Determinants of Analysts' Forecast Accuracy

Empirical Evidence from Sweden
Thanks

The authors of this thesis would like to express their sincere gratitude to their examiner Natalia Semenova and supervisor Damai Nasution for making this thesis possible. Your guidance and valuable feedback has greatly improved our work.

The authors would also like to thank their opponents for helpful comments and pleasant company at the seminars.

____________________                                         ____________________
Sofie Areskoug                                                  Niklas Karlén
Abstract

Bachelor Thesis, Program of Master of Business and Economics, 15 hp
School of Business and Economics – Linnaeus University in Växjö
2FE30E:3 Spring, 2017

Authors: Sofie Areskoug and Niklas Karlén
Supervisor: Damai Nasution
Examiner: Natalia Semenova

Keywords: Financial Analyst, Gender, Determinants of forecast accuracy, Sweden

Title: Determinants of Analysts' Forecast Accuracy – Empirical Evidence from Sweden

Background: The search of finding analysts who make the best forecasts has been an ongoing process since the 1930's. Determinants that can help predict the forecast accuracy of the analysts are in the interest of both investors and brokerage houses. Newer research in this area has taken gender of the analyst into consideration. Women are widely under-represented in the analyst occupation and there is evidence that investors are apprehensive toward women in the financial sector.

Purpose: The aim of this thesis is to examine determinants of forecast accuracy regarding analysts covering Swedish companies. The authors have confidence in the research to benefit investors in their decisions on the Swedish stock market. In addition, the authors aim to shed light on the unequal gender representation of female analysts.

Method: This thesis has examined 519 individual scores of forecast accuracy from 284 financial analysts covering stocks on the Swedish Index OMXS30. The forecasts are from the years 2016 and 2017. This study has a quantitative strategy and the data have been tested by an OLS estimates regression.

Results: The empirical evidence shows that being a female analyst have a statistically significant positive effect on forecast accuracy. Female analysts covering Swedish stocks seem to outperform their male colleagues. Furthermore, insignificant results were found for firm complexity, industry complexity, brokerage house and analyst experience.
Table of Contents

1 Introduction ............................................................................................................................................. 1
  1.1 Background........................................................................................................................................... 1
  1.2 Problem Discussion and Research Gap ................................................................................................. 3
  1.3 Research Question ............................................................................................................................... 4
  1.4 Purpose ............................................................................................................................................... 4
  1.5 Disposition ......................................................................................................................................... 6

2 Methodology ............................................................................................................................................. 7
  2.1 The Authors’ Background .................................................................................................................... 7
  2.2 Research Approach ............................................................................................................................... 7
  2.3 Research Philosophy ............................................................................................................................ 8
    2.3.1 Epistemological Consideration ...................................................................................................... 9
    2.3.2 Ontological Consideration ........................................................................................................... 10
  2.4 Research Strategy ............................................................................................................................... 11
  2.5 Ethical Principles .................................................................................................................................. 12
  2.6 Research Design .................................................................................................................................. 13

3 Theoretical Framework ............................................................................................................................. 15
  3.1 Literature Review .................................................................................................................................. 15
  3.2 Institutional Background ........................................................................................................................ 16
  3.3 The Efficient Market Hypothesis (EMH) ............................................................................................ 17
  3.4 Behavioral Finance ............................................................................................................................... 18
  3.5 Determinants of Differential Forecast Accuracy in Prior Literature ................................................... 20
    3.5.1 The Existence of Differential Forecast Accuracy ........................................................................... 20
    3.5.2 The Individual Ability of the Analyst ............................................................................................. 21
    3.5.3 Task Complexity ............................................................................................................................ 21
    3.5.4 Work Environment ....................................................................................................................... 22
    3.5.5 Summary of Prior Research ........................................................................................................ 22
  3.6 Gender Characteristics and its Effect on Prediction Accuracy............................................................... 23
  3.7 Hypothesis Development ..................................................................................................................... 24

4 Empirical Method ..................................................................................................................................... 26
  4.1 Data Collection ..................................................................................................................................... 26
    4.1.1 Primary and Secondary Data ......................................................................................................... 26
    4.1.2 Thomson Reuters ........................................................................................................................... 26
    4.1.3 OneSearch .................................................................................................................................... 27
Table of Figures

Figure 1 Research Strategy ........................................................................................................ 12
Figure 2 An Illustration of the Relationship Among the Three Types of Information ... 18

Table of Tables

Table 1 An Illustration of Findings in Prior Research............................................................... 22
Table 2 Descriptive Statistics .................................................................................................. 39
Table 3 Pearson Correlation Coefficient ................................................................................. 39
Table 4 Model Summary........................................................................................................... 41
Table 5 ANOVA Table ............................................................................................................ 41
Table 6 Multicollinearity ........................................................................................................ 42
Table 7 Coefficients of the Regression Model........................................................................ 43
Table 8 Compilation of Hypotheses and Results .................................................................... 44
1 Introduction

The introduction chapter includes a background for financial analysts and previous studies about analysts' forecast accuracy. The situation for women in finance is also discussed. A problem discussion is presented in this chapter which derives into a research question. Finally, the purpose of the thesis is discussed followed by a disposition.

1.1 Background

No one can be certain of what will happen in the future. Therefore, forecasts and predictions are important to help plan for future events. In the financial sector, forecasts done by analysts are predictions of companies’ performance and future values. The forecasts are used as a platform for investors to guide them in the right direction in their decision-making process, where better forecasts lead to better investments. The accuracy of the forecast is measured by how close the analyst's earnings forecast for the company is to reality when the company discloses their results. Elton, Bruber and Gultekin (1981) presented evidence that forecasts in line with the consensus are incorporated into the company price and that the market reacts to expectational forecasts that differ from the consensus. This suggests that forecasts made by financial analysts may affect the market and therefore the accuracy of their forecasts is important.

There are two different types of financial analysts, buy-side- and sell-side analysts. A sell-side analyst usually works for a brokerage house, making forecasts for external clients while the buy-side analyst usually works for a pension fund or a mutual fund and make recommendations to their money managers. According to Norby (1972), the description of a security analyst is someone who analyses securities and thereafter makes forecasts. Securities are instruments traded on the financial market, such as stocks, options and warrants. These denote an ownership interest. This thesis is solely based on sell-side analysts. There is a broad usage of the term financial analyst and because of that, the term analyst will hereby be used synonymously to a financial sell-side security analyst, throughout this thesis.
The analyst occupation is a competitive profession and they are accustomed to being measured by their performance. A financial analyst's abilities are measured by their forecast accuracy. The search for how to find the best analysts started already during the 1930's when Cowles (1933) asked the question "Can Stock Market Forecasters Forecast?". This search is still an ongoing process. The determinants of differential forecast accuracy have been widely discussed in the past. In the 1990’s, the discussion was whether differential analyst forecast accuracy actually existed. There was no consensus that determinants improving forecast accuracy by analysts truly existed until the later part of the decade and the beginning of the 21st century. Brown and Mohd (2003) created a model with different characteristics regarding individual factors of the analysts', task complexity and work environment. The aim of their research was to develop a model with determinants explaining forecast accuracy of the analyst. Brown and Mohd’s (2003) model has been the basis for later research in this area and their research is the main article for this thesis. Most of the researchers studying financial analysts and the determinants of their forecast accuracy have used either an American or an unspecified sample.

Newer research has taken gender into consideration when studying financial analysts. There is evidence that female and male analysts differ in forecast accuracy. Kumar (2010) argues that female analysts make bolder and more accurate forecasts. A study by Kiecker, Palan and Areni (2000) argue that females are more comprehensive processors as well as sensitive and attentive to details. This would suggest that women are desirable as employees by brokerage firms, but real world statistics shows us differently. It is recognized that there has been a historical under-representation of women in high profile and lucrative careers (Green, Jegadeesh and Tang, 2009). This is also true for the financial analyst occupation where less than 17 percent of the chartered financial analysts in the U.S. are women (Pavlenko Lutton and Davis, 2015). A report by Shandilya and Duncan (2015) state the picture is roughly the same in Europe. The organization for financial analysts in Sweden CFA Society Sweden state "Despite strong evidence that a more gender-balanced investment industry would benefit investors, women are still under-represented throughout our profession" (CFA Society Sweden, 2016).

Betz, O’Connell and Shepard (1989) argue that different genders have different values and traits and will therefore result in different decisions, judgments and
behaviors. A study by Bharadwaj Badal (2014) shows that gender diversity is vital to any workplace. The advantages of a well-diversified workplace are many. To name some, diversified workplaces perform better than non-diversified workplaces due to different viewpoints, ideas, market insights, easier to serve a diverse customer base and most importantly a gender-diverse company can easier attract and retain talented women. Bharadwaj Badal's (2014) study indicates that the financial analyst occupation is missing out on a lot of talent, which most likely leads to less accurate forecasts.

Sweden is known for its development when it comes to gender equality and is often ranked in the top when compared to other countries. The European Institute for Gender Equality has ranked Sweden to be the most gender equal country in the European Union 2005, 2010 and 2012 (European Institute for Gender Equality, 2017). Due to its high levels of equality, Sweden is a relevant country to study for differences in forecasting accuracy between the genders.

By extending prior research in this field and shedding light to the under-representation of female analysts in the financial sector, the authors aim to contribute to a more gender-diverse financial sector. The authors also claim there to be limited research in this area done with a Swedish sample. This research benefit investors on the Swedish stock exchange since it is of importance to understand what determinants of financial analysts will generate the best forecast accuracy.

1.2 Problem Discussion and Research Gap

Prior studies have found multiple determinants that can help investors to understand what to look for when searching for analysts who make accurate forecasts. Brown and Mohd (2003) found task complexity for the analyst, timing of the forecast and work environment to all have a significant effect on the analyst's forecast accuracy. These results have been confirmed by later research (Clement, Koonce and Lopez, 2007). Prior literature has examined analysts from different countries and different markets, but there is little research to be found where Swedish companies are included in the sample. The authors of this thesis have not found any reasons to expect prior researchers’ findings to differ for analysts following Swedish companies. However, this has not been examined earlier and therefore a research gap
is identified. Arguments can also be made that prior research is outdated since the samples in the prior studies consist of forecasts done by analysts before the 21st century. Information is moving faster in today's society and the landscape of financial services is changing, which may have changed the way analysts work (Höbe, 2015).

Prior research that examines the relationship between female analysts and their forecast accuracy compared to male colleagues argues that female analysts make bolder and more accurate forecasts (Kumar, 2010). Atkinson, Baird and Frye (2003) have found evidence that investors are skeptical toward women in the financial sector and state "we find significantly lower net asset flows into mutual funds managed by women compared with funds managed by men". The authors also argue that this may explain why there are so few female analysts and they write that employers "may be apprehensive to hire a female manager if they fear that investors will prefer male-managed funds" (Atkinson, Baird and Frye, 2003). This suggests that females are having a hard time getting into the financial sector and getting lucrative positions in male dominant occupations.

1.3 Research Question

This thesis aims to examine if determinants found in prior research that has proved to affect analysts forecast accuracy, also can explain forecast accuracy for analysts covering Swedish companies. The determinants examined are individual factors of the analyst, task complexity for the analyst as well as work environment. The authors also examine whether any differences can be found in forecast accuracy depending on the analyst's gender.

*RQ: Are there determinants that affect forecast accuracy for financial analysts covering Swedish companies?*

1.4 Purpose

The purpose of this thesis is to shed light on gender equality and to provide empirical evidence of determinants that affect the forecast accuracy of analysts covering Swedish companies.
The authors state it to be of importance to examine determinants of analysts' forecast accuracy for both brokerage houses and investors, particularly regarding analysts covering Swedish companies since research about it are very limited. According to Clement, Koonce and Lopez (2007), investors would be able to better predict when analysts make more accurate forecasts if they got a better understanding of the skills and experiences that are relevant to analysts' performance. This study is also of importance for employers hiring analysts. A better understanding of what leads to high levels of analyst performance may influence how employers select and train their employees.

Evidence that female analysts are similar or better than male analysts would help women build credibility. Hopefully, this leads to more female analysts getting hired which would diversify the workforce for this occupation. A more diversified workforce leads to better results which are important for investors who seek the best forecast accuracy possible from the analysts.

The purpose of this research is also to contribute to the behavioral finance theory. Contrary to the behavioral finance theory, the efficient market hypothesis argues that all available information about companies is already interpreted in a correct way by investors. This suggests that there would be no room for analysts' forecasts. Evidence that there are determinants that have different effect on analysts' forecast accuracy would be in favor of the behavioral finance theory.
1.5 Disposition

• This chapter consists of a brief background of financial analysts and their forecast accuracy. Later there is a problem discussion that develop in to this thesis research questions. The purpose of the thesis is also presented.

• In this chapter the thesis choice of methodology is carefully explained. It consist of the authors background and the choice of research approach, philosophy, strategy and design. Ethical principles are also discussed.

• This chapter includes a review of prior literature and economic theories. The chapter derives into a hypothesis development.

• This chapter explains how the data have been collected. The sample, variables and empirical model are described.

• The results are presented and analyzed in this chapter.

• In the last chapter the authors refer back to the research question and draw conclusions from the results. The authors also discuss the implications for this research. There will also be suggestions for further research.
2 Methodology

This chapter focuses on the authors’ prerequisites, what research approach and what philosophy the thesis is built on. The research strategy is also explained in this chapter.

2.1 The Authors’ Background

The two authors of this thesis are students at the Linnaeus University in Växjö, Sweden. The thesis is written during their sixth semester. Both students are studying to become a Civilekonom, which is title earned after graduating from a specific magister degree program in business and economics. The program consists of courses in business administration, economics, finance and statistics.

They both did their fifth semester abroad in the U.S. to be able to continue their studies toward a Civilekonom-degree with a financial orientation. In the U.S. they studied relevant courses which gave them deeper international knowledge in accounting and finance. The authors have both shown an interest in the financial world and more specifically the equity market. Gender diversity is a topic the two authors find fascinating. They have noticed an unequal situation in the financial analyst occupation which they find interesting to analyze.

The authors have worked together on projects prior to this thesis with great results. They know each other well and they make a superb team.

2.2 Research Approach

There are two main approaches that can be used when writing a thesis, deductive and inductive. The first research approach is deduction. This approach starts with theories and thereafter makes empirical predictions. These are later on verified by the collection of data. The deduction approach is frequently used in the field of nature science (Bryman, 2012, p. 24).

Deduction has a number of different characteristics. One characteristic is the search to explain the causal relationship between concepts and variables. Another is that concepts need to be operationalized, so that facts can be measured in a quantitative
way. The third characteristic of a deductive approach is reductionism, which means that if reducing problems to the simplest possible, they are better understood. Finally, generalization is the last characteristic of a deductive approach and refers to generalizing a carefully selected sample of sufficient size (Saunders, Lewis and Thornhill, 2016, p. 146).

The second research approach is inductive approach and is based on observations and theories that are presented at the end of the research process. The approach works backward, in comparison to a deductive approach, to form a hypothesis. This way of tackling the methodology is very common for social science (Björklund and Paulsson, 2012, p. 64). By starting with observing the sample and thereafter forming a hypothesis, the inductive approach leaves room for alternative explanations, in contrast to a deductive approach (Saunders, Lewis and Thornhill, 2016, p. 147).

This thesis will be using the deductive approach when examining analysts' forecast accuracy since the applied theory comes from prior research. There are already existing theories about forecast accuracy that is the base of this research, which is the authors' motivation for choosing the deductive approach. The research idea is to derive a conclusion based on observations and the collection of data. To construct the association between theory and hypothesis the authors need to develop hypotheses based on prior research to examine the research question: Are there determinants that affect forecast accuracy for financial analysts covering Swedish companies? As a result, the authors motivate a deductive approach since the thesis is based on already existing theories and the process for the research is linear. This means the thesis starts with a research question with empirical predictions and ends with a conclusion based on collected and analyzed data.

The authors claim that an inductive approach is inappropriate for the thesis since there are theories and hypotheses presented in the beginning of the research and the paper is not as process-oriented as an inductive approach would require.

2.3 Research Philosophy

In research, there are two philosophical viewpoints. There is the epistemological issue which concerns the question of what is to be regarded as acceptable knowledge
(Bryman, 2012, p. 27) and there is the ontological issue which is concerned with the nature of social entities (Bryman, 2012, p. 32).

2.3.1 Epistemological Consideration

There are three main epistemological positions, these are positivism, interpretivism and realism. Positivism is the position that affirms the significance of imitating natural sciences for application to social reality. Positivism means that only phenomena that have been confirmed by the senses can be looked upon as knowledge (Bryman, 2012, p. 28). It is important for the positivism position that science is conducted in an unbiased way. The position of positivism entails elements of both deduction, the purpose of the theory is to generate hypotheses that can be tested, and induction, knowledge arrives at the gathering of facts (Saunders, Lewis and Thornhill, 2016, p. 136-137).

Realism, the second position of epistemology, is similar to positivism in some ways. Both share the view that there is a reality separated from our description of it. Realism differs from positivism in the way that realism argues that the scientist's conceptualization of reality is just a way to explain that reality, while a person with the positivism position claim that the scientist's conceptualization of the reality directly reflects that reality (Bryman, 2012, p. 29).

Interpretivism is the third position and it is the opposite of positivism. The interpretivism position argues that the subject matter of social sciences is essentially different from natural sciences (Bryman, 2012, p. 28).

The authors argue that there are elements of positivism in this thesis. First, the thesis is based on a scientific method and the data collected is quantitative, which means it is observable and measurable. Since the thesis is analyzing a large amount of data this assumption is suitable. Second, due to the data being collected from Thomson Reuters, it is to be considered as unbiased and value-free research. In addition, the science is conducted in an objective way where the authors are neutral and independent of the research.

Positivism is commonly used in combination with a deductive approach, which is the case for this thesis. Only knowledge that has been confirmed by prior research is warranted as knowledge acceptable for this thesis. These elements describe an
approach imitating those of natural science and therefore the authors claim this thesis epistemological position to be positivism. Realism does not suit the authors’ purpose of their research even though the close connections to positivism would suggest that it could be of interest for the authors. What makes it unsuitable for this thesis is the realism’s conceptualization of reality, that there is a reality separate from our description of it. The interpretivism position, that requires a strategy which respects the differences between people and objects, do not suit this type of research either. It requires primary data which may lead to biased results, which would hurt the aim of this research of being objective.

2.3.2 Ontological Consideration

The central point of ontology is whether social entities should be considered as objective entities with their own reality external from social actors, or if they are to be considered as social constructions built from the perceptions of social actors. The two different viewpoints can be categorized into objectivism and constructionism (Bryman, 2012, p. 32).

Saunders, Lewis and Thornhill (2016, p. 128) describe the ontological assumption objectivism as where social phenomena have their own meaning independent from social actors. It can be described by an organization as a tangible object. An organization represents a social order with its own set of rules and regulation. The rules exert pressure on the individuals who inhabit it to learn and apply them. People therefore follow what is soon standardized procedures, because if they do not they might be reprimanded or fired. Bryman (2012, p. 32) argues, that by the objectivism standpoint the organization is therefore seen upon as a constraining force that acts on its members.

An alternative ontological position is constructionism, which asserts that social phenomena are being accomplished by social actors. According to Bryman (2012, p. 33) constructionism implies that social phenomena is in a constant state of revision and not only produced through social interactions.

The research questions, different determinants’ association with analysts’ forecast accuracy will be examined by analyzing secondary data with a quantitative strategy. By doing so, using objectivism as the thesis ontological consideration is suitable. In
addition, results from the thesis can be looked at objectively and the author's opinion will not change the outcome which also motivates an objective approach. The authors will in an objective way research the social phenomena of analysts. Financial analysts are a part of a formal structure and the authors assume that this structure is similar to all organizations. The authors disregard any beliefs they have made from previous interactions with financial analysts, to steer clear of experiences affecting any conclusions made in this thesis, which motivates an objectivistic way.

The opposite position, constructionism, where the researcher presents a specific version of social reality does not fit this type of research. Therefore, the ontological consideration in question for this thesis will be objectivism.

2.4 Research Strategy

In order to answer the research question of the thesis, a suitable research strategy must be chosen. The two methods that can be used when writing a thesis are qualitative or quantitative. It is important to know the differences between the two research strategies in order to recognize which one is most suitable for the chosen research question.

According to Bryman (2012, p. 36), a qualitative research is a strategy that emphasizes words rather than quantification when collecting and analyzing data. A qualitative research is commonly described as the generation of theories rather than testing of theories. Saunders, Lewis and Thornhill (2016, p. 568) explain that qualitative data results in non-standardized data and that the analysis is based on conceptualizations.

A quantitative research, on the other hand, is a research strategy that emphasizes quantification in the collection of data (Bryman, 2012, p. 35). The quantified data collected for a quantitative research is usually larger than for a qualitative research. The data is later summarized in a statistical form so that the hypotheses can be tested and analyzed. Saunders, Lewis and Thornhill (2016, p. 568) explain that quantitative data is based on meanings derived from numbers and results in numerical and standardized data. They also state that the analysis is based on diagrams and statistics.
The authors have decided to use the quantitative research strategy for this thesis. To be able to distinguish the determinants of forecast accuracy, a large sample is desirable. The larger sample size will give this study a higher statistical reliability which motivates the use of a quantitative method. A quantitative research strategy will also add to the results' objectivism. Another reason for why a quantitative approach is suitable for this thesis is the use of numbers rather than words, which will be easier to repeat and copy for others. The authors argue that a quantitative research method is suitable for the thesis since the aim is to confirm hypotheses using data and statistical results. Furthermore, the authors run a multiple regression model and conduct the analysis from diagrams and statistics. A qualitative method could not be applied because the authors do not seek to explain, explore or understand hypotheses and the analyzed data will not be in words or images, but numbers.

![Figure 1 Research Strategy](image)

### 2.5 Ethical Principles

There are four ethical principles to consider when doing a social research. The four principles are according to Bryman (2012, p. 135):

1. – Whether there is harm to participants.
2. – Whether there is a lack of informed consent.
3. – Whether there is an invasion of privacy.
4. – Whether deception is involved.

Since this thesis only focuses on secondary data, the authors do not feel the need to discuss principle one and two any further. What is interesting for this research from an ethical perspective is principle three and four.

The right to privacy is a principle that is strong for most people in society (Bryman, 2012, p. 142). This thesis is interested in a broad perspective where there is no room
to point out the analysts' individual performance specifically, the results will only present their performance as a group. The names of the analysts are used for identifying the gender and their names will not be presented in this thesis. By using secondary data, implying that the analysts have already given their consent to have their information available to the public, and by not specifying analysts individually, there is no need to suspect that the analysts used for this research would feel that the authors are invading their privacy.

Deception is the last ethical principle to consider. It occurs when "researchers present their work as something other than what it is" (Bryman, 2012, p. 143). The principle is important since this thesis methodology is set to be objective. It is therefore important that the authors are careful when selecting the sample and discussing the result to not present it in a biased way.

2.6 Research Design

According to Bryman (2012, p. 46), there are some prominent criteria for the evolution of research, which are ability and validity. These criteria are central of judgment for the quality of the research (Saunders, Lewis and Thornhill, 2016, p. 202).

Reliability refers to the possibility of replicating the research (Saunders, Lewis and Thornhill, 2016, p. 202). Bryman (2012, p. 46) claims that reliability can be a problem when using a quantitative strategy. This is due to uncertainty in measuring the stability of data. This thesis uses secondary data and the variables are explained in such a way that any researcher should be capable of repeating the process. By doing so, the authors claim the research to be reliable.

The second criterion of research is validity. This criterion is concerned with the conclusions that are being made in a research, that the appropriate measurements have been used and that the analysis of the results is accurate (Bryman, 2012, p. 47 and Saunders, Lewis and Thornhill, 2016, p. 202). There are different aspects of validity. The ones that are relevant to this thesis are measurement validity, internal validity and external validity.

For a quantitative strategy, it is common to apply measurement validity. The focus is on whether the measures that are being used for the research reflect the right concept
(Bryman, 2012, p. 47). The variables used in this thesis are all inspired by prior research. Therefore, the authors of this thesis argue they will not have any problems regarding measurement validity.

Internal validity is accepted when the research accurately describes a causal relationship (Saunders, Lewis and Thornhill, 2016, p. 203). It is common to refer to how the independent variables affect the dependent variable when discussing causality (Bryman, 2012, p. 47). This concept can be associated with casual or explanatory studies which are usually true for quantitative research and positivism (Saunders, Lewis and Thornhill, 2016, p. 203). By using the right methods to analyze the data, which are derived from prior literature in this area, the authors claim to have taken internal validity into consideration.

The third aspect of validity is external validity. It is concerned with how the findings of the research can be applied and generalized to other relevant groups (Bryman, 2012, p. 47 and Saunders, Lewis and Thornhill, 2016, p. 204). The authors aim to apply and extend prior research to a Swedish sample and argue that the results could be of interest to countries with both geographical- and cultural similarities to Sweden.
3 Theoretical Framework

This chapter includes a description of the theory the thesis is based on. The theoretical framework contains the research objectives that are needed to understand the factors behind financial analysts and forecast accuracy.

3.1 Literature Review

This thesis has support from several articles where Brown and Mohd (2003) and Kumar (2010) are used as key literature. Both articles are well recognized within their field and have been referred to in many studies. For this thesis, they will provide a foundation to the hypothesis development. Brown and Mohd (2003) created a model to explain analysts' forecast accuracy and by doing so they found multiple determinants that have a significant effect on forecast accuracy. Kumar (2010) found in his research evidence that female analysts make bolder and more accurate forecasts, in comparison to their male colleagues.

What stand out with these studies are the large sample sizes. Brown and Mohd created a sample of 172,837 observations, taken from forecasts made 1987-1999 and the sample of Kumar’s (2010) study consists of 2,200,758 forecasts, 17,240 analysts and 13,636 stocks. Data from 1983 to 2006 is analyzed and interpreted in Kumar’s (2010) study. Kumar (2010) explains in his article how he had to manually identify the analysts’ gender after their names since there is no information about gender in Thomson Reuters. The authors of this thesis had to go through the same procedure in order to sort analysts by gender.

To gain robust results, Kumar (2010) considers quarterly forecasts, even though the main focus is on yearly forecasts. To assure robustness, the study makes several tests. For example, a re-estimation of the accuracy regression is made and the timeframe is expanded. Another example is that Kumar (2010) refers to Clement and Tse (2005) when estimating a specific regression, where only the last analyst forecast for a company within a year is considered. Kumar (2010) also tests for robustness by adding skewness as a control variable, to control for male and female analysts response to skewness in realized earnings. All these precautions prove how thorough Kumar’s (2010) work is. Brown and Mohd (2003) do not consider any robustness
test. However, several sensitivity analyses were made which gives their research more creditability.

When looking at prior literature in the research area of analysts and their forecast accuracy, it is rare to include economic theories. However, the authors of this thesis find it necessary for the research question and hypotheses stated to understand the theory behind the stock market. Therefore, relevant economic theories are presented in this thesis.

3.2 Institutional Background

The financial sector has gone through tremendous amount of changes influenced by the aftermath of the economic crisis that erupted in 2007-2008 (Hõbe, 2015). New legislation and innovations in a technology intensive sector, as the financial sector is, has radically changed traditional banking (Hõbe, 2015).

The financial analyst occupation in Sweden is a competitive profession. To be accepted as an analyst a license is required that is provided by the Swedish company SwedSec Licensiering AB. The organization makes sure people working in the financial sector have the right education and knowledge, by providing a certain license. The purpose of the license is to maintain the public's trust of the companies in the financial sector (SwedSec, 2017).

A financial analyst is usually employed by a brokerage firm, where they make predictions of companies’ future earnings and growths, and sometimes also place buy, sell or hold recommendations for securities. Every year, the Swedish weekly business magazine Affärsvärlden together with Financial Hearing rank the financial analysts and brokerage houses operating in Sweden. The rating system considers three perspectives which are industry knowledge, forecast reliability and recommendations (Financial Hearing, 2016). In the top of the ranking for brokerage houses for the year 2016 big players like Carnegie, Nordea and SEB can be found.

There is also a rating for the top ten most highly rated individual analysts. Ranked number one for the year of 2016 is a female analyst, while the remaining nine analysts are males.

There are two major organizations for financial analysts operating in Sweden, these are Sveriges Finansanalytikers Förbund and CFA Society Sweden. Both
organizations strive to offer valuable services to their members that can help them in their professional careers (Sveriges Finansanalytikers Förbund, 2017 and CFA Society Sweden, 2015). In the annual report for 2015/2016 from CFA, the organization discusses the lack of women in the financial sector. The organization state "Despite strong evidence that a more gender-balanced investment industry would benefit investors, women are still underrepresented throughout our profession" (CFA Society Sweden, 2016). They refer to a report done by Shandilya and Duncan (2015) which shows that only 19 percent of the analysts are women, even though women represent 50 percent of the workforce and college graduates with a business degree. The study was done in the U.S. but Shandilya and Duncan (2015) state that the picture is roughly the same in Europe.

3.3 The Efficient Market Hypothesis (EMH)

To get an understanding of financial analysts and their relevance in the financial sector, a central economic theory called the efficient market hypothesis (EMH) needs to be explained. An efficient market is a market where the price of the shares fully reflects all the information available (Fama, 1969). An important part of the EMH is that the share price of companies instantly reacts to news and adjusts correctly to reflect the new information. If the price of a share adjusts slowly or if it overreacts to the new information, it is not an efficient market (Hillier et al., 2010, p. 351). The EMH means that the investors can only expect a normal rate of return since new information is immediately reflected in the share price and leaves no time for the investors to trade on the news. Fama (1969) claims there is enough evidence to argue that the EMH stands up well. However, the theory is controversial and therefore the authors of this thesis also discuss a contradictory concept called behavioral finance.

By the definition of an efficient market, there would be no need for a financial analyst since all the information is already considered to reflect the price of the shares, which would make the analyst's forecasts to be superfluous. Fama (1969) argues that the market reacts differently to different types of information. He classifies these types of information and their effect on the market efficiency into three different categories: The weak form, the semi-strong form, and the strong form. The different market forms are illustrated in figure 2.
The market is considered to be of a weak form if it fully incorporates the information of past share prices. This type of market is called the weak form since historical price information is the easiest type of information about a company's equity there is to acquire (Hillier et al., 2010, p. 353). A market with the semi-strong form of efficiency is a market where the price reflects all historical information and all public available information. The next form of market efficiency is the strong form which is defined as a market where the price reflects all types of information, both public and private (Hillier et al., 2010, p. 353).

Figure 2 An Illustration of the Relationship Among the Three Types of Information

3.4 Behavioral Finance

In a world where market efficiency exists, there would be no use for a financial analyst. That is, if the market never is under- or overpriced there is no point for financial analysts in trying to predict the market. If that were to be true, there would be no need for this thesis and prior research to identify determinants of the analysts' forecast accuracy. However, the theory of behavioral finance takes stand in investors’ irrationality and the non-randomness of its deviation (Hillier et al., 2010, p. 361), which makes room for forecasts by financial analysts. This theory suggests
that analysts have an important role in the financial market to lead investors in the right direction by making accurate forecasts.

By applying a behavioral finance theory, a description of how human behavior influence individuals, actors and financial markets can be made (Shleifer, 2000, p. 2). Shleifer claims the theory to expect systematic and significant deviation from efficiency and that humans have fallibility in competitive markets (Shleifer, 2000, p. 24).

Behavioral finance is a broad field that proposes complex psychological-based theories that affect investors. For this thesis, the authors need to narrow this theory down to see how behavioral finance is based on the belief that the three conditions supporting market efficiency cannot hold in reality. The conditions are:

- Rationality
- Independent deviation from rationality
- Arbitrage

Shleifer questions peoples' rationality and state that it is common for investors to not diversify sufficiently. Instead, Shleifer (2000, p. 24) refers to the theory of investor sentiment, which is based on how investors shape their beliefs and valuations. Even though the behavioral finance does not view all investors as irrational, it presupposes that many are (Hillier et al., 2010, p. 361).

Behavioral finance questions whether deviations from rationality truly are random. According to Hillier et al. (2010, p. 361), there are two principles that can explain people’s deviation from rationality. The first principle is called representativeness and suggests that when an investor sees a sector on the market with a short historical growth of high revenue, they are under the impression that the growth will continue. The price will eventually stall and fall down. Shleifer (2000, p. 113) defines representativeness as the belief to see a pattern in truly random sequences. The second principle is conservatism and argues that the market seems to adjust slowly when receiving new information. Shleifer (2000, p. 113) suggests that conservatism is the lengthy update of models in the face of recent evidence. Hillier et al. (2010, p. 362) claim that conservatism is due to that investors are slow when calibrating their emotions and beliefs.
Concerning arbitrage, behavioral finance states the opposite of the Efficient Market Hypothesis. Arbitrage is usually referred to as a risk-free investment, it concerns a situation where an investor know that an equity is mispriced and act on that information (Hillier et al., 2010, p. 352). Shleifer (2000, p. 24) argues that in the real world, arbitrage is not remotely perfect. This is due to a lot of securities not having any good substitute which makes arbitrage fundamentally risky and limited. The result of a limited arbitrage describes why prices not always respond to information by the right quantity. According to Hillier et al. (2010, p. 362), the theory of arbitrage may not hold in reality. Even though the market may be over- or underpriced, there is no guarantee that stock prices will move a certain way. They claim investors to be irrational and since amateur investors may be taking opposite positions, market price will only adjust to correct levels if the position size of the professional investors is larger than the amateur investors. Therefore, arbitrage strategies may involve too large of a risk for eliminating market efficiency.

3.5 Determinants of Differential Forecast Accuracy in Prior Literature

3.5.1 The Existence of Differential Forecast Accuracy
The first question that has to be answered when researching determinants of analysts' forecast accuracy, is whether differential forecast accuracy does exist among analysts. In early studies of this issue, a lot of research suggested that there was no evidence of individual financial analysts that could provide consistently more accurate forecasts over multiple years (Sinha, Brown and Das, 1997). Butler and Lang (1991) did not find differential forecasting abilities in their sample when they examined individual analysts. They argue the differences they witnessed in prediction accuracy to depend on the persistence of the individual analyst’s optimism or pessimism. However, a study done by Sinha, Brown and Das (1997) found their results to be contrary to prior findings. What they discovered was that prior research failed to find determinants of differences in analysts forecast accuracy due to a lack of control for the forecasts' recency. Brown and Mohd (2003) argue that these findings of contrary results led to more studies that examined and identified different determinants which could distinguish the good analysts from the bad ones.
3.5.2 The Individual Ability of the Analyst

Similar to Sinha, Brown and Das (1997), Mikhail, Walther and Willis (1997) conducted a study to control for other variables that could have affected earlier studies' insignificant results. They found evidence that analysts’ forecast accuracy increased together with the analysts’ firm experience. A study by Clement (1999) examined other determinants that could explain forecast accuracy. He argues that forecast accuracy increase with experience, both firm experience and the work experience of the analyst. Clement (1999) compares the analyst labor market to a tournament, where only the stronger performers continue and the weaker are forced out of the occupation. This could be one explanation of why more experienced analysts seem to outperform those with less experience. The firm-specific skills the analyst acquires over time seem to also play a role in the analyst’s forecasts accuracy. Clement (1999) argues that over time, the analyst gets a better understanding of the company’s reporting practices and might also establish relationships with company insiders. These information advantages lead to more accurate forecasts. Clement’s (1999) findings have been confirmed by later research (Brown and Mohd, 2003., and Clement, Koonce and Lopez, 2007).

3.5.3 Task Complexity

Clement (1999), Brown and Mohd (2003) and Clement, Koonce and Lopez (2007) all studied the task complexity of the analyst's work and its effect on forecast accuracy. Task complexity is measured by the number of firms the analyst follows and the number of industries the analyst covers. Clement (1999) and Clement, Koonce and Lopez (2007) found significant evidence that both firm complexity and industry complexity affect the analyst’s forecast accuracy negatively. Clement (1999) argues that larger portfolios allow the analyst to spend less time and attention to each firm and industry. Brown and Mohd (2003) also found industry complexity to have a negative effect on the analyst forecast accuracy. However, they did not find any significant relationship between firm complexity and forecast accuracy.
3.5.4 Work Environment

Clement (1999) wanted to examine whether the size of the brokerage house had any effect on forecast accuracy. He argues that a larger brokerage house where more analysts are employed should have superior resources compared to smaller brokerage houses, which should contribute to better forecast accuracy. Clement (1999), Brown and Mohd (2003) and Clement, Koonce and Lopez (2007) found significant evidence in their studies that an analyst working at the top decile brokerage houses makes more accurate forecasts.

3.5.5 Summary of Prior Research

Table 1 illustrates how forecast accuracy is affected by independent variables in prior studies. An empty box shows that the independent variable was not a part of the research. A plus sign in the table indicates that the independent variable had a positive significant effect on forecast accuracy while a minus sign shows that the independent variable had a negative significant effect on forecast accuracy. A zero indicates that there was no proven significant relationship between the independent variable and forecast accuracy. The table also illustrates how the research in this area has developed over time by researchers adding variables.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Forecast Age</th>
<th>Analyst experience</th>
<th>Firm Experience</th>
<th>Firm Complexity</th>
<th>Industry Complexity</th>
<th>Brokerage Size</th>
<th>Forecast Frequency</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sinha, Brown &amp; Das (1997)</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mikhail, Walther &amp; Willis (1997)</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clement (1999)</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brown &amp; Mohd (2003)</td>
<td>-</td>
<td>+</td>
<td>0</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Clement, Koonce &amp; Lopez (2007)</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Kumar (2010)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>

*Table 1 An Illustration of Findings in Prior Research*
3.6 Gender Characteristics and its Effect on Prediction Accuracy

According to the gender socialization theory, male and female characteristics are the reason there is a difference in male and female behavior (Betz, O’Connell and Shepard, 1989). The article has confidence in that decisions and behavior derives from values and traits brought by gender. The gender socialization theory originates from the psychoanalysis of Freud and social learning theory. While Freud argues gender behavior being biologically determined, the social learning theory claims gender behavior being connected to gender roles produced by social factors (Betz, O’Connell and Shepard, 1989). The point of view for the social learning theory is that gender characteristics and behaviors are learned, in opposite to Freud’s theory of biological determination.

A study done by Sapienza, Zingales and Maestripieri (2009) states that women are on average less prone to take on risk when it comes to financial decisions, compared to men. Niederle and Vesterlund (2007) claim women to be less competitive. Assuming they are, the consequences are fewer women entering and winning competitions. As a result, women's chances of being successful in competitions for promotions and lucrative positions is decreasing. This may be one explanatory factor for the under-representation of women in the financial sector.

Kiecker, Palan and Areni, (2000) argue that there is a difference how males and females process information. In their study, they created a test to analyze how men and women differ when they were given a coding assignment involving written narratives. The authors argue that due to the character of females’ knowledge structure, they are more comprehensive processors. Also, females are more sensitive and attentive to details. It is further argued that this comprehensive and detailed processing leads female to identify more themes and concepts, compared to males (Kiecker, Palan and Areni, 2000). The results from the study are relevant and can be connected to the financial analyst occupation.

Kumar (2010) wanted to expand an area of research he witnessed to be emerging in accounting and finance literature, the importance of gender. He argues that there was a research gap to fill when examining financial analysts and the gender equality issue. In his research, he proposes that due to a perception of discrimination in the male dominant analyst occupation, only the most competitive group of women is likely to pursue a career. This would suggest that female analysts are unlikely to be
representative of the female population. Kumar (2010) states that due to this self-selection among women in the analyst occupation, female analysts is a group that is superior in comparison to their male associates.

Kumar (2010) claims to have found evidence that female analysts issue bolder and more accurate forecasts than their male colleagues. The results from his study proved to be robust and can therefore not be explained by a non-random selected sample.

### 3.7 Hypothesis Development

Prior studies prove there are determinants that can help explain financial analysts’ forecast accuracy. However, even though researchers have been able to establish variables that can help to explain forecast accuracy, they have never been tested and analyzed with only a Swedish sample. This thesis will state five different hypotheses in order to answer the research question. All the hypotheses derive from prior literature.

Prior literature has found inconclusive results regarding the forecast accuracy by financial analysts and the number of firms they are covering. Clement (1999) and Clement, Koonce and Lopez (2007) found evidence of a negative relationship between forecast accuracy and an increasing number of firms covered by the analyst. Brown and Mohd (2003) found no significant result when analyzing the same relationship. The authors of this thesis expect firm complexity to have a negative effect on forecast accuracy regarding analysts covering Swedish stocks:

**H1:** Forecast accuracy decreases with the number of firms followed by the analyst.

According to Clement (1999) and Brown and Mohd (2003), industry complexity of the analysts' work has a negative effect on forecast accuracy. By increasing the number of industries followed by the analyst, their expertise in each of the industries suffers (Clement 1999). The authors of this thesis test if the same negative association can be found for financial analysts covering Swedish companies:

**H2:** Forecast accuracy decreases with the number of industries followed by the analyst.

Prior findings agree on a positive association between the analysts working at a top decile brokerage house and the forecast accuracy (Clement, 1999., Brown and Mohd,
2003 and Clement, Koonce and Lopez, 2007). Clement (1999) suggests that the reason analysts working for a larger brokerage house outperform analysts working for a smaller brokerage house is mainly due to superior resources that can contribute to better forecast accuracy. The authors of this thesis test if the same association can be found for financial analysts covering Swedish companies:

H3: Analysts who are employed by large brokerage houses make more accurate forecasts than analysts employed by small brokerage houses.

Clement (1999) was first in the area to add the work experience of the analyst as an independent variable for forecast accuracy, with a predominately positive result. Clement, Koonce and Lopez (2007) have also done research in the field and by analyzing an unspecified sample they found a positive association for analyst experience and forecast accuracy. It is reasonable to believe that the analysts, who have worked the longest and gained the most experience, perform better and make more accurate forecasts than analysts who have less work experience. The authors of this thesis test if the same positive result can be found with analysts covering Swedish companies:

H4: Forecast accuracy increases with the analyst’s forecasting experience.

Kumar (2010) argues that female analysts outperform their male colleagues. In a different study, Kiecker, Palan and Areni, (2000) claim that females are more comprehensive and detailed when processing information, which may be an explanatory factor for Kumar's (2010) findings. Supported by prior research, the authors of this thesis aim to clarify if female analysts outperform their male colleagues in Sweden and thus, gender can be a determinant for forecast accuracy:

H5: Forecast accuracy increases if the analyst is of female gender.
4 Empirical Method

This chapter includes data collection, sample and empirical models. The chapter also entails a variable description. Finally, the methods for testing the data are explained.

4.1 Data Collection

4.1.1 Primary and Secondary Data

Data analysis can refer to primary or secondary data. According to Bryman (2012, p. 13), primary data is described as when the researcher both collects and analyzes the data first hand. Secondary data, on the other hand, has already been collected by someone else. The benefits of using secondary data are its time efficiency and the high quality of data which is generally true if it is collected from a reliable source (Bryman, 2012, p. 312-313). For these reasons, this thesis is based on secondary data.

The authors of this thesis use caution when selecting the secondary sources. The authors stress the importance of a critical approach when collecting the data, to assure the quality of the thesis. Therefore only reliable and unbiased data will be used. The secondary data collected for this thesis will be analyzed with IBM SPSS software. The authors are experienced using this software which is acknowledged for statistical research.

4.1.2 Thomson Reuters

Secondary data will be collected from the Thomson Reuters Institutional Brokers Estimates System (I/B/E/S) database, which is consistent with Brown and Mohd's (2003) and Kumar's (2010) studies. Thomson Reuters is an acknowledged data supplier which has more than 100 years of experience. Thomson Reuter has a strict regulation and can therefore assure reliable and unbiased data (Thomson Reuters, 2017). Furthermore, they supply data from 930 brokers. The database provides information regarding forecasts for earnings, cash flows, revenues and stock recommendations for more than 22,000 companies.
4.1.3 OneSearch

Relevant articles for this thesis have been collected from OneSearch, provided by Linnaeus University. OneSearch is Linnaeus University’s search engine. It consists of a wide range of published articles, books, magazines and theses. The authors have selected articles that have been published in relevant journals with high rankings.

When searching for published articles, following key words were used:

- Determinants of analyst forecast accuracy
- Financial analyst
- Financial analyst characteristics
- Forecast accuracy
- Gender
- Gender differences
- Sell-side analyst

The carefully selected articles are used throughout the introduction- and theory chapter, with the purpose of supplying a foundation to the thesis. Prior research will also be referred to in the fourth chapter when this thesis research model is explained. Finally, to make conclusions from this research, the results need to be compared to prior findings.

4.2 Sample

4.2.1 Population and Sample

A population is the total amount of units selectable for the research area of interest. From the population, a sample is selected that is examined in the research and the goal is to use a sample representable for the full population (Bryman, 2012, p. 187). Bryman (2012, p. 197) argues that the larger the sample, the more precise the result. In contrast to Brown and Mohd (2003) who used a U.S sample and Kumar (2010) who used an unspecified sample for their research, this thesis study's the Swedish market. It means that the sample size is considerably smaller compared to Brown and Mohd's (2003) and Kumar's (2010) studies.

The sample in this research consists of earning forecasts from the years of 2016 and 2017 by analysts covering companies listed on the Swedish Index OMXS30.
OMXS30 is an index that consists of the 30 stocks with the highest turnover on the Swedish stock exchange. The companies included in OMXS30 can be found in appendix 1.1. The reason to limit the sample to companies on one single index instead of all listed companies in Sweden is mostly due to time limitations when writing a bachelor thesis. To access all information needed for this research, the authors had to go through every analyst in the sample individually. This is an issue prior literature did not have to cope with.

The authors’ arguments to use companies from the OMXS30 are that large companies often are more predictable and consistent in their earnings than small companies. In addition, large companies disclose more detailed information and more frequently than small companies (Wolters Kluwer, 2016). Finally, there is more information to gather for analysts covering larger companies.

The authors started off with a sample consisting of 634 individual analyst forecast scores. The first step in the data sorting process was to sort data by gender and to delete the observations lacking gender information. Thereafter, the authors excluded data with missing values for other variables. Finally, five cases of extreme outliers were observed and deleted. In total, an amount of 115 observations had been deleted from the sample. The final sample consists of 519 individual forecast scores for 284 analysts. 8.3 percent of the forecasts were made by female analysts.

4.2.2 Outliers

Outliers are defined as observations that are significantly different from the other observations. What distinguishes an outlier is the extreme high or low value for the variable (Hair et al., 2014, p. 62). The univariate method is a way of detecting outliers and examines the distribution of each variables observations. Hair et al. (2014, p. 65) describe when outliers should be retained or deleted from the sample. It is suggested that only outliers where there is evidence that they are genuinely aberrant should be deleted. If not deleted, the extreme values can influence the result in an incorrect way (Hair et al., 2014, p. 63).

Brown and Mohd (2003) deleted observations where the forecasts of the companies’ earnings were outside of the 1st and the 99th percentile. Kumar (2010) made a similar choice for his study when he decided to delete all data in the independent variables
that were outside of the 0.5\textsuperscript{th} and the 99.5\textsuperscript{th} percentile. This thesis observes outliers for each of the variables by using a boxplot. However, in line with Hair et al. (2014, p. 65) only extreme outliers will be deleted to ensure the sample's generalizability to the entire population. Outliers are identified as observations outside both the boxplot and its whiskers. To find outliers the inner range, the difference between quartile 1 and quartile 3, is multiplied by 1.5. The score is subtracted to the first quartile and added to the third quartile. Any observation outside this range is called an outlier. However, this thesis only deletes extreme outliers. These are found by multiplying the inner range with 3, the score is deducted from the first quartile and added to the third quartile. Any values outside this range are defined as extreme outliers.

The sample of this thesis consisted of five extreme outliers. One observation had an extreme outlier regarding the number of firms followed by the analyst. Analyst experience had four observations with extreme outliers. The five observations with extreme outliers were deleted from the sample and are not included in the descriptive statistics presented in this thesis.

4.2.3 Missing Values

It is common for data to contain missing values which raise an issue for the generalization of the results (Hair et al., 2014, p. 40). Missing values has different impacts on data. First, missing values reduce the number of observations and can make an adequate sample inadequate. Second, results based on a nonrandom missing data process can be biased. Hair (2014, p. 46) argues for deletion of the observations if a variable has missing values equal to or lower than 15 percent. The authors will based on this delete all observations with missing values since none of the variables have more than 15 percent of missing values.

Hair et al. (2014, p. 53) describes balanced data as a research strategy. The strategy suggests deletion of all observations consisting of a missing value. An advantaged of balanced data is the simplicity to implement and it is suitable for a large sample size where there are low levels of missing values. On the contrary, a disadvantage of the strategy is that it may result in a great reduction in sample size. A balanced data approach is selected for the thesis, all observations with missing values are deleted.
4.3 Empirical Model

This thesis is based on the model created by Brown and Mohd (2003). Their model is acknowledged in the research field of analysts’ forecast accuracy and most of the variables included in their model have proved to have a significant effect on forecast accuracy. This motivates the authors to use Brown and Mohd's (2003) model as a base for their research.

4.3.1 Brown and Mohd's Empirical Model

\[ PMAF = \beta_0 + \beta_1 \cdot DAGE + \beta_2 \cdot DFEXP + \beta_3 \cdot DNCOS + \beta_4 \cdot DNSIC2 + \beta_5 \cdot DTOP10 + \beta_6 \cdot DFREQ + \mu \]

The aim of Brown and Mohd’s (2003) research was to create a model which could explain forecast accuracy better than the one variable forecast age. The model consists of the dependent variable forecast error and six different independent variables. The independent variables are forecast age, firm experience, firm complexity, industry complexity, brokerage size and forecast frequency. Every independent variable except firm complexity proved to have a significant effect on the analyst's forecast accuracy.

Brown and Mohd (2003) make a suggestion for future research to identify additional determinants that enhance analysts’ forecast accuracy (Brown and Mohd, 2003).

4.3.2 Regression Model

\[ PMAF = \beta_0 + \beta_1 \cdot DNCOS + \beta_2 \cdot DNSIC2 + \beta_3 \cdot DTOP10 + \beta_4 \cdot DGEXP + \beta_5 \cdot GEND + \mu \]

The authors of this thesis aim to recreate Brown and Mohd's (2003) model and to use it for a Swedish sample. However, this thesis will not be able to control for forecast age which has proved to be the most important independent variable in prior studies. This is due to a lack of available data. Two other variables were also dropped because of the same reason, these are firm experience and forecast frequency. The authors are aware of the problem with omitting variables, which may affect the model in a negative way.
Two independent variables have been added to Brown and Mohd's (2003) model to identify additional determinants for forecast accuracy. A variable measuring the analyst’s work experience is included from Clement, Koonce and Lopez’s (2007) research study. Since one of the hypotheses aims to examine if the gender of the analyst has an effect on forecast accuracy, the authors have decided to also include a gender dummy variable.

Each independent variable in the regression model answers each one of the five hypotheses stated. None of the models mentioned have been tested on a Swedish sample and the authors are the first to test determinants of forecast accuracy for analysts covering Swedish companies.

4.3.3 Variable Description
Explained by Bryman (2012, p. 48) a variable is an attribute of a case that varies. If the case does not vary it means it is a constant, which are rarely of interest to researchers. There are different types of variables. First, the differences between dependent variables and independent variables need to be distinguished. Independent variables are deemed to have a causal effect on the dependent variable (Bryman, 2012, p. 48).

To be able to answer the research question of this thesis, several dummy variables were created. A dummy variable is dichotomous which means it has only two different values (Bryman, 2012, p. 335). A dummy is created when a non-metric variable needs to be converted to a metric variable.

The data and information for all the variables are collected from the I/B/E/S database. All of the variables have been used in prior research. The way of measuring the analysts' forecast accuracy in this thesis is however not consistent with prior literature, due to limitations of available data. Prior research expressed forecast accuracy as absolute forecast error. The absolute forecast error is measured as the difference between the predicted earnings and the actual earnings of the company. In this thesis, the absolute forecast error is part of the calculation done by Thomson Reuters to give the analyst a justified score depending on their forecast accuracy. The authors argue that the method used in this thesis is an adequate substitute since the accuracy score of the forecasts is a thorough calculation done by Thomson Reuters.
Following is a description of the variables used in the model of this thesis:

- **Forecast Accuracy (PMAF) - Dependent Variable**

  Analyst forecast accuracy is calculated as a score between 1-100, where 1 is the worst possible score and 100 is the best possible score. A score of 50 is the mean score of forecast accuracy for all analysts covering the company. The base of each analyst score is a calculation depending on four factors: the company's actual reported earnings, the analyst's absolute forecast error, the analyst's forecast error compared to the forecast error of other analysts following the company, and the variance of the forecast errors. The score is presented on a summary basis, which is an aggregate of all the analysts' scores for the four most recent quarters reported.

- **Firm Complexity (DNCOS) - Independent Variable**

  The number of firms the analyst followed during the four quarters that is the basis of the analyst's forecast accuracy score.

- **Industry Complexity (DNSIC2) - Independent Variable**

  The number of industries the analyst followed during the four quarters that is the basis of the analyst's forecast accuracy score. Industry is defined by Thomson Reuters in the I/B/E/S database and can be found under the analyst's performance scorecard.

- **Brokerage Size (DNTOP10) - Independent Variable**

  Dummy variable that is equal to 1 if the analyst worked at a top decile brokerage house and 0 if otherwise. A top decile brokerage house is ranked by measuring the number of employees that work for the broker. Those in the top ten percent are considered to be in the top decile of brokerage houses. The information of the number of employees for each of the brokerage houses is collected from their annual reports and by direct contact.

- **Analyst Experience (DGEXP) - Independent Variable**

  The forecasting experience of the analyst. It is calculated by the number of years the analyst has been registered in the I/B/E/S database.
• Gender (GEND) - Independent Variable

Dummy variable that is equal to 1 if the analyst's gender is female and 0 if otherwise. The gender of the analysts has been determined by identifying them after their names. The names are provided by Thomson Reuters and manually sorted by the authors.

4.4 Method of Analysis

There are several different techniques for analyzing quantitative data. It is important to decide on what technique to use when designing the sample selection. This is because not all techniques can be applied to any variable and because the nature and size of the sample will inflict limitations of usable techniques (Bryman, 2012, p. 330).

Multivariate analysis is statistical techniques that simultaneously analyze more than two variables (Hair et al., 2014, p. 4). To decide what type of multivariate analysis technique to use, Hair et al. (2014, p. 11) suggest the researcher to make three judgments about the research objective and the nature of data. The first judgment is to decide if one or several variables can be identified as a dependent variable. The aim of this thesis is to find determinants for forecast accuracy, which means one dependent variable is identified. The next step is to determine how many of the variables will be treated as dependent variables. The research question in this thesis is answered by examining one dependent variable. The last step is to identify how the variables, both dependent and independent, are measured (Hair et al., 2014, p. 11-12). In this thesis, all the variables are either metric by nature or transformed into metric variables. After making the three judgments, the authors of this thesis have decided that a multiple regression is the appropriate method of multivariate analysis.

Most of the prior literature in this research area have tested their data by using the multivariate analysis technique multiple regression (Clement, 1999., Brown and Mohd, 2003., Clement, Koonce and Lopez, 2007 and Kumar, 2010). The authors of this thesis run their regression by using ordinary least squares (OLS) estimators. It is the most powerful and popular method of regression analysis (Gujarati, 2004, p. 58) and all prior researchers, except Kumar (2010) have analyzed their data using OLS estimates.
4.4.1 The Assumptions of Multivariate Analysis

Multivariate techniques are based on a set of fundamental assumptions that are representing the requirements of the underlying statistical theory (Hair et al., 2014, p. 69).

Normality is the most important assumption in multivariate analysis. According to Hair et al. (2014, p. 69), "If the variation from the normal distribution is sufficiently large, all resulting statistical tests are invalid". To measure the normal distribution of the data, two measurements called kurtosis and skewness are used. Kurtosis measure the peakedness or flatness of the distribution and skewness is used to measure the balance of the distribution (Hair et al., 2014, p. 69). To test the normality of the distribution, the authors will look at the values of skewness and kurtosis. A normal distribution should have a skewness value of zero and a kurtosis value of three (Gujarati, 2004, p. 148).

Homoscedasticity is the second assumption in multivariate analysis. In a sample with a homoscedastic relationship, the dependent variable display equal levels of variance across the range of independent variables (Hair et al., 2014, p. 72). Homoscedasticity is preferable because the variance of a dependent variable should not be explained by a limited range of independent values (Hair et al., 2014, p. 72). Hair et al. (2014, p. 73) suggest researchers to examine homoscedasticity graphically by using box plots where the length of the box and its whiskers represent the variation of the data. If the box plots do not differ too much in size and shape between the groups of variables, homoscedasticity can be assumed.

The third assumption is linearity. This is an implicit assumption since correlations only represent the linear relation between variables. The easiest way to identify linearity is to examine scatterplots of each of the variables (Hair et al., 2014, p. 74).

Absence of correlated errors is the last assumption. It means that the researcher should attempt to ensure that the prediction errors are uncorrelated with each other (Hair et al., 2014, p. 74). To identify if correlated errors exist, the researchers have to identify patterns in the error terms. If patterns can be identified, the researchers cannot be certain that the prediction errors are independent.
The authors of this thesis test the data for normality. Hair et al. (2014, p. 75) argue that a sample size bigger than 200 observations usually is not affected too much by non-normal distribution. For the other assumptions, the authors will assume that the variables are in line with the requirements for statistical theory.

4.4.2 Multicollinearity
When testing research with a multiple regression, there is always a risk of correlation among the independent variables. The ideal situation of a multiple regression model is that there is a high correlation between the independent variables and the dependent variable, but with little correlation among the independent variables themselves (Hair et al., 2014, p. 196). The correlation among the independent variables is called multicollinearity. When there is multicollinearity, the independent variables capability of predicting the dependent variable decreases (Hair et al., 2014, p. 197).

A way to measure multicollinearity is the variance inflation factor (VIF). VIF is expressed in a number, where a higher VIF value indicates a higher degree of multicollinearity. It is complicated to determine what an acceptable degree of multicollinearity is. Hair et al. (2010, p. 200) suggest that for a large sample size, a VIF score above ten indicate high multicollinearity. The authors of this thesis examine the degree of multicollinearity for each of the independent variables and allow a maximum VIF value of ten.

4.4.3 Significance Level
Whether the hypotheses stated in this thesis will be rejected or not depends on the level of statistical significance. This is called the alpha level (Gujarati, 2004, p. 136). Alpha indicates where researchers set the limit of error, which is the probability of rejecting the null hypothesis when it is actually true (Hair et al., 2010, p. 9).

Prior studies use different levels of statistical significance. Clement (1999) use an alpha of 0.1 while Brown and Mohd (2003) use levels of 0.05 and 0.01. A lower level of alpha gives the author an assurance of not rejecting the null hypothesis when it is actually true. However, a lower level of alpha also reduces the statistical power. With a low level of alpha, the chances of finding a significant result decreases (Hair
et al., 2014, p. 9). The authors of this thesis have decided to present all the three significance levels (1%, 5% and 10%) for the correlation test and to use the conventional alpha level of 5% for the multiple regression analysis.

4.4.4 Correlation Test
In addition to a multiple regression, the authors examine the variables by using a Pearson correlation test. Correlation is explained as the strength of a relationship between two variables. The correlation is measured by a scale of +1 to –1. A value of +1 is defined as a perfect positive correlation, which means the variables move identically. A correlation coefficient of –1 means that the variables move in totally opposite directions. A value of, or close to 0 suggest a weak or nonexistent correlation between the variables (Saunders, Lewis and Thornhill, 2016, p. 545).

Saunders, Lewis and Thornhill (2016, p. 546) claims that Pearson’s correlation coefficient can be used to test the strength of a relationship between two variables when the variables consist of numerical data. Other correlation tests are Spearman’s rank correlation coefficient and Kendall’s rank correlations coefficient, both tests are for samples containing ranked data.

4.4.5 The Quality of the Regression Model
To get a measurement of how well the variance of the dependent variable is predicted by the regression model, coefficient of determination, $R^2$, is observed. The $R^2$ value is expressed as a score between 0 – 1 where 1 equals a perfect fit of the sample regression line (Gujarati, 2004, p. 81).

There is a problem that occurs regarding the $R^2$ value when running a multiple regression. The $R^2$ value always increases when variables are added to the model, even if the variables are not significant. To avoid this problem, a better measurement for the goodness of fit can be used, this is called the adjusted $R^2$. This measurement adjusts for the number of independent variables relative to the sample size. While the $R^2$ value can only increase when variables are added to the model, the adjusted $R^2$ value can either increase or decrease when variables are added to the model depending on how relevant they are (Hair et al., 2014, p. 189). It is difficult to
determine if the adjusted $R^2$ value is high or not. The authors of this thesis are using the adjusted $R^2$ value to compare with prior research.

When running a regression, an ANOVA test is done simultaneously. The ANOVA test gives information about the whole regression model. One of the information given by the ANOVA test is the F ratio. The F ratio is used to test if the regression model is statistically significant. Statistical significance of the regression model is desirable since it means that the researchers can be confident that the model is not specific just to the sample tested (Hair et al., 2014, p. 188).

Since there are variables from prior research that will not be included in the regression model of this thesis, the authors expect a different goodness of fit than prior studies.
5 Results and Analysis

*In this chapter, the authors present the results and their analysis. The chapter includes descriptive statistic, Pearson's correlation coefficient and results from the regression model. The chapter ends with a summary of the findings.*

5.1 Descriptive Statistic

As shown in table 2, the average analyst covering companies listed on OMXS30 has a forecast accuracy score of 52.09, follows 13.35 companies and 7.08 industries. Furthermore, the average analyst covering OMXS30 does not work at a top decile brokerage house, has 5.18 years of experience and is male.

The standard deviation of forecast accuracy is 21.431, firm complexity is 6.398 and industry complexity is 4.305. Also, the analyst's experience has a standard deviation of 5.123. The standard deviation of the dummy variables brokerage house and gender will not be interpreted due to its irrelevance.

The forecast accuracy has a skewness of -0.227, which is close to a normal distribution, 0, merely with a distribution skewed slightly to the right. Firm complexity has a skewness of 1.045, industry complexity has a skewness of 0.984 and analyst experience has a skewness of 1.764. These numbers are interpreted as a skewed left distribution. The balance of the distribution for the dummy variables brokerage house and gender will not be interpreted due to its irrelevance.

The kurtosis of forecast accuracy is –0.104 and its distribution is therefore interpreted as flatter than a normal distribution, which is 3. For firm complexity, the kurtosis is 1.060 and for industry complexity, 0.958. Thus, the distribution for firm complexity and industry complexity is interpreted as flat. On the other hand, the distribution of analyst experience has a peak with a kurtosis of 3.460. The peakedness of the distribution for the dummy variables brokerage house and gender will not be interpreted due to its irrelevance.

The authors recognize that the variables are not normally distributed. Due to the large sample size and the small deviation from a normal distribution, the authors have decided not to mean adjust the data to obtain a more normal distribution. This is contrary to prior research.
Table 2 Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>PMAF</th>
<th>DNCOS</th>
<th>DNSIC2</th>
<th>DNTOP10</th>
<th>DGEXP</th>
<th>GEND</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>52.090</td>
<td>13.350</td>
<td>7.080</td>
<td>0.140</td>
<td>5.180</td>
<td>0.080</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>21.431</td>
<td>6.398</td>
<td>4.305</td>
<td>0.348</td>
<td>5.123</td>
<td>0.279</td>
</tr>
</tbody>
</table>

Distribution

<table>
<thead>
<tr>
<th></th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.227</td>
<td>1.045</td>
</tr>
</tbody>
</table>

Percentiles

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Median</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>40.00</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 3 Pearson Correlation Coefficient

The results from the correlation test are mostly inconclusive with Brown and Mohd's (2003) findings. In their study, they found forecast accuracy to have a significant positive correlation with brokerage size and significant negative correlation with
industry complexity. In this thesis, brokerage size has a negative correlation with forecast accuracy and industry complexity has a positive correlation with forecast accuracy, but these correlations are insignificant.

Consistent with Brown and Mohd (2003), this thesis found a positive insignificant correlation between forecast accuracy and firm complexity. The positive correlation between forecast accuracy and analyst experience is consistent with Clement's (1999) findings. However, Clement found the correlation to be statistically significant which this thesis did not.

Kumar (2010) did not present a correlation test in his research. The significant correlation between forecast accuracy and gender found in this thesis can therefore not be compared with any prior research.

There are several independent variables that have significant correlations among each other. Firm complexity is significantly correlated at the one percent level with both firm complexity and analyst experience, and is correlated at the five percent level with gender. Additionally, industry complexity is significantly correlated at the one percent level with analyst experience, and is significantly correlated with gender at the five percent level. These significant correlations among the independent variables might indicate high levels of multicollinearity.

5.3 Multiple Regression

5.3.1 Model

The adjusted $R^2$ for this thesis model is 0.009, which can be found in table 4. This suggests that the model is weak when describing the dependent variable, forecast accuracy. This result is in line with Brown and Mohd's (2003) model which had an adjusted $R^2$ value of 0.013. The lower level of adjusted $R^2$ in this thesis may be due to the omitted variables. The two added variables, analyst experience and gender, do not explain forecast accuracy better than the three omitted variables forecast age, firm experience and forecast frequency.
<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.136 a</td>
<td>0.019</td>
<td>0.009</td>
<td>21.335</td>
</tr>
</tbody>
</table>

a. Predictors: (Constant), GEND, DGEXP, DNTOP10, DNSIC2, DNCOS

Table 4 Model Summary

The results from the ANOVA table 5 show that the model is insignificant at the five percent level, which is the level of significance this thesis applies to the regression model. The model would however be significant if the authors applied a ten percent level of significance.

The insignificant result indicates that the model used for this research is not a good fit for the data. The model does not explain the dependent variable better than the intercept. It is reasonable to believe that the insignificance of the model is due to the omitted variables from prior research.

ANOVA a

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Regression</td>
<td>4417.914</td>
<td>5</td>
<td>883.583</td>
<td>1.941</td>
</tr>
<tr>
<td></td>
<td>Residual</td>
<td>233499.647</td>
<td>513</td>
<td>455.165</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>237917.561</td>
<td>518</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. Dependent Variable: PMAF
b. Predictors: (Constant), GEND, DGEXP, DNTOP10, DNSIC2, DNCOS

Table 5 ANOVA Table

As expected from the correlation test, there is some multicollinearity among the independent variables. The VIF is varying from 1.015 to 2.326, the values can be found in table 6. The authors decide to not regard this as a problem since the VIF values are relatively low. The authors also stated earlier that they would accept multicollinearity up to VIF values of 10.
5.3.2 Firm Complexity
The firm complexity of the analyst's work has a negative effect on forecast accuracy. This means that adding another firm to the analyst’s portfolio negatively affects the analyst’s forecast accuracy. The findings are however not significant at the five percent level which is consistent with Brown and Mohd's (2003) findings. The results are inconclusive with Clement (1999) and Clement, Koonce and Lopez (2007) who found firm complexity to have significant negative effect on forecast accuracy. Due to the insignificant result, the authors fail to reject the null hypothesis for H1.

5.3.3 Industry Complexity
Industry complexity of the analyst's work has a positive effect on forecast accuracy. This means that adding another industry to the analyst's portfolio positively affects the analyst's forecast accuracy.

There is no significance for industry complexity at the five percent level and the result for the variable is inconclusive with prior studies (Clement, 1999., Brown and Mohd, 2003 and Clement, Koonce and Lopez, 2007) who found significant negative results. The authors fail to reject the null hypothesis for H2 because of the lack of significance.

5.3.4 Brokerage Size
If the analyst works at a top decile brokerage house, the forecast accuracy is affected negatively. Brokerage size has an insignificant effect on forecast accuracy, at the five
percent level, which is inconclusive with prior research's findings (Clement, 1999., Brown and Mohd, 2003 and Clement, Koonce and Lopez, 2007). Prior studies have found brokerage size to have a significant positive result on forecast accuracy. Since brokerage size is statistically insignificant, the authors fail to reject the null hypothesis for H3.

5.3.5 Analyst Experience
The analyst's experience has a positive effect on forecast accuracy. This result suggests that an analyst who has more years of work experience makes more accurate forecasts than an analyst with fewer years of work experience.

Analyst experience is not significant at the five percent level, which contradicts Clement's (1999) and Clement, Koonce and Lopez’s (2007) previous studies where they found significant positive results for analyst experience on forecast accuracy. Due to the insignificant result of analyst's experience, the authors fail to reject the null hypothesis for H4.

5.3.6 Gender
Gender has a significant positive effect on forecast accuracy at the five percent level. A forecast made by a female analyst is more accurate than a forecast made by a male analyst. These findings are consistent with Kumar (2010). Since a statistically positive significance is found, the authors reject the null hypothesis and accept H5.

<table>
<thead>
<tr>
<th></th>
<th>-</th>
<th>+</th>
<th>-</th>
<th>+</th>
<th>+</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNCOS</td>
<td>-0.086</td>
<td>0.395</td>
<td>-1.182</td>
<td>0.069</td>
<td>9.270</td>
</tr>
<tr>
<td>DNSIC2</td>
<td>0.700</td>
<td>0.227</td>
<td>0.667</td>
<td>0.716</td>
<td>0.006</td>
</tr>
</tbody>
</table>

*Table 7 Coefficients of the Regression Model*
The results for the hypotheses are summarized in Table 8:

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1: Forecast accuracy decreases with the number of firms followed by the analyst.</td>
<td>Failed to reject the null hypothesis</td>
</tr>
<tr>
<td></td>
<td>H1 is not accepted at the five percent level</td>
</tr>
<tr>
<td>H2: Forecast accuracy decreases with the number of industries followed by the analyst.</td>
<td>Failed to reject the null hypothesis</td>
</tr>
<tr>
<td></td>
<td>H2 is not accepted at the five percent level</td>
</tr>
<tr>
<td>H3: Analysts who are employed by large brokerage houses make more accurate forecasts than analysts employed by small brokerage houses.</td>
<td>Failed to reject the null hypothesis</td>
</tr>
<tr>
<td></td>
<td>H3 is not accepted at the five percent level</td>
</tr>
<tr>
<td>H4: Forecast accuracy increases with the analyst’s forecasting experience.</td>
<td>Failed to reject the null hypothesis</td>
</tr>
<tr>
<td></td>
<td>H4 is not accepted at the five percent level</td>
</tr>
<tr>
<td>H5: Forecast accuracy increases if the analyst is of female gender.</td>
<td>Reject the null hypothesis</td>
</tr>
<tr>
<td></td>
<td>H5 is accepted at the five percent level</td>
</tr>
</tbody>
</table>

*Table 8 Compilation of Hypotheses and Results*
6. Conclusion and Implications

This chapter includes conclusions based on the results from the previous chapter. Implications affecting the thesis will be discussed and suggestions for future research are presented.

6.1 Conclusion

Determinants of financial analysts' forecast accuracy have been in the interest of both investors and brokerage houses for many years. The purpose of this thesis was to fill a research gap where determinants of analysts' forecast accuracy found in prior research have been applied to examine analysts covering Swedish companies. The authors also aimed to highlight the unequal gender diversity that exists in the financial analyst occupation, where female analysts are widely under-represented.

This research have used a quantitative research strategy to answer the research question "Are there determinants that affect forecast accuracy for financial analysts covering Swedish companies?". By examining prior research, the authors found the task complexity of the analyst's work, the analyst's experience, the work environment and the gender of the analyst were relevant factors to research further.

By running a multiple regression using OLS estimates, this thesis has found both significant and insignificant results. This thesis contradicts prior research by getting insignificant results for firm complexity, industry complexity, size of the brokerage house and analyst experience. Only Brown and Mohd (2003), who also found insignificant results for firm complexity, are consistent with this thesis.

Analysts covering Swedish companies seem to differ from analysts in prior research regarding the effect industry complexity and the size of the brokerage house has on the analysts' forecast accuracy. The firm complexity and experience of the analysts covering Swedish companies affect forecast accuracy similar as prior research found for their sample.

This research suggests it may be valuable for investors searching for analysts with high forecast accuracy, to gather information about the analyst. Analysts covering Swedish companies make better forecasts when they follow more industries and are more experienced. Their accuracy decreases when they add firms to their portfolio.
and if they work at a top decile brokerage house. The results from this model are however statistically insignificant and can therefore not be relied on.

The main contribution of this thesis is the statistically significant result for the gender variable. A positive significant result suggests that the forecast accuracy of female analysts is higher than for male analysts. This is consistent with Kumar's (2010) findings using an unspecified sample. The result is of relevance to investors, whom by recognizing this will have a higher probability of getting more accurate forecasts if they follow female analysts.

This thesis has not examined the reason why female analysts seem to outperform their male colleagues. It is therefore difficult to determine if the under-representations of female analysts are justified or not. However, this result strengthens the argument that brokerage firms are missing out on a lot of talent by not having a gender-diverse workforce.

The authors aimed to find support for the behavioral finance theory by finding significant evidence of determinants that could explain the analysts' forecast accuracy. Therefore, by finding a significant result for genders effect on forecast accuracy, the thesis contributes to the behavioral finance theory.

6.2 Implications

The findings in this thesis differ significantly from prior research. The authors suggest there are several underlying reasons for this. The first reason is the omitted variables and the absence of control for forecast age. To control for age of the forecast were known already in the 1990's to have the capability of changing the entire outcome of this type of research. The added variables, gender and analyst experience, did not have more explanatory power for forecast accuracy than the omitted variables, which resulted in a weaker model.

Second, the authors expect the big difference in sample size for prior research and this thesis to be another reason for the high level of insignificant variables. Brown and Mohd's (2003) sample consisted of 172,837 forecasts, in comparison to the 519 observations for this thesis. The difference in sample size depends on the authors’ difficulties accessing data.
The last reason the authors claim the results to differ from prior research is the different measurement of analyst forecast accuracy. This thesis has relied on Thomson Reuter’s measurements of forecast accuracy where each of the analysts is given a score between 1-100 depending on their performance estimating companies’ future earnings. Prior research had access to each of the analysts' forecasts and could therefore measure the analysts' accuracy performance by the absolute forecast error.

The model used in this thesis proved not to fit the data, which makes the results unreliable. Therefore, the authors cannot be certain that the results from the regression are to be trusted.

6.3 Future Research

This thesis tests the hypotheses by running a multiple regression model. The authors assume that all of the assumptions of multivariate analysis are true. Therefore, a suggestion for further research is to test for linearity between the variables, homoscedasticity and absence of correlated errors, which have not been tested for.

Since this thesis found a significant relationship between female analysts and forecast accuracy, an additional suggestion may be to examine what explanatory factors lay behind the difference between male and female analyst forecast accuracy. It would be interesting to examine if Kumar's (2010) theory of stronger self-selection among women is the underlying reason.

Furthermore, a recommendation for future research is to control for forecast age, firm experience and forecast frequency since they contribute to a model that better explain forecast accuracy.
References


CFA Society Sweden, 2015. (Online) Available at: https://www.cfasociety.org/sweden/Pages/AboutUs.aspx (accessed 26 Apr. 2017)


Appendix

1.1 OMXS30

<table>
<thead>
<tr>
<th>Stocks Included in OMXS30</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABB</td>
</tr>
<tr>
<td>Alfa Laval</td>
</tr>
<tr>
<td>Assa Abloy</td>
</tr>
<tr>
<td>AstraZeneca</td>
</tr>
<tr>
<td>Atlas Copco A</td>
</tr>
<tr>
<td>Atlas Copco B</td>
</tr>
<tr>
<td>Boliden</td>
</tr>
<tr>
<td>Electrolux</td>
</tr>
<tr>
<td>Fingerprint Cards</td>
</tr>
<tr>
<td>Getinge</td>
</tr>
<tr>
<td>Hennes &amp; Mauritz</td>
</tr>
<tr>
<td>Investment Kinnevik</td>
</tr>
<tr>
<td>Investor</td>
</tr>
<tr>
<td>Lundin Petroleum</td>
</tr>
<tr>
<td>Nokia</td>
</tr>
</tbody>
</table>