economics focus on a learning structure where agents gather information only by observing the actions or choices of other agents and there is no direct communication between them. An essential line of research has however also developed models where information aggregation is facilitated by agents communicating with each other and receiving advice via word of mouth communication. Even though there is a commonly held general belief that actions speak louder than words, in a social learning setting the two forms of information transmission can in most cases considered to be comparable. In addition, recent experimental evidence to support the notion that in some cases advice and word of mouth can be even more influential in affecting decisions than observing actions as demonstrated e.g. by Çelen, Kariv & Schotter (2010).

The relative importance of actions vs. words in aggregating private information in all probability depends on the theoretical and experimental setting, which will be discussed more in depth later on when some of the experimental evidence on the matter is reviewed. For now I simply follow the approach of Chamley (2003), who noted, "When agents learn from actions, these actions are the "words" for communication." (p. 4). Based on this, in the context of this thesis, the information value and effectiveness in facilitating learning can with fair degree of generalizability be seen roughly equal, if not noted otherwise. This justifies comparison between different models of social learning that feature either actions or words for information transmission in their learning mechanism. While the body of literature dealing with social learning is rather extensive, the variety utilizing specifically word of mouth communication in a social learning mechanism is alas still rather narrow. This in turn further stresses the importance of comparability of models with other forms of observable behavior such as actions discussed above.

The first work to explicitly introduce word of mouth communication to the social learning theory literature was the seminal contribution of Ellison & Fudenberg (1995), who consider a simple model of competing products with naïve and boundedly rational agents utilizing a myopic decision rule and information transmitted through word of mouth communication in determining their choices. They find that, while certainly being a plausible outcome under some light premises such as limited sample sizes, the realization of asymptotic learning depends closely on the structure of the word of mouth communication process. In addition, their findings also implicate that a smaller amount of communication, or a smaller sample size in the context of their model, makes conformity and thus herding more likely.

An alternative view on social learning via word of mouth is taken by Banerjee & Fudenberg (2004), who in turn examine a model of word of mouth learning incorporating rational agents sampling N other agents under a Bayesian learning mechanism to establish the convergence properties of such a system. The conclusions of Banerjee & Fudenberg are somewhat similar to the conclusions of Ellison & Fudenberg (1995) in the sense that they find that, under suitable assumptions such as proportional sampling, there is a strong possibility for asymptotic learning to result. In contrast however, their results implicate that large samples of agents from which agents receive communication are more favorable to herding than small samples.

The disparity between the results of the two studies highlights the enormous sensitivity of learning outcomes in social learning models to the structure of the model and its underlying assumptions. Just as most theories and models produced on the subject differ in their assumptions on the structure of the world, they also differ in their findings under these assumptions. In following sections I go through some of most vital assumptions and modeling choices applied in relevant social learning literature regarding sampling rules used by agents and the structure of communication in order to draw conclusions on their implications on any possible asymptotic learning outcomes.

2.2 Learning mechanisms

The models presented by Ellison & Fudenberg (1995) and Banerjee & Fudenberg (2004) can be classified as representing the two different main branches of social learning research in respect to the learning mechanism they incorporate. The former features boundedly rational agents utilizing a myopic decision making rule, while the latter assumes rational agents able to make use of Bayesian learning abilities in their decision making. The above distinction is not definitive in the sense that some social learning models combine both Bayesian and myopic elements, and that there are several other important model characteristics that can be used to classify social learning models. This particular type of classification into rational and non-rational models is however undoubtedly useful, as the learning methods applied by agents usually have repeating effects on outcomes through other model specifications and partially determine other influential factors such as e.g. the sampling rules used by agents and the structure of the communication process among others.

2.2.1 Bayesian learning

A large proportion of the models suggested in relevant social learning literature, the seminal works of Banerjee (1992) and Bikhchandani et al. (1992) among others, feature some form of a rational Bayesian learning mechanism to facilitate learning. The term Bayesian learning refers to an environment where agents act according to a decision rule where they incorporate Bayesian inference and Bayes' rule to constantly and rationally update their beliefs when new information becomes available. Formally Bayes' rule can be is defined as:

$$p(A|B) = \frac{p(B|A)*p(A)}{p(B)}, \tag{2.1}$$

where $p(A|B)$ represents the posterior probability assigned to A after B has been observed, $p(B|A)$ represents the likelihood, or in other words the probability, of observing B after observing A and $p(A)$ and $p(B)$ respectively represent the independent prior probabilities of observing A or B. In Bayesian inference, the posterior probability for A determines the probability assigned by an agent to the state of the world A occurring after having observed the relevant set of data in the form of observing B. All this simply means that agents constantly update their assessments of the probabilities for different states of the world, probabilities derived via Bayes' rule, and, as soon as new information becomes available, agents immediately incorporate it to their assessments on which they base their decisions.

The decision rules utilized in the models of Bayesian social learning vary, but they generally tend to follow the same pattern of agents first rationally forming beliefs and then choosing the optimal choice from some group of alternatives according to a decision rule determined by model construction. As discussed more in depth below, many researchers that have opted to model social learning with boundedly rational agents take the Bayesian world as their point of departure and introduce bounds on agent behavior based on their modeling needs. The most common restrictions on agent rationality include e.g. myopia in the sense that agents only maximize their current utility and not their lifetime utility and restricting agents' to forming beliefs only on the behavior of the subset of other agents who they can observe and not the entire population. The exact decision rules that agents are assumed to utilize nevertheless vary between models and are usually determined jointly with other key model characteristics such as the sampling rule used to determine the observations of agents and the structure of communication and information transmission.

2.2.2 Myopic learning

The other important class of learning mechanisms in social of research, apart from the Bayesian models is boundedly rational models that restrains from the heavy assumptions of Bayesian rationality and rely on agents having simpler naïve learning rules guiding their actions. The question of how rational can economic agents assumed to be is of course a difficult question and a question that's definitive answer is outside the main scope of this thesis. Nonetheless, considering the world we live in and the decision environments of real word economic agents, the boundedly rational models of social learning certainly seem as plausible proposals in many instances. In the real world it is usually plausible to presume at least some degree of boundedness in the rationality of agents, if nothing else stemming from our bounded cognitive abilities. No conclusion on whether or not assuming bounds on rationality is the right approach for modeling agent behavior in social learning models has yet been reached, and at the moment both Bayesian and myopic models can be validly seen as relevant predictors of agent behavior, depending on the situation.

While there exists a wide array of different myopic decision rules applied to social learning context the seminal illustration of a boundedly rational decision rule in social learning literature is given by Ellison & Fudenberg (1993), who present two models for technology adaption with boundedly rational agents. For example, the first of their two models incorporates homogenous agents utilizing a naïve and myopic rule of thumb decision rule that ignores all historical data and considers only the optimal choice of the previous period when making the decision on which of the two possible technologies to adopt. The evolution of the system at issue is then described by:

$$E(x_{t+1}|x_t) = (1-\alpha)x_t + \alpha p, \tag{2.2}$$

Where x_t and x_{t+1} represent the fraction of the population using technology A in periods t and t+1 respectively. Naturally, the fractions using technology B are then respectively given by (1- x_t) and (1- x_{t+1}). α represents the fraction of the population chosen randomly that are allowed to revise their choices at each period and p is simply the probability that using technology A gives agents a higher payoff than using technology B.

The naïve decision rule of agents in Ellison & Fudenberg (1993) outlined above states that all agents allowed to revise their choice of technology choose the technology that had the higher payoff in the preceding period after observing the average payoffs of the two choices in the previous period. As the population is homogenous, the users of each technology receive the same payoff as all others using that technology. The payoff parameter p describing the relative payoffs of the two technologies is defined by a stochastic process consisting of the sum of a constant, but unknown, parameter θ and an independently and identically distributed shock parameter $\varepsilon_{t,}$ with a zero mean. The role of the shock parameter $\varepsilon_{t,}$ is to bring variance to the relative payoffs of the two technologies. Depending on the intrinsic payoffs of the technologies the effects of the changes in $\varepsilon_{t,}$ may or may not be enough to affect the ordering of the payoffs.

A notable feature of Ellison & Fudenberg (1993) that separates their models from some of the other similar models brought forward is that periodically some of the agents are allowed to re-assess their decision between the technologies and are not constrained to making a once and for all decision. This is one of the main drivers in their results as demonstrated by their first proposed model described above in which the speed of technology adoption is correlated with the level of the payoff difference between the two technologies, which in turn is determined by their intrinsic quality and an exogenously determined technological shock parameter.

Although the welfare implications of social learning models are examined more thoroughly in a subsequent section, I note that according to Ellison & Fudenberg (1993) the practical implication of their simple model discussed above is that technologies that have a small probability of resulting in a big payoff improvement and a large probability in resulting in a small loss will in all likeliness be adopted relatively slowly. If true, this can have substantial effects on a societal level as e.g. vaccines and seatbelts can be classified as technologies possessing such traits.

In conclusion, both the models based on Bayesian learning and the "rule of thumb" models featuring non-Bayesian learning have their place in social learning literature. The relative abundance of Bayesian models compared to the more myopic models is most likely a signal of most researchers seeing the world of fully rational agents as their starting point in modeling, and depending on their perspective and the exact target of research introduce bounds on agents' rationality when required. Although the structure of learning is one of the most vital elements in modeling social learning, the equilibrium outcomes are generally a sum of numerous factors. An important determinant of model specification closely related to agents' learning mechanisms is the sampling rule that is used in forming the samples that agent's observe. The observed samples are one of the main inputs of the learning mechanisms and decision rules described above and hence warrant a closer look in the next section.

2.3 Sampling rules

In tight relation to learning mechanisms in determining the attributes of social learning are the sampling rules agent's use to form their beliefs. These sampling rules can be in regard to observing the actions of other agents, their payoffs or receiving information in another form such as advice or word of mouth. In addition to the structure of the learning mechanism itself, the choice of sampling rule is one of the most important determinants of social learning models and is usually considered to be the key determinant of system behavior and convergence. Sampling rules determining what agents observe are related to both the form of information they receive, in other words whether they observe actions or receive communication etc., and the amount and variety of information agents receive in terms of the size of the group of agents they sample

The sampling rules incorporated into social learning models range from the simple proportional sampling rules, such as observing the behavior of all agents, to more selective sampling rules constraining observations to a small group of agents or even to a single agent. Broadly speaking, in relation to the chosen sampling rule, social learning models can be divided into two different categories. In the first category there are models with proportional sampling and samples give an accurate picture of the whole population. In the other category there are models that use un-proportional or biased sampling and hence oversample a certain group of individuals. This second category of models with oversampling can be further divided into models where decision makers observe only the behavior of single other decision maker, to models where they observe the behavior of a small group of individuals determined by a specific observational rule such as their neighborhood or the existence of a group of extensively influential group of individuals and to models where all the behavior of all agents is observed but an un-proportionally large part of the sample is formed by observations of the behavior or the prominent group.

2.3.1 Proportional sampling

The most common proportional sampling rules applied are either letting agents observe all previous actions, such as in Bikhchandani et al (1992), or independently drawing the observations from a probability distribution, such as in Banerjee & Fudenberg (2004), or letting agents observe a sufficient random sample of the population as in Ellison & Fudenberg (1995). When many of the early contributions to social learning theory, such as Bikhchandani et al. and Ellison & Fudenberg brought forward models with proportional sampling, the current focus of the field has evidently shifted towards models with non-proportional sampling and their implications. Proportional sampling models have however remained relevant, and other notable examples of models featuring proportional sampling are given by Vives (1997), Smith & Sørensen (2000). Each of these models differs on their characteristics and also on the proportional sampling rule applied.

The exact definition for what is meant with proportional sampling has as of yet not been uniformly defined and definitions used vary between scholars. It can mean either agents observing all behavior preceding their decision point or just a sufficiently representative sample of that behavior. In the context of this thesis I consider, both samples containing observations of all previous behavior as well as all samples that authors have labeled as proportional to be adequately representative of the underlying population and qualify as proportional sampling. Even if easy to model and intuitively straightforward, proportional sampling rules can sometimes lack in some of the modeling flexibility that the un-proportional sampling rules considered next possess.

2.3.2 Oversampling

Given that the modeling properties and requirements of proportional sampling might be less cumbersome than biased sampling, there can nevertheless be benefits in applying an insightful un-proportional sampling rule. Considering models with oversampling of some proportion of population allows the researcher to focus on the effects of a smaller group of individuals on the outcome of the model and is in most cases also an empirically more plausible solution. After all, we as humans are usually bounded in our samples to observing the behavior of those individuals that it is actually possible for us to observe, rather than receiving accurate information on the behavior of the whole relevant population. In terms of the frequency of sampling utilized in models, sampling usually goes hand in hand with decision points in the sense that social learning is facilitated with allowing agents some form of information sampling before requiring them to make a decision.

A special class of oversampling models is formed by maybe the simplest form of oversampling in which agents observing only the behavior of a single agent, their immediate predecessor and is considered by e.g. Banerjee (1992) and Çelen & kariv (2004a). This class of models is related to other theories of Bayes rational sequential decision making and the unique feature of restraining the sample to include only a single observation diminishes the information set available to decision makers compared to larger samples. Furthermore, Çelen & Kariv note that, as the amount of information available to decision maker increases with the number of observations in his sample, an agent immediately succeeding a deviator in sequential decision making is less likely to follow the deviator as his sample size grows. The results of Çelen & Kariv are supported by similar results reached by Banerjee & Fudenberg (2004), who proceed to argue that larger samples are more favorable to system convergence and hence asymptotic learning as they contain more information and rule out the possibilities of receiving very extreme samples as in the case of sampling only one individual.

Models allowing agents to observe only the behavior of their predecessor are an important part of the theory of social learning and are used in modeling decision making e.g. in some voting situations where decision are announced sequentially. The part of social learning literature using oversampling that is most relevant to word of mouth facilitated social learning literature has nevertheless concentrated on oversampling a small but nonsingular group of individuals. This sort of a small group that is observed by all agents in their sample has been descriptively described by Bala & Goyal (1998) as the model possessing a royal family. The natural extension of this in the context of word of mouth learning is to consider such a group of influential individuals to be some sort of opinion leaders whose behavior and decisions are observed by all agents for some reason or another.

Additional models further considering different forms of oversampling of a certain group of agents include Ellison & Fudenberg (1993) in a boundedly rational environt and Banerjee & Fudenberg (2004) in a Bayesian environment. They both reach similar conclusions in terms of sufficiently structured oversampling including popularity weighing or extreme payoffs can assist in facilitating asymptotic learning.

Building on their basic model of word of mouth facilitated social learning with proportional sampling, Banerjee & Fudenberg (2004) also consider alternative sampling rules of agents experiencing perception biases or reporting biases affecting their samples. Their results indicate that that oversampling of the more efficient choice may lead to the system to be more likely to converge towards it. This result driven by biased sampling rules is important when considering the fact that, as noted by Anderson (1998) among others, word of mouth communication is produced relatively more by individuals who have received extreme payoffs, positive- or negative-, than by individuals who have received payoffs close to average. The joint implication then is that word of mouth communication by agents is likely to drive the system to converge to one of the extremes via facilitating asymptotic learning.

2.3.3 Social networks

An assumption often made in social learning literature is that the signals received from agents via sampling are of uniform value to the decision maker. In the context of word of mouth communication, a more empirically plausible approach would be to consider signals coming from different agents not to be of equal value as in reality people tend to value more the opinion of e.g. the so called experts or in many instances of people to whom we relate to more closely than others. One way of incorporating such properties is to consider the implications of modeling learning through an explicitly mapped social network connecting the agents, rather than e.g. drawing observation samples from a random distribution. Especially in recent years have models utilizing social network mapping been gaining momentum in studying social learning.

This is not surprising as the recent advantages in computer technology has allowed us to utilize ever more computational power in making calculations, which of course can be very advantageous in mapping social networks and computing outcomes. Network modeling has been utilized in both models with Bayesian and non-Bayesian learning mechanisms. The benefits of considering explicit social networks in learning can be quite distinct in terms of understanding between agent communication and its effects. One of the major advantages network modeling is that it allows us to consider the different types of signals agents receive through communication.

In reality, there are numerous factors influencing how an agent perceives the signal he/she receives. For example some of the major factors affecting signal strength, quality and effectiveness relate to the distance between the two agents communicating as well as the identity of the communicators and the strength of their social tie. Assuming all word of mouth to be equal can be advantageous from a modeling perspective, as it can greatly simplify the model and reduce some of the computational burden it might include. Nonetheless, from an empirical perspective, assuming equal signals across agents can be seen as a somewhat naïve assumption and a possible bias producing factor when the real world performance of the model is considered. A closer look on the empirical and experimental evidence gathered on social learning models will be taken in the next chapter.

One of the most prominent examples of a Bayesian model of network facilitated social learning has been brought forward by Acemoglu, Dahleh, Lobel, & Ozdaglar (2011), who examine a Bayesian equilibrium in a sequential learning model utilizing a general form of a social network. Their work can be partly seen as extending the seminal work of Banerjee (1992) and Bikhchandani et al. (1992) to cover more general social networks. Whereas Banerjee (1992) assumes individuals only observe the behavior of the individual immediately preceding them and Bikhchandani et al. assume individuals to observe the full history of play, i.e. all previous actions of each decision maker, Acemoglu et al. postulate that agents to observe the past actions of a stochastically-generated neighborhood of individuals. Depending on some of the additional assumptions made on the private beliefs held by individuals and the possible oversampling of some group of agents, the main findings of Acemoglu et al. indicate a strong probability for an asymptotic learning outcome in the stochastic networks described by their model.

Among the main findings of Acemoglu et al. (2011) is also the fact that, while the presence of a royal family, or an excessively influential group of individuals, can significantly hinder and slow down asymptotic learning, it cannot prevent it completely. According to Acemoglu et al. this result is bound to hold as long as private beliefs are bounded in the sense that there is a bound for the amount of information in a private information signal, and the group of influential individuals does not constitute the full set of observations sampled by the other agents. Another example of a network learning model in a Bayesian rational setting is given by Gale & Kariv (2003), who examine a social network where agents can only observe the actions of the other agents that they are connected to via the network. They find that over time the initial diversity in the network driven by diverged private information is eventually replaced by conformity in actions, if not in beliefs, hence giving support to the findings of Acemoglu et al.

In the setting of boundedly rational social learning models with social networks, some of the important models suggested include Bala & Goyal (1998), DeMarzo, Vayanos & Zwiebel (2003) and Golub & Jackson (2007). The model of Bala & Goyal focuses on examining a connected society where agents are considered to be neighbors when they have access each other's' complete histories of actions and outcomes, and vice versa. In the model of Bala & Goyal this structure leads to their most important finding of payoff equality between neighbors. The result is driven by idea that all agents making a decision need to receive at least the same payoff as their predecessor did, as they can always copy the choice of their predecessor. In the long run this means that all agents will obtain the same payoffs and, if different choices have different payoffs, all agents will in the long run converge to making the same choice. This payoff equality result has also garnered support in the studies of Gale & Kariv (2003) and Acemoglu et al. (2011) in the Bayesian rational modeling framework.

One of the key attributes separating studies on social learning in networks is their findings on the so called "royal family" effect first introduced by Bala & Goayal (1998). In many of the models utilizing a boundedly rational learning mechanism, e.g. Bala & Goyal and Golub & Jackson (2007) to name a few, the existence of a prominent influential group of agents observed by all tends to prevent sufficient information aggregation to facilitate asymptotic learning. In contrast to this, the results of the Bayesian model of Acemoglu et al. (2011) are that, in the Bayesian framework agents constantly update their beliefs and also form beliefs on the unobserved agents in the population. This means that the presence of a prominent group does not prevent asymptotic learning as long as the said group does not form the entire sample of observations for all individuals. The effects of prominent groups can be very important for studying the effects of word of mouth communication. It is more than plausible to assume at least some sort of a prominent group of individuals consisting of so the called experts having a greater influence on opinion formation than other network members due to the better quality of their advice or their greater visibility.

One of the big challenges in modeling social learning in social networks is related to the stochastic nature of social networks, as they hardly ever remain stable through time. The formation of social networks is in part driven by individuals themselves, and it is highly plausible that an individual receiving a bad payoff from acting on the information received from his/her current social network takes action to alter his/her network in order to receive more useful signals in the future. This of course produces a challenge to all static models of social networks. That being said, modeling always requires some sort of simplification when compared to real life. Perhaps future models utilizing ever more the possibilities opened by computer aided network mappings and computational power can give us more insight on the matter.

2.4 Efficiency and welfare

After reviewing some of the key features and determinants of the theoretical models on social learning, it is time to take a look at their welfare implications in a broader context. The modeling properties discussed above are important, but they are not the sole determinants of social learning model outcomes and behavior. It is therefore now time to take a look at some of the most important properties of social learning model equilibriums in terms of efficiency- and welfare implications. As the main focus of social learning models has tended to be on identifying the possible drivers for asymptotic learning and herd behavior, I start by briefly discussing some of the results that reached on the convergence and stability properties of asymptotic learning, after which the focus is turned towards examining the welfare implications of social learning models and asymptotic learning equilibriums.

2.4.1 Efficiency of asymptotic learning

The efficiency of social learning outcomes is of course a key interest when it comes to social learning models. Because such a wide variety of different social learning models each with their assumptions and characteristics have been suggested, as of yet, there are not that many universal conclusions and implications to be drawn as no single model can be seen to have risen above others in terms of its explanatory power and behavior describing abilities. There are however some general key points across the different models to be discussed. As with social learning models in general, these findings are usually very model specific and parameter sensitive and great care needs to be taken when making broader conclusions based on them.

The convergence of social learning models to an asymptotic learning outcome is a natural point of interest in social learning models. Generally speaking, most of the social learning models brought forward in literature do converge to an asymptotic equilibrium for at least some set and values of parameters. The main points of discussion in the literature have mainly been more on the specific model characteristics that produce such an asymptotic learning outcome. Considering that the initiating driver for the whole field of social learning literature in economics, and indeed the focus of most of the models that belong this body of literature, has been on explaining the observable real world outcome of herd behavior in its various forms, it is not surprising that researchers have presented models that do in fact result in herding. In terms of learning outcomes one can even go as far as generalizing that most models discussed in social learning literature converge to a stable equilibrium over time. Only in some special cases, such as of a sequential decision models similar to Banerjee (1992) is there a credible possibility for multiple opposite informational cascades to arise in such a way that the system there is not sufficient convergence in the long term, but the was left indeterminably hovering between steady states.

Additional points of consideration, related to the speed of information aggregation and model convergence arising from the literature are the notions of results of Vives (1997) and Acemoglu et al. (2011). Vives argues in his paper that the speed of information aggregation in a market based social learning model of the type referred to above does not differ from the speed of socially optimal convergence. This in turn implicates that any welfare loss resulting from herd behavior should be attributed rather to the equilibrium end result of the process than to the speed that the convergence it is attained with. Acemoglu et al. on the other hand raise the point of the existence of a prominent group of individuals that is observed by others to hinder the convergence speed of social learning. In their network model the presence of such a group they describe as informational leaders will according to their results seriously slow down the speed of convergence of asymptotic learning, but will not be able to prevent it in the long run.

In terms of sampling, proportional sampling and larger sample sizes generally speaking tend to be more beneficial for the existence of a social learning equilibrium. As implied by Banerjee & Fudenberg (2004), the likelihood for equilibrium to arise is larger with large proportional samples than biased sample sizes or the extreme sample size of a single observation. This is based on the idea that larger sample sizes and proportional sample sizes allow for more information to be transmitted, making it easier for agents to aggregate information between them, allowing all agents to discover choice that is most efficient or has the highest payoff. Combined with the results of Vives (1997) this leads to the conclusion that larger and more proportional sample sizes are, ceteris paribus, better on a societal level from a welfare point of view as well.

Nevertheless, the above result drawn from the Bayesian world of social learning needs to be taken with a grain of salt when extended into the world of myopic learning mechanisms. In contrast to the above results Ellison & Fudenberg (1995) share the results of many Bayesian models in terms that smaller sample sizes contribute positively to the likelihood of herding. That being said, they also continue to argue further that, at least in their myopic model of word of mouth facilitated learning, the learning outcomes are most efficient when there is a limited amount of communication and information aggregation between the agents. This is a good example of variance in the implications of social learning models and their sensitivity to the discretion and assumptions of the modelers.

One more important point regarding the efficiency of social learning discussed Ellison & Fudenberg (1993), but not widely observed in literature, is the assumption of a heterogeneous population in a technology adoption framework. In the second model introduced in Ellison & Fudenberg it is presumed that there are two competing technologies, each optimal to a different group of agents, and the question of interest is whether the right technology is adopted by the right agents. The important results are that, no amount of popularity weighing will lead to an efficient outcome, which is in contrast to the homogenous population models, where popularity weighing tends to slow down the rate of convergence, but does not prevent it completely.

The issue of heterogeneous agents is very important, as in reality gents do differ in many aspects e.g. in their preferences. Studying a learning environment with truly heterogeneous agents would be very insightful for drawing real world implications for social learning models, but in turn also be more difficult to model than the usual homogenous agent environments. Hopefully we will see this line of research explored more in depth as computer assisted models of explicit social network modeling develop further in the future.

2.4.2 Welfare implications

Considering that the welfare effects of substantial herd behavior can be considerable, the welfare implications of social learning models still need more research for any definitive conclusions to be drawn. Many researchers do touch upon the welfare implications implied by their model, but the conclusions to be drawn are usually limited to the acknowledgement of the possibility of either a positive or a negative equilibrium, e.g. in the form of adopting an inefficient technology, to emerge as a result of information aggregation process. An important caveat is the fact that, while herd behavior is usually discussed in the context of having a negative effect on welfare, herding and asymptotic convergence as the result of social learning in itself says nothing about the efficiency and welfare properties of the solution. At least in theory, it is just as possible for herd behavior to produce positive welfare effects on individual- and social levels as it can produce negative effects.

Perhaps the most thorough discussion on the welfare implications of social learning models is offered by Vives (1997), who compares the welfare effects produced by market based learning to optimal learning in smooth and noisy versions of a statistical prediction model, where the smoothness of the model refers to the continuity of payoffs and action spaces. To interpret the welfare effects of social learning Vives defines herd behavior as agents deviating from socially optimal weighing of private information and public information. The socially optimal outcome is then defined as the solution to a team optimization problem, where the decision rules of agents minimize the discounted sum of their prediction errors.

The results of Vives' (1997) welfare analysis indicate that the socially optimal team solution will succeed in accumulating more information and lead to a socially more desirable outcome than the market based solution representing a standard Bayesian social learning model for information aggregation. Furthermore, Vives argues that, though there will be convergence for both the market based- and socially optimal solution, the convergence to an equilibrium will be slow and that the welfare loss arising from the under aggregation of information present in the market based solution can be quite substantial.

The main contributor to the inefficiency of the market based solution when compared to the socially optimal solution according to Vives (1997) lies in the informational externality produced by agents and their behavior in the market based model. When agents respond to their private information they reveal it at least partly the other agents assisting the aggregation of information in the economy. Agents do not however take this positive information externality that increases the value of public information into account when making their individual decisions. According to Vives this externality is internalized in the socially optimal model, which increases the value of publicly available information and decreases the probability of agents refraining to use only their own private information.

As highlighted in Banerjee & Fudenberg (2004), the appearance of a sequence of agents who find the publicly available information to be of such low information value that they are forced to act on their private information can disrupt an ongoing herding process and ultimately prevent the asymptotic learning equilibrium that the system would have otherwise converged to. This aspect is also closely related to the "royal family" effects discussed above. In a situation where the existence of a prominent group of agents reduces the value of the information acquired by agents via sampling by so much that it is of little value to them, they might be forced to rely solely on the information provided to them by their own private signals. In turn, under different assumptions, it is also possible for the prominent group to dominate agents' samples in such a way that the system will converge to the behavior of these so called opinion leaders.

The area of interest where the welfare implications of social learning are naturally at their largest is the adoption of new customs innovations and technologies. As has been argued in social learning literature and discussed above, there are many factors that influence the diffusion of new ideas. The main conclusion to be made when considering the welfare effects of social learning and herd behavior is the need to identify the different factors that affect learning process and outcomes. The asymptotic learning outcomes, i.e. herding behavior, produced in the models reviewed above are quite robust, at least in a technical sense. They do not however say much about the determinants the actual targets of herding. In many herding models the direction of herding can even be completely arbitrarily determined, and as a result the likeliness of an efficient technology being adopted over an inefficient one can come down to mere chance.

An important caveat to keep in mind when generalizing the results of social learning models to broader implications is that the models are best used to describe the conceivable convergence properties of the system and to describe the learning process resulting in herd behavior. These need nonetheless combined with careful analysis of the actual origins of the herd behavior, such as the actions of the first individual in a sequential decision model producing informational cascades to determine the direction of herding and the broader implications of such behavior. To further sharpen the picture on the effects of social learning, some of the welfare effects of social learning observed in real life situations are discussed further in the next chapter, which focuses on presenting and reviewing the empirical and experimental evidence on herd behavior and social learning.

3. Evidence for social learning

The empirical feasibility of social learning models has been extensively studied in recent years, both experimentally and by utilizing the tools of modern econometric analysis. A large variety of experimental settings have been used to study social learning in a laboratory, particularly to identify the conditions that are favorable for informational cascades to develop. In this chapter, I first take a look at some of the important studies on informational cascades utilizing an experimental laboratory setting, after which I briefly discuss the substantial body of literature testing the social learning herd behavior hypotheses in the financial markets. The focus is then turned towards the role of advice in social learning and how it can affect social learning outcomes in the laboratory. Finally, this chapter finishes with the examination the issue of how do word of mouth and advice contribute to social learning and what implications does this have on behavior in the product markets.

3.1 Herd behavior and rationality

The seminal contribution to the empirical testing of social learning hypotheses on informational cascades and herd behavior is given by Anderson & Holt (1997), who in their laboratory experiments and subsequent econometric analysis test empirically the validity of the early theoretical contribution of Bikhchandani et al. (1992). Anderson & Holt's experiments were one of the first attempts to test the feasibility of social learning theories of informational cascades in a controlled laboratory environment, and much of the subsequent empirical analysis of social learning has concentrated on replicating their original work and its different variations.

The basic setting in Anderson & Holt involves asking a random sequence of individuals to correctly predict, based on their own private signal and the observation of the choice of the previous subject, which of the two urns containing different sets of colored balls was a ball drawn. The experimental setting intends to mimic the situation where subjects must predict which of two equally likely events has taken place after receiving a private signal implicating the relative probabilities and observing prediction of the previous subject. As in Bikchandani et al. (1992), the subject who is first to act must act based solely on his private signal and thus reveal it to the other subjects. As one can see, the experimental setting matches closely the theoretical model of Bikhchandani et al. and allows the researchers to observe the formation of cascades, if they are subject to rise as the theory suggests.

The findings of Anderson & Holt (1997) are very much in line with what the theory of Bikchandani et al (1992) predicts. In the experiments of Anderson & Holt, informational cascades were found to develop more than two thirds of the occasions when possible and in about one half of those cases the cascades led the subjects to herd towards the incorrect choice, acting on a misrepresentative private signal of the subject in the sequence. This result speaks strongly in favor of informational cascades being a considerable driver behind herd behavior, as in a laboratory environment other influential factors such as the subject's inherent need for conformity among the group is minimized. In addition, in a study closely related to Anderson & Holt (1997), Çelen & Kariv (2004b) found that, while cascades tended to develop about one third out of times when possible, almost every time when they appeared they resulted in agents herding towards the correct payoff maximizing choice. Additional support for the findings of Anderson & Holt is also offered by Hung & Plott (2001), who vary the institutional setup to cover majority rule and conformity rewarding institutions.

In the special setting of the financial markets, herding in the forms of booms and busts in the asset markets has been frequently observed throughout time. This could lead one to presume that for example bubbles and market crashes were caused by the same mechanisms of informational cascades and herd behavior. Although some studies have voiced concerns for the possibility of informational cascades to arise in a market setting, in the light of the resent findings the contemporary general consensus seems to be that, as argued by e.g. Drehmann, Oechssler & Roider (2005), the presence of a flexible market price is usually strong enough force to prevent herding. This view is also supported by the experimental tests of Cipriani & Guarino (2005) among others.

Nevertheless there has been a tremendous amount of studies on herd behavior in the financial markets. The reason for this could be the huge financial effects that herding might produce or the fact that the findings on whether there are true herd behavior and informational cascades in the markets is still somewhat mixed. It is also important to remember that even in the case of no asset price speculation or herding, changes in asset price fundamentals can cause enormous movements in their pricess in a short amount of time. The separation of speculative or herd behavior driven price changes from those caused by fundamentals continues to be been a key challenge in this field.

The empirical evidence from laboratory experiments seems to show strong support for herd behavior and the informational cascades theories. The evidence is much more inconclusive on the nature of the learning mechanism that real life agents utilize. The findings of Anderson & Holt (1997), and especially of Hung & Plott (2001), seem to validate the assumption of Bayesian rationality of subjects at least to an extent. Yet in their econometric analysis Anderson & Holt find indications that about a third of the subjects deviated from Bayes' rule and used simple calculations to make inferences. This in turn made these subjects to rely more on their private signals compared to the public information offered by the predictions by other subjects.

If Anderson & Holt found that a third of the subjects tended to deviate from Bayesian rational actions, Çelen & Kariv (2004b) go one step further and attribute the boundedly rational actions of the early movers in the behavioral sequence as the main reason for the development informational cascades in a laboratory environment. According to Çelen & Kariv the agents moving early in the sequence tended to overvalue their private information, which lead to it being more representative in the public information samples of subjects acting after them. Çelen & Kariv also found that subjects tended to put excessive weight on their private signals and private information compared to the publicly available set of information, but that in the long run the subjects' actions tended to converge more towards what is predicted by assuming Bayesian inferences.

Further support for overvaluing of private information and hence non-Bayesian decision rules is given by Weizsäcker (2010), who performed meta-analysis of 13 different social learning experiments to concur that in his sample the subjects tended to deviate from the actions suggested by their private information only, if the public information sample suggested a very high probability for their private sample to be misleading. In terms of agents making Bayesian inferences, theory would predict that such deviations would tend to occur already when the probability the private information being misleading reaches the level where the sample is right or wrong with equal probabilities. These findings of course raise doubts over the Bayesian rationality of the participating agents. As an explanation for these findings, Kübler & Weizsäcker (2004) offer egocentric bias, as in their experiments subjects tended to be overconfident in the sense that they attributed a higher error rate to the predictions of others then to their own predictions.

When testing for the bounds of agent's rationality, researchers have found some interesting details. In their study Kübler & Weizsäcker (2004) modified the model of Anderson & Holt by introducing a small cost for agents to receive signals. The results were that, compared to the perfect Bayesian equilibrium of the first agent purchasing the signal and every subsequent agent acting based on it, too many agents positioned at later stages of the sequence decided purchase a signal and violate Bayesian rationality. Based on this Kübler & Weizsäcker argue that the equilibrium of a game with costly signals is bound to depart from its Bayesian equilibrium more than a game where signals are costless. This implies that in the real world agents tend to be less than perfectly rational in many cases, as there are a number of applications where acquiring signals involves some kind of a cost.

Exact conclusions on whether agents tend to implement a Bayesian rational or a boundedly rational decision rule are hard to make with the light of the current experimental evidence. The evidence does seem to imply that there are at least some bounds to the rationality of agents at least as a whole. An interesting and seemingly very robust finding has been that allowing subjects to offer verbal information in the form of advice increases rationality of agents. This factor will be explored more in the next section where learning from advice and word of mouth is discussed.

3.2 Learning from advice

As was briefly mentioned in chapter 2, word of mouth communication and advice can have significant effects on social learning outcomes. As previously discussed, word of mouth has been used in theoretical models to successfully facilitate information aggregation between agents to a great extent. In the above discussions on asymptotic learning it was assumed that the informational value of observing the actions of previous agents to be roughly equal to the information value of receiving verbal advice from those agents and that social learning can be facilitated with both methods of information transmission. This has also been the common approach in the theoretical social learning literature. Keeping this in mind, some of the recent laboratory experiments on social learning have attained some interesting results on how agents actually seem to value the information they receive through word of mouth advice compared to the information they obtain through observing the choices of their peers.

3.2.1 Power of advice

From a theoretical standpoint the value of advice of course depends on its information content and the value of that information. The receivers of advice are usually unable to determine the value of the information they receive before putting the advice into action and acting on it. Therefore the ex-ante value that agents assign to advice they receive is in reality affected by the credibility of the advice giver. In real life situations people tend to, at least in theory, put a higher value to the advice given by experts who are known to be knowledgeable on the subject at hand. This could plausibly lead to the advice of these experts carrying more weight compared to the advice given by someone who is considered to possess an ordinary level of knowledge, and thus also to carry more weight than observing the actions of these other naïve agents. The theory however also predicts that after controlling for the information value of observing actions and observing advice the two should carry equal weight in the decision making process' of agents.

Yet, some recent studies on social learning in a laboratory environment have found that agents actually tend to ceteris paribus weigh more heavily the advice they receive from their fellow naïve participants than the information they receive from observing their actions. This is a rather interesting result considering that the informational value of the two should be equal by experiment design. In their recent study on the effects of advice on social learning in the laboratory Çelen et al. (2010) construct three different types of simple social learning experiments with sequential actions to test for the subjects' willingness to follow the actions and/or advice of their predecessors and the possible payoff increasing or decreasing effects of receiving advice.

The main findings of Çelen, et al. (2010) are that subjects appear to be more willing to follow the advice of their predecessor than to copy their action. Moreover, the presence of advice seems to have significant positive effect on the subjects' payoffs and welfare. What makes this result somewhat surprising is that the experimental setting of Çelen et al. was such that both the advice given by predecessors and their actions observed by the successors were both in binary form, resulting in making the informational value of observing advice and observing actions to in all relevant aspects identical. All participating subjects were also considered not to be experts on the subject and thud to give naïve advice that should ex-ante be no more informative than the advice given by any other agent, including advice that would have been given by the agent receiving the advice in question.

Çelen et al. (2010) also found some noteworthy effects in terms of welfare implications as well. The presence of advice was found to considerably improve the accuracy of decision makers and to increase their payoffs compared to situations where they didn't receive any advice. The reason for this as argued by Çelen et al. was that the subjects tended to disagree less with their predecessors advice than their actions, increasing the number of subjects following the advice of their predecessors in situations where the signal received from the predecessor conflicted with the subjects own private signal. According to Çelen et al. the implication of this was that in their sequential model over time the individuals would start to rely more on the advice they received then to their own private signals.

Considering the impact of advice giving on the rationality of the subjects' behavior, the presence of advice has been found to induce the participants to act more according to economic theory than those participants who didn't receive advice. This result has been argued at least by Schotter (2003), who reviews a set economic games played by the subjects across a set of experiments. The reason for this improved rationality was argued by Schotter to be the fact that the process of giving or receiving advice forces the subjects to assess their positions differently. The subjects seemed to learn better, i.e. make better predictions, when they were forced to either advice someone or to receive advice before making a decision. The exact cognitive process of how including the advice factor into the decision problem enhanced the learning of agents is not clear, but one plausible explanation is that it made the agent's re-asses their first hand opinions and to therefore spend more cognitive effort on making the decision.

Most of Celen et al.'s (2010) results receive strong support from the earlier paper by Schotter (2003), who surveyed previous laboratory experiments on social learning and arrived to similar conclusions in terms that the subjects did tend to be more willing follow the advice of naïve advisors than to copy their actions. Schotter even goes as far as arguing that, based on the experiments he analyzes, agents generally prefer to rather receive advice from the individual preceding them than to receive the same set of information that the preceding individual had when he or she had made the recommendation. This result is intriguing in many respects. Individuals are usually thought to be more prone to being egocentric and overconfident on their own abilities than the exact opposite suggested by this finding. At the very least, a rational agent would in all likeliness put at least as much faith in his/her own conclusions on the agent's optimal as to those made by another agent.

Why then would an individual think that another ex-ante similar individual would be able to make a better decision than he/she himself/herself could make? The exact answer for this is not yet completely clear. Çelen et al. (2010) suggest that this might be because the receiver of the advice assumes that the advisor has put more effort on formulating the advice, e.g. for reputational reasons than an individual contemplating his/her own action would. Other reasons might lie in the unknown parts of our cognitive structures that make us weigh verbal information in the form of advice more heavily in comparison obtaining information through observing actions. Future research can hopefully share more light into exactly are the driving forces behind the power of advice.

3.2.2 Social ties

An important consideration regarding Schotter's (2003) results on the enhanced rationality of behavior when the subjects must give or receive advice is that, in the particular experimental setting forming the basis of the result, the both advisor and receiver of advice were both simultaneously present. This could have had an effect on the results as both sides might then consider reputational issues, which could have affected their decisions and increased the effort level exerted by the subjects. The quality of learning by the advisor, and hence the quality of his/her advice, might have been worse if there were no personal contact or asocial tie between the advisor and receiver of the advice. In the modern age of internet where word of mouth advice from sources that we have nonexistent social ties with is more and more abundant as well as important, this is one of the aspects of word of mouth facilitated social learning that is increasing in its significance, but still remains largely undefined.

The generally held belief on the importance of social ties to word of mouth is that closer the social tie, more influence will the advice or word of mouth between the agents have. Historically the influence of word of mouth that we produce has been mostly restricted to our family, friends and acquaintances that we deal with in our daily interactions. With the rise of the internet, this has however changed completely. Through the internet and the through the various channels and platforms it facilitates people are now able to reach and share messages with literally millions of people. The implications of this to e.g. the marketing- and public relations practices of companies are tremendous, and, due to the rapid rise of these new technological possibilities firms, consumers and academics alike are only beginning to realize the full potential of these new mediums.

As of yet, there has not been much research done on comparing the effects of technology facilitated advice, such as advice given via the internet ,and more traditional forms of word of mouth advice. One of the few studies that have taken a look at the matter is Steffes & Burgee (2009), who compared the effects of receiving advice on which professor to choose for a collage course from different sources. Their somewhat surprising results were that, when choosing a professor, students tended to weigh most, not only their personal experiences, but also advice they received from an anonymous source on a website. These two sources were found to be more important than receiving traditional forms of word of mouth. What makes the results so interesting is the fact that students put more weight on the anonymous word of mouth on the internet than to traditional word of mouth from, not only people they have weak social ties to, such as their academic advisor, but also from people they have strong social ties with, such as their friends.

The results of Steffes & Burgee (2009) are interesting because the traditional theoretical view on the effects of social ties on the rating of advice, as expressed in Brown & Reingen (1987) among others, is that a stronger social tie between individuals results in a bigger ability for one to influence the other. This might very well be true in the setting of traditional face to face interaction, and Steffes & Burgee did find that traditional word of mouth from individuals that the subjects had close social ties carried more weight than word of mouth from individuals with weak ties. The interesting factor is the high relative appreciation for internet word of mouth. In a related study on the effects of online reviews on book sales, Chevalier & Mayzlin (2006) found that written user reviews had more power in determining sales than simple average rankings. This would imply that, while the weight carried by online word of mouth can be considerable, the impact of the message is correlated with its resemblance to the traditional forms of word of mouth.

3.3 Product markets

The importance of word of mouth for promotion and product sales has been recognized in marketing for decades now. Adding the recent findings on the power of naïve advice discussed above, it is clear that word of mouth in its various forms has great implications for companies and their marketing strategies. Personal advice has always been an important factor in human decision making, whether in determining which restaurant to go to, which school to enroll your children in or which neighborhood to move into. In recent years the importance of naïve advice has also grown immensely through the development of the internet and range of possibilities it offers.

Internet has not only opened new possibilities for marketers and product manufacturers to spread information about their products, but it has also made possible for people to acquire relevant information on different products and to share their views to the masses e.g. through different kinds of online product reviews and recommendations. The usefulness of word of mouth to the diffusion of new technologies and new products has also been recognized in marketing literature, and word of mouth effects are for instance an important part of most widely accepted product diffusion models, such as the Bass' (1969) infamous s-shaped diffusion model. That being said, even if the existence of word of mouth influence has been recognized, it has proven to be very hard to quantify these effects and their exact influence on sales.

3.3.1 Word of mouth and sales

The exact relationship between word of mouth and product sales has proven to be hard to quantify. One of the main reasons for this lies in the inherent endogeneity between the two. Just as it is plausible to assume that positive word of mouth has a positive effect on the demand for a product, it is also plausible that strong sales figures themselves can help to create positive word of mouth, hence creating a cycle that blurs observations on the true direction of the causality. Further challenges to the matter brings the difficulties in measuring word of mouth delivered in traditional conversations and determining the aspect, in terms of e.g. quality or quantity, of communication to measure.

Presumably due to the difficulties in measuring and quantifying word of mouth and linking it into product sales outcomes, there has not been much research produced on it, at least in the field of economics. Thankfully with the surface of the internet and its many channels for word of mouth communication it has become easier to measure at least the kind of word of mouth it facilitates. The empirical findings on the effects of online word of mouth on product sales have been mixed. Some studies on the subject have found that positive word of mouth does in fact have a positive effect on product sales, while others have found no considerable effect to be present. The determination of such causalities can naturally be complicated as there are numerous factors besides word of mouth that can affect product sales.

One of the more convincing cases for word of mouth influencing product sales is made by Chevalier & Mayzlin (2006), who examined the effects of online book reviews on their sales figures in online bookstores. Their findings were that, as one would suspect, having a higher rating had a positive effect on the relative sales of the book and that negative ratings carried more weight in relation to positive ones. As mentioned before, Chevalier & Mayzlin also found written reviews by naïve advisors and other customers to be more influential than simple rankings based on user ratings. One possible explanation for this could be that the written reviews either carried more information or allowed the readers to assimilate themselves with the advice givers, increasing the influence of the advice and making the online word of mouth more similar to traditional word of mouth.

Some reservations on the results of Chevalier & Mayzlin (2006) is required by the fact that in the underlying data the customers would receive the word of mouth information only after going to the bookstore website and hence were probably already considering purchasing at least some of the products for sale. This means that it remains unclear whether word of mouth was indeed the driving force behind the customer's decision to purchase a book in the first place or just a parameter of their decision rule on determining which kind of book they were going to buy, i.e. the brand of the product they were going to purchase. It is a topic in need of some further research and further evidence on how much of an influence can word of mouth have on sparking the initial purchase intention or should it be seen as more directing the demand already formed toward particular types and brands.

The results of Chevalier & Mayzlin (2006) are contrasted by the findings of Duan, et al. (2008), who found that, after accounting for possible endogeneity in word of mouth and demand, the ratings given in online movie reviews had no significant effect on the movie's box office demand. Duan et al. did however find that the box office revenues were positively affected by the sheer volume of the online word of mouth. This sort of awareness effect is an important finding, not only because it verifies the common marketing knowledge of boosting demand by creating knowledge and "buzz" on the product, but also because it is a step in the right direction in determining the most important demand affecting components of word of mouth.

Although the conventional wisdom, and a view supported by Chevalier & Mayzlin (2006), is that negative word of mouth is weighed more heavily than positive, in their study of the effects of negative- and positive word of mouth on reported brand purchase probabilities East, Hammond & Lomax (2008) found positive word of mouth to carry relatively more weight. They also found that the effects of word of mouth on the probability of purchase were greatly dependent on the underlying probabilities. The result of positive word of mouth being more influential than negative East et al. attribute to the familiarity of the brand in question and argued that in relation positive word of mouth offers more diagnostic value than negative word of mouth, because the familiarity of the product attenuates the diagnostic value of negative information.

The results achieved by East, Hammond & Lomax (2008) on the relative impacts of positive and negative word of mouth are certainly debatable. There is however also another key concept in their work that deserves attention. In the work of East et al. there were significant signs of anchoring effects in the terms of resisting word of mouth that countered the receiver's current opinions. This implicates that all word of mouth is not created equal. Not only in its objective quality, but also in terms of how well it fits the opinions of the receiver. Furthermore, it implies that word of mouth could be better used in enhancing current customer loyalty rather than to successfully acquire new customers with it.

3.3.2 Firm performance

In terms of company performance, there has not been much successful work in quantifying the costs and benefits of word of mouth. When comparing the financial effects of different customer loyalty metrics, Morgan & Rego (2006) find that the quantity of word of mouth in terms of the average number of recommendations does have a positive impact on the future market share of the company, but that it also has a negative impact on the company's future gross margins. As a reason for this Morgan & Rego suggest the large financial costs related to generating positive word of mouth. More and more companies do however engage proactively in word of mouth generation nowadays and try to seed the markets with positive information. This can of course be a successful strategy for some firms, but the industry wide effects and financial returns on word of mouth are yet to be clearly determined.

As the company wide and industry wide performance effects of word of mouth remain somewhat ambiguous, establishing a clear link between word of mouth and company financial performance remains as an intriguing avenue for further research, one which is unfortunately outside the scope of this thesis. One of the most appealing characteristics of word of mouth generated consideration and awareness is its low cost structure in terms of possibly attracting large amounts of consumer attention for relatively small amounts monetary investments. The topic of how to seed a market with word of mouth communication to create positive product hype is an active topic among marketing researchers and has also been incorporated in some product- and technology diffusion models. Even if the benefits of positive word of mouth seem to be clear for everyone, the hard part remains to be linking the small micro level effects of word of mouth distribution to company performance and even further to market level industry performance.

In conclusion, there is still a lot of work to be done in identifying the effects of word of mouth on demand behavior and product sales. Some reasonable attempts have been made, but the exact causalities between word of mouth, product sales, and firm performance still continues to elude us. Due to the numerous difficulties that are involved in isolating these effects, there are still challenges to be met in the future. Hopefully the ever growing amount of data coming from online platforms allows us to gain further insight on the matter. Still, further attempts should be made to gain more insight on the effects of word of mouth advice on social learning, and the implications that it might have in real world markets. The rest of this thesis aims to contribute to our understanding of word of mouth by examining the effects of word of mouth learning in the context of the smartphone market via examining the determinants of the effectiveness of word of mouth at the individual consumer level and investigating the market share effects of positive word of mouth at the market level.

4. Smartphone market data

This chapter lays out the foundations for the empirical analysis of this thesis by pointing out some of the key characterizing determinants of the smartphone industry and by describing the smartphone market data used in the empirical analyses of this thesis. The empirical section of the thesis is based on the analysis of data from two different datasets, one consisting mainly from recent sales figures of smartphone devices and the other consisting of survey data gathered via a consumer survey on related brand awareness. This chapter proceeds with a description of the smartphone industry and its key characteristics. As the industry is in many ways different from other more traditional industries, a brief description of some of its key demand determining characteristics will help to give context to the subsequent analysis. After the above mentioned industry description, the datasets used in the empirical analysis of the thesis are described in more detail.

4.1. Market description

The rise of the mobile industry has been one of the most prominent economic developments in the world during the last two decades. The advances in mobile technology are widely seen as one of the main drivers and facilitators of economic growth in the near future. The roots of the contemporary smartphone industry are rooted in the mobile phone industry that preceded it and still somewhat continues to coexist with it today. The shift from the earlier mobile phones, aptly described as “feature phones” due them mostly being a sum of their features, to devices labeled as “smartphones” with more cross functional usability has been gradual and relatively seamless. The feature phone market is still vibrant in many countries, and many of the same companies producing smartphones also produce feature phones as well. As new technological advances are discovered the center of gravity of the industry will however keep shifting more and more towards these the so called smartphones and other smart devices.

The significance of smartphones and their economic impact is not limited to the industry itself as they can well be seen as contributing to economic growth in additional ways via facilitating communication and information sharing. As modern economies move more and more heavily into information based economies with ever bigger share of the economy consisting of the services sector, smartphones can be seen as facilitating the development of new business models that utilizing them and acting as production inputs for many other firms and industries. As of today, smartphones can credibly be seen as the focal point of the technology industry and their significance is generally believed to only grow in the future as the devices gain more and more users.

The boundaries of the smartphone industry are extremely hard to classify for a number of reasons, not least for there being a huge amount of overlap with many other categories of products with similar uses and functionality. On one hand, the smartphone market overlaps with the mobile phone market as both types of devices can be used to make phone calls and also share many of the same basic functionalities. On the other hand, the smartphone market also heavily overlaps with the tablet market, and to a degree even with the laptop market, as well, as both types of devices have relatively similar in functionality, with the main differences coming from device size, screen size and performance related issues.

Furthermore, as mentioned, it is highly plausible that in some markets smartphones compete of the same customers as notebooks and other types of personal computers. This can be seen to be a plausible scenario especially in the developing world, where due to income constraints households might consider buying only a single device to use for communication and connecting to the internet. In addition to significant amount of overlap among between these closely related markets there are other factors including constant market expansion and uncertainties about future developments that complicate the estimating the size of smartphone market. To give a rough idea of the market's size and significance, according to the technology research firm Gartner (2013), more than 200 million smartphones were sold worldwide in the last quarter of 2012 alone and globally the size of the market is expected to grow substantially, especially in the emerging markets.

As stated before, the smartphone market is very unique in many of its characteristics. Smartphones as products can be seen as simultaneously possessing the characteristics of multiple products sold in a bundle. The final user experience that the buyer purchases is distinctly determined jointly by the hardware, software and the ecosystem of the product in question. Some of the industry participants are vertically integrated and control all three of the main aspects of their product. Others have paired- or grouped up to jointly contribute to the production of the end product, and hence do not have complete control over their eventual user offering. Further complicating the situation is the fact that the quality of an operation system is undoubtedly tied to the quality of its ecosystem, which is in part determined by the amount and quality of applications produced by third party developers.

The degree of vertical integration and the degree of control over the end product is one of the important separating factors among smartphone producers. Some producers, such as Apple, are fully vertically integrated in terms of providing the hardware, the software and the ecosystem, while others like Nokia and Samsung mainly produce only the hardware component of the product and rely on software and ecosystems provided by Microsoft and Google respectively. While the data used in the econometric work of this thesis is based on sales figures of smartphone handsets, which can be seen as being a bundle of the handset, the operating system and the ecosystem it possesses, the handset sales data can be seen as accurately representing the sales numbers of smartphone operations systems as well, which, rather than the handset market shares, are the factor of interest in the subsequent empirical analysis. The relative weighing of the significance of each of the parts of the smartphone product bundle of course varies from one consumer to another, but the choice of the combination of operating system and the ecosystem that is tied to is without a doubt an important enough piece of the puzzle to merit further empirical analysis.

On the device manufacturer side of the market there are numerous players producing different handsets, but from an operating system standpoint practically the whole market can be seen as utilizing one of the four most popular operation systems that according to information technology research firm IDC (2013) are Google's Android, Apples iOS, Microsoft's Windows Phone (WP) or BlackBerry made by the similarly named company. Each of these operating systems comes with their own ecosystem in terms of connectivity and applications and represents a distinctly different product offering. Importantly, they are also limited to be sold with certain set of handsets depending on the operation system. When Google's android and Microsoft's Windows Phone operating systems can be purchased with a wide variety of handsets from different manufacturers, at the other end of the spectrum iOS and BlackBerry operating systems can only be purchased with a bundle with a device from a certain manufacturer, Apple and BlackBerry respectively.

Another major factor in the nature of competition in smartphone market and perhaps something that draws it apart from some of the more conventional industries is the hyper competitive nature of competition in which the winner captures a lion's share of the market. There are of course other markets that can be described as an oligopoly with few of the largest players holding most of the market, but what makes the smartphone market unique is the speed that changes in the market structure can and frequently do occur. The hypercompetitive nature of the market guarantees that the margin of errors for the producers of handsets and operation systems is extremely slim and the cutthroat competition ensures that no means of competition from aggressive pricing to patent lawsuits are being left unused.

The hypercompetitive nature of the smartphone markets implies a high sensitivity in market outcomes to small changes in outcome determinants such as word of mouth. Due to the hypercompetitive winner takes all nature of the market, even slight increases in positive word of mouth recommending a product could have large effects on market outcomes. The bundle-like nature of smartphones as a product and the significant network effects that are related to their ecosystems, both in terms of connecting users with other people as well as locking users into using products in a specific ecosystem, make word of mouth recommendations an important factor in determining shifts in product sales and the possible long term market equilibrium. Especially in new markets can word of mouth have a strong influence on product market shares via affecting the initial purchase decisions of consumers who are deciding on which of the available ecosystems to enter. After purchasing a device and entering an ecosystem there can be considerable costs involved with switching to using a device from another ecosystem.

The relative market shares of the four main smartphone operating systems vary from market to market and in most markets there are also some niche products being sold with other operation systems. There is also significant variance between the market shares of the operation systems between different price points. Some operation systems such as iOS are also simply not available in all markets and all price points. Nevertheless, the recent sales of other operation systems apart from the four most popular operations systems have been of negligible importance at the market level. According to the information technology research firm IDC (2013) the four main operation systems made up more than 96.9 percent of the smartphone operation systems sold in the last quarter of 2012, implying that the four biggest players in the operation system market hence make up the whole market to a degree that they can be seen as accurately representing the whole market in the context of this thesis. In more depth the smartphone market data used in the empirical analysis for this thesis and its sources are examined in the next section, which concentrates on data and its sources.

4.2 Data sources

The empirical part of this thesis is based on the analysis of data originally coming from two different sources, each dataset describing the different aspects of the smartphone market. The first set of data, market tracking dataset, consisted mainly of observations of smartphone sales figures in different markets. The second set of data, Brand Relationship Tracker survey data, on the other hand consisted of consumer survey data on brand awareness and brand loyalty. For the empirical analyses of the thesis the Brand Relationship Tracker data was first utilized separately for the analysis of the effectiveness of word of mouth advice, after which the Brand Relationship Tracker datasets was combined with operation system market share information calculated from the market tracking dataset to produce a panel dataset for the analysis of operation system market shares.

The final panel dataset used in analyzing changes in operation system market shares consisted of strongly balanced panel dataset with monthly observations from 14 different countries representing a well-balanced sample of different smartphone markets over the time period of 18 months spanning from July 2011 to December 2012. The same set of countries and same time period were also used for the advice rating analysis that utilized only the survey data from the Brand Relationship Tracker dataset. A list of countries and the number of observations from each country used in the advice rating and market share analysis are presented in appendix 1 in Table 6 and Table 7 respectively. Table 6 represents the final dataset for the advice rating analysis via the Heckman selection model constructed from the Brand relationship tracker data. Table 7 represents the final panel dataset for the market share analysis, which included data from both the Brand Relationship Tracker data and the market tracking data. The two sources of data from which the data for this thesis was gathered are described in more detail below.

4.2.1 Market tracking data

The dataset containing market tracking data of smartphone sales was collected by the GfK group, one of the largest market research groups in the world, and is based on tracking smartphone handset sales to consumers in retail outlets in the selected markets across the globe. In practice the final monthly sales numbers included in the data are calculated for each smartphone handset through extrapolation of observed sales figures. As tracking every single individual retail outlet is somewhat impossible due to practicality restrictions, GfK can only partially observe the actual sales numbers in each market. The retail sales coverage of observations also varies from one market to another. To reach the market representative final sales volumes for each handset in each market the observed sales volumes are extrapolated to market representing volumes through GfK’s estimated coverage of the retail network in the specific country.

In addition to the sales numbers of different models and defining model characteristics, the market tracking data also included the estimated nonsubsidized retail prices for the different models smartphones and the operation system of each device. For the purpose of being included in the operation system market share analysis the sales monthly device sales figures of handsets were aggregated to operation system and country level to calculate the market shares of the four operation systems used in the analysis in each of the 14 countries in each of the 18 periods from July 2011 to December 2012. As previously explained, to facilitate market share analysis, the aggregated sales data from the market tracking dataset was subsequently added to the Brand Relationship Tracker dataset.

4.2.2 Brand Relationship Tracker data

The Brand Relationship Tracker dataset consisted of data gathered via a consumer survey and was collected by the device manufacturer Nokia, one of the world’s biggest smartphone manufacturers and a key participant in the smartphone market. The Brand Relationship Tracker survey is collected monthly by Nokia and includes questions related to brand awareness and preferences for consumers who either own a smartphone or a more traditional mobile phone and or are determined to purchase one in the next 12 months. The dataset used for the analysis of the effectiveness of advice with the Heckman selection model only included observations of consumers who already owned a smartphone. The data on the Brand Relationship Tracker dataset was collected via a consumer survey that was carried out 25 different markets across the globe, 14 of which are featured in the analyses if this thesis. As previously mentioned, both the Brand Relationship Tracker dataset as well as the market tracking dataset included observations from the time period of 18 months between July 2011 and December 2012.

In markets where the proportion of the population with online access is representative of the general population in the country the survey interviews were conducted online. In countries where the online population is not representative of the general population the survey was conducted either via face to face interviews or a mixture of face to face interviews and an online survey. To facilitate the generalization of the survey results to apply to the general population of the country and the market, in the analyses of this thesis the survey results were also weighed with population representation weights to make the data reflect population characteristics and to make the results applicable to the whole markets and their participants as a whole.

The survey questions in the Brand Relationship Tracker survey that were used to measure consumer experiences on receiving advice on a smartphone purchase and form the basis of the data used for the Heckman selection model used to for the empirical analysis of advice ratings are presented in appendix 2. The characteristics of the data that are relevant in terms of the empirical analysis of this thesis will be discussed more in the next chapter, which will outline the empirical framework of the thesis.

5. Empirical framework

This chapter begins by defining the thesis' main empirical research questions derived from theory and by relating them to their broader context of social learning theories and empirical literature presented in the previous chapters. After defining the main research objectives for the empirical analysis the focus is turned towards econometric modeling and the main models for the empirical analysis part of this thesis are presented. The Heckman selection model that is used to study the effectiveness of advice is presented first, after which the panel data models used to study the effects of word of mouth on operation system market shares are presented. The purpose of this chapter is to define the appropriate empirical framework used to study the main research questions of the thesis and to construct a base on which the empirical results of the analyses can be discussed in the next chapter.

The paramount economic theories of social learning that were laid out in chapter 2 were mostly concerned with the features and characteristics facilitating herd behavior. As discussed, one of the possible means of facilitating information aggregation among agents is word of mouth advice. The different theoretical settings where word of mouth can facilitate asymptotic learning were already discussed in chapter 2, but the direct implication of the models is that under suitable circumstances word of mouth can cause significant herding among agents towards a common choice when faced with a similar decision problem. The conditions for herding, efficient or inefficient in terms of agent payoffs and welfare, to arise were discussed in depth in chapter two and, for the empirical analysis section of the thesis the interest is concentrated on the real world implications of the numerous theoretical models of social learning.

As pointed out by the previous empirical and experimental literature on social learning and the role of word of mouth advice in experimental and real world situations discussed in chapter 3, the herding outcomes of social learning theories also appear to be robust in real life as well. This being said, the body of empirical literature on word of mouth facilitated social learning applied in a real life market situation is still scarce. The objective of this thesis is hence to contribute to this body of research by studing the effects of word of mouth advice and word of mouth facilitated learning in the real world setting of smartphone markets. The smartphone markets can be considered very suitable for such an analysis due to their high level of competition, rapid market dynamics and the large quantities of word of mouth advice between consumers for facilitating information aggregation.

5.1. Research questions

The main empirical research questions of the thesis are two folded. First, this thesis seeks to clarify the factors that determine the influence of word of mouth advice and answer the question of where does the most valuable word of mouth advice come from. Second, in addition to examining these micro level effects of word of mouth, this thesis also seeks to determine some of the broader market level implications of word of mouth and answer the question of how does word of mouth advice affect the relative market shares of smartphone operation systems. The empirical analysis part of the thesis will begin by examining the individual consumer aspect of word of mouth and in regard to the effectiveness of word of mouth advice, from where the natural progression will be to examine the aggregated effects of the individual level effects on the scale of the whole market.

The first of the main research questions is interested with the effects of word of mouth advice from the consumer's perspective. It intends to clarify the determinants of valuable advice in terms of the sources and channels from which word of mouth is received. In the sample of consumer survey data of smartphone owners used for empirical analysis almost two thirds of the respondents reported their choice of smartphone being influenced by word of mouth advice. In the empirical analysis sections of the thesis I seek to determine the effects of sources and channels from where word of mouth advice is received on the valuations given to that advice. I seek to find what type of word of mouth consumers perceive as most valuable in providing them useful information for their purchase decision, where did it come from, from whom did it come from, and whether or not the advice was actively sought.

The second the main research question deals with the aggregated market effects of word of mouth facilitated social learning in a real world market by seeking to understand the effects of word of mouth facilitated social learning on the market shares of smartphone operations systems. The intention is to shed light to any possible effects of word of mouth on the market equilibrium and dynamics. As previous empirical research on this topic is so limited the emphasis of this study is on defining and understanding the basic underlying principles of word of mouth advice's market wide effects. One of the key aspects in the second research question is whether or not word of mouth advice will facilitate the kind of herding towards a particular choice of brand that is implied by many social learning theories.

To seek answers to the research questions laid out above, the research questions are studied through the empirical analysis of relevant data related to the smartphone market. Other relevant factors of interest in the context of this thesis will be discussed in later chapters as the need arises. In the next sections devoted to methodology I will present the empirical models of the thesis and describe how they can be used to study the thesis' research questions and the other relevant factors of interest. The purpose of the methodology sections is to describe the empirical methods of the thesis in enough depth to facilitate the discussion of the results of the analysis in the next chapter.

As the research questions of the thesis are two folded in terms of examining the micro level and market wide effects of word of mouth, and the datasets utilized in their analysis differ, the methodology of the thesis involves two different classes of models. The determinants for the individual level effectiveness of word of mouth advice are examining through the well-known Heckman selection model, while the panel data analysis of the market share effects of word of mouth are is conducted utilizing an ordinary least squares specification, a fixed effects specification and a random effects specification. Starting with the Heckman selection model, these model specifications will be presented in the remaining sections of this chapter.

5.2. Heckman selection model

This section specifies the structure and methodology for the Heckman selection model of the thesis. It also describes the key variables of interest in the model and discusses the important factors related to observing the marginal effects of these variables as well as other relevant modeling related factors. The purpose of the Heckman selection model is to facilitate the analysis of how ratings given for word of mouth advice are affected by a group of independent variables related to the source of the advice and the setting of information aggregation.

An important attribute of the Brand Relationship Tracker dataset is that the data was generated through the means of a survey questionnaire, and not all participants had received advice in their purchase decision when purchasing their smartphone. Due to this there is a distinct possibility for a sample selection bias to severely hinder the estimation of the effectiveness of word of mouth advice in the whole population. Because only the ratings for advice of those subjects that actually did receive advice are observed, it is impossible to know the ratings that the part of the population that did not receive advice would have given if they in fact had received advice. This means that, if not corrected for this sample selection bias, the estimation results for are not generalizable to the whole population surveyed or to the general public.

This kind of problem with sample selection bias is typical in empirical work dealing with either non-representative or truncated samples. Thankfully there has been different ways for controlling this sample selection bias been presented by in econometrics literature. A widely accepted method for controlling for sample selection bias called the Heckman selection model, also known as the Heckman correction, was developed by the economics Nobel laureate James Heckman in the later part of the 1970's and has been applied extensively in empirical research ever since. The idea behind the Heckman selection model as laid out in e.g. Heckman (1979) is to deal with the sample selection bias as an omitted variable problem and to divide the analysis of behavioral relationships into two parts.

In the first part of the two stage estimator, a selection model for the probability of an observation being included in the sample is estimated, typically through a probit regression. In the second stage of the two stage estimator the selection bias is corrected via transforming the probabilities of being included in the sample into an additional explanatory variable trough the inverse Mills ratio, also known as Heckman's lambda. The inverse Mills ratio, defined as the ratio of the probability density function to the cumulative distribution function of a distribution is then used as an additional independent variable in the estimation of the outcome model to determine true behavioral relationships in the population. As is evident, the inverse Mills ratio included in the outcome equation of the estimator represents the effects of an omitted variable on the independent variable of interest, namely the effects probability of being included in the selected sample.

5.2.1 Model specification

In the context of this thesis the Heckman selection model is used to study how the subjects rated the advice they received when searching for their current smartphone device. Drawing from Heckman (1979), a formal description of the two part Heckman selection model to be estimated is given by following of set of equations:

$$y_1^* = \beta_1 x_1 + \varepsilon_1 \tag{5.1}$$

$$y_2^* = \beta_2 x_2 + \varepsilon_2, \tag{5.2}$$

where $x_1$and x_2 are vectors of independent variables and $\varepsilon_1$and ε_2 are jointly bivariate normally distributed error terms. Equation (5.1) and (5.2) above respectively represents the selection equation and the outcome equation of the Heckman selection model. In the selection equation (5.1) y_1^* indicates whether the person has received advice, in which case $y_1^* > 0$, or whether the person has not received any advice, in which case $y_1^* = 0$. The main variable of interest is y_2^*, the dependent variable in the outcome equation, which denotes the ratings given to the received advice. An important notion is that the main variable of interest y_2^* is only observable and included in the observation if the person has in fact actually received any advice and $y_1^* > 0$.

To estimate the model the selection equation outlined above is estimated using a probit regression of the form:

$$P(y_1 = 1|x_1) = \Phi(\beta_1 x_1), \tag{5.3}$$

where y_1 =1 indicates that the person has received advice, $x_1$is a vector of explanatory variables, β_1 is a vector of unknown parameters and Φ is the cumulative distribution function of the standard normal distribution After the estimation of the selection equation the inverse Mills ratio λ can then be calculated as the ratio of the probability density function to the cumulative distribution function:

$$\lambda = \frac{\phi(\beta_1 x_1)}{\Phi(\beta_1 x_1)}, \tag{5.4}$$

where ϕ and Φ denote the probability density function and cumulative distribute function respectively.

After the calculation of the inverse Mills ratio, λ is then added as an independent variable to the outcome equation (5.2). Finally, the outcome equation is then estimated trough a maximum likelihood ordered probit procedure. According to Verbeek (2008), in this setting the Maximization of the loglikelihood function for the outcome with respect to the unknown parameters will lead to consistent and asymptotically efficient estimators that have an asymptotic normal distribution. Verbeek also points out that not observing the actual correction term λ is not a threat to the validity model as the only unknown element in λ, β_1, can be estimated consistently by the probit maximum likelihood estimation of the selection equation.

The estimation for the outcome equation in the second step of the Heckman selection method describes the conditional expected value of y_2^* given the vector of independent variables x_2 and given that y_1^* =1, i.e. the expected rating for the advice received given that the person has actually received any advice. As the meaningful interpretation of the regression coefficients given by this ordered probit model can be rather complicated, to facilitate proper analysis the marginal effects for the probabilities for the lowest and for the highest rating outcomes are then calculated based on the estimates given by the ordered probit procedure for the outcome equation. The marginal effects for a discrete change in an independent variable are reached by partially differentiating the expected value of the dependent variable y_2^* with respect to the vector of dependent variables x_2 as given by:

$$\partial P(y_2^* = \alpha_j) \;/\; \partial x_2, \tag{5.5}$$

where $y_2^*$is the independent variable of the outcome equation, $x_2$is the vector of independent variables in the outcome equation and $\alpha_j = 1, 5$.

The complete Heckman selection model outlined above was estimated using the full set of Brand Relationship Tracker data described in Table 6 in appendix 1. The estimation results for the model are examined in the next chapter, which discusses the results of the empirical models of the thesis. However, for further insight on the model, the regression variables included in the estimations of the Heckman selection model are discussed in the next section.

5.2.2 Variables

The variables used in estimating the Heckman selection equation and their definitions are displayed in Table 8 in appendix 3 and some of the key summary statistics for these variables are displayed in Table 9 in appendix 3. To illustrate the use of estimation variables, the final empirical specification for the selection equation of the selection model of this study may be expressed as

$$advice = \alpha + \beta_1 * gender + \beta_2 * developedmarket + \beta_3 age25to39 + \beta_4 age40to54 + \beta_5 * age55plus + \beta_6 coreconsumer + \varepsilon. \quad (5.6)$$

The empirical specification for the outcome equation on the other hand may be expressed as:

$$rating = \alpha + \beta_1 * active + \beta_2 * children + \beta_3 * friends + \beta_4 * colleagues + \beta_5 * parents + \beta_6 * spouse + \beta_7 * siblingsandcousins + \beta_8 * gender + \beta_9 * developedmarket + \beta_{10} * age25to39 + \beta_{11} * age40to54 + \beta_{12} * age55plus + \beta_{13} * coreconsumer + \beta_{14} * via_{face2face} + \beta_{15} * via_{phone} + \beta_{16} * via_{some} + \beta_{17} * via_{email} + \beta_{18} * via_{instmes} + \beta_{19} * invmills + \varepsilon. \quad (5.7)$$

A notion meriting some discussion is the conventional assumption that, as noted by both Cameron & Trivedi (2005) as well as Verbeek (2008), the accurate identification of the selection equation of the Heckman model is greatly enhanced by adding at least one identification variable not included in the outcome equation as an independent variable in the selection equation, while otherwise including the same set of independent variables. In the setting of the Heckman selection model of this thesis the identification variable should be a variable that affects the probability of receiving advice, but does not affect any of the ratings given to the advice received.

Both Cameron & Trivedi (2005) as well as Verbeek (2008) also further argue that such exclusion restrictions as adding variables to the selection equation that are not included in the outcome equation that are plausibly defendable are hard to make. As no potential variables for justifiable theory based exclusion restrictions are present in the data, the full set of variables included in the selection equation is also included in the outcome equation. This means that the identification of the Heckman selection model is facilitated through the nonlinearity of the functional form and the nonlinearity of the inverse mills ratio, which makes possible separate identification of the selection equation and the outcome equation.

The natural main variable of interest in the Heckman selection model is the dependent variable in the outcome equation of the model (*rating*), the rating given for the advice that participants had received when they were purchasing their smartphone handset. As can be seen the (*rating*) variable ranges between 1 and 5 on an ordinal integer value scale. Due to this dependent variable of the outcome equation being defined on an ordinal scale, an ordered probit procedure via the maximum likelihood method is used to estimate the outcome equation. The other dependent variable of the model, the one included in the selection equation, is the dummy variable (*advice*), which is either one or zero depending on whether the person has received advice. Due to the binary nature of (*advice*) the selection equation of the Heckman selection model is estimated with trough a probit regression where the probability of receiving advice is estimated via regressing (*advice*) with respect to the independent variables of the selection equation specification.

The variables included in the estimations were either formed by utilizing the set of questions displayed in appendix 2, were based on the demographic properties of the respondents or were calculated based on the two. As can be seen from Table 9, practically all of the independent variables in this model are of binary nature. Apart from the variables (*gender*) and (*developedmarket*) a positive outcome is noted by the binary variable taking the value 1. I.e. the variable (*age25to39*) taking the value 1 implies the respondent was aged between 25 and 39. The variable (*gender*) takes the value 1 if the respondent was a male and (*developedmarket*) takes the value of 1 if the respondent was living in the developed world. For further description the full set of independent variables utilized in the estimations and their definitions are presented in Table 8, and their key summary statistics are presented in Table 9 in appendix 3.

5.3 Panel data models

As the Heckman selection model presented in section 5.2 was concerned with the effects of word of mouth advice on an individual agent level, the focus is now turned towards the market level effects of word of mouth. In empirical literature, there has so far been little previous research done on the market share effects of word of mouth. The main reasons for this are most likely the elusive nature of word of mouth as a medium communication that can sometimes be hard to observe and the large number of factors that can be seen has effecting the market shares of market participants in a given market.

To overcome these traditional challenges, the approach taken in this thesis is to utilize the large amount of data on word of mouth and sales available from the different smartphone markets across the globe and the powerful empirical methods made available by modern econometrics and panel data estimation methods. As previously referred, the estimation methods utilized in the analysis of the panel dataset consist of the ordinary least squares method, the fixed effects method and the random effects method. Additionally, specification test for testing the validity of these model specifications are also performed. The next sections outline the methodology for performing these estimations and specification tests.

5.3.1 Ordinary least squares model

As a starting point, the first estimation method utilized for analyzing the relationship of word of mouth advice and operation system market shares is a standard ordinary least squares estimation of the relationship between period t market shares and the ratio of word of mouth advice in favor of an operation system present in the population at time t-1. The ordinary least squares method, or OLS method, is a common, standard and relatively simple method for estimating the unknown parameters in a linear regression model. Despite its relatively simplicity, OLS can still have strong explanatory power in many applications. The functioning of the OLS method is based on minimizing the squared distances vertical between the values observed in data and values predicted by a linear approximation. A simple ordinary least squares model for estimating the i^{th} observation of the dependent variable y can be expressed as:

$$y_i = \beta x_i + \varepsilon_i, \tag{5.8}$$

Where y_i is the i^{th} observation of the dependent variable, x_i is the i^{th} observation of the vector of independent variables and ε_i is the unobserved error term, accounting for the differences between the observed responses y_i and the estimated values predicted by βx_i, associated with the i^{th} observation.

In the context of this thesis the utilized ordinary least squares model can be stated as:

$$logOSMS_{it} = \beta * WOMratio_{it} + \varepsilon_{it}, \tag{5.9}$$

where ($logOSMS_{it}$) is the logarithmic market share of operation system i at time t and ($WOMratio_{it}$) is the share of positive word of mouth for operation system i out of the total amount positive word of mouth for all operation systems during the previous month at time t. ε_{it} is the error term. To control for individual operation system specific effects, the OLS specification of equation (5.8) is estimated separately for each of the 4 operation systems in each of the 14 markets.

As with all panel data estimations of the thesis, the dependent variable ($WOMratio_{it}$) was formed on the basis of a consumer survey where the respondents, were asked, if a certain brand had been recommended to them by someone during the last month. The responses were then aggregated for each operation system for each month in each of the countries to produce the relative share of word of mouth as a share of the total amount of word of mouth for each operation system during the previous month in the specific country. ($WOMratio_{it}$) then represents the share of positive word of mouth for operation system i during the month before t out of the total amount of positive word of mouth during that period. As the amount of word of mouth and product sales are likely correlated in such a fashion that they both have an influence on each other, constructing the word of mouth regressors from the ratios of word of mouth is very intuitive and helps in controlling for the possibility of reverse causality affecting the empirical estimations.

As one can see from equation (5.9) the dependent variable in the specification is the logarithmic market share of operation system i in country j and the sole independent variable is the ratio of word of mouth recommendations in favor of the operation system i in the previous month. The choice of using the log transformed market share as the dependent variable is driven by the high likeliness of a nonlinear relationship between market shares and word of mouth advice. As transmitted word of mouth is likely to result in more word of mouth being transmitted by the initial receivers, and as one person can in principle affect several others, it is theoretically likely for word of mouth to experience increasing returns to scale as captured by the log-linear functional form specification used in the panel data estimations.

Obtaining the right variable specification for empirical estimations is usually not as straight forward as one could hope for. Cases in favor of specifications different from the specifications chosen for this thesis can certainly be made. To dispel any uncertainty relating to the validity of the chosen functional form and the chosen independent and dependent variables, the validity of the chosen variable specification is tested by running additional regressions utilizing different variable specifications. The results of these estimations are presented in Table 10 and Table 11 in appendix 4. Table 10 presents the ordinary least squares estimation results for the four operation systems estimated in the United Kingdom market with the four different variable specifications and Table 11 presents the fixed effects estimations for the same variable specifications in the same market. The UK market was chosen for the testing due to its good representativeness of balanced market characteristics.

In specification number 1 the dependent variable is the logarithmic market share of an operation system, (*logOSMS*) and the independent variable is the operation systems ratio of word of mouth, (*WOMratio*). In specification number 2 the independent variable is (*logOSMS*) and the two dependent variables are (*WOMratio*) and the one period lag of (*logOSMS*). In specification number 3 the dependent variable is again (*logOSMS*), but now the sole independent variable is (*WOMratio*) normalized by dividing it with the previous period's market share. In specification number 4 the independent variable is the change in (*logOSMS*) between the current period and the previous period, and the independent variable is (*WOMratio*).

Looking at the results in appendix 4, one immediately notices the high coefficients of determination in specifications 2 and 3. The high R-squared values are due to high correlation between this period's and the previous period's market shares, as in specification 2 the previous periods market share is used as an independent variable and in specification 3 (*WOMratio*) is normalized with the previous period's market share. The unfortunate effect of this is that it makes these R-squared values incomparable to the other two specifications. Judging by the estimation results of both the OLS estimates and the fixed effects estimates as a whole, it is specification 1 that seems to perform best out of out of the four different possible specifications in terms fitting the data and producing meaningful statistically significant estimates. Specification 1 is therefore chosen as the specification to be used in the panel data estimations.

As the homoscedasticity of the error terms and the normality of their distribution is a major concern for a model specification such as that of equation (5.9), it is likely that the normal standard errors produced by OLS estimations are biased. To prevent this, the OLS specifications are estimated with heteroskedasticity robust standard errors. Another major cause of concern regarding the results from the estimations of an ordinary least squares model, such as the one presented above, is the large probability that it experiences some form of omitted variable bias, as there is only a single independent variable in the estimated specification. To minimize the possibility for producing biased estimates for the effects of word of mouth, the above OLS specification is estimated not only separately for each of the

14 countries, but also estimated separately for each of the four operation system in each of the countries.

Estimating separate models is however not the only method for controlling for omitted variable bias. Typically in the analysis of panel data, the possibility for omitted variable bias is controlled for by estimating a model that has time invariant individual specific effects included in model. Following this approach, in the next section a fixed effects model with operation system specific fixed effects is presented.

5.3.2 Fixed effects model

One of the important benefits of utilizing a panel dataset for empirical analysis is the ability of utilizing some of the powerful panel data estimation methods that can utilize both the spatial as well as temporal dimensions of the data. One of the most used panel data estimation methods in empirical economics literature is the fixed effects estimation method. The so called fixed effects models are widely utilized in empirical literature to control for any unobserved heterogeneity in the data and or possible omitted variable bias. In contrast to a standard ordinary least squares model, in a fixed effects model, also known as individual specific effect models, the time invariant component of the error term is allowed to be correlated with the dependent variables of the model. The idea behind the fixed effects method is that given that the unobserved heterogeneity in the data is constant over time it can be removed trough differencing that removes the time invariant components of the model.

To illustrate the fixed effects estimation method, consider the following linear unobserved effects model:

$$y_{it} = \beta x_{it} + u_{it}. \tag{5.10}$$

The error term u_{it} in equation (5.10) can be divided into two different components as follows:

$$u_{it} = \alpha_i + \varepsilon_{it}, \tag{5.11}$$

where ε_{it} is the idiosyncratic component of the error term that is uncorrelated with x_{it} and α_i is the time invariant component that is allowed to correlate with x_{it}. According to Cameron & Trivedi (2009) it is then possible to consistently estimate β for the time varying regressors in x_{it} by subtracting the corresponding model for individual means:

$$\left(y_{it}-\overline{y}_i\right)=\beta(x_{it}-\overline{x}_i)+(\varepsilon_{it}-\overline{\varepsilon}_i), \quad (5.12)$$

which eliminates the time invariant component α_i of the error term u_{it}. in equation (5.10).

The remarkable effect of this relatively simple procedure is to remove the unobserved and individual specific time invariant effects from the estimation results allowing the modeler to focus on the time variant effects on the dependent variable. The obvious drawback of the fixed effect method is that the model cannot include any time invariant independent variables as they would disappear during the differencing procedure. This restriction in the usability of the fixed effect model is nonetheless compensated by the increased ability for controlling for omitted variable bias. As in most study's featuring non-experimental data there is a large possibility for some of the important covariates that have an influence the dependent variable not to be included in the data, the fixed effects method allows to control for the bias the non-inclusion of these variables might have on the estimations produced by a standard ordinary least squares estimation.

In the context of this thesis, the fixed effects model utilized in the estimation of operation system market shares can be expressed as:

$$logOSMS_{it}=\beta*WOMratio_{it}+\alpha_i+\varepsilon_{it}, \quad (5.13)$$

where ($logOSMS_{it}$) is the logarithmic market share of operation system i at time t and ($WOMratio_{it}$) is the share of positive word of mouth for operation system i out of the total amount positive word of mouth for all operation systems during the previous month at time t. ε_{it} represents the idiosyncratic component error term and α_i represents the time invariant component of the error term. The above fixed effects model is estimated with the panel dataset described in Table 7 in appendix 1. Estimating the fixed effects specification separately for all the 14 markets allows for controlling for any specific time invariant market effects and allows the estimated coefficients to differ from the estimates reached via implementing these controls via an ordinary least squares estimation method.

As mentioned in Cameron & Trivedi (2009), the standard errors produced by fixed effects models are likely to be correlated and hence some correction to the default ordinary least squares standard errors is warranted. As the default standard errors produced by Stata 12 used to estimate the model are of the standard ordinary least squares type, the approach of

Cameron & Trivedi is followed and cluster robust standard errors with the operation system as the cluster variable are used in the estimations of the fixed effects model. The cluster robust standard errors allow the error terms to be correlated within the clusters, as in within the operation systems, but not between them.

The main driver behind the utilization of the fixed effects model is to achieve higher estimator efficiency compared to the ordinary least squares estimations, while still simultaneously retaining control over variety of sources for omitted variable bias, especially considering the limited number of possible independent variables available in the data. The use of a fixed effects method for analysis can be seen as warranted by the large possibility for an omitted variable bias in estimates, if the individual operation system specific effects such as inherent product characteristics are not controlled for. As it is more than probable that there are other factors besides word of mouth that affect market shares, the fixed effects method is utilized to minimize the effects of time invariant product characteristics related to the quality of the smartphone product and its various determinants and the user experience that the final product provides. As product quality can be seen as correlating only with the time invariant component of the error term, controlling for the product inherent fixed effects should in this case produce better estimates than a standard ordinary least squares approach.

The validity of the assumption of operation system specific time invariant fixed effects is tested by utilizing the popular Wald test, developed by Wald (1943) and aptly described by Cameron & Trivedi (2005) as the preeminent hypothesis test in microeconometrics. In practice, the Wald test is applied to test the validity of the fixed effects assumption via including dummy variables for each operation system in an ordinary least squares model, to account for any possible fixed effect, and testing post estimation for the joint hypothesis of all these dummy variables being not statistically significantly different from zero. The Wald test described is applied for all 14 markets studied and its results are discussed in the next chapter along with other estimation and specification test results.

In search for possible further improvement on the estimations reached with the ordinary least squares and fixed effects approaches, in the next section another type of model with individual specific effects, the random effects model, is explored.

5.3.3 Random effects model

In addition to the fixed effects model, another type of model that can be used to model individual operation system specific effects on the dependent variable model is the random effects model. The two different types of models with individual specific effects are relatively similar in their construction and estimation procedures, but the models and their interpretations do differ in some of their key aspects. Whereas in the fixed effects models the variation across the cross-sectional units is assumed to be caused by an individual specific and time invariant fixed effect, as described above, in the so called random effects models the variation is assumed to be random and hence uncorrelated with the independent variables in the model.

Technically, the main difference between fixed effects models and random effects models relates to their assumption on the composition error term of the individual specific effects model in equation (5.11). As described by Cameron & Trivedi (2005) among others, whereas the fixed effects specification presumes α_i to be time invariant and correlated with x_{it}, the random effects specification assumes α_i to be an independent and identically distributed random variable. with $\alpha_i \sim [0, \sigma_\alpha^2]$. α_i is also assumed to be independent from ε_{it} ,and both α_i as well as ε_{it} are assumed to be independent from the vector of independent variables x_{it}. The estimated random effects model is then easily obtained from equation (5.13) by making the appropriate interpretations and assumptions on the structure of the error terms.

One of the benefits for the random effects specification when compared to the fixed effects specification is that the random effects specification allows for time invariant independent variables, such as gender, to be used as independent variables as they are not removed by differencing. On the other hand, fixed effects models are less prone to experience omitted variable bias than random effects specifications and, according to Cameron & Trivedi (2005), fixed effects model also have the attraction of allowing for the establishment of causation under weaker assumptions than the random effects model. The choice between a fixed or a random effects specification can often be tricky, but fortunately ways to test for the right specification have been developed.

The most commonly utilized test for testing the validity of the random effects specification, the Hausman test, was developed by Hausman (1978) and is easily applicable for most panel data. As is done in the context of this thesis, the test can be applied by first separately estimating the model with fixed effects and then with random effects and then utilizing the Hausman specification test to test for the appropriateness of the random effects specification. The null hypothesis of the the Hausman test is that the individual specific effects are uncorrelated with the independent variables of the model, corresponding with the random effects specification.

According to Cameron & Trivedi (2009), under the null hypothesis that the individual specific effects are random, the two estimators should be similar due to the consistency of both estimators. If this is the case, according to Cameron & Trivedi the random effects specification should be favored as in the case of a fixed effects specification the use of only within variation leads to less efficient estimates than with a random effects specification. On the other hand if the two estimators substantially differ the random effects estimator is inconsistent and a fixed effects specification should be used. Consequently a Hausman test statistic large enough for the appropriate confidence level, will lead to the rejection of the null hypothesis and to the fixed effects model to be favored over a random effects specification.

Where the Hausman test can be used to differentiate between the fixed effects and the random effects model, it does not give information regarding the validity of the random effects model when compared to the simple ordinary least squares estimator. Luckily again, a method for differentiating between the random effects specification and a simple OLS model has been proposed by Breusch & Pagan (1980) and is called the Breusch Pagan Lagrange multiplier test. The Lagrange multiplier test can be used to test for whether there is variance across entities that would imply a panel effect in the data and strongly favor a random effects specification over OLS. The null hypothesis of the test is that there is no significant difference across units and no heteroskedasticity is present. The null hypothesis is then rejected if the test statistic is large enough for the appropriate confidence level. Rejection of the null hypothesis would then lead to the conclusion that the variance across units is significant and a random effects specification should be favored over the ordinary least squares model, and vice versa.

To test for the right empirical specification for the analysis of the effects of word of mouth on market shares both the Hausman test and the Breusch Pagan Lagrange multiplier test are performed for each of the 14 markets under analysis. As described, the Hausman test is performed after estimating both the fixed effects and random effects specifications and the Breusch Pagan test is performed after the estimation of the random effects specification.

Along with other estimation results, the results of the specification tests are discussed in the next chapter, which focuses on the results of the empirical analysis performed under the framework laid out in this chapter.

6. Empirical results

This chapter presents the empirical findings of the thesis and relates them to the theoretical framework laid out in the preceding chapters, as well as contrasts the findings with those of previous empirical research. The chapter begins by a presentation of the estimation results for the Heckman selection model presented in chapter 5.2. The estimation results for the Heckman selection model and its marginal effects are presented and discussed. After the Heckman selection model the focus of this chapter will turn to presenting and discussing the results of the panel data analysis and the findings of the three empirical models, the ordinary least squares model, the fixed effects model and the random effects model. To determine the performance of the panel data model specifications, the results from the three different model specification tests are also analyzed. As the estimation results are presented their statistical, as well as practical, significance is highlighted and their theoretical, as well as practical, implications are discussed.

6.1 Estimation results for the Heckman selection model

The estimation results for the outcome equation of the Heckman selection model are presented in Table 1, while the results for the selection equation can be found in Table 12 in appendix 5. Looking at the estimation results in Table 1 one immediately notices that all independent variable coefficients are found to be significant at the 1 percent significance level, except the variable 55plus which is found to be significant at the 5 percent significance level. Of great interest is of course the significance of the coefficient estimated for the inverse mills ratio, which is estimated to be significant at the one percent significance level. The negative and highly significant inverse mills ratio signals that a selection bias is apparent in the data and estimating the model without correcting for the selection bias would have produced downwardly biased estimates. Indeed, all of the independent variable coefficients estimated in the selection equation are significant at 1 percent significance level showing further evidence for selection bias in the data warranting the use of the Heckman selection model procedure.

Table 1. Heckman outcome equation estimates

Independent variable	Coefficient	Standard error
active	0.0870***	0.0188
children	0.264***	0.0131
friends	0.0317***	0.0116
colleagues	0.0483***	0.0113
parents	0.0516***	0.0129
spouse	0.137***	0.0106
siblingsandcousins	0.0585***	0.0117
gender	0.423***	0.0859
developedmarket	1.214***	0.149
age25to39	0.142***	0.0336
age40to54	0.204***	0.0695
age55plus	0.204**	0.0858
coreconsumer	0.249***	0.0325
via_face2face	0.299***	0.0161
via_phone	0.128***	0.0139
via_some	0.193***	0.0156
via_email	0.368***	0.0216
via_instmes	-0.100***	0.0174
invmills	-3.815***	0.802
Model statistics		
Number of observations	42407	
Log likelihood	-58844.993	
Chi2	6267.64	
Prob > chi2	0	
Pseudo R2	0.0506	

*** p<0.01, ** p<0.05, * p<0.1

The control group, and hence the base case, for the estimated coefficients is represented by a consumer who is a female, aged between 16 and 24, lives in the developing world and is not considered to be a "core consumer". The control group for the advice sources in terms of people is represented by advice classified in the data as having been received from "other person" and the control group in terms of mediums is represented by advice classified in the data as being received through "other medium".

As explained in the previous chapter when the Heckman selection model was presented, the practical interpretation of the coefficients produced by estimating the ordered probit model can be rather complicated. For greater clarity and to facilitate the analysis of the results, the marginal effects of changes in the independent variables of the outcome equation on the probabilities of the highest and the lowest rating outcome, and their standard deviations, were calculated and are presented in Table 2. As expected, the estimated marginal effects also turn out to be highly significant. The estimated marginal effects presented in Table 2 represent the marginal effect of a discrete change in the value of an independent variable to the probability of a certain rating to be given for the experience of receiving advice on their smartphone purchase by the person receiving advice. The last column of Table 2 reports the estimated partial effects on the probability of receiving advice.

Table 2. Estimated marginal effects in the Heckman model and their standard deviations

Variable	Marginal effect on rating		Standard deviation		Partial effect on P(advice)
	Rating 1	Rating 5	Rating 1	Rating 5	
active	-0.00124***	0.0186***	0.000284	0.00394	
children	-0.00304***	0.0618***	0.000205	0.00331	
friends	-0.000435***	0.00686***	0.000162	0.0025	
colleagues	-0.000652***	0.0105***	0.000155	0.00247	
parents	-0.000680***	0.0114***	0.000168	0.00288	
spouse	-0.00186***	0.0298***	0.000172	0.00231	
siblingsandcousins	-0.000778***	0.0128***	0.000158	0.0026	
gender	-0.00635***	0.0901***	0.0015	0.0179	-0.0736***
developedmarket	-0.0106***	0.344***	0.00129	0.0493	-0.1274***
age25to39	-0.00189***	0.0312***	0.000451	0.00747	-0.0278***
age40to54	-0.00239***	0.0474***	0.00071	0.0172	-0.0597***
age55plus	-0.00223***	0.0486**	0.000753	0.0222	-0.0724***
coreconsumer	-0.00264***	0.0601***	0.0003	0.00864	0.0248***
via_face2face	-0.00437***	0.0639***	0.000338	0.00339	
via_phone	-0.00159***	0.0291***	0.000178	0.00327	
via_some	-0.00224***	0.0449***	0.000194	0.00388	
via_email	-0.00341***	0.0939***	0.000226	0.00631	
via_instmes	0.00151***	-0.0210***	0.000297	0.00347	
invmills	0.0518***	-0.831***	0.0112	0.175	

(*** p<0.01, ** p<0.05, * p<0.1)

Interpreting the results of the calculated partial effects, all partial effects for the dependent variables in the selection equation are all highly significant and, apart from the variable (*coreconsumer*), strictly negative, implying these variables having a negative effect on the probability of receiving advice. The variable (*coreconsumer*), which is estimated to have a positive partial effect, is a dummy variable for the consumer being defined as a core consumer on the basis of his or her socio-demographic attributes. In the current context the core consumer classification can be seen as a valid proxy for the consumer being a technological forerunner influential in new technology diffusion. It is only natural that such consumers are probable to exert more cognitive resources and to consume more word of mouth when making technology related purchase decision, resulting in a higher ex ante probability of them having received advice.

In terms of the estimated marginal effects on ratings, while all of the estimated marginal effects are significant at least at the 5 percent significance level and hence can be considered highly significant, the estimation results are also much in line with what one could expect based on theory economic theory. Unsurprisingly, actively seeking word of mouth advice increases the probability of the advice being given the highest rating by 1.9 percentage points and decreases the probability of it being given the lowest rating by 0.1 percentage points. The results is unsurprising as intuitively active search for word of mouth implies a higher cognitive invest from the individual, making him or her inclined to rate the result of their search favorably to enhance the value of their effort in the form of invested time and recourses.

In terms of sources for word of mouth and in terms of people, as can be seen from Table 2, by far the biggest marginal effect for an increase in the probability of a high rating is found to be for advice received from children, which was estimated to increase the probability of the advice being rated as five by 6.2 percentage points when compared to the control group. Similarly, the advice received from children is also the least probable to receive the poorest rating when compared to the control group. After children the second highest marginal effect for a positive rating is found for advice received from one’s spouse, which is estimated to increase the probability of the highest rating by 3 percentage points. Friends, colleagues, parents and siblings as well as cousins all have relatively smaller marginal effects similar to each other.

These results for the Heckman selection model imply that the ratings given are strongly correlated with closeness of the relationship of the advisor and the receiver of the advice. A stronger social tie between the two implicates a higher rating for the advice by the receiver. This result, while somewhat contradicting the findings of Steffes & Burgee (2009), who found that a stronger social tie does not necessarily implicate a higher rating for advice, can be considered rather intuitive. The result is also in line with the results of Brown & Reingen (1987), who found that strong ties were perceived as more influential, and the traditional theoretical view of the people closest to us having the most influence over us.

Considering the demographic variables that were included in both the selection equation and the outcome equation, the biggest positive effect on probability for the highest rating is estimated for the consumer living in a developed market, which increases the probability for the highest rating by 31.2 percentage points. The possible reasons for this might be related to either cultural differences or to consumers in developed countries on average purchasing more expensive smartphones than consumers in developing countries and thus experiencing the device offering a higher value better matching their expectations, which in turn results in relatively higher ratings. The fact that the marginal effect of the consumer being a core consumer, and thus a so called technological forerunner, is high as well also speaks in favor of this interpretation.

Other notable results related to the demographic variables is also that the probability of rating advice highly seems to increase with age as all age groups have a positive marginal effect that increases with the age groups. Also, males were estimated to being more prone towards giving higher ratings as they are 9 percentage points more likely than women to rate the advice they receive with the highest rating. Both results the probability for a high rating increasing with age and males more probably giving higher ratings than females can be seen stemming from the same underlying product knowledge related factor.

As younger people tend to be more technologically savvy then the old, and men tend to be more interested in technology than women, the underlying factor behind these results might be greater technological knowledge driving the word of mouth consumption process of these individuals, resulting in higher ratings than those given by people with less technological knowledge. A similar view is supported by the fact that advice received from children was estimated to be over 5 percentage points more likely to be given the highest rating than the advice from parents. As younger generations in general possess a better technological knowledge than older generations, this also speaks in favor of underlying product knowledge being one of the drivers behind valuing advice more highly.

In terms of the medium of transmission for word of mouth, the highest probability for receiving the highest rating when compared to the control group is estimated for advice received through email followed by advice received as traditional face to face advice and advice received through social media. While estimated as being statistically significant at the 1 percent significance level, the robustness of the result of advice received via email having the highest probability of receiving the highest rating is somewhat debatable. The empirical interpretation could be that people only receive advice via email form people that they have strong social ties to, resulting in such a high estimate. On the other hand, the proportion of consumers in the data who had received advice through email was relatively low, which questions the validity of generalizing the result to hold for larger populations.

Interestingly receiving advice through instant messaging was estimated to have a negative effect on the probability of a high rating and a positive effect on the probability of a low rating. As with advice received through email, the proportion of consumers who had received advice through instant messaging was rather low, and the possible reasons for such an unintuitive result could relate to the fact that classification of instant messaging can cover not only messages between persons known to each other, but also chats between complete strangers. If the effect of the social tie between the advice giver and receiver is as strong as estimated, the social tie effect can dominate the channel effect resulting in advice received from persons with whom the receiver has weak social ties with as poor disregarding the channel.

In general the estimation results for the effects of the channel that the word of mouth is transmitted with seem to reflect the results estimated for the human sources of word of mouth. In both cases closeness of social tie and the degree of human interaction associated with the advice seem to be positively correlated with the ratings given to advice. As written reviews can be perceived as emulating traditional word of mouth more closely than simple ordinal ratings, the results of this thesis are therefore similar to those found by Chevalier & Mayzlin (2006), who found that written reviews outperformed ordinal ratings in explanatory power when determining sales figures.

The results of the Heckman selection model suggest that the closer the social tie and bigger the resemblance of the medium of transmission to traditional face to face transmission, the higher the probability of the advice being rated highly. The implications of these results to social learning outcomes in the context of the smartphone market will be examined in the next sections, as the panel data estimation results for the effects of word of mouth on market shares are discussed.

6.2 Estimation results for the ordinary least squares model

The estimation results for the ordinary least squares model are presented in Table 3. As can be seen, the ordinary least squares specification was first estimated jointly for all operation systems in each of the 14 markets and then estimated separately for each operation system in each country. Table 13 in appendix 6 presents the bivariate correlations between the dependent variable logistic market share and the independent variable word of mouth ratio for each operation systems in each of the 14 markets. As can be seen from Table 13, the correlation coefficients vary greatly across countries and across operation systems. In some countries very high positive correlations are found, while in others the observed correlation is even found to be negative. The highest positive correlations are in most markets found for the BlackBerry operation system. Although there is large variance in the magnitude and direction of the correlations implied by the bivariate correlation coefficients, in general there is a strong message of a tight relationship between the two variables.

Moving from the bivariate correlations to the estimation results from the ordinary least squares model, the results of the OLS estimations are presented in Table 3. A simple ordinary least squares estimation estimated jointly for all operation systems finds word of mouth to have a positive effect on market share in all markets apart from Brazil and India, where the coefficient for (*WOMratio*) is for some reason estimated to be negative. The highest effect for word of mouth is estimated for Indonesia, where an increase of one percentage point in the positive word of mouth ratio is estimated to increase an operation systems market share by over 8.6 percent. In 9 out of the 14 markets a one percentage point increase in an operation systems share of word of mouth is estimated to increase its market share by more than 3 percent.

Interestingly, in contradiction to most markets, in the three continental Europe markets of France, Germany and Italy the coefficient for word of mouth was estimated to be only 1.1, 2 and 1.3 respectively, likely indicating some region specific structural market characteristics that lessen the impact of word of mouth. When examining the results, it needs to however be noted that in terms of the magnitude of the effect of word of mouth on market shares, these results need to be taken with a grain of salt as in the estimations performed jointly for all operation systems there is a large possibility for individual operation system specific effects biasing the estimation results. Furthermore, while most of the coefficients estimated for word of mouth are significant at the 1 percent significance level, for some markets the R-squared values reported for each of the OLS regressions performed without discriminating between the operation systems are rather poor. It is likely that bundling up the four operation systems together hides much of the inherent variation between them.

Table 3. Ordinary least squares estimates

Ordinary least squares estimations for all operation systems														
Country	Australia	Brazil	China	France	Germany	India	Indonesia	Italy	Mexico	Nigeria	Russia	Saudi Arabia	UK	USA
Dependent variable	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS
WOMratio	3.404***	-0.0925	3.448***	1.147*	2.005***	-4.838*	8.628***	1.314**	6.214***	3.743***	5.116***	3.804***	3.051***	6.872***
	(0.465)	(0.938)	(0.953)	(0.659)	(0.498)	(2.683)	(1.079)	(0.549)	(1.936)	(0.317)	(1.076)	(1.144)	(0.806)	(0.634)
Constant	-3.142***	-2.405***	-4.473***	-2.236***	-2.769***	-1.163	-4.679***	-2.352***	-3.955***	-3.679***	-4.310***	-3.409***	-2.613***	-4.119***
	(0.269)	(0.369)	(0.557)	(0.296)	(0.284)	(0.781)	(0.343)	(0.27)	(0.681)	(0.236)	(0.483)	(0.484)	(0.336)	(0.281)
Observations	72	72	72	72	72	72	72	72	72	72	72	72	72	72
R-squared	0.221	0	0.05	0.029	0.077	0.059	0.312	0.036	0.135	0.46	0.127	0.119	0.146	0.405
Ordinary least squares estimations for Android														
Country	Australia	Brazil	China	France	Germany	India	Indonesia	Italy	Mexico	Nigeria	Russia	Saudi Arabia	UK	USA
Dependent variable	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS
WOMratio	-4.013	0.252	-0.117	-0.181	-0.177	0.145	-2.084*	0.368	2.460*	0.911	-0.0953	6.253**	0.783	1.187
	(2.65)	(0.285)	(0.454)	(1.381)	(0.731)	(0.995)	(0.998)	(0.297)	(1.253)	(2.392)	(0.0992)	(2.169)	(1.266)	(0.94)
Constant	0.00349	-0.214***	-0.0968	-0.423**	-0.314***	-0.305	-0.536**	-0.435***	-0.825***	-1.531***	-0.160***	-1.315***	-0.855***	-0.865***
	(0.335)	(0.0613)	(0.0668)	(0.17)	(0.0944)	(0.188)	(0.204)	(0.0451)	(0.219)	(0.148)	(0.0229)	(0.251)	(0.149)	(0.178)
Observations	18	18	18	18	18	18	18	18	18	18	18	18	18	18
R-squared	0.136	0.027	0.008	0.001	0.003	0.001	0.158	0.06	0.168	0.015	0.024	0.191	0.021	0.088
Ordinary least squares estimations for BlackBerry														
Country	Australia	Brazil	China	France	Germany	India	Indonesia	Italy	Mexico	Nigeria	Russia	Saudi Arabia	UK	USA
Dependent variable	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS
WOMratio	18.98***	10.32***	15.86	6.974***	3.009	13.38***	0.0254	6.910***	6.827***	0.244	3.549	-0.404	5.174***	13.64***
	(3.168)	(3.421)	(9.189)	(2.138)	(2.474)	(3.763)	(0.868)	(1.426)	(1.038)	(0.321)	(3.319)	(1.052)	(1.44)	(2.671)
Constant	-6.869***	-5.822***	-10.29***	-3.942***	-3.945***	-5.887***	-0.559	-4.282***	-4.048***	-0.556*	-7.275***	-0.894**	-3.234***	-6.388***
	(0.539)	(0.819)	(1.560)	(0.595)	(0.41)	(1.109)	(0.363)	(0.282)	(0.325)	(0.284)	(0.477)	(0.388)	(0.434)	(0.517)
Observations	18	18	18	18	18	18	18	18	18	18	18	18	18	18
R-squared	0.493	0.259	0.137	0.434	0.08	0.375	0	0.505	0.629	0.043	0.031	0.004	0.355	0.492
Ordinary least squares estimations for Windows Phone														
Country	Australia	Brazil	China	France	Germany	India	Indonesia	Italy	Mexico	Nigeria	Russia	Saudi Arabia	UK	USA
Dependent variable	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS
WOMratio	5.264	0.802	-4.384	7.062*	-4.489**	21.06***	12.59**	3.881	13.97	0.156	7.140*	-3.96	5.66	2.194
	(4.872)	(3.048)	(5.763)	(3.766)	(2.078)	(5.961)	(5.477)	(7.466)	(18.56)	(1.534)	(3.734)	(5.066)	(4.733)	(3.584)
Constant	-4.003***	-3.528***	-3.696***	-4.306***	-3.109***	-7.268***	-5.786***	-3.610***	-6.851***	-4.865***	-3.611***	-5.585***	-4.374***	-4.393***
	(0.437)	(0.415)	(0.747)	(0.421)	(0.312)	(1.181)	(0.564)	(0.968)	(2.061)	(0.219)	(0.451)	(0.413)	(0.532)	(0.377)
Observations	18	18	18	18	18	18	18	18	18	18	18	18	18	18
R-squared	0.039	0.004	0.038	0.12	0.28	0.42	0.171	0.017	0.026	0	0.204	0.032	0.04	0.016
Ordinary least squares estimations for iOS														
Country	Australia	Brazil	China	France	Germany	India	Indonesia	Italy	Mexico	Nigeria	Russia	Saudi Arabia	UK	USA
Dependent variable	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS	logOSMS
WOMratio	-1.453	-2.480**	0.548	-0.855	0.0197	4.598	-1.094**	-1.053	-2.785	1.525*	-0.323	1.227**	1.900***	1.967***
	(1.527)	(1.081)	(2.025)	(1.211)	(0.92)	(7.546)	(0.485)	(0.794)	(2.625)	(0.773)	(0.887)	(0.531)	(0.648)	(0.657)
Constant	-0.183	-1.693***	-2.744**	-1.169*	-1.525**	-5.285*	-3.662***	-1.052**	-1.17	-4.446***	-2.160***	-2.859***	-2.104***	-1.908***
	(0.957)	(0.475)	(1.133)	(0.619)	(0.557)	(2.552)	(0.146)	(0.415)	(0.936)	(0.148)	(0.455)	(0.25)	)(0.3	(0.339)
Observations	18	18	18	18	18	18	18	18	18	18	18	18	18	18
R-squared	0.058	0.175	0.003	0.027	0	0.03	0.067	0.063	0.072	0.19	0.004	0.092	0.233	0.351
Robust standard errors in parentheses														*** p<0.01, ** p<0.05, * p<0.1

To further investigate the effects of word of mouth on market shares, the results for the ordinary least squares estimations performed separately for each operation system in each of the markets are presented in Table 3 below the estimates for the all operation systems combined. The separate estimations for each operation system bring out some of the variation between them as the estimated coefficients for (*WOMratio*) experience huge variations across different operation systems and different markets. For Android the OLS estimator seems to perform poorly as the estimated coefficients for are found to be statistically significant at least at the 10 percent significance level only in two of the 14 markets and the estimated coefficients of determination, R-squared values, are rather low.

On the other hand, for BlackBerry the simple OLS estimator seems to perform much better in explaining the variance in market shares. The estimated coefficients for (*WOMratio*) for BlackBerry are found to be significant at the 1 percent significance level for 8 out of the 14 markets, with the coefficients of determination for these markets attaining decent values. Between Android and Blackberry, the two ends of the estimation results spectrum, are iOS and Windows Phone for which both the estimation results in terms of coefficients estimated for (*WOMratio*) and the reported regression R-squared values vary greatly between markets.

While in some markets the coefficient for word of mouth is estimated to be even negative, though usually not at any reasonable level of statistical significance, the highest coefficient for word of mouth in these operation system specific OLS estimations is estimated for Windows Phone in India, where a one percentage point increase in Windows Phone's word of mouth ratio is estimated to increase its market share by 21 percent. Very high coefficient estimates are also found for BlackBerry in many markets with, e.g. in Australia a one percentage point increase in BlackBerry's word of mouth share is estimated to increase its market share by almost 19 percent. Similarly high and statistically significant at the 1 percent significance level coefficients for (*WOMratio*) are also estimated for BlackBerry in many of the markets.

The general Implication of the results for the ordinary least squares estimations is very similar to what can be reached from examining the correlation coefficients presented earlier. Correlation and in some cases even causality of varying strength between word of mouth and market shares is implied by the results. In some markets and for some operation systems such as BlackBerry, and to an extent iOS, the simple OLS estimator seems to perform well in linking word of mouth and market shares. For some operation systems in some of the markets a valid case for positive word of mouth having a positive effect on market shares can be made. Yet, in some cases the OLS estimation results leave a lot to be desired in terms of robustness and explanatory power of the model and results. To further examine the relationship between word of mouth and market shares, the results for the estimated fixed effects model are presented in the next section.

6.3 Estimation results for the fixed effects model

The estimation results for the fixed effects model are presented in Table 4. As one can see, the results for the fixed effects model are rather different from the results of the similar ordinary least squares model estimations including all of the operation systems. The coefficient for word of mouth is found to be significant only in two of the fourteen markets. In India (*WOMratio*) is found to be significant at one percent significance level and in the United Kingdom (*WOMratio*) is found to significant at the 5 percent significance level What is notable in the estimation results for India is the size of the estimated coefficient. According to the estimation results, a one percentage point increase in the relative word of mouth share for an operation system in India will result in an increase of 10.5 percent increase in its market share. In the UK on the other hand, a one percentage point increase in (*WOMratio*) is estimated to increase an operation systems market share by 3.6 percent.

Table 4. Fixed effects model estimates

Country	WOMratio		Constant		Observations	R-squared
Australia	2.013	(3.823)	-2.794*	(0.956)	72	0.017
Brazil	2.819	(3.581)	-3.133**	(0.895)	72	0.045
China	1.869	(3.12)	-4.079**	(0.78)	72	0.008
France	2.494	(2.531)	-2.572**	(0.633)	72	0.066
Germany	-0.245	(1.449)	-2.207***	(0.362)	72	0.001
India	10.53***	(2.744)	-5.005***	(0.69)	72	0.18
Indonesia	1.707	(2.902)	-2.949**	(0.725)	72	0.015
Italy	1.67	(2.231)	-2.441**	(0.558)	72	0.02
Mexico	3.918	(3.057)	-3.381**	(0.764)	72	0.021
Nigeria	0.652	(0.393)	-2.906***	(0.0983)	72	0.017
Russia	0.895	(1.02)	-3.255***	(0.255)	72	0.008
Saudi Arabia	0.508	(0.898)	-2.585***	(0.224)	72	0.003
UK	3.634**	(1.128)	-2.758***	(0.282)	72	0.057
USA	4.453	(2.882)	-3.514**	(0.721)	72	0.163

Robust standard errors in parentheses — *** p<0.01, ** p<0.05, * p<0.1

The results for the fixed effects model are rather interesting in the sense that, while the coefficient for word of mouth is estimated to be significant in only two out of the fourteen markets, in the two markets where (*WOMratio*) is estimated to be significant it is not only significant but also rather large. Also very interestingly, although not statistically significant, the estimated coefficient for (*WOMratio*) in Germany is negative even after accounting the individual operation system specific effects. Low values for the coefficients of determination are typical for panel data models and fixed effects estimations as in such models the explanatory power of the intercepts are not included in the coefficients. Nevertheless, the R-squared values reported in Table 4 indicate poor fit for the fixed effects model. In the face of such mixed and unintuitive results validity of the model specification must hence be analyzed and further inspected.

To compare the fixed effects model and the ordinary least squares model and to test whether there are operation system specific fixed effects in the data the previously introduced Wald test for the joint hypothesis of all operation system dummies being zero. The results of the Wald tests are reported in Table 14 in appendix 7. As one can see from the Wald test results reported in Table 14, the F-statistic values greatly exceed the commonly used critical values for the rejection of the null hypothesis. The null hypothesis of dummy variable coefficients being zero is thus firmly rejected in all markets, implying the presence of operation system specific effects.

The joint implications of the results of the fixed effects model and the results of the Wald test are that, though according to the Wald test there are definitely operation system specific effects inherent in the data, these effects might not be properly captured by the ordinary least squares method. Due to the relatively poor explanatory power of the fixed effects model a random effects specification was also estimated. To see, if the estimated random effects specification performs better than the fixed effects specification in capturing the variance in market shares caused by word of mouth, the results of the random effects model are presented in the next section.

6.4 Estimation results for the random effects model

The estimation results for the random effects model presented in Table 5 are somewhat similar to the estimation results from the fixed effects model and most of the estimated coefficients are close to those produced by the fixed effects estimations. While the estimated (*WOMratio*) coefficient in the United Kingdom is still significant at the 1 percent significance level, in India the coefficient is now significant only at the 10 percent significance level. In Nigeria the coefficient for (*WOMratio*) was not statistically significant in the fixed effects estimation, but in the random effects estimation it is now estimated to be significant at the 5 percent significance level.

As can be seen from Table 5, the estimation results from the random effects specification are mixed. Even though positive and significant at the 1 percent significance level in the OLS estimation, the estimated coefficient for (*WOMratio*) in Germany is still negative, albeit not statistically significant, just as in the fixed effects model. In terms of magnitudes of the effects of word of mouth, the coefficients estimated for UK and India are close those estimated with the fixed effects specification.

Table 5. Random effects model estimates

Country	WOMratio		Constant		Observations
Australia	2.243	(3.359)	-2.852*	(1.496)	72
Brazil	2.718	(3.549)	-3.107**	(1.326)	72
China	1.957	(3.108)	-4.101*	(2.227)	72
France	2.414	(2.461)	-2.553**	(1.01)	72
Germany	-0.173	(1.4)	-2.225***	(0.863)	72
India	9.722*	(5.793)	-4.803**	(2.032)	72
Indonesia	1.96	(3.008)	-3.012**	(1.499)	72
Italy	1.638	(2.108)	-2.432**	(0.99)	72
Mexico	4.144	(3.2)	-3.437**	(1.461)	72
Nigeria	0.811**	(0.341)	-2.946***	(1.051)	72
Russia	0.964	(1.051)	-3.272**	(1.553)	72
Saudi Arabia	0.592	(0.753)	-2.606**	(1.068)	72
UK	3.561***	(1.324)	-2.740***	(0.885)	72
USA	4.57	(2.79)	-3.543***	(1.296)	72

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

For Nigeria the random effects estimation estimates the coefficient for (*WOMratio*) to be rather low at 0.8. This result coupled with the random effects estimation results for India would implies that word of mouth is almost twelve times as effective in affecting market shares in India as it is in Nigeria, as a one percentage point increase in (*WOMratio*) is estimated to increase an operation systems market share 0.8 percent in Nigeria and 9.7 percent in India. The validity of this result naturally requires some reservations as such a difference between markets in the effects of word of mouth can be seen as rather large. A possible explanation for the relatively high coefficient estimates for India might relate to the country's culture focusing on close personal ties, which as previously shown can greatly increase the effectiveness of word of mouth.

What is also interesting across the estimation results for all three different panel data estimation methods is the robustness of the estimated market share increasing effects of positive word of mouth in the United Kingdom market. The coefficient for (*WOMratio*) is estimated to be positive and significant at the 1 percent significance level all of the three panel data models including the full group of operation systems. The probable reason for such robust estimates is the balanced nature of the high level competition in the UK market producing variance in market shares facilitating more precise and robust estimates.

As the estimates produced by the random effects estimations are very close to those produced by the fixed effects estimations, the validity of the random effects specification compared to the fixed effects specification is tested via the Hausman test introduced in the previous chapter. The results of the Hausman test are reported in Table 15 in appendix 7. As can be seen from the results, the null hypothesis of the individual specific effects being uncorrelated with (*WOMratio*), the independent variable of the model, is not rejected in any of the 14 markets at any of the commonly used critical values. The implication of this result is that the random effects estimator is consistent and in this case should be favored over the fixed effects estimator as argued by Cameron & Trivedi (2009).

Although superior to the fixed effects specification according to the Hausman test results, the explanatory power of the random effects specification seems to still leave a lot to be desired. The appropriateness of the random effects specification is therefore compared also to the ordinary least squares specification with the previously introduced Breusch Pagan Lagrange multiplier test. The results for the Breusch Pagan Lagrange multiplier test are reported in Table 16 in appendix 7. The test results indicate that the null hypothesis of the variance between entities being zero and no panel effect being present in the data is firmly rejected in all markets for all commonly used significance levels. The results therefore indicate there to be significant variance across entities in the data, implying the presence of individual operation system specific panel effect. Combining the results of the Hausman test and the Breusch Pagan Lagrange multiplier test leads to the conclusion of the data experiencing individual operation system specific effects of the random effects type.

As a whole, the estimation results for the panel data estimations presented above signal mixed results. The random effects specification seems to perform best out of the three different model specifications considered, but the results are it produces are not conclusive. At the market level there seems to be a strong implication of correlation between word of mouth and market shares, but in the light of the current data establishing robust causality between the two has proven to be far more difficult In terms of whether word of mouth causes herd behavior in the sense that positive word of mouth leads to consumers herding towards a certain brand in their consumption choices is not yet fully clear in the light of the current data. Some of the results, such as those found for the UK market, are however suggestive of such an outcome being a valid possibility. The key drivers of the empirical findings presented in this chapter as well as their broader implications are discussed more profoundly in the next chapter, which presents the main conclusions to be drawn from the thesis.

7. Conclusions

The purpose of this thesis was to examine the effects of word of mouth facilitated social learning in the context of the smartphone markets. Although the phenomenon of social learning has been recognized long ago, the exact mechanism how social learning facilitates information aggregation and what are the factors that contribute favorably to extreme learning outcomes are not yet clear. After reviewing the relevant body of social learning literature, the effects and effectiveness of word of mouth were examined in the context of the smartphone market, on an individual level and at a market wide level. This chapter concludes the findings of the thesis, discusses its limitations and explores its implications for further research in the field.

7.1. Findings

The main findings of this thesis can be divided between the results from the individual level analysis and the results of the market level analysis. At the individual level, the effectiveness of word of mouth is found to correlate closely with strength of the social tie between the advisor and the receiver of the advice. This view is further supported by the results of how the medium of advice transmission effects the valuation of word of mouth advice, as advice received via more personal type mediums was rated more highly than advice received via impersonal mediums. A large rating increasing effect was also found for the consumer participating in active search for advice, most likely reflecting the rating increasing effect of cognitive resource investment.

Other factors, such as demographic factors, seem to have an influence as well with most notably consumers living in the developed word to be much more probable to rate the advice they receive highly when compared to consumers living in the developing world. The possible reasons for this is might well be related to cultural differences or hte the relatively higher product quality of the smartphones purchased by developed country consumers when compared to their developing country counterparts. Besides place of residency, results regarding the effects of other demographic factors such as age and gender seem to reflect the underlying technological proficiency of the consumer, as being a male, being young and being a technological forerunner all increased the probability of rating received advice highly when making a decision on a smartphone purchase.

At the market wide level the main interest of this thesis was to study the effects of word of mouth on smartphone operation system market shares and to determine whether word of mouth advice could result in social learning facilitated herd behavior resulting in consumers herding towards a particular brand in their consumption decisions. While the empirical results suggest that there is definitely a strong correlation between the two, the strength of the relationship and the existence and strength of causality between word of mouth and market shares is not yet fully clear. The results suggest that in some markets positive word of mouth does have a positive effect on an operation system's market share. That being said, the estimation results for some of the markets are a lot less conclusive. What is clear is that the results imply there to be large differences between markets and operation systems in word of mouth's effects on market shares.

Considering the market share effects of word of mouth, the results of this thesis can be summarized as implying positive word of mouth having a positive effect on a product's market share, but due to mostly data related limitations the exact isolation and quantification of this effect still requires further research. The somewhat robust results for word of mouth driving market shares the United Kingdom market are especially interesting. In the UK positive word of mouth is estimated to have a positive and a highly significant effect on market shares with all three panel data estimation methods. The United Kingdom market is an important example in the sense that the four players each have a significant share of the market, which is not the case in all markets. Some players having stable dominating position in some of the markets naturally brings some challenges to modeling market share changes due to their stabile nature.

Based on the results of the thesis' empirical analysis it can be seen as plausible that word of mouth does affect market shares, at least to a degree. This market share effects is however highly dependent on the multiple market related factors such as market characteristics and factors influencing the social learning process. Additionally, based on the empirical analysis it can be seen as possible for word of mouth to result in social learning via aggregation of information. Herd behavior in consumption choices resulting in consumers herding towards a particular brand is hence a valid possible market outcome. Nevertheless, as with all social learning situations, the end result is dependent on numerous market and situation specific factors. The results of this thesis are encouraging for linking word of mouth with sales and ultimately with company performance, but more specific future research on the subject will hopefully help in isolating the effects.

7.2. Limitations

When examining the results of this thesis and generalizing them into a broader context, there are several provisions and limitations to take into account. Some the limitations of this thesis are obvious, whereas some of them merit further inspection. As with any empirical study utilizing empirical data, the nature and structure of the data available bring bounds on the empirical analysis. As the Brand Relationship Tracker data utilized for empirical analysis was collected via survey methods, the question of how well the recorded responses actually reflect the true opinions and valuations of subjects. The estimations performed with the Heckman selection model were rather simple in terms of identifying the sources and effects of word of mouth. Future research with a more diverse and flexible dataset could bring more light to the different factors determining the effectiveness of word of mouth, especially in terms of channel effects for technology facilitated word of mouth.

The lack of suitable identification variables for the Heckman selection model, and especially the lack of suitable control variables for the panel data estimations, complicated the empirical analysis and could have hindered the isolation of word of mouth effects. Considering the market share effects of word of mouth, the generalizability of the results reached here needs to be carefully considered as, while covering a geographically diverse set of markets, the performed analysis was based on data from only a single industry. Moreover, as the available data enabled the use of only a single independent variable in the panel data estimations, the robust identification of causality between word of mouth and market shares provides further challenges.

The nature of word of mouth as a highly variable measure in terms of experiencing considerable variance between markets, products and time periods brings about considerable challenges to its analysis via econometric techniques. The empirical analysis in this thesis attempted to control for some of these effects via imposing controls for some of these factors. Nevertheless, the time variant effects of product characteristic of the operation systems can be seen as possible factors influencing the estimation results, complicating the identification of the market share effects of word of mouth. As the product characteristics of the operation systems may experience considerable variation in time, the explanatory power of the models that do not take this into account can be seen as questionable. Hopefully future research utilizing methods that allow for the time variant effects to be controlled aid in further identification of the product market effects of word of mouth.

7.3. Implications

The field of social learning is still a young field in economics and new theoretical as well as empirical research is constantly expanding the bounds of our knowledge. One of the functions of this thesis is to act as a platform helping to guide future research in the field. As the precise effects of word of mouth advice in social learning are not yet conclusive, more research is needed to validate social learning theories in a real world environment. Based on the findings of this thesis, there is a variety of factors that can be seen as beneficial in aiding future research. Many of the biggest limitations of this thesis are related to the data available for empirical analysis. Richer datasets including larger amount of variables available for estimations than the datasets utilized here will without a doubt prove more beneficial in accurately isolating the effects of word of mouth and controlling for time variant variable properties among other relevant factors.

In terms of the findings of this thesis, its contributions can be seen as one of the first attempts in empirical modeling the market effects of word of mouth. Armed with more diverse and suitable data, future research in the field will hopefully both broaden our understanding on how word of mouth facilitates social learning and what are its quantifiable real world implications. One of the most interesting avenues for further research is how the medium of transmission affects the effectiveness of word of mouth advice. As the internet, together with platforms such as social media, facilitate more and more advice transmission, the effects of these new channels to the age old phenomenon of social learning within a group invite further examination.

Identifying the determinants of the effectiveness of word of mouth in affecting individual decision making and aggregating information between agents is essential for determining its broader effects. The effective identification of word of mouth's effect on the market behavior of consumers and companies alike will require more research to enable us to fully understand the mechanism of how the two are linked. As mentioned before, the availability of sufficient data to allow for diverse modeling will be of essence. Establishment of proof for robust causality between word of mouth and product sales will be the first step in this endeavor. From an economics perspective one of the important points of interest will also lie in how word of mouth affects company performance and market conditions. These two factors will hopefully be examined further in future research in the field.

The objective of this thesis has been to contribute to the body of social learning literature by examining the effects of social learning and herd behavior by reviewing the extensive body of relevant literature and empirically investigating the determinants of the real world effects of word of mouth in social learning. Moreover, the thesis has especially aimed to contribute to the relative scarce body of literature empirically examining the real world market effects of word of mouth advice. As shown by both theoretical and empirical research, the social learning effects of word of mouth are extremely sensitive to numerous situation and learning environment specific factors. More research on the different determinants of learning outcomes in terms of sources of word of mouth and its delivery conditions is needed to broaden our understanding of word of mouth's implications on individual agent behavior and on market outcomes.

References

Literature references

Acemoglu, D., Dahleh, M. A., Lobel, I., & Ozdaglar, A. (2011). Bayesian learning in social networks. *The Review of Economic Studies*, *78*(4), 1201-1236.

Anderson, E. W. (1998). Customer satisfaction and word of mouth. *Journal of Service Research*, *1*(1), 5-17.

Anderson, L. R., & Holt, C. A. (1997). Information cascades in the laboratory. *The American economic review,87*(5), 847-862.

Bala, V., & Goyal, S. (1998). Learning from neighbours. *The Review of Economic Studies*, *65*(3), 595-621.

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

Banerjee, A., & Fudenberg, D. (2004). Word-of-mouth learning. *Games and Economic Behavior*, *46*(1), 1-22.

Bass, F. M. (1969). A new product growth model for consumer durables. *Management Science. 15*(5), 215-227.

Becker, G. S. (1991). A note on restaurant pricing and other examples of social influences on price. *The Journal of Political Economy, 99*(5), 1109-1116.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *The Journal of political Economy*, *100*(5), 992-1026.

Breusch, T. S., & Pagan, A. R. (1980). The Lagrange multiplier test and its applications to model specification in econometrics. *The Review of Economic Studies*, *47*(1), 239-253.

Brown, J. J., & Reingen, P. H. (1987). Social ties and word-of-mouth referral behavior. *Journal of Consumer Research*, *14*(3), 350-362.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and applications.* New York, NY: Cambridge University Press.

Cameron, A. C., & Trivedi, P. K. (2009*). Microeconometrics using stata*. College Station, TX: Stata Press.

de Caritat, M. J. A. N. (1785). Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. Paris, France: L'imprimerie royale.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. Journal of marketing research, *43*(3), 345-354.

Çelen, B., & Kariv, S. (2004a). Observational learning under imperfect information. Games and Economic Behavior, *47*(1), 72-86.

Çelen, B., Kariv, S. (2004b). Distinguishing informational cascades from herd behavior in the laboratory. The American Economic Review *94*(3), 484-498.

Çelen, B., Kariv, S., & Schotter, A. (2010). An experimental test of advice and social learning. Management Science, *56*(10), 1687-1701.

Chamley, C. P. (2003). Rational herds: Economic models of social learning. West Nyack, NY, USA: Cambridge University Press.

Cipriani, M., & Guarino, A. (2005). Herd behavior in a laboratory financial market. American Economic Review, *95*(5), 1427-1443.

Clement, M., Proppe, D., & Rott, A. (2007). Do critics make bestsellers? Opinion leaders and the success of books. Journal of Media Economics, *20*(2), 77-105.

DeMarzo, P. M., Vayanos, D., & Zwiebel, J. (2003). Persuasion bias, social influence, and unidimensional opinions. The Quarterly Journal of Economics, *118*(3), 909-968.

Drehmann, M., Oechssler, J., & Roider, A. (2005). Herding and contrarian behavior in financial markets: An internet experiment. The American economic review, *95*(5), 1403-1426.

Duan, W., Gu, B., & Whinston, A. B. (2008). Do online reviews matter?—An empirical investigation of panel data. Decision Support Systems, *45*(4), 1007-1016.

Gale, D., & Kariv, S. (2003). Bayesian learning in social networks. Games and Economic Behavior, *45*(2), 329-346.

Galef Jr, B. G., & Laland, K. N. (2005). Social learning in animals: empirical studies and theoretical models. Bioscience, *55*(6), 489-499.

Golub, B.,& Jackson, M. O. (2010). "Naïve Learning in Social Networks and the Wisdom of Crowds." *American Economic Journal: Microeconomics, 2*(1), 112-149.

East, R., Hammond, K., & Lomax, W. (2008). Measuring the impact of positive and negative word of mouth on brand purchase probability. *International Journal of Research in Marketing*, *25*(3), 215-224.

Ellison G. and Fudenberg D. (1993) Rules of Thumb for Social Learning. *The Journal of Political Economy, 101*(4), 612-643.

Ellison, G., & Fudenberg, D. (1995). Word-of-mouth communication and social learning. *The Quarterly Journal of Economics*, *110*(1), 93-125.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, *46*(6), 1251-1271

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica, 47*(1), 153-161.

Hung, A. A., & Plott, C. R. (2001). Information cascades: Replication and an extension to majority rule and conformity-rewarding institutions. *American Economic Review, 91*(5), 1508-1520.

Kübler, D., & Weizsäcker, G. (2004). Limited depth of reasoning and failure of cascade formation in the laboratory. *The Review of Economic Studies*, *71*(2), 425-441.

Machiavelli, N. (1988). *The Prince.* (Q. Skinner & R. Price, Trans.). Cambridge: Cambridge University Press. (Original work published 1514)

Morgan, N. A., & Rego, L. L. (2006). The value of different customer satisfaction and loyalty metrics in predicting business performance. *Marketing Science*, *25*(5), 426-439.

Morgan, T. J. H., Rendell, L. E., Ehn, M., Hoppitt, W., & Laland, K. N. (2012). The evolutionary basis of human social learning. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1729), 653-662.

Schotter, A. (2003). Decision Making with Naive Advice. *The American Economic Review*, *93*(2), 196-201.

Smith, L., & Sørensen, P. (2000). Pathological outcomes of observational learning. *Econometrica*, *68*(2), 371-398.

Steffes, E. M., & Burgee, L. E. (2009). Social ties and online word of mouth. *Internet Research*, *19*(1), 42-59.

Verbeek, M. (2008). *A guide to modern econometrics* (3rd ed.). Chichester; England: Wiley.

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

Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, *54*(3), 426-482.

Weizsäcker, G. (2010). Do we follow others when we should? A simple test of rational expectations. *American Economic Review*, *100*(5), 2340–2360.

Electronic references

Gartner. (2013, February 13). Gartner Says Worldwide Mobile Phone Sales Declined 1.7 Percent in 2012 [Press release]. Retrieved from http://www.gartner.com/newsroom/id/2335616

IDC. (2013, February 14). Android and iOS Combine for 91.1% of the Worldwide Smartphone OS Market in 4Q12 and 87.6% for the Year, According to IDC [Press release]. Retrieved from http://www.idc.com/getdoc.jsp?containerId=prUS23946013

Appendix 1. Countries and observations

Table 6. Observations utilized in the estimation of the Heckman selection model

Country	Freq.	Percent	Cum.
Australia	8439	12.47	12.47
Brazil	2447	3.62	16.09
China	6203	9.17	25.26
France	7359	10.88	36.13
Germany	6581	9.73	45.86
India	2876	4.25	50.11
Indonesia	2948	4.36	54.47
Italy	6668	9.86	64.32
Mexico	2923	4.32	68.64
Nigeria	272	0.4	69.05
Russia	2928	4.33	73.37
Saudi Arabia	4673	6.91	80.28
UK	8402	12.42	92.7
USA	4941	7.3	100
Total	67660	100	

Table 7. Observations utilized in the estimation of panel data models

Country	Freq.	Percent	Cum.
Australia	72	7.14	7.14
Brazil	72	7.14	14.29
China	72	7.14	21.43
France	72	7.14	28.57
Germany	72	7.14	35.71
India	72	7.14	42.86
Indonesia	72	7.14	50
Italy	72	7.14	57.14
Mexico	72	7.14	64.29
Nigeria	72	7.14	71.43
Russia	72	7.14	78.57
Saudi Arabia	72	7.14	85.71
UK	72	7.14	92.86
USA	72	7.14	100
Total	1008	100	

Appendix 2. Brand Relationship Tracker questions

"When looking around for and choosing your current phone was your choice influenced by any of the following people…"

- Your children,
- Your friends
- Your colleagues
- Your parents
- Your spouse or partner
- Your brothers, sisters or cousins
- Other

"How would you rate your experience of getting guidance from other people when looking around for and choosing your current phone?"

1 = Poor, 2 = Fair, 3 = Good, 4 =Very good, 5 =Excellent, 98 = Don't know

"Did you actively seek advice and guidance from people you know when choosing your current phone?"

- Yes
- No

"How did you seek this advice and guidance?"

- Face to face
- Over the phone
- Online using social media (e.g. Facebook, Twitter etc.)
- By email
- Online using instant messaging
- Other

Appendix 3. Heckman selection model variables

Table 8. Definition of variables used in the Heckman selection model

Variable	Definition
rating	Rating given for advice
advice	Dummy variable indicating whether advice has been received
active	Dummy variable indicating if advice was active sought
children	Dummy variable indicating if advice was received from children
friends	Dummy variable indicating if advice was received from friends
colleagues	Dummy variable indicating if advice was received from colleagues
parents	Dummy variable indicating if advice was received from parents
spouse	Dummy variable indicating if advice was received from ones spouse
siblingsanandcousins	Dummy variable indicating if advice was received from siblings or cousins
gender	Dummy variable for participants gender (male=1)
developedmarket	Dummy variable for participants market (developed=1)
age25to39	Dummy variable for the participant to be aged from 25 to 39
age40to54	Dummy variable for the participant to be aged from 40 to 54
age55plus	Dummy variable for the participant to be aged 55 or over
coreconsumer	Dummy variable for the participant to be classified as a "core consumer"
via_face2face	Dummy variable indicating if advice was received via face to face communication
via_phone	Dummy variable indicating if advice was received via phone
via_some	Dummy variable indicating if advice was received via social media
via_email	Dummy variable indicating if advice was received via email
via_instmes	Dummy variable indicating if advice was received via instant messaging
invmills	Inverse Mills ratio calculated based on the selection equation

Table 9. Summary statistics for variables used in the Heckman selection model

Variable	N	Mean	Std. Dev.	Min	Max
rating	42859	3.6725	0.9211	1	5
advice	67660	0.6467	0.4780	0	1
active	67660	0.4120	0.4922	0	1
children	67660	0.1453	0.3524	0	1
friends	67660	0.3710	0.4831	0	1
colleagues	67660	0.2440	0.4295	0	1
parents	67660	0.1260	0.3319	0	1
spouse	67660	0.3092	0.4622	0	1
siblingsanandcousins	67660	0.1953	0.3964	0	1
gender	67660	0.5203	0.4996	0	1
developedmarket	67660	0.6956	0.4602	0	1
age25to39	67660	0.4282	0.4948	0	1
age40to54	67660	0.2377	0.4257	0	1
age55plus	67660	0.1167	0.3210	0	1
coreconsumer	67660	0.0861	0.2805	0	1
via_face2face	67660	0.3458	0.4756	0	1
via_phone	67660	0.1082	0.3107	0	1
via_some	67660	0.0974	0.2965	0	1
via_email	67660	0.0397	0.1953	0	1
via_instmes	67660	0.0497	0.2173	0	1
invmills	67660	0.5576	0.1126	0.3059	0.7394

Appendix. 4 Variable specification tests

Table 10. Ordinary least squares variable specification tests

1. Dependent variable (*logOSMS*); Independent variable (*WOMratio*)				
OS	Android	BlackBerry	WP	iOS
Dependent variable	logOSMS	logOSMS	logOSMS	logOSMS
WOMratio	0.783	5.174***	5.66	1.900**
	(-1.326)	(-1.742)	(-6.929)	(-0.861)
Constant	-0.855***	-3.234***	-4.374***	-2.104***
	(-0.152)	(-0.531)	(-0.771)	(-0.413)
Observations	18	18	18	18
R-squared	0.021	0.355	0.04	0.233
2. Dependent variable (*logOSMS*); Independent variables (*WOMratio*) and lagged (*logOSMS*)				
OS	Android	BlackBerry	WP	iOS
Dependent variable	logOSMS	logOSMS	logOSMS	logOSMS
WOMratio	1.506	0.548	-4.484	1.597
	(-1.241)	(-1.101)	(-3.719)	(-1.022)
Lagged logOSMS	0.668***	0.837***	0.964***	0.185
	(-0.194)	(-0.129)	(-0.138)	(-0.241)
Constant	-0.413*	-0.475	0.449	-1.736**
	(-0.209)	(-0.49)	(-0.783)	(-0.632)
Observations	17	17	17	17
R-squared	0.498	0.836	0.786	0.219
3. Dependent variable (*logOSMS*); Independent variable normalized (*WOMratio*)				
OS	Android	BlackBerry	WP	iOS
Dependent variable	logOSMS	logOSMS	logOSMS	logOSMS
Normalized WOMratio	-0.441	-0.597***	-0.0958***	-0.0233
	(-0.62)	(-0.109)	(-0.0147)	(-0.155)
Constant	-0.658***	-0.734***	-3.041***	-1.150***
	(-0.151)	(-0.179)	(-0.158)	(-0.251)
Observations	17	17	17	17
R-squared	0.033	0.668	0.738	0.001
4. Dependent variable one period change in (*logOSMS*); Independent variable (*WOMratio*)				
OS	Android	BlackBerry	WP	iOS
Dependent variable	deltalogOSMS	deltalogOSMS	deltalogOSMS	deltalogOSMS
WOMratio	1.382	-0.276	-4.859	0.69
	(-1.317)	(-0.903)	(-3.326)	(-1.284)
Constant	-0.143	0.0418	0.628	-0.329
	(-0.146)	(-0.274)	(-0.372)	(-0.62)
Observations	17	17	17	17
R-squared	0.068	0.006	0.125	0.019
Standard errors in parentheses			*** p<0.01, ** p<0.05, * p<0.1	

Table 11. Fixed effects variable specification tests

Specification	1. Ratio	2. Lag	3.Normalized	4. Changes
Dependent variable	logOSMS	logOSMS	logOSMS	deltalogOSMS
WOMratio	3.634**	-0.573		-1.104
	(-1.128)	(-1.491)		(-1.468)
Lagged logOSMS		0.890***		
		(-0.0553)		
Normalized WOMratio			-0.0968***	
			(-0.00172)	
Constant	-2.758***	-0.0416	-1.586***	0.296
	(-0.282)	(-0.474)	(-0.00457)	(-0.367)
Observations	72	68	68	68
R-squared	0.057	0.739	0.666	0.018

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Appendix 5. Heckman selection equation estimates

Table 12. Heckman selection equation estimates

Variable	Coefficient	Standard error
gender	-0.219***	0.0105
developedmarket	-0.361***	0.0113
age25to39	-0.0814***	0.0129
age40to54	-0.171***	0.015
age55plus	-0.204***	0.0197
coreconsumer	0.0742***	0.0179
Constant	0.877***	0.0125
Model statistics		
Number of observations	67050	
Log likelihood	-39753.921	
Chi2	1743.93	
Prob > chi2	0	
Pseudo R2	0.0215	

*** p<0.01, ** p<0.05, * p<0.1

Appendix 6. Market share and word of mouth correlations

Table 13. Correlations between market shares and word of mouth

Correlations between (*logOSMS*) and (*WOMratio*) in each country for each operation system				
	Android	BlackBerry	WP	iOS
Australia	-0.368	0.702	0.198	-0.241
Brazil	0.164	0.509	0.063	-0.419
China	-0.088	0.371	-0.194	0.057
France	-0.037	0.659	0.347	-0.164
Germany	-0.056	0.284	-0.529	0.006
India	0.034	0.613	0.648	0.173
Indonesia	-0.398	0.007	0.414	-0.260
Italy	0.245	0.711	0.132	-0.251
Mexico	0.410	0.793	0.161	-0.269
Nigeria	0.125	0.208	0.006	0.435
Russia	-0.153	0.175	0.451	-0.061
Saudi Arabia	0.437	-0.061	-0.178	0.303
UK	0.146	0.596	0.200	0.483
USA	0.297	0.702	0.127	0.592

Appendix 7. Model specification tests

Table 14. Wald test results

Country	F-statistic	Prob > F
All	552.420	0.000
Australia	125.24	0.000
Brazil	182.15	0.000
China PRC	671.5	0.000
France	311.63	0.000
Germany	401.06	0.000
India	106.68	0.000
Indonesia	376.39	0.000
Italy	135.78	0.000
Mexico	79.27	0.000
Nigeria	139.11	0.000
Russia	850.71	0.000
Saudi Arabia	439.96	0.000
UK	117.6	0.000
USA	155.88	0.000

Table 15. Hausman test results

Country	Chi-squared	Prob>Chi-squared
Australia	0.09	0.764
Brazil	0.09	0.758
China	0.02	0.894
France	0.07	0.792
Germany	0.17	0.681
India	2.57	0.109
Indonesia	0.56	0.456
Italy	0.01	0.942
Mexico	0.05	0.828
Nigeria	1.81	0.179
Russia	0.12	0.726
Saudi Arabia	0.15	0.696
UK	0.01	0.911
USA	0.16	0.686

Table 16. Breusch Pagan Lagrange multiplier test results

Country	Chi-bar-squared	Prob>Chi-bar-squared
Australia	430.19	0.000
Brazil	472.31	0.000
China	570.76	0.000
France	526.74	0.000
Germany	539.06	0.000
India	326.94	0.000
Indonesia	524.87	0.000
Italy	442.49	0.000
Mexico	358.63	0.000
Nigeria	386.82	0.000
Russia	573.33	0.000
Saudi Arabia	543.75	0.000
UK	422.36	0.000
USA	452.29	0.000