Information and Financial Markets

Stefan Anchev
To my parents Marija and Gligor
Abstract

The results in this thesis are consistent with the hypotheses that: 1) the incomplete dissemination of information across investors helps in explaining the occurrence and the persistence of cross-sectional stock return anomalies, 2) the properties of the investor base of a stock have implications for the informativeness of the stock's price and 3) a greater quantity of firm disclosure places less sophisticated investors at an information disadvantage. Overall, the thesis provides new empirical evidence about the role of information in financial markets.

Investor Base and Stock Return Anomalies

After controlling for market capitalization, the predictability of future stock returns associated with each of the earnings-to-price ratio, the book-to-market ratio, the past return, the total volatility of returns and the return on assets is more pronounced among stocks with smaller total and/or institutional investor bases. These results appear even after controlling for several other stock characteristics and potential risk factors and they are both statistically and economically meaningful. Thus, they are consistent with the hypothesis that the incomplete dissemination of information across investors helps in explaining the occurrence and the persistence of cross-sectional stock return anomalies.

Investor Base and Stock Price Informativeness

The relative idiosyncratic volatility of future stock returns is: 1) negatively associated with the absolute size of the total and the institutional investor base, 2) positively associated with the institutional ownership, 3) negatively (positively) associated with the average stock portfolio size (Herfindahl index) of the investor base and 4) positively associated with the indirect (i.e., through nominees) ownership. These results appear after controlling for several other stock characteristics and they are both statistically and economically meaningful. Thus, they are consistent with the hypothesis that the properties of the investor base of a stock have implications for the informativeness of the stock’s price.

Individual Investors and Quantity of Firm Disclosure

When the amount of information disclosed by a firm is greater (or increases), the stock portfolio weights that individual investors allocate (through trading) to that firm’s stock are lower (or decrease) and suboptimal. The former result is less pronounced or nonexistent for more financially competent individuals and for positions in firms with a poorer information environment. When they do allocate greater portfolio weights to the stock of a firm that discloses more, individuals, regardless of their financial competence, earn lower returns. Overall, these results are consistent with the hypothesis that a greater quantity of firm disclosure places less sophisticated investors at an information disadvantage.
# Table of Contents

I. Economics of Information 1  
   A. Information as a Good 2  
   B. Information in Relation to the State of Agents 4  

II. Role of Information in Financial Markets 8  
   A. Informational Efficiency of Security Prices 8  
   B. Disclosure of Information by Firms 10  

III. Purpose and Hypotheses 17  
   A. Investor Base and Stock Return Anomalies 17  
   B. Investor Base and Stock Price Informativeness 18  
   C. Individual Investors and Quantity of Firm Disclosure 19  

IV. Data 20  
   A. Investor Bases of Stocks 20  
   B. Disclosure Scores of Firms 22  
   C. Stock Portfolios of Individual Investors 23  
   D. Other Characteristics of Stocks and Firms 23  
   E. Other Characteristics of Individual Investors 23  

V. Methodology 23  
   A. Portfolio Sorts 24  
   B. Fama and MacBeth (1973) Regression Models 25  
   C. Cross-Sectional Time-Series Regression Models with Fixed Effects 26  

VI. Conclusion 26  

References 30
Information and Financial Markets

STEFAN ANCHEV*

ABSTRACT

The results in this thesis are consistent with the hypotheses that: 1) the incomplete dissemination of information across investors helps in explaining the occurrence and the persistence of cross-sectional stock return anomalies, 2) the properties of the investor base of a stock have implications for the informativeness of the stock’s price and 3) a greater quantity of firm disclosure places less sophisticated investors at an information disadvantage. Overall, the thesis provides new empirical evidence about the role of information in financial markets.

When I was around 13 years old, my father gave me a book on one of the wealthiest families in the world, the Rothschild family. It was then when I first learned about the famous story of how Nathan Rothschild used his early knowledge of the outcome of the Battle of Waterloo in 1815 to earn a fortune. Namely, using his couriers to deliver the message from the battlefield, Nathan knew entire 48 hours before anyone else in London that the British and the Prussian armies had defeated the French army under the command of Napoléon Bonaparte. With this information at his disposal, it is said that Nathan immediately started buying British bonds and ultimately earned millions. Some aspects of this story, however, have been proven to be incorrect. Nevertheless, although I could not exactly articulate it at the time, the idea that a single piece of information can be so consequential was and still is fascinating to me. Later on, throughout my education, I learned that the role of information in financial markets and the economy in general is essential.

I. Economics of Information

What is information? As pointed out in Bates (1988 and 1990), this question has plagued researchers for decades. While the answer to it differs across disciplines, the focus in this thesis is on information from the perspective of the economists (who have laid out the foundations for the research on the role of information in financial markets in both finance and accounting). Even among only them, however, one of the few issues related to the economics of information on which a consensus has been reached is that information is a peculiar asset. Indeed, most prior efforts in this area have been directed towards the measurement of the value (see, e.g., Machlup (1962) and Porat (1978)) or the flow (see, e.g., Pool (1983)) of information, rather than towards its definition. Hence, as noted in Bates (1988 and 1990), given the absence of agreement with respect to what information is and how to measure it, its treatment in economics has, thus far, been predominantly theoretical.

In economic theories, information has been treated in two different ways (Bates (1990)). First, it has been treated as a good (i.e., the object of transactions between agents) and, therefore, as an object of the theories. Second, it

*Umeå School of Business, Economics and Statistics, Umeå University.
has been treated in relation to the state of agents (i.e., how informed they are) and, hence, as another feature of the theories. Both of these treatments are described below.

A. Information as a Good

Information has been treated as a good simply because it has some of the properties of a good. Indeed, it can have a utility (i.e., usefulness), a value, a production cost, an opportunity cost and it can be exchanged (Bates (1990)). Information, however, as noted in Shapiro and Varian (1998) and Varian (2001), has also several other, more peculiar properties. One such property is that information must be consumed before its value can be determined with certainty. Another one is that information typically has a high cost of production, but a very low cost of reproduction. Lastly, in its consumption, information is usually nonrivalrous and, sometimes, it is nonexcludable. Because of these properties, therefore, the treatment of information as a good has been problematic.

As pointed out in Bates (1988 and 1990), Arrow (1962) is probably among the first to note that the complete value of information cannot be determined with certainty before the information is purchased and used (i.e., information is a so-called experience good). Indeed, a piece of information cannot by evaluated in the same way as, for example, a chair (Stiglitz (2000)). Before purchasing and/or using a chair, it is possible to just look at it to reliably evaluate it. In contrast, a piece of information must be purchased even before it can be known what it is (i.e., information is unpredictable; otherwise, there is no need to pay for it) and, since the circumstances of its use can affect its value, it must be used before it can be evaluated with confidence (Arrow (1962)). With respect to this latter insight, the value of information can be influenced by its production and opportunity costs and by, for example, its certainty, diffusion, applicability, content and decision-relevance (Hirshleifer (1973)), as well as the structural and the political circumstances of its use (Black and Marchand (1982)). Thus, as argued in Hirshleifer (1971) and Marshall (1974), the value of information is likely to arise primarily from its future utility or usefulness (i.e., from its ability to allow agents to make decisions that yield higher future payoffs).

This conceptualization of the value of information, which is not much different from that of other goods, and the fact that the same information can be used multiple times under different circumstances suggest that its value can be variable and, hence, uncertain (Bates (1988)). Most traditional economic theories, however, have been of deterministic nature and they have not been able to easily cope with uncertainty.

It has been possible, however, to overcome these concerns by moving from a deterministic to a probabilistic theoretical framework of analysis (Bates (1988 and 1990)). In this regard, the developments in decision and probability theories have been of particular importance. A key concept in these theories is the expected value of a random variable, which is a weighted mean of all the possible values of that variable, where their respective probabilities of

---

1This subsection is largely based on Bates (1988 and 1990).
occurrence are used as weights. Thus, the expected value of a certain information can be expressed as a weighted mean of all the possible values of that information, where the probabilities of the occurrence of all the respective uses of the information are used as weights. With this approach, hence, as noted in Bates (1988 and 1990), it is possible to account for the potential variation in the value of the same information and to treat the value of information as a constant. This means that, by substituting its true value with its expected value, information can be more easily treated. Moreover, since the use of expected values in relation to information is similar to their use with respect to other goods, as suggested in Coase (1974), at least with respect to its value, information can be treated in the same way as any other good.

Unlike a good, however, information typically has a high fixed cost of production, which is also sunk, but a very low variable cost of reproduction, which tends towards zero (Shapiro and Varian (1998) and Varian (2001)). Indeed, whereas the reproduction of a good, such as a car, is quite costly since it requires considerable resources (e.g., labor), the reproduction of a piece of information, such as a stock price forecast, can be done by just memorizing or recording (e.g., on paper or electronically) and, thus, it is almost costless. From this, as pointed out in Shapiro and Varian (1998) and Varian (2001), it follows that the marginal cost of information also tends towards zero.

The marginal cost of information and one of the necessary conditions for the optimal allocation of resources in perfectly competitive markets, that the marginal cost of a certain good must be equal to the marginal benefit from the transaction of that good, further suggest that the marginal benefit from the transaction of information in such markets must tend towards zero as well (Bates (1990)). Under these conditions, hence, it has been thought that agents cannot recover their fixed costs of information production, which means that they do not have incentives to produce any information (Shapiro and Varian (1998) and Varian (2001)). As a result, information has frequently been treated as a good in the context of imperfectly competitive markets, such as monopolies (see, e.g., Admati and Pfleiderer (1986)).

Regardless of the type of market, as noted in Shapiro and Varian (1998) and Varian (2001), information is usually nonrivalrous, which means that its use by one agent does not reduce its availability for use by other agents, and, sometimes, it is nonexcludable, which means that one agent cannot exclude other agents from using it. These properties of information imply that its value to a certain agent can be determined not only by its use by that agent, but by its uses by other agents as well (Bates (1988)). Moreover, the properties also suggest that agents may not be able to fully appropriate the future payoffs from their information (Stiglitz (2000)). Furthermore, while public goods typically have both of these properties, depending on the legal regime (e.g., the use of laws for protecting intellectual property), information, as private goods, can be excludable. Hence, as pointed out in Bates (1988 and 1990), information has been treated as both a public good (see, e.g., Samuelson (1954 and 1958)) and a private good (see, e.g., Boulding (1966) and Marshall (1974)).
Almost without exceptions, the early treatments of information in relation to the state of agents have involved the assumption that it is perfect. According to this assumption, all agents obtain all available information instantaneously and at no cost. Economists, however, as pointed out in Stiglitz (2000), seem to have long been aware that this assumption may be problematic (see, e.g., Smith (1776), Mill (1848), Marshall (1920) and Weber (1925)). For example, Smith (1776) recognizes that the borrowers of highest quality may leave the market when lenders increase interest rates, which should not occur if lenders know the riskiness of each borrower and if each borrower is required to pay a premium on the basis of her riskiness. Another example is Marshall (1920), who realizes that the productivity of employees can increase if they are paid higher salaries, which implicitly shows his understanding that, perhaps due to the difficulties of observing the activities performed by employees, their salaries oftentimes do not correspond to their activities. Thus, while many earlier studies have made similar observations, they have not identified a problem related to information as one of the causes for them (Stiglitz (2000)). Indeed, even the mainstream economic theories have ignored the potential problems related to the assumption of perfect information and, instead, they have assumed that, as long as information is not too imperfect, such problems are less severe.

As Stiglitz (2000) points out, for example, in the highly influential competitive general equilibrium theory of Arrow and Debreu (1954) and McKenzie (1954), it is assumed that the contracts between agents are complete (in the sense that agents anticipate all the future states of the world) and that there are no impediments to their enforcement. Although other transaction costs have been suggested as an explanation (see, e.g., Williamson (1979)), it seems reasonable that the frequent incompleteness of agents’ contracts observed in the real world arises largely because, when contracting, they cannot always anticipate all the future states of the world (Stiglitz (2000)). Moreover, it also seems plausible that law enforcement agents (e.g., courts) cannot always observe the activities of other agents, which is the reason why they often find it difficult to determine if their contracts are enforced.

At the same time, there has been hope that the understanding of agents’ allocation of resources to the production of information would improve as the field of economics in general progresses and that, ultimately, the findings that emerge when information is assumed to be perfect, would also emerge when it is assumed to be imperfect (Stiglitz (2000)). It has, therefore, been assumed that, when making decisions, agents equate the marginal benefits to the marginal costs of information production. While Stigler (1961) does not examine the consequences for the equilibrium that occurs under this condition, he is the first to explicitly recognize that information is costly. This, however, as noted in Stiglitz (2000), does imply that, in equilibrium, there can be a price distribution (i.e., a violation of the law of one price). Stigler (1967) further

---

2This subsection is largely based on Stiglitz (2000).
argues that information costs, just as any other transaction costs, exist and that they can be treated in the same way as them. Moreover, he suggests that many of the seeming imperfections in financial markets can be explained by the existence of such costs. Modern economic theories, however, have shown that the recognition of even small information costs can have considerable consequences for many of the other assumptions that arise when information is assumed to be perfect.

For example, as pointed out in Stiglitz (2000), while it is considered that investors who own equity securities share the risks associated with those securities to a greater extent than investors who own debt securities, firms usually obtain considerably more financing by issuing the latter (Mayer (1990)). One potential explanation for this can be that, relative to investors, firms’ managers have information about the value of the securities of their firms that is of a higher quantity and/or quality. Their willingness to issue equity securities, hence, can be interpreted by investors as an information that, on average, those securities are overpriced (i.e., their prices are higher than their true values) and that their future prices will decrease (Greenwald, Stiglitz and Weiss (1984) and Myers and Majluf (1984)), which can be the reason why the incentives of firms for issuing such securities are weaker. Thus, firms can prefer to issue debt securities, which commonly increases the probability of their financial distress or bankruptcy (Stiglitz (2000)). Consequently, instead of being risk-neutral and considering only the net present values (NPVs) of the available projects when selecting among them as assumed in neoclassical economic theories (e.g., Modigliani and Miller (1958)), firms can be risk-averse and they can even forgo investing in projects with positive NPVs because of considerations for their financing.

Apart from having potential consequences for many of the other assumptions that arise when information is assumed to be perfect, the recognition of information costs can give rise to additional problems. Hayek (1945), for example, is the first to realize that (decentralized) prices contain information only about the scarcity of resources and that this information is sufficient for their optimal allocation. However, as Stiglitz (2000) suggests, this may not be true because, apart from information about the scarcity of resources, agents can consider information about the characteristics of resources and/or the behavior of other agents (e.g., employers are probably interested in knowing the strengths and the weaknesses of their employees and/or their productivity). If this information is costly, however, then there can be problems related to adverse selection and/or incentives (i.e., moral hazard).

Problems related to adverse selection can arise when, between the agents involved in the transaction of a resource, there are information asymmetries (i.e., differences in the quantity and/or quality of information) with respect to its characteristics. The relevance of such information asymmetries can be easily illustrated with the example of Akerlof (1970). Consider a market for used cars, in which some sellers sell good cars and some sellers sell bad cars (so-called lemons). In addition, assume that, while the sellers know the characteristics of their cars, the buyers do not have such information and, hence, they cannot and do not distinguish between the good and the bad cars.
Knowing this, therefore, the buyers' best guess for a given car can be that its characteristics and true value are those of the average car. This means that the buyers can average the prices of all cars (i.e., they can underprice the good cars and overprice the bad cars) and that they may not be willing to pay more than the average price for any car. As a result, the sellers of the good cars may not be willing to sell their cars and they can leave the market.

Akerlof (1970), however, does not consider the possibility that some agents (either the sellers or the buyers in the example above) can obtain and/or produce additional information. Stiglitz (2000) points out two streams of literature in which such a possibility has been considered.

The first stream focuses on agents' self-selection, which refers to the processes in which agents can reveal information about their own characteristics through the choices that they make (see, e.g., Spence (1974) and Rothschild and Stiglitz (1976)). It is important to note that, in this stream, agents do not necessarily need to know their own characteristics and that, instead, it is sufficient that agents with different characteristics can make different choices (Stiglitz (2000)). One of the major findings under these conditions has been related to the theory of the firm. For example, in the theory of Rothschild and Stiglitz (1976), a firm can maximize its profits, but conditional on the actions of the other firms.

In the second stream, agents can invest resources in obtaining and/or producing information (see, e.g., Arrow (1973) and Stiglitz (1975)). Their incentives to do so, however, are likely to be affected by one of the key issues in this stream (which is also mentioned in the prior subsection), that agents may not be able to fully appropriate the future payoffs from their information (Stiglitz (2000)). As a result, one of the most important findings under these conditions has been that agents who have more and/or better information are not necessarily better off than agents who do not have such information (i.e., the future payoffs of the former are not always higher).

Further, problems related to incentives can arise when, between the agents involved in the transaction of a resource, there are information asymmetries with respect to their behavior. The importance of such information asymmetries can be easily demonstrated with an example similar to that of Arrow (1964 and 1971). Consider a market with buyers and sellers of insurance policies for cars against the risk of theft. Additionally, assume that, while the buyers know all of their behavior with the cars, the sellers cannot and do not have such information and, thus, they cannot and do not stipulate it in the insurance policies. As a result, the buyers can behave in a manner that increases the risk of theft (e.g., they can leave their cars unlocked).

Considering these problems, Ross (1973), Stiglitz (1974) and Jensen and Meckling (1976) argue that such problems arise and are very pronounced in modern firms between their investors and managers. Stiglitz (1974) also proposes two ways in which agents can mitigate such problems. These involve monitoring the activities of agents and payment structures that strengthen their incentives. Although a large stream of literature has focused on the latter (see, e.g., Nalebuff and Stiglitz (1983)), as pointed out in Stiglitz (2000), the recognition of the problems related to incentives, the various ways in which
agents attempt to mitigate these problems and the advantages and the disadvantages of those ways have probably had the most important consequences for many of the findings that arise when information is assumed to be perfect.

Indeed, while Stigler (1961) analyzes how much information agents decide to obtain conditional on the price distribution of a good, Diamond (1971) demonstrates that, if agents have to pay even arbitrarily small information costs, then the good need not have a distribution of prices and, in equilibrium, it can have one price, that is higher than its marginal cost (i.e., a price that arises in imperfectly competitive markets, such as monopolies). Several subsequent papers (e.g., Salop and Stiglitz (1977 and 1982), Varian (1980) and Stiglitz (1985)), however, find that, even if agents have to pay information costs and if these costs differ across agents, then, in equilibrium, there can be a price distribution. As Stiglitz (2000) emphasizes, these results have shown that information costs can have significant consequences for the nature of the highly influential competitive general equilibrium of Arrow and Debreu (1954) and McKenzie (1954). In addition, he notes that the results reflect the understanding that all agents involved in the transaction of a resource can obtain and/or produce additional information, which means that it has been recognized that both the information that agents have and the information that they do not have can be endogenous (see, e.g., Salop (1977) and Edlin and Stiglitz (1995)).

In addition to these, information costs have been found to have major potential consequences for many other findings that arise when information is assumed to be perfect. As already mentioned, the cost of producing information is typically fixed, which is the reason why it has been established that its value increases with the scale of its production. Hence, the value of information can be nonconvex (Radner and Stiglitz (1984)). In contrast to the assumption in traditional economic theories that the preferences of agents are convex and continuous, however, this suggests that their preferences for information can be nonconvex and discontinuous (see, e.g., Radner (1968) and Arnott and Stiglitz (1988)). With such discontinuities in the behavior of agents, Rothschild and Stiglitz (1976) demonstrate that it is possible that an equilibrium in competitive markets does not even exist.

Apart from the examples of the potential consequences from the developments in the treatment of information in relation to the state of agents described above, there are many other examples. Indeed, as documented in Stiglitz (2000), these developments have had potential consequences for the fundamental theorems of welfare economics, the relevance of (nonmarket) institutions, the efficiency of markets and the relevance of the financing and the governance of firms. A review of all of these consequences, however, is beyond the scope of this thesis. Nevertheless, what is important to keep in mind is that the recognition of information costs and the problems related to adverse selection and/or incentives that arise from them have had a profound influence on the way we think about many different phenomena in the economy (Stiglitz (2000)).
II. Role of Information in Financial Markets

Although there is a very large literature about the role of information in financial markets, this thesis involves two of the most important streams of that literature, which are related to the informational efficiency of security prices and the disclosure of information by firms. Both of these streams of literature are described below.

A. Informational Efficiency of Security Prices

As pointed out in Stiglitz (2000), one of the greatest developments in the history of economic thought has been related to the informational efficiency of security prices. According to the Efficient Market Hypothesis (EMH), which is formalized in Fama (1970), the prices of securities always fully reflect all of the available information. This hypothesis is associated with another one, the so-called random walk hypothesis, according to which the changes in the prices of securities (i.e., their returns) are independent and random. The logic behind this idea is that, if there are no impediments to the dissemination of information across investors and over time and if the prices of securities always fully reflect all of the available information, then their future returns should reflect only future information and they should be independent of their current and/or past returns. Moreover, since information is by definition unpredictable (as explained in the prior section), security returns should also be unpredictable and, hence, random. One of the most important implications of this rationale is that the informationally efficient price of a certain security should imply an expected return on that security that reflects the risk associated with the security. Therefore, even if they use technical and/or fundamental analyses (i.e., analyses of the past prices and/or the fundamental characteristics of securities), investors in securities with such prices should not be able to earn returns on those securities which, after accounting for their exposure to risk, are abnormal.

Fama (1970) recognizes that the EMH is an extreme hypothesis (see also Grossman and Stiglitz (1980)) and, depending on the type of information involved, he further proposes three different forms of it. The so-called weak form of the EMH suggests that the prices of securities always fully reflect all of the publicly available information about the past prices of the securities. In addition to this information, according to the so-called semi-strong form of the EMH, the prices of securities always fully reflect all of the other publicly available information (e.g., earnings and/or interest rates). That the prices of securities always fully reflect all of the publicly and privately available information is indicated by the so-called strong form of the EMH. Note that, while the semi-strong form encompasses the weak form of the EMH, both of these forms are encompassed by the strong form of the EMH. It is important to note, however, that, regardless of the form of the EMH, this hypothesis in itself “is so general that it has no empirically testable implications” (Fama (1970, p. 384)).

Indeed, as pointed out in Fama (1970), when testing the EMH, one must
take a stance with respect to the relation between the expected returns on securities and their risks. That is, the EMH must be tested in the context of an equilibrium model of expected returns. This, however, leads to the so-called joint hypothesis problem, which means that, if the result from a test of the EMH suggests that it should be rejected, then it is difficult, if not impossible, to determine whether the result is such because the hypothesis is truly incorrect or because the assumed equilibrium model of expected returns is incorrect (Campbell, Lo and MacKinlay (1997)). Nevertheless, in this regard, the Capital Asset Pricing Model (CAPM) of Sharpe (1964), Lintner (1965) and Mossin (1966) has been of a particular importance.

The CAPM is built upon the normative theory of Markowitz (1952 and 1959), who shows how investors should form portfolios of securities in order to achieve an optimal performance in terms of risk and return. In his theory, investors are assumed to be risk averse and, when selecting among the available portfolios, they consider only the means and the variances of the returns on the portfolios over a single holding period. As a result, investors form portfolios with a maximum expected return for a given variance of returns or portfolios with a minimum variance of returns for a given expected return. Portfolios formed in this way are said to be mean-variance efficient.

In the CAPM, the portfolio optimization principles in Markowitz (1952 and 1959) are combined with several additional assumptions. First, all securities are infinitely divisible and investors do not face transaction costs, taxes, impediments to taking short positions and impediments to borrowing or lending at a riskless rate. Second, all investors obtain all available information instantaneously and at no cost. Third, an identical set of securities is available to all investors. Finally, all investors agree on the parameters of the true return distribution of each security (i.e., they have homogeneous beliefs with respect to the true means, variances and covariances of the returns on all securities). Under these conditions, in the CAPM, the market portfolio is a weighted average of investors’ portfolios and, therefore, it is also mean-variance efficient. This implies that differences across securities’ expected returns should be completely explained by differences in their exposure to systematic risk (i.e., their market betas).

Despite this prediction and its wide use by both academic and nonacademic researchers, approximately 450 other variables (so-called anomalies) have been discovered that seem to have the ability to explain or predict the future returns on stocks.³ The literature, however, has, so far, provided just a few incomplete explanations for these anomalies that are unrelated to the dissemination of information across investors.

The seeming predictability of the future returns on stocks has also raised questions about the informativeness of their prices (see, e.g., Malkiel (2003)). However, while even the psychological biases of investors have been suggested as explanations for it (see, e.g., Thaler (2005)), the identification of the determinants of the informativeness of stock prices has, thus far, also been incomplete.

B. Disclosure of Information by Firms

In financial markets, firms attempt to obtain financing for their ideas and investors are willing to provide such financing in the expectation of earning positive future returns on their investments. Although both firms and investors have incentives to collaborate, the optimal allocation of resources between them can often be hindered by problems that arise when information is costly, such as the previously described problems related to adverse selection and/or incentives.

Problems related to adverse selection arise because, relative to investors, firms’ managers typically have information about the values of the securities of their firms that is of a higher quantity and/or quality. Such information asymmetries usually limit the ability of investors to accurately estimate the values of firms’ securities. This problem is further aggravated by the incentives of managers to exaggerate the values of the securities of their firms in order to attract more financing. Therefore, if investors cannot distinguish between securities of higher and lower values, they may underprice the securities with higher values and overprice the securities with lower values. Ultimately, the problems related to adverse selection can lead to a market failure (Akerlof (1970)).

Further, problems related to incentives arise because, apart from the information asymmetries between the managers of firms and their investors, the ownership and the control of firms are typically separated (Berle and Means (1933)). In particular, once investors provide financing to a certain firm, they are usually not involved in the management of that firm. Managers, however, can have incentives to make decisions that are not always in the best interest of investors (Jensen and Meckling (1976)). For example, they can pay themselves an excessive compensation or they can issue more senior debt securities.

As noted in Healy and Palepu (2001) and Beyer et al. (2010), the combined effects of the problems related to adverse selection and incentives give rise to investors’ demand for the disclosure of information by firms (hereafter, firm disclosure) and to the institutions that facilitate it. Firms disclose information to investors through mandatory disclosures, such as annual reports, which contain, for example, management discussions and analyses, financial reports and footnotes to these reports. In addition to the information in these disclosures, many firms frequently disclose other information to investors voluntarily, through, for instance, conference calls, press releases and/or their websites. Lastly, third parties, such as information intermediaries (e.g., financial analysts and/or the financial media) can provide investors with important information about firms. Despite the efforts of these intermediaries and the recent technological advancements, however, the most important sources of information for investors, as pointed out in Merton (1987) and Beyer et al. (2010), are likely to remain to be the firms themselves.

While many of the decisions with respect to the policies of information disclosure of firms lie within the domain of their managers’ responsibilities, the problems related to adverse selection and incentives have led researchers
to recognize that the information environment of a firm is determined endogenously (see, e.g., Dye (2001), Verrecchia (2001) and Beyer et al. (2010)). Considering this, the so-called unraveling result of Grossman and Hart (1980), Grossman (1981), Milgrom (1981) and Milgrom and Roberts (1986) suggests that if: 1) firms can disclose their information without incurring any costs, 2) firms can disclose their information in a credible way, 3) firms cannot commit to a certain information disclosure policy, 4) the objective of the managers of firms is to maximize the prices of their securities, 5) investors know that firms have information and 6) all investors interpret the disclosure or the nondisclosure of any information by firms in the same manner and firms know that manner, then firms have incentives to always voluntarily disclose all of their information. The reason for this result is that investors are assumed to interpret any undisclosed information by firms as being unfavorable with respect to the values of their securities.

Although the unraveling result has been regarded as a seminal work that has inspired a substantial amount of literature, due to violations of its premises, it has failed in explaining the observed behavior of firms in the real world in relation to their disclosure of information.\(^4\) Moreover, while most of the theoretical and the empirical literature on firm disclosure has maintained that firms can benefit from a greater disclosure in the form of, for example, an improvement in the liquidity of their securities and/or a reduction in their cost of capital, many of them do not voluntarily disclose all of their information (Fields, Lys and Vincent (2001) and Beyer et al. (2010)).

Therefore, the existence of firms with different amounts of voluntarily disclosed information has offered conditions where the regulation of firm disclosure can be justified (Beyer et al. (2010)). It is important to note that, in contrast to the theories on voluntary firm disclosure, currently, there is no comprehensive normative theory or paradigm on mandatory firm disclosure (Dye (2001) and Verrecchia (2001)). Instead, researchers have, thus far, focused on the potential rationales for it.\(^5\)

As described in Beyer et al. (2010), four such rationales have, so far, been identified. First, when disclosing information about their own financial circumstances, firms can also reveal information about the financial circumstances of other firms. This kind of externalities can lead to policies of information disclosure that are not socially optimal and, hence, mandatory firm disclosure is expected to improve social welfare (Dye (1990) and Admati and Pfleiderer (2000)). Second, one firm’s disclosure of information may indirectly affect the decisions of other firms. For example, the macroeconomic or revenue expectations of a firm can influence the strategic or operational decisions of other firms, such as their decisions to enter a new market (Pae (2000)) or to develop a new product (Hughes, Kao and Williams (2002)). Thus, by making firm disclosure mandatory, more informed decision-making by firms can be achieved. Third, although Beyer et al. (2010) cast doubt on the claims that the previously described problems related to incentives can lead to welfare

\(^4\)For a more extensive discussion of the lack of empirical support for the unraveling result, see Dye (2001) and Beyer et al. (2010).

losses, some researchers argue that mandatory firm disclosure can enable more efficient contracting between agents (see, e.g., Holmström (1979)). Lastly, as suggested in Dye and Sunder (2001), a common regulation of firm disclosure may reduce investors' costs of information production through an enhanced comparability of firms.

In contrast to the potential benefits of mandatory firm disclosure, there is almost an equal amount of literature that offers arguments against the regulation of firm disclosure. For example, mandatory firm disclosure can reduce the extent to which the risks associated with securities are shared among investors (Hirshleifer (1971), Diamond (1985) and Dye (1990)) or it can reduce the informativeness of their prices (Diamond (1985), Fischer and Stocken (2010) and Guttman (2010)). Hence, an agreement among researchers with respect to the need for regulation of firm disclosure has, thus far, not been reached. Nevertheless, in most financial markets around the world, the disclosure of information by firms has been highly regulated.

For example, in the European Union (EU), since 2005, firms have been required to prepare their financial reports according to the International Financial Reporting Standards (IFRSs), which are developed by the International Accounting Standards Board (IASB). According to the IASB’s *Conceptual Framework*:

“The objective of general purpose financial reporting is to provide financial information about the reporting entity that is useful to existing and potential investors, lenders and other creditors in making decisions about providing resources to the entity. Those decisions involve buying, selling or holding equity and debt instruments, and providing or settling loans and other forms of credit” (IASB (2013a, p. 195)).

This statement represents the overarching objective of financial reporting from the perspective of the IASB. It is the foundation for the remaining parts in the *Conceptual Framework* and it represents the starting point in the IASB’s process of developing accounting standards. For these reasons, it is expressed in quite broad terms and it is not specific to any particular context. Hence, a few elements of it need further elaboration.

The IASB argues that the financial reports of firms should provide information about their financial positions and information about the events that change those positions. While investors can have other information needs, both of these types of information should be useful for them when they make decisions about investing in securities. The IASB, however, recognizes that firms’ financial reports do not and cannot provide all of the information that existing and potential investors may need. Additionally, it is acknowledged that the information disclosed in the financial reports of firms is not designed to show the values of their securities, but rather to enable investors to estimate

---

6Similarly, in the United States (US), the Generally Accepted Accounting Principles (GAAPs), which are designed by the Financial Accounting Standards Board (FASB), govern mandatory firm disclosure.

7The objective of financial reporting defined by the FASB is very similar.
them.

Further, since the objective of providing information that is useful for making investment decisions is quite broad, the IASB provides additional guidelines as to what constitutes useful information. In particular, to be useful, information must be both relevant and faithfully represented (i.e., reliable).\(^8\)

According to the IASB, information is assumed to be relevant if it can affect investors’ decisions about investing in securities. Moreover, information can make a difference in the decisions made by investors even if some of them choose not to act on it or if they are already aware of it from other sources. To be relevant, information should also have predictive value, confirmatory value or both. If information can help investors in predicting future outcomes, it is said that it has predictive value. In contrast, information is said to have confirmatory value if it can be used to evaluate the accuracy of prior predictions. A closely related concept to the relevance of information is the concept of its materiality, which is applied when deciding about the level of detail of the information to be disclosed. The materiality of information, however, is seen as a firm-specific aspect of its relevance and it refers to the relevance of a piece of information in the context of a firm’s financial report.

The IASB defines faithfully represented information as information that is complete, neutral and free from error. Complete information regarding a certain economic phenomenon should encompass all of the necessary aspects of that phenomenon that would enable investors to understand it. For example, when disclosing information about a particular group of assets owned by a firm, the disclosure can include a description of the nature of those assets, a numerical representation of the assets and its meaning, the method of calculating the numerical depiction and a description of the current and the future factors and circumstances that may have an effect on the assets. Moreover, information neutrality means absence of bias when deciding on the information to be disclosed or the presentation format to be used in the disclosure. Finally, information is considered to be free from error if no errors are made in the presentation of an economic phenomenon, the selection of the process used to produce the information or its application.

Further, investors represent one of the largest, if not the largest, group of users of firm disclosures. Indeed, the IASB explicitly recognizes that existing and potential investors are some of the primary users to whom the financial reports of firms are directed.\(^9\) This category of users is assumed to range from individuals with relatively limited resources (e.g., technology and/or cognitive ability) to more resourceful institutions, such as mutual and/or hedge funds. As already discussed, all of these investors may face problems related to adverse selection and/or incentives. As a result, they are likely to be interested in information that can help them to more accurately estimate the values of securities in order to make more informed investment decisions.

During the last five decades, numerous so-called value relevance studies

\(^8\)Comparability, verifiability, timeliness and understandability are the qualitative characteristics that also enhance the usefulness of information.

\(^9\)Other users, such as firms’ managers and/or regulators, may also find firms’ financial reports useful; however, these reports are not primarily directed to those users.
have been conducted with the aim of determining if and to what extent have the IASB’s and the FASB’s objectives been achieved, making this probably the most popular question in accounting research.\textsuperscript{10} Most of this literature has focused on the role of firm disclosure in financial markets, where researchers adopt the perspective of investors. Although the intended aim of the literature has been to evaluate the usefulness of accounting information for investors when they estimate the values of securities, Barth, Beaver and Landsman (2001) argue that, because the accounting standard-setting boards do not specifically stipulate how much relevance and/or faithful representation is sufficient for certain information to be useful, the usefulness of information is not a properly defined concept in accounting research. The usefulness of accounting information, therefore, cannot be evaluated from the literature on its value relevance. Instead, as pointed out in Barth, Beaver and Landsman (2001), this literature is designed to evaluate if a particular accounting item reflects information that is used by investors when they estimate the values of securities.

According to these authors, moreover, empirical tests of the value relevance of accounting information are joint tests of its relevance and faithful representation, where the information is considered to be value relevant if it has a predicted association with security prices and/or returns.\textsuperscript{11} This, however, has not been a stated objective of the accounting standard-setting boards. As a result, Barth, Beaver and Landsman (2001) contend that tests of the value relevance of accounting information represent only one approach to empirically operationalize the accounting standard-setting boards’ stated criteria of the relevance and the faithful representation of information.

In relation to this, Francis and Schipper (1999) discuss four possible operationalizations of the construct value relevance of accounting information.\textsuperscript{12} According to their first operationalization, the value relevance of accounting information is measured as the return from an investment strategy based on that information (see, e.g., Fama and French (1992), Lakonishok, Shleifer and Vishny (1994), Chan, Jegadeesh and Lakonishok (1996), Sloan (1996) and Bernard, Thomas and Wahlen (1997)). The second operationalization stipulates that accounting information is value relevant if it can be used in models for estimating the values of securities or if it is helpful in predicting the variables that are used in such models (see, e.g., Ou and Penman (1989) and Lev and Sougiannis (1996)). The vast majority of the literature studying the value relevance of accounting information has used the third and the fourth operationalization described in Francis and Schipper (1999). The studies that use the former typically examine the returns on securities over a shorter period of time surrounding the dissemination of a certain piece of accounting information by firms themselves, where statistically significant returns are interpreted as an

\textsuperscript{10}For a more detailed review of this literature, see Barth, Beaver and Landsman (2001), Healy and Palepu (2001), Holthausen and Watts (2001), Kothari (2001) and Dumontier and Raffournier (2002).

\textsuperscript{11}The definitions in Ohlson (1999), Barth (2000) and Holthausen and Watts (2001) are similar.

\textsuperscript{12}A similar categorization is provided in Holthausen and Watts (2001) and Dumontier and Raffournier (2002).
indication that investors, indeed, use that information when they estimate the values of those securities (see, e.g., Ball and Brown (1968) and Beaver (1968)). In these so-called information content studies, therefore, the value relevance of accounting information is determined by its capacity to alter the overall quantity and/or quality of information among investors. Finally, the idea behind the fourth operationalization is to investigate the association between certain accounting information and the returns on securities over longer time periods, where the information is considered to be value relevant if it can affect the security returns (see, e.g., Collins, Maydew and Weiss (1997), Francis and Schipper (1999) and Lev and Zarowin (1999)). A distinct characteristic of this approach is the assumption that the source of information does not have a significant role in the relation between the variables of interest. Therefore, unlike the third operationalization, this operationalization does not require the financial reports of firms to be the earliest source of information.

As can be seen, a common feature of the literature on the value relevance of accounting information is the use of aggregated variables, such as security prices and/or returns, to examine the use of such information by investors. As noted in Holthausen and Watts (2001), however, the prices of and/or the returns on securities reflect the aggregation of their estimated values by investors and the information that they use to estimate those values. Since these variables have been widely used in the studies on the value relevance of accounting information, as pointed out in Dumontier and Raffournier (2002), these studies represent an indirect attempt to determine if such information is used by investors when they estimate the values of securities. In fact, as already mentioned, while the literature that uses the third operationalization described above makes an assumption with respect to investors’ use of firm disclosures, the literature that uses the fourth operationalization does not even attempt to do so. Perhaps the most explicit concern about the use of aggregated variables in this context was made long before studies on the value relevance of accounting information were widely conducted, which can be seen in Beaver (1972) who points out that:

“... it is important to distinguish between the securities market and the individual investors that compose the market, because the role of (accounting) information can be vastly different in each context. To a certain extent, the distinction is artificial, in the sense that the aggregate actions of the individuals determine market behavior. However, the process of aggregation is often deceptive, and if we fail to make the distinction, we may be subject to any one of a number of fallacies of composition. In many cases, what is “true” for the group as a whole is not “true” for any individual of that group, and conversely” (Beaver (1972, p. 408)).

Using a similar reasoning, more recently, Holthausen and Watts (2001) point out that regulators and the accounting standard-setting boards are concerned with knowing whether and how different groups of investors or even individual investors use firm disclosures. These concerns stem from the fact that investors differ in terms of their resources and many investors are not in a
position to require firms to disclose information directly to them. This, in turn, is likely to lead to different costs of obtaining and/or producing information across investors, which, as already discussed, can give rise to problems related to adverse selection and/or incentives. As argued in Holthausen and Watts (2001), therefore, since regulators and the accounting standard-setting boards do not prefer such conditions, it seems reasonable to assume that they are not particularly interested in the use of firm disclosures by investors as a group.

Barth, Beaver and Landsman (2001), however, do not agree with these claims and they argue that nowhere in the Conceptual Framework of the IASB and the FASB is written that they are interested in the use of firm disclosures by different groups of investors or individual investors. Rather, as further argued, the attention of these organizations is directed towards investors as a group of users of firm disclosures.

Almost a decade after this disagreement, regulators and the accounting standard-setting boards in the EU and the US have, in fact, expressed an interest in knowing more about the use of firm disclosures by different groups of investors. For example, the IASB, the FASB and the Securities and Exchange Commission (SEC) have sought suggestions from both individual and institutional investors for the improvement of the disclosure of information by firms and for the development of new and existing accounting standards (see, e.g., SEC (2007), IASB (2010) and FASB (2012)).

In the EU, the IASB is currently undertaking a long-term project to revise its Conceptual Framework with one of the most important areas under revision being firm disclosure. In relation to this, in 2012, the IASB and the IFRS Advisory Council discussed the feasibility of a short-term project to improve the existing requirements for firms related to their disclosure of information and, since most prior efforts in this area had resulted in more such requirements, the IASB was strongly advised against such a project. In 2013, therefore, the IASB hosted a public forum with various users of the information disclosed by firms, its preparers, auditors, regulators and members of the accounting standard-setting boards to better understand the problems related to the existing requirements for firm disclosure.

As summarized in IASB (2013b), from the users’ perspective, the following conclusions were made. The inappropriate application of the concept of materiality and an inflexible financial reporting framework (with a focus on compliance instead of communication) lead to an excessive amount of disclosed information. To prevent this, the participants in the forum recommended that: 1) it needs to be clear who the primary users of firm disclosures are, 2) those users need to use professional judgement when demanding information (i.e., they should ask for more relevant and material information instead of just more information) and 3) the key concepts of the financial reporting framework should be more firm-specific. It is worth noting that an excessive amount of disclosed information was also one of the main problems voiced by its preparers. More specifically, they claimed that firms disclose more information than is necessary because, for example, their cost of not disclosing information has increased over time.

In the US, during 2012, the FASB initiated a project with the aim of im-
proving the effectiveness of the disclosure of information in the notes to the financial reports of firms. The FASB believes that these notes should clearly communicate the most important information, which should reduce the overall amount of disclosed information (FASB (2012)). In relation to this, even more explicit concerns have been expressed by the Chair of the SEC:

“When disclosure gets to be “too much” or strays from its core purpose, it could lead to what some have called “information overload” - a phenomenon in which ever-increasing amounts of disclosure make it difficult for an investor to wade through the volume of information she receives to ferret out the information that is most relevant” (SEC (2013)).

Considering these recent developments and the ability of the previously discussed aggregated variables, such as security prices and/or returns, to capture the use of firm disclosures only by investors as a group, during the last decade, researchers have focused on the use of firm disclosures by different groups of investors, and particularly by the group of individual investors, who are typically presumed to be less sophisticated than institutional investors (see, e.g., Barber and Odean (2008)). Nevertheless, many interesting questions of interest to, for example, researchers, investors, regulators and the accounting standard-setting boards are yet to be answered. As can be seen, one such question that has recently been raised is whether and how the amount of information disclosed by firms (i.e., the quantity of firm disclosure) affects less sophisticated investors, such as individual investors.

III. Purpose and Hypotheses

The purpose of this thesis is to provide new empirical evidence about the role of information in financial markets (i.e., in determining the characteristics and the behavior of security prices and investors). For this purpose, the thesis consists of three different papers, which involve three different hypotheses that arise from previously developed economic theories. These theories and hypotheses are presented below.

A. Investor Base and Stock Return Anomalies

The first paper, which is titled Investor Base and Stock Return Anomalies (see Anchev (2018a)), involves the theory of Merton (1987). This theory is motivated primarily by the empirical observation, such as that in Friend and Blume (1975) and Blume and Friend (1978 and 1986), that both individual and institutional investors own portfolios with only a fraction of the available securities. While Merton (1987) recognizes that these findings can be attributed to other factors as well (e.g., transactions costs, taxes or impediments to taking short positions), he also argues that they can be related to information costs. For instance, if private information about a particular security can be produced at a cost, then information asymmetries can arise between the investors who pay that cost and the investors who cannot afford to do so, which can cause
the latter to take neither long nor short positions in the security. As further argued in Merton (1987), however, while information costs of this type can be relevant, there is another cost that logically precedes them. Namely, before they can pay any other information costs, investors must first pay the cost of learning that the security exists.

Thus, Merton (1987) develops a version of the CAPM with the distinctive assumption that all investors must incur a fixed cost to learn that a particular security exists and that most of them pay this cost for only a subset of the available securities. It is important to note that, in Merton’s (1987) model, it is possible for some investors to know about all the available securities. When this is true for all investors, however, the model is equivalent to the CAPM. Further, although the subsets of the available securities differ across investors, as in the CAPM, all investors who know about a certain security know and agree on the true parameters of its return distribution (i.e., they have conditional homogeneous beliefs), which is the reason why trading in each security takes place only between equally informed investors. In this setting, hence, the majority of investors form their optimal portfolios using only the subset of the available securities in their own information sets, which is the reason why these portfolios are not mean-variance efficient with respect to all of the available securities.

Indeed, Merton (1987) demonstrates that, for the markets for securities known to a smaller number of investors to clear, the investors’ portfolios that contain these securities must be considerably overweight in them (i.e., the weights on these securities in the investors’ portfolios must be much greater than their weights in the market portfolio, which are also the optimal weights in the CAPM). In his economy, thus, the risk-return performances of investors’ portfolios are particularly suboptimal among portfolios with less-known securities.

Since the market portfolio is a weighted average of investors’ portfolios, in Merton’s (1987) equilibrium, it is also not mean-variance efficient with respect to all of the available securities and its suboptimal risk-return performance is also driven by less-known securities. This, in contrast to the CAPM, implies that differences across securities’ expected returns from differences in their exposure to factors other than systematic risk should arise and persist predominantly among securities known to a smaller number of investors, as predicted by Merton (1987). In his model, because all investors who know about a certain security take long positions in that security, securities known to a smaller number of investors are securities with smaller investor bases. One hypothesis that arises from the theory of Merton (1987), therefore, is that:

*The incomplete dissemination of information across investors helps in explaining the occurrence and the persistence of cross-sectional stock return anomalies.*

**B. Investor Base and Stock Price Informativeness**

The second paper, which is titled *Investor Base and Stock Price Informativeness* (see Anchev (2018b)), involves the theory of Peress (2010), which
combines those of Merton (1987) and Verrecchia (1982). In this theory, the informativeness of the price of a given stock is defined in terms of the amount of information contained in the price about the future payoffs of that stock and it is affected by the size and/or the composition of the stock’s investor base via several different channels whose relative importance determines the net effect. First, a larger investor base deteriorates the informativeness of a stock price because a tradeoff arises between the extent of risk sharing among investors and their private information production. Indeed, when the number of investors is greater, risk is better shared because each investor, on average, owns a smaller fraction of the outstanding shares. However, because this lower scale of investment decreases the potential benefits from private information, each investor’s incentives for investing resources in researching stocks are reduced and, thus, such information is produced less. Second, by contributing with potentially more informed opinions, an investor base in which investors have such opinions improves the informativeness of a stock price. Lastly, the informativeness of a stock price is directly and indirectly influenced by the average risk tolerance and/or the average propensity for liquidity trading of the investor base. On average, more risk tolerant investors bear more risk. While this directly increases their incentives for producing private information, it indirectly makes less risk tolerant investors disproportionately less interested in producing such information. Further, the presence of investors who are more prone to trade for liquidity reasons directly decreases the amount of private information contained in the price of a stock. Indirectly, however, their presence motivates the production of private information by other investors because, with a noisier stock price, such information is easier to conceal when trading (see also Grossman and Stiglitz (1980) and Verrecchia (1982)). Overall, when the net effect of all these mechanisms is positive, stock prices should be more informative. One hypothesis that arises from the theory of Peress (2010), thus, is that:

\[
\text{The properties of the investor base of a stock have implications for the informativeness of the stock's price.}
\]

C. Individual Investors and Quantity of Firm Disclosure

The third paper, which is titled Individual Investors and Quantity of Firm Disclosure (see Anchev (2018c)), involves the theories of Indjejikian (1991) and Kim and Verrecchia (1994). In contrast to the well-known argument that the publicly disclosed information by firms mitigates information asymmetries and generally makes investors better off (see, e.g., Diamond (1985), Amihud and Mendelson (1986 and 2000), Merton (1987), Diamond and Verrecchia (1991) and Easley and O’Hara (2004)),\(^{13}\) these theories suggest that such information can also increase the information asymmetries between investors who differ in terms of their sophistication. Kim and Verrecchia (1994), for example,

\(^{13}\)As pointed out in Leuz and Verrecchia (2000), the information asymmetries in this literature can arise either between firms and their investors or across investors.
demonstrate that, by offering more opportunities for private information production, more firm disclosure gives rise to information asymmetries between the market participants who exploit these opportunities and those who are less able to bear the cost of doing so (a similar logic can be found in Indjejikian (1991)). Because of greater firm disclosure, thus, less able investors can face more severe adverse selection problems that arise from trading with better-informed parties, which, as shown in Glosten and Milgrom (1985) and Kyle (1985), can lead to economic losses. This means that more firm disclosure can be of a differing value to investors, with less sophisticated investors, such as individual investors, benefiting from it relatively less than sophisticated ones. One hypothesis that arises from the theories of Indjejikian (1991) and Kim and Verrecchia (1994), hence, is that:

A greater quantity of firm disclosure places less sophisticated investors at an information disadvantage.

IV. Data

Each of the three papers in this thesis involves rarely available and highly detailed data from several different sources on the characteristics of stocks, firms or individual investors from Sweden during the period from the end of 1999 to the end of 2008. These data are described below.

A. Investor Bases of Stocks

In Anchev (2018a and 2018b), the data on the investor bases of stocks are from the Swedish Central Securities Depository (SCSD). As described in Eckbo, Paone and Urheim (2011), apart from the registration and the safekeeping of all equity and debt securities issued in a dematerialized form (i.e., securities that are recorded electronically in a book-entry system) in Sweden, the SCSD provides clearing and settlement of the transactions in these securities, as well as, a variety of additional services to their issuers and owners (e.g., the execution of dividend payments). As the only organization authorized to perform these functions in Sweden, the SCSD operates in accordance with the Swedish Financial Securities Act (1998:1479) and the Swedish Securities Market Act (2007:528), under the supervision of the Swedish Financial Supervisory Authority (SFSA) and the Swedish Central Bank (SCB).

As stipulated in the Swedish Companies Act (2005:551), when private or public firms issue dematerialized equity or debt securities in Sweden, they must register the complete issues of those securities with only one authorized

---

14 Additionally, Kim and Verrecchia (1991) show that, when public information is more precise, informed investors increase the precision of their private information more than uninformed investors. Lundholm (1988), moreover, demonstrates that public and private information signals with sufficiently correlated errors can complement each other. Another study suggesting a positive relation between public firm disclosure and investors’ private information production is that of McNichols and Trueman (1994), according to which informed investors with a shorter investment horizon produce more private information in the expectation of public firm disclosure or an increase in its precision.
central securities depository. Firms can do so, either directly or through agents, by first applying for an affiliation with the SCSD. After successfully completing a review process with respect to their legal, organizational and financial circumstances, firms enter into an affiliation agreement with the SCSD, which then, in its electronic system, opens accounts for the firms that show the total amounts of the issued securities (e.g., the total number of outstanding shares for a stock issue). In the final stage of the issuing process, the SCSD transfers the securities from the accounts of the firms to the accounts of the investors in the securities.

As described in Eckbo, Paone and Urheim (2011), investors in a stock issued in Sweden can keep their shares of that stock in accounts at the SCSD opened either in their own names or in the names of their nominees (e.g., brokers). If investors keep their shares in their own accounts, the investors’ data (i.e., their names, identification numbers, postal addresses, bank accounts and the number of shares and voting rights owned) are registered in the electronic system of the SCSD. When investors choose to keep their shares in their nominees’ accounts, however, it is the data on the nominees, not the investors, that are registered in the SCSD’s electronic system, supplemented with a note that the shares are kept on behalf of a third party. Thus, rather than showing the number of shares owned by each of the nominee’s clients, a nominee’s account shows the total number of shares owned by all of the nominee’s clients. Nevertheless, as required by the SCSD, nominees must, in their own electronic systems, continually register their clients’ data (i.e., the same data as that registered in the investors’ own accounts) and they must periodically provide these data to the SCSD.

For all stocks registered in its electronic system, the SCSD uses the data from the investors’ and the nominees’ accounts to continually maintain registers of investors (Eckbo, Paone and Urheim (2011)). Since these registers do not include data on the nominees’ clients, at the end of each calendar quarter and before the annual general meetings of firms (or at their request), the SCSD obtains these data from the nominees and it produces expanded registers of investors. These registers are the source of the data on the investor bases of stocks used in Anchev (2018a and 2018b).

Namely, my empirical strategies involve data on the investor bases of all common stocks issued in Sweden (i.e., stocks with ISINs that begin with the ISO country code for Sweden “SE”) which were traded on either the Stockholm

---

15 At the same time, the SCSD assigns unique International Securities Identification Numbers (ISINs) to the securities. The ISINs represent 12-character alphanumeric codes which are constructed according to the International Standard 6166 of the International Organization for Standardization (ISO) and which consist of 2-letter codes for the countries where the issuing firms are legally registered or have their headquarters, 9-digit basic numbers and 1-digit control numbers.

16 To open accounts at the SCSD, investors and nominees must appoint authorized account operators who act as intermediaries between them and the SCSD; of course, this is not necessary when the investors and the nominees themselves are account operators. Only institutions, such as financial institutions (e.g., banks), that fulfil certain organizational, legal, technical and capital requirements can be authorized as account operators by the SCSD. As such, they have the right to open accounts for investors and nominees in the SCSD’s electronic system and to register new or amended data in those accounts.
Stock Exchange (SSE), the Nordic Growth Market (NGM) or Aktietorget at any point in time during the period from the end of 1999 to the end of 2008. At the end of each second and fourth calendar quarter during this period, for a given stock, the data set contains all investors who own the shares of that stock and the number of shares each of them owns. The precision of the data allows me to not only observe the investors and their long positions in a particular firm, but also in the specific stock of that firm. This is an important advantage of these data since many firms with common stock issued in Sweden have multiple classes of it which typically differ in terms of their voting rights. I can also observe whether the investors are individuals or institutions, Swedish tax payers or not and whether they keep their shares in their own accounts or in the accounts of their nominees. Finally, while the data set contains the postal codes of the investors with a registered address in Sweden, I can observe only the countries of those investors without such an address.

B. Disclosure Scores of Firms

In Anchev (2018c), the disclosure scores of firms are obtained from the “Best Annual Report” competition (a competition for the most informative annual report) organized by the Swedish Stockholders’ Association (SSA) since the early 1990s (except in 2006). The SSA was founded in 1966 and it is operating as an independent organization representing the interests of individual investors in Sweden. Its monthly magazine Aktiespararen, in which the disclosure scores are published (typically in each August edition), is one of Sweden’s most read business magazines.

In the SSA’s competition, the annual reports of all firms listed on the SSE and the NGM are scrutinized manually and firms earn points if a certain information item (e.g., dividend policy) is disclosed, regardless of its quality. The SSA awards points to a firm’s annual report according to a pre-selected

---

17All of these stock exchanges operate under the supervision of the SFSA. With the implementation of the Markets in Financial Instruments Directive (MiFID) in the EU, since 2007, the main segments of the SSE and the NGM have been having a legal status of a “regulated market”, while their alternative segments and Aktietorget have been having a legal status of a “multilateral trading facility” (MTF). Both before and after the implementation of the MiFID, the trading venues in the two categories were and still are almost identical in terms of their market infrastructures, but not in terms of their listing requirements. One such requirement, which has differed even across venues in the same category and which is relevant in Anchev (2018a and 2018b), is related to the firms’ minimum number of investors. During the observation period, for example, while primary listings on the main segment of the SSE required at least 500 investors, those on the main segment of the NGM required at least 300 investors (Eckbo, Paone and Urheim (2011)). This one and the other listing requirements, most of which concern firms’ legal, financial and operating circumstances, expected market capitalizations, distributions of voting rights, governance structures and information disclosures, have been generally stricter for listings on the regulated markets.

18Before 2004, firms with multiple common stock classes usually had one class carrying one vote and another one carrying one thousandth of a vote. According to the Swedish Companies Act (2005:551), however, a firm’s stock class cannot carry votes which are more than ten times greater than those of any other stock class issued by the same firm, which is the reason why, since 2004, firms with multiple classes of common stock typically have been having one class that carries one vote and another one that carries one tenth of a vote.
list of information items that are allocated weights. In most years, the information items related to a firm’s strategy, product or service, business model, operations, markets and competitors are allocated the most weight. These are followed by information items concerning past and current economic performance, future risks and corporate governance.

C. Stock Portfolios of Individual Investors

In Anchev (2018c), I also use data on the stock portfolios of all individual investors born in 1963 or 1973 who reside in Sweden. The data on these investors’ long positions in stocks with ISINs that begin with “SE” are from the previously described data set from the SCSD and the data on their long positions in other stocks are obtained from Statistics Sweden. While, as already explained, the former data are available until 2008, the latter data are available only until 2007.

D. Other Characteristics of Stocks and Firms

Apart from the data described above, in each of the three papers in this thesis, the data on the other characteristics of stocks and firms are obtained from Thomson Reuters Datastream, Worldscope, the Applied Quantitative Research (AQR) Capital Management’s website and Direkt (a business news database in Sweden).

E. Other Characteristics of Individual Investors

In Anchev (2018c), the stock portfolios of individual investors are matched with a comprehensive set of their other socio-economic and demographic characteristics. In particular, disaggregated data on individuals’ wealth are provided by the Swedish Tax Agency, while the data on their other socio-economic and demographic characteristics are obtained from the “Longitudinal integration database for health insurance and labor market studies” (“LISA” by Swedish acronym) constructed by Statistics Sweden.

V. Methodology

The data on the characteristics of stocks, firms and individual investors described in the prior section are observed over time. As discussed in Hsiao (2003), Baltagi (2005) and Wooldridge (2010), such cross-sectional time-series data (so-called longitudinal or panel data) have advantages and disadvantages.

One of the advantages of using cross-sectional time-series data is that these data can be used to control for the unobserved heterogeneity of the units of observation (e.g., stocks, firms or individual investors) and/or the time periods. Not considering these types of heterogeneity in the empirical analyses can lead to biased results. Hence, one of the greatest advantages of using cross-sectional

19The 1963 and 1973 cohorts were initially selected for the purposes of a different research project.
time-series data is their ability to prevent an omitted variable bias of the kind discussed here. Moreover, because cross-sectional time-series data contain information regarding the variation in the characteristics of the units of observation from two different sources (i.e., the variation across the units of observation at a certain point in time and the variation for the same unit of observation over time), these data are more informative. Using such data, therefore, usually provides less biased and more efficient results, as well as opportunities to examine the behavior of the units of observation at a certain point in time or over time.

Despite these advantages of using cross-sectional time-series data, these data are often obtained through repeated surveys of individuals and/or institutions (using questionnaires or interviews), which are typically quite costly. In addition, there are many problems related to the execution of such surveys (e.g., problems related to the determination of the sample of individuals and/or institutions to be surveyed, their lack of responsiveness and/or their self-selection) that usually prevent researchers from creating data sets of a higher quality. As already explained, however, the data described in the prior section are not obtained through surveys and, therefore, none of these problems appear in the papers in this thesis.

In each of these papers, previously developed methods that have been established and exploited in the prior related literature are used. These methods and their advantages and disadvantages are described below.

A. Portfolio Sorts

As pointed out in Fama and French (2008), one of the two commonly used methodological approaches in the prior literature for identifying patterns in the cross-section of stock returns (see, e.g., Nagel (2005)), which is also used in Anchev (2018a), involves examining the returns on portfolios of stocks, formed by sorting stocks on the basis of their characteristics. Although the returns on all formed portfolios are typically presented and examined, the focus of the researchers’ analyses is usually on the returns on hedged portfolios, which are formed by taking long and short positions in the stocks from the extreme quantiles. The main advantage of this approach is the simplicity of both its implementation and the interpretation of the obtained results. One disadvantage, however, is that hedged portfolio returns can be significantly influenced by the returns on certain stocks which, in terms of their market capitalizations, are substantially not representative of the relevant population of stocks. For example, the returns on equally weighted hedged portfolios can be considerably influenced by the returns on the relatively large number of small-cap stocks. The first reason for this is that, even though small-cap stocks represent only a fraction of the total market capitalization of all stocks, their returns receive equal weights as the returns on the other stocks in the portfo-

---

20For the reasons why, rather than studying the returns on individual stocks, it is preferable to analyze the returns on portfolios of stocks, see Blume (1970), Friend and Blume (1970), Black, Jensen and Scholes (1972), Fama and MacBeth (1973) and Fama and French (2004).
lios. Second, small-cap stocks ordinarily constitute a relatively large portion of the stocks in hedged portfolios because they are commonly quite numerous in the population of stocks and/or because the cross-sectional variation in the sorting stock characteristics is generally greater among them. In an attempt to circumvent these problems, researchers usually examine the returns on value (i.e., market capitalization) weighted hedged portfolios. The returns on these portfolios, however, can be strongly influenced by the returns on the relatively small number of large-cap stocks, which can also lead to biased inferences. Another disadvantage of using sorts is that they cannot separately identify the marginal effect that each sorting stock characteristics has on the portfolio returns. In addition, by only sorting stocks, it is difficult to draw precise inferences with respect to the functional forms of the relations between the sorting stock characteristics and the portfolio returns. Finally, since forming portfolios with a reasonably sufficient number of stocks is normally feasible by sorting stocks on the basis of only three to four stock characteristics at a time, the number of control variables that can be considered simultaneously is severely limited.

B. Fama and MacBeth (1973) Regression Models

The second methodological approach noted in Fama and French (2008) that is commonly used in the prior literature for identifying patterns in the cross-section of stock returns involves estimating regression models as in Fama and MacBeth (1973). These models have also been commonly used in the prior literature that studies the determinants of the informativeness of stock prices in the cross-section of stocks (see, e.g., Ferreira and Laux (2007)). Hence, they are used in Anchev (2018a and 2018b).

In the prior literature related to these papers, when estimating Fama and MacBeth (1973) regression models, researchers have cross-sectional time-series data and they first estimate regressions at the end of each time period, with one characteristic of a certain stock as the dependent variable and the other characteristics of that stock as independent variables. At the end of each time period, therefore, the estimated coefficient on each independent variable can be different. The reported coefficients, however, are typically the equally weighted time-series means of these coefficients. Moreover, the standard errors are based on the time-series of the estimated coefficients over all time periods.

These standard errors are adjusted for the cross-sectional correlation of the error terms, but they are not adjusted for their autocorrelation. Therefore, Newey-West standard errors, adjusted for heteroskedasticity and an autocorrelation with a maximum time lag order of one or more time periods, are commonly used. As shown in Petersen (2009), these standard errors are biased, but the bias is quite small.

Fama and MacBeth (1973) regression models have other advantages and disadvantages as well. These regressions can provide marginal effects and they can do so for a relatively large number of independent variables. Moreover, by examining the estimated residuals from the regressions, relatively precise inferences about the functional forms of the relations between the variables
of interest can be made. Nevertheless, the results from the regressions estimated with equally weighted observations can be significantly influenced by the relatively large number of small-cap stocks, which usually have more extreme characteristics. It is, however, possible to estimate value weighted regressions, which can mitigate the problems of this kind. Finally, given that the characteristics of certain stocks can be extreme, the results from estimating regression models as in Fama and MacBeth (1973) can be quite sensitive to the presence of stocks with such characteristics.

C. Cross-Sectional Time-Series Regression Models with Fixed Effects

Following the prior literature on the implications of the characteristics of stocks and/or firms for individual investors (see, e.g., Calvet, Campbell and Sodini (2009) and Lawrence (2013)), in Anchev (2018c), cross-sectional time-series regression models with fixed effects are used.

As pointed out in Hsiao (2003), Baltagi (2005) and Wooldridge (2010), when estimating regression models using cross-sectional time-series data, it is reasonable to expect that some of the assumptions of the ordinary least squares (OLS) model are violated. Indeed, because the units of observation are repeatedly observed over time, it is no longer appropriate to assume that the observations of their characteristics are independent either across the units at a certain point in time or for the same unit over time. Instead of using OLS models, therefore, one commonly used approach is to estimate cross-sectional time-series regression models with fixed effects, which involve certain assumptions with respect to the effects of the unobserved unit-specific characteristics and time-specific factors on the dependent variable.

When estimating regressions with fixed effects, these characteristics are considered to be permanent and the factors are considered to be temporary. Moreover, no conditions are imposed regarding the relations between them and the independent variables. The purpose of these regressions, therefore, is to control for the unobserved permanent characteristics of the units of observation that can affect their behavior differently at a certain point in time (but uniformly over time) and/or for the unobserved temporary factors that can affect their behavior differently over time (but uniformly at a certain point in time). One way of achieving this involves estimating OLS models with indicator variables for each unit of observation and/or each time period.

Although the possibility of controlling for the unobserved heterogeneity of the units of observation and/or the time periods is quite appealing, the use of regression models with fixed effects comes at a cost. Indeed, in the approach above, there is usually a quite large loss of degrees of freedom, which may lead to less efficient estimated coefficients. Moreover, due to perfect multicollinearity, the coefficients on any independent variables that do not vary over time (e.g., the place of an individual’s birth) cannot be estimated.

VI. Conclusion

In Anchev (2018a), I find that, after controlling for market capitalization, the predictability of future stock returns associated with each of the earnings-
to-price ratio, the book-to-market ratio, the past return, the total volatility of returns and the return on assets is more pronounced among stocks with smaller total and/or institutional investor bases. These results appear even after controlling for several other stock characteristics and potential risk factors and they are both statistically and economically meaningful. Thus, they are consistent with the hypothesis in Merton (1987) that the incomplete dissemination of information across investors helps in explaining the occurrence and the persistence of cross-sectional stock return anomalies.

With this interpretation in mind, Anchev (2018a) contributes to the literature which has, thus far, discovered around 450 stock return anomalies (see Hou, Xue and Zhang (2017)), but which has provided just a few incomplete explanations for them that are unrelated to the relative size of the investor base of a stock (see, e.g., Campbell (2000), Fama and French (2004), Cochrane (2011) and Nagel (2005)). It is important to recognize, however, that the anomalies examined in Anchev (2018a) are also not completely explained. Indeed, in some of the empirical tests, some of the anomalies appear in a statistically meaningful way even among stocks with larger investor bases.\(^{21}\) Additional research on this topic is, therefore, needed.

Further, the results in Anchev (2018a) suggest that some stock return anomalies, even from the perspective of Merton’s (1987) version of the CAPM, reflect a mispricing. Hence, it seems that, among stocks with smaller investor bases, there are greater arbitrage opportunities for investors, in the sense that they can earn higher future returns by taking long positions in the underpriced stocks and short positions in the overpriced stocks. Depending on whether and how the information costs proposed in Merton (1987) vary over time, however, the occurrence of such arbitrage opportunities in the future and the time period over which they can persist are an empirical issue.

Regardless of the nature of the stock return anomalies, the implication of the results documented in Anchev (2018a) for firms seems to be that they should invest resources in the expansion of their investor bases. For example, firms can invest in advertising in the financial media, they can list their securities on one or more exchanges, they can have their securities graded and promoted by investment banks or they can invest in making their securities an eligible investment for investors who are constrained by requirements related to prudent investing (Merton (1987)). By doing so, firms can reduce their cost of equity capital (through the improvement in the extent to which their idiosyncratic risk is shared among their investors) or they can prevent or reduce the mispricing of their stocks.

\(^{21}\)It is worth noting, however, that, when the portfolio returns are value weighted, the results in Anchev (2018a) suggest that one of the most puzzling anomalies, the momentum anomaly (which is related to the past return), does not appear in a statistically meaningful way neither among stocks with larger nor among stocks with smaller investor bases. Hence, when controlling for the relative size of the investor base of a stock and when portfolio returns are value weighted, this anomaly seems to be nonexistent. This finding is consistent with that of Hong, Lim and Stein (2000) who use the market capitalization of a stock and the number of financial analysts that follow it as measures of the speed of information dissemination across investors and who find that the momentum anomaly is more pronounced among small-cap stocks and stocks with a lower analyst coverage.
Further, since the relative size of the investor base of a stock can partially or completely explain the occurrence and the persistence of several cross-sectional stock return anomalies, for researchers who study the predictability of future stock returns in the cross-section of stocks, the implication of my findings is that they should consider the role of this variable in their empirical analyses. Finally, considering that the previously observed results represent an addition to the broad collection of prior empirical observations in support of Merton’s (1987) version of the CAPM (see the opening section in Anchev (2018a)), perhaps the most intriguing implication of Anchev (2018a) is that the explanation for the patterns in the cross-section of stock returns that are commonly thought of being anomalous need not require a radical departure from the paradigm that investors are rational.

Further, in Anchev (2018b), I find that the relative idiosyncratic volatility of future stock returns is: 1) negatively associated with the absolute size of the total and the institutional investor base, 2) positively associated with the institutional ownership, 3) negatively (positively) associated with the average stock portfolio size (Herfindahl index) of the investor base and 4) positively associated with the indirect (i.e., through nominees) ownership. These results appear after controlling for several other stock characteristics and they are both statistically and economically meaningful. Thus, they are consistent with the hypothesis of Peress (2010) that the properties of the investor base of a stock (i.e., its size, informedness, average risk tolerance and/or average propensity for liquidity trading) have implications for the informativeness of the stock’s price.

With this interpretation in mind, to my knowledge, Anchev (2018b) is the first to provide empirical evidence consistent with the theory of Peress (2010). Indeed, the paper provides a description of the properties of the investor bases of stocks and their associations with the relative idiosyncratic volatility of future stock returns. An empirical examination that can be more informative of the causal relations between the variables of interest, however, requires a more careful consideration of their endogenous nature. Nevertheless, Anchev (2018b) contributes to the prior literature on the implications of the properties of the investor bases of stocks (see, e.g., Bodnaruk and Ostberg (2009) and Jankensgård and Vilhelmsson (2018)) and on the determinants of the relative idiosyncratic volatility of future stock returns in the cross-section of stocks (see, e.g., Ferreira and Laux (2007)). Finally, and perhaps most importantly, given that there is some empirical evidence which suggests that they invest considerable resources in the management of their investor bases (see, e.g., Beyer, Larcker and Tayan (2014a and 2014b)), the findings in Anchev (2018b) can be relevant for firms.

Lastly, in Anchev (2018c), I find that, when the amount of information disclosed by a certain firm is greater (or increases), the stock portfolio weights that individual investors allocate (through trading) to that firm’s stock are lower (or decrease) and suboptimal. The former result is less pronounced or nonexistent for more financially competent individuals and for positions in firms with a relatively poor information environment. When they do allocate greater portfolio weights to the stock of a firm that discloses more, however,
individual investors, regardless of their financial competence, are found to earn lower returns.

Although inconclusive about the underlying mechanism generating the observed results, these findings can be interpreted in terms of a theoretical model in which a greater quantity of firm disclosure places less sophisticated investors at an information disadvantage. Anchev (2018c), thus, offers new empirical evidence consistent with the models in which more firm disclosure increases the information asymmetries between market participants who differ in terms of sophistication (see, e.g., Indjejikian (1991) and Kim and Verrecchia (1994)). In addition, the findings documented in Anchev (2018c) represent a contribution to the nascent empirical literature on the implications of firm disclosure for, presumably less sophisticated, individual investors (see, e.g., Lawrence (2013)).

Further, the evidence presented in Anchev (2018c) can be seen as a timely and valuable input in the current debate on a potential disclosure overload (see FASB (2012), IASB (2013b) and SEC (2013)) and, thus, should be of interest to regulators and the accounting standard-setting boards. Indeed, considering the increased complexity of the information environment, these stakeholders have recently expressed an interest in understanding whether the current firm disclosures are excessively voluminous for investors to process relevant information. Because of the inherent problem of precisely identifying firm disclosure that is excessive, however, I leave it for future research to explore the implications of this construct for investors.

Given that agreement on the need for mandatory firm disclosure is yet to be reached (see Beyer et al. (2010)) and the innate uncertainty involved in setting the minimum or optimal level of such disclosure, an important caveat of Anchev (2018c) is that it does not consider situations in which firms disclose below the mandatory requirements. Examining investors’ behavior under such conditions, hence, represents another venue for future research.

Finally, it is important to note that the analyses in Anchev (2018c) are limited to the broader, but economically less influential group of individual investors. When deciding on future policy changes, regulators and the accounting standard-setting boards should consider the needs of many other constituent groups (e.g., institutional investors, lenders or suppliers). Reconciling the interests of all groups, however, will probably be challenging because, as Verrecchia (2001, p. 163) points out, “… if one makes market agents … sufficiently diverse, it is difficult, if not impossible, for disclosure to yield a positive benefit for everyone”.

Overall, each of the three papers in this thesis (i.e., Anchev (2018a, 2018b and 2018c)) provides new empirical evidence about the role of information in financial markets.
REFERENCES


Anchev, Stefan, 2018a, Investor base and stock return anomalies, Working paper, Umeå University.


Anchev, Stefan, 2018c, Individual investors and quantity of firm disclosure, Working paper, Umeå University.


Beyer, Anne, David F. Larcker, and Brian Tayan, 2014a, Study on how investment horizon and expectations of shareholder base impact corporate decision-making, Working paper, National Investor Relations Institute and Rock Center for Corporate Governance at Stanford University.
Beyer, Anne, David F. Larcker, and Brian Tayan, 2014b, Does the composition of a company’s shareholder base really matter?, Working paper, National Investor Relations Institute and Rock Center for Corporate Governance at Stanford University.


Myers, Stewart C., and Nicholas S. Majluf, 1984, Corporate financing and investment decisions when firms have information that investors do not
Smith, Adam, 1776, An inquiry into the nature and causes of the wealth of nations, in Edwin Cannan, ed. (Methuen, London, 1904).
Subrahmanyam, Avanidhar, 2010, The cross-section of expected stock returns: what have we learnt from the past twenty-five years of research?, European Financial Management 16, 27–42.