Good From Far or Far From Good?

- An Experimental Study on Financial Experts Forecasting Ability

Andreas Malmquist[£]
Major in Finance
Stockholm School of Economics

André Rhodin[$]
Major in Economics
Stockholm School of Economics

Abstract

Are stock market professionals better than novices in forecasting stock prices? Forecasting stock prices is a highly complicated assignment which encompasses a large number of tasks. Still, much research points to the fact that highly knowledgeable stock market professionals fail to outperform novices in stock price forecasting. In our thesis, 75 participants performed one-month stock price forecasts for stocks listed on the OMXS30-index. Participants consisted of both stock market professionals and novices. Our results were: (i) experienced stock market professionals make less accurate forecasts than novices; (ii) stock market professionals are not more confident than novices; (iii) confident individuals make less accurate forecasts than less confident individuals; (iv) confidence level and forecasting behaviour cannot be influenced by training.

[£]20727@student.hhs.se [$]20749@student.hhs.se

Master's Thesis in Finance

Tutors: Prof. Högfeldt, Dr. Andersson

Dissertation: 10.15, 2007-09-10

Venue: See Student Web

Discussant: Gyllenhammar, C., Kvick, F., Mertzig, M.

ACKNOWLEDGEMENT

The authors would like to thank Prof. Högfeldt for his inspiring comments and Dr. Andersson for several hours of brainstorming and careful proofreading. We would also like to thank all of our respondents who provided us with the data.

TABLE OF CONTENTS

1. INTRODUCTION....4
2. LITERATURE REVIEW AND HYPOTHESES....6
2.1 The Efficient Market Hypothesis....6
2.2 Factors That Affect the Forecasting Ability....6
2.3 The Experience Dimension....7
2.4 The Confidence Dimension....9
2.5 The Training Dimension....12
2.6 Hypotheses....13
3. METHOD AND DATA....15
3.1 Participants....15
3.2 The Design of the Questionnaire....15
3.3 Incentive Scheme....16
3.4 Independent Variables....17
3.5 Performance Measures....18
3.6 Benchmarking Profiles....21
4. EMPIRICAL FINDINGS – HYPOTHESES TESTING....22
4.1 Descriptive Statistics....22
4.1 Hypothesis 1....24
4.2 Hypothesis 2....27
4.2 Hypothesis 3....28
4.4 Hypothesis 4....28
4.5 Hypothesis 5....30
4.6 Robustness Tests....31
5. DISCUSSION AND SUMMARY....36
6. CRITICISM....38
7. CONCLUSIONS....39
APPENDIX – INDUSTRY REGRESSIONS....45
APPENDIX – THE QUESTIONNAIRE....48

1. INTRODUCTION

Announcements of Mergers and Acquisitions (M&A) and Initial Public Offerings (IPO) are made with great fanfare, and solemn promises are given that the synergies and the good strategic fit will increase the shareholder value. Consequently, stock market professionals alter their buy or sell recommendations and the price of the company's shares adjusts according to the market's expectations. When a company appoints a new CEO the share price might also change drastically. If the previous CEO has performed badly the board of directors wants to communicate that they have the power to act. This has an impact on the stock market professionals who revise their recommendations. However, it is not certain that any of these examples motivate an increase or decrease of the company's intrinsic value and thus the fair price of the shares.

During the autumn of 2000, Handelsbanken published a prospect for an IPO of the IT-company Adcore and later a buy recommendation linked to Adcores acquisition of Cell Strategy. However, in both cases Handelsbanken strongly overestimated the value of the synergies.[1] Even though this example is extraordinary and the prediction came out during the IT-era. It still reflects how difficult it can be to make accurate stock market predictions and how wrong the analysts were, despite their high level of confidence.

Stock market professionals are constantly exposed to situations where they need to make judgements based on uncertain information. Some factors are considered to influence the outcome and the overall accuracy of their judgements. Factors such as experience in the decision field and knowledge of probability theory should lead to improved judgement skills. Also, an increased complexity of the task and limited time should result in a lower quality of their judgements (Gigerenzer & Todd, 2000). However, the correlation between these aspects and the outcome of the judgements is not always as transparent as one might think. This is evident from the stock market where the prices of stocks are governed by earnings and dividends. The size of future earnings is obviously uncertain and the task of stock market professionals is hence, to some extent, to form judgements on the basis of these uncertainties. Our thesis focuses on the forecasting ability of stock market professionals. This subject contains a number of different dimensions and has been investigated by several researchers (see Wärneryd, 2001). The thesis focuses on the three dimensions; experience, confidence and training. We use a questionnaire to collect the data. This is not a very common approach within finance; nevertheless it is often used within financial psychology. Some dimensions are simple to analyse with data from a questionnaire while others can be difficult. We have chosen these three as they can be analysed with the data from this setup. They have also been used in earlier studies (e.g. Yates, McDaniel, Brown, 1991), thus our

[1] The title for the IPO-prospect was: "Adcore. Strongest Buy in the history of strong buys!". A few months later Adcore paid SEK 370 million for the company Cell Strategy. Handelsbanken expected huge synergies from the acquisition and repeated its strong buy recommendation. This time they wrote "WE ARE CONVINCED THAT THIS IS A GREAT DEAL FOR ADCORE!!!". They sat the price target to SEK 100 which equalled a three times price increase over the current stock price. The same year Adcore became the worst performing stock on the Stockholm Stock Exchange and declined to SEK 30 öre (Ågerup, 2002).

results can be validated and compared to other research. It should be stressed that we examine the forecasting accuracy of individuals. This setup does not necessarily allow us to draw conclusions about the financial markets, as these are largely affected by other factors than the behaviour of stock market professionals.

If experience in the decision field would indeed improve a person's ability to form correct judgements, we would expect to see that stock market professionals are superior to novices in forecasting. However, much research has concluded that this is not the case (Andersson, 2004). Hence, our research question is; *are financial experts better than novices in forecasting stock prices?* It should be noted that stock market professionals may be called financial experts regardless of evidence that they perform well (Andersson, 2004). We will from now on use this terminology and, accordingly denote stock market professionals as financial experts.

The last decades of research has not only focused on the forecasting ability of financial experts, but also on their characteristics and the actual process of forming stock markets predictions. The financial psychology strand studies the psychological mechanisms that underlie many of the judgements that people form. One such psychological mechanism is how confidence affects a person's ability to form judgements under uncertainty. According to Lichtenstein, Fischhoff and Phillips (1982) people are in general subject to some degree of overconfidence, meaning that the individual overestimates his/her ability. History proves that the financial expert in the Adcore case were highly overconfident. The overconfident individual is prone to wishful thinking and self-enhancement, and thus forms probability judgements that are distorted by these self-serving motivations (Griffin & Brenner, 2004). Applied to the financial markets, one proposed effect of overconfident traders is that they will make the market overreact to less relevant, more attention-grabbing information such as a rumour or a prominent news article with strong human interest, while under-react to important information such as corporate earnings (Odean, 1998). Based on these findings our second research question is; *are financial experts more confident than novices and has confidence a deteriorating effect on the forecasting accuracy?*

As in all other areas, the base for expertise in the financial markets should consist of talent and training. However, the effects from training in forecasting stock price developments has often shown to be weak (Staël von Holstein, 1972). One could of course speculate how the predictions of Adcore's stock price had differed if someone had informed them about the difficulties involved in accurate stock market forecasting and overconfidence. Our final research question is hence; *is it possible to change the participant's answers through training?*

There are mainly three aspects that distinguish our thesis from earlier research. Firstly, earlier studies used graduate students as financial professionals (e.g. Yates et al., 1991), whereas we use financial experts with minimum one year of professional experience. Secondly, analysing the effects of training Swedish financial experts has not been done before in a similar context. Finally, the knowledge about the difficulties involved in forecasting has increased and the market conditions have changed significantly, since the early pioneers first presented their research (e.g. Cowles, 1933). Therefore, our goal is also to give the reader an up-to-date review of earlier findings.

2. Literature Review and Hypotheses

In this section we present the basic metric and terminology used throughout our thesis. Firstly, we present the efficient market hypothesis and general factors that affect the forecasting ability. Secondly, we present the relevant research concerning our three research dimensions: experience, confidence and training, as mentioned in the introduction and in the final part of the chapter we present our hypothesis.

2.1 The Efficient Market Hypothesis

Several authors argue that it is difficult to correctly forecast stock price movements (e.g. Yates et al., 1991). One proposed explanation for this modest level of accuracy is that the markets for such securities are efficient and thus stock prices follow a random walk (Fama, 1970; 1991). The stock price is assumed to reflect the intrinsic value of a company. The intrinsic value is the rational expectation of the company's net present value of future free cash flows. If the financial markets indeed are efficient then the stock prices should reflect all public information. When new information is revealed, the prices will immediately adjust so as to reflect this information. This implies that unless a person is an insider in every company and has information which is not known to the public; even the most skilled financial expert could not make accurate stock market forecasts over an extended period of time (Muradoglu & Önkal, 1994).

However, there is disagreement on whether the financial markets are efficient or not. Several phenomena and empirical data have led researchers such as Shefrin and Statman (1994) to challenge the efficient market theory. They assume a financial market where both rational traders and noise traders participate. The rational traders collect information rationally whereas the noise traders commit cognitive errors in the judgement process and thus do not act rationally. Had there only been rational traders in the market, the prices would have been efficient. However, the activity of noise traders causes phenomena such as turn-of-year effects and excessive volatility. Wärneryd (2001) argues that investors do not form their judgements independently and for example display herding behaviour. Moreover, there could be delays in how stock prices reflect information. Thus, stock prices might be efficient from a long-term horizon but not in the short-term.

Above mentioned arguments are examples that could make the market theoretically predictable for a rational trader. Thus, it is interesting to see if financial experts can form better judgements than novices as well as investigate how educational information affects the accuracy of their predictions.

2.2 Factors That Affect the Forecasting Ability

It is obvious that complexity and time constraint increases the difficulty of making correct probabilistic judgements (Gigerenzer & Todd, 2000), which in turn increases the difficulty of making forecasts. The design of the feedback, and the time horizon for receiving the feedback, has also been shown to affect the accuracy of the forecast.

In a much cited experiment, Murphy and Brown (1985) examined the accuracy of weather forecasts. Weather forecasting is characterized by immediate and reliable feedback after the judgement. This feedback enables the forecaster to draw valid conclusions and thus obtain a high forecasting accuracy. Similar results were reached by Johnson and Bruce (2001) who performed a non-experimental test in a naturalistic environment on racetrack bettors in the United Kingdom. The analysis displayed that racetrack bettors, who by definition receive regular and unequivocal feedback of their judgements, can make accurate forecasts as a group. A study performed by Andersson (2007) examines the calibration of bookmakers at a gaming company. Andersson concludes that the bookmakers, in general, are well calibrated and capable of making accurate predictions of future outcomes.

In sum, we conclude that the forecasting accuracy is affected by several factors. These factors are present to a varying degree in different professions. Consequently, a great forecasting accuracy in weather forecasting does not imply similar high accuracy in stock market forecasting, regardless of the financial expert's actual judgement and knowledge.

2.3 The Experience Dimension

When a person gains relevant experience one intuitively expects him/her to become more skilled in their profession. Our first research question asks whether stock market professionals are better than novices in forecasting stock prices. Table 1 presents a summary of the relevant literature that we discuss bellow. As one can see, much research within financial psychology concludes that the correlation between experience and accuracy is in fact poor.

Table 1
Summary of previous research and findings of forecasting ability of financial experts.

Authors	Main findings relevant for our thesis
De Bondt (1991)	▪ weak correlation between experience and accuracy
Staël von Holstein (1972)	▪ financial experts are not able to provide accurate forecasts
Yates et al. (1991); Muradoglu and Önkal (1994)	▪ financial experts are worse than novices in forecasting the stock market

Andersson (2001) examined the effect of experience on decision behaviour in lending to small firms. The participants consisted of business students, junior loan officers and senior loan officers. The results from the computerised experiment showed for example that senior loan officers acquired significantly more information cues than the other two groups when forming their judgements. Moreover, the senior loan officers tended to be much more cautious than the novices in the decision process. These findings are in line with Shanteau (1992) who suggests that financial experts are more efficient than novices in identifying relevant information and patterns. Experts are also more skilled in developing decision strategies which help them to systemize their decision making. It should however be stressed that these skills do not automatically translate into improved forecasting skills.

The weak correlation between experience and accuracy for stock brokers, as depicted by Shanteau (1992), is further supported by De Bondt (1991). The author examined approximately 5 400 individual forecasts of the S&P index, made by stock market professionals, between 1952 and 1986. The forecasts had either a three or seven-month time horizon. De Bondt (1991) concluded that the average forecast had no predictive power and was useless for purposes of investment strategy.

Staël von Holstein (1972) examined the ability of financial experts in forecasting the stock price movements over ten two-week intervals. The 72 participants with different occupations were presented with five scenarios for the price change for each stock, and were asked to assess the probabilities for each of these scenarios. However, they were not able to provide accurate forecasts. Furthermore, he compared the results to that of a uniform forecaster. A uniform forecaster is a mechanical forecasting scheme that sets the same probabilities for all intervals. He found that only three (bankers), out of the 72 participants were able to beat this scheme.

As we can see from several studies (Staël von Holstein, 1972; De Bondt, 1991) the correlation between experience and forecasting ability is weak or non-existent. Yates et al. (1991) argued that there exists a negative correlation between experience and forecasting ability, an "inverse experience effect". They performed their study on undergraduate and graduate students in finance, where the latter group was assumed to be more experienced. The participants made probabilistic forecasts of the quarterly changes in the stock prices and earnings for 31 companies listed on the New York Stock Exchange. Consistent with Staël von Holstein (1972), the overall accuracy of both price and earnings forecasts were modest. Moreover, undergraduate students were more accurate than the graduate students. Their explanation for the inverse experience effect was that experts weight more irrelevant information into their predictions. As a person acquires more experience within a domain, he/she expands his/her belief about which signals are predictive. As some of these signals in fact have little or no effect on stock prices and/or earnings, his/her forecasts will be based on more irrelevant information. The inverse experience effect was also found by Muradoglu and Önkal (1994) who used a similar method to Yates et al. (1991) and studied the forecasting accuracy of experts and novices for currency fluctuations.

As shown above the empirical and experimental evidence is not clear-cut but it seems to suggest that the correlation between experience and accuracy is week or even negative, which is contrary to our first research question. It is hence of interest to analyse the financial experts' performance in our context, the Swedish stock market, and compare our results to previous research.

2.4 The Confidence Dimension

Confidence in oneself, or one's ability to perform specific tasks, is a quality with both benign and harmful sides. Taylor and Brown (1988) argue that confidence can lead to higher motivation and persistence which clearly is a positive consequence. However, it is also easy to overestimate one's abilities. When the average confidence level is above the level of performance, it is instead referred to as overconfidence (Andersson, Edman, & Ekman, 2005). Our research question is; *are financial experts more confident than novices?* Table 2 shows a summary of previous research of overconfidence. As one can see, a few authors conclude that experts in fact are more overconfident than novices.

Table 2
Summary of previous research of expert's level overconfidence.

Authors	Main findings relevant for our thesis
Lichtenstein et al. (1982); Fischhoff and McGregor (1982); Keren (1991); Gigerenzer et al. (1991); Odean (1998)	▪ people are systematically overconfident
Glaser et al. (2005); Bhandari and Deaves (2006)	▪ professionals are more overconfident than novices
Odean (1998)	▪ for trivial tasks, underconfidence appears to be dominating
Gervais and Odean (2001)	▪ overconfidence develops dynamically with experience, i.e. it does not follow a linear path
Russo and Schoemaker (1992)	▪ overconfidence does not seem to be prevalent in trivial or familiar tasks but only in complex ones
Dittrich et al. (2005)	▪ overconfidence seems to increase with the difficulty of the task

A large body of research agrees that people, in general, are systematically overconfident about the accuracy of their judgements and knowledge (Lichtenstein et al., 1982; Fischhoff & McGregor, 1982; Gigerenzer, Hoffrage, & Kleinböltin, 1991; Keren, 1991; Odean, 1998). In other words, people are prone to believe that the accuracy of their judgements is more correct than it actually is. Partly, it is a result of that people tend to attribute too much of their success to their own ability rather than to uncontrollable factors and chance, and blame uncontrollable factors when they fail (Odean, 1998; Gervais & Odean, 2001). Overconfident people also tend to believe that their abilities are above average (Svenson, 1981; Taylor & Brown, 1988). Furthermore, they display an unrealistic optimism and a belief that they can control random events (Langer & Roth, 1975; Weinstein, 1980). All these aspects help to distort the accuracy of their judgements.

The complexity of the judgement does not only affect the quality of the forecasts as described in section 2.2, it also determines the forecaster's confidence level. Overconfidence does not seem to be prevalent in trivial or familiar tasks but only in complex ones (Russo & Schoemaker, 1992; Griffin & Brenner, 2004).

Moreover, overconfidence seems to increase with the difficulty of the task (Gigerenzer et al., 1991; Dittrich, Güth,, & Maciejovsky, 2005). For trivial tasks, underconfidence instead appears to be dominating (Odean, 1998).

One might believe that overconfident people who fail more often than expected should become less confident over time (Hirshleifer, 2001). This view is also advocated by Gervais and Odean (2001) who found support for the fact that overconfidence develops dynamically, i.e. it does not follow a linear path. The authors examined overconfidence with stock market traders and concluded that overconfidence initially increases as a person gains experience. However, throughout their career investor's inevitability face failures which will reduce ones confidence (Gervais and Odean, 2001).

Glaser, Langer and Weber (2005) conducted a study where they asked their participants to perform four different tasks. In the first session 33 German bankers participated. The task was to make confidence intervals for 20 knowledge questions (both general knowledge and finance related). Secondly, they were asked to make a self-assessment of their own performance and compare their estimated performance to the other participants. The third task was to make 15 stock market forecasts by assigning confidence intervals. Finally, they were asked to make trend forecasts for the stock market using confidence intervals. The authors also performed a second session for a larger group, consisting of 90 investment bankers and 76 students. The results from both sessions indicated that the judgements of both the professionals and the students were biased by overconfidence, but the degree of overconfidence was much higher with the professionals. The positive relationship between experience and overconfidence receives further support from other researchers (Bhandari & Deaves, 2006).

As Table 2 shows, many researchers argue that people in general are overconfident, but overconfidence does not seem to be prevalent in trivial tasks but only in complex ones. As stock price forecasting is indeed highly complex, we expect this task to give rise to overconfidence. Although the research literature provides a mixed picture of how confidence develops with experience, we expect financial experts to be more confident than novices, as depicted in our second research question.

The Effects of Overconfidence: As Table 3 shows there are several implications caused by overconfident agents on the financial markets. However, we have limited our thesis to only analyse how the level of confidence affects the forecasting accuracy. Our research question asks if confidence has a deteriorating effect on the forecasting accuracy. In agreement to our research question, increased levels of overconfidence seem to have a detrimental effect on the quality of stock price forecasts.

Table 3
Summary of previous research and findings of the effects of overconfidence.

Authors	Main findings relevant for our thesis
Odean (1998); Sniezek et al. (1990)	▪ increased levels of overconfidence have a detrimental effect on the quality of forecasts
Griffin and Brenner (2004)	▪ overconfidence often causes overprediction and/or overextremity in forecasts
Menkhoff et al. (2006)	▪ overconfidence tends to increase risk taking
Griffin and Tversky (1992)	▪ overconfident people tend to disregard how valid the information is
Glaser et al. (2004)	▪ the overconfident trader is prone to overestimate the precision of private information

In a financial market context there are several important effects of overconfidence. Overconfidence often causes over-prediction and/or over-extremity (Griffin & Brenner, 2004). Over-prediction means that people overestimate the probability for a specific outcome. Consequently, their probability judgements are consistently too high. Over-extremity means that the probabilities tend to be too close to either 0 or 100%, i.e. too extreme (Griffin & Brenner, 2004). Menkhoff, Schmidt, & Brozynski (2006) argue that this translates into overconfident stock market professionals who are prone to take on more risk.

A large body of research shows that overconfident stock market traders cause an increased trading volume (Odean, 1998; Benos, 1998; Gervais & Odean, 2001). Furthermore, increased levels of overconfidence seems to have a detrimental effect on the quality of stock price forecasts (Sniezek, Paese & Switzer III, 1990; Odean, 1998) and give rise to various mispricing phenomena such as the size- and the book-to-market effect (Daniel, Hirshleifer, & Subrahmanyam, 2001; Daniel, Hirshleifer, & Teoh, 2002). This should further increase the trading volume, as inaccurate forecasts must subsequently be corrected through additional trading (Glaser, Nöth, & Weber, 2004). These findings are interesting from a stock owner perspective as the large trading costs associated with increased trading, tends to decrease the portfolio owner's wealth (Barber & Odean, 2000).

Finally, overconfident people place too much weight on salient, anecdotal, and less relevant information (Kahneman & Tversky, 1973) and tend to disregard how valid the information is (Griffin & Tversky, 1992; Bloomfield, Libby, & Nelson, 2000). The overconfident trader is also prone to overestimate the precision of private information (Glaser et al., 2004). Hence, the presence of overconfident traders causes the market to overreact to information of low validity and underreact to abstract, statistical, and highly relevant information (Odean, 1998).

In accordance with our second research question we expect to find that increased levels of confidence have a detrimental effect on the forecasting accuracy.

2.5 The Training Dimension

Probability assessments often play a key role in financial decision making, as the possible outcomes often are uncertain. Given the difficulties involved in the judgement process, it is of interest to see if one can alter the decision maker's behaviour in a way that reduces the error term. Our final research question asks if it is possible to change the participant's answers through training. Table 4 presents a summary of the relevant literature that we discuss bellow. As one can see much research shows that training can improve the accuracy as well as decrease the level of overconfidence.

Table 4
Summary of previous research and findings of the effects of training.

Authors	Main findings relevant for our thesis
Gaeth and Shanteau (1984), Hershey and Walsh (2001)	▪ training can improve the overall accuracy
Lichtenstein and Fischhoff (1980); Koriat et al. (1980); Arkes et al. (1987); Kunda (1990); Block and Harper (1991); Trafimow and Sniezek (1994)	▪ training can reduce overconfidence
Staël von Holstein (1972)	▪ insignificant or no general improvements on overall accuracy from training
Fischer (1982)	▪ insignificant or no general improvements on overall accuracy from training

Various approaches to reduce the error term have been proposed in the last decades of research (Staël von Holstein, 1972; Lichtenstein & Fischhoff, 1980; Gaeth & Shanteau, 1984; Camerer & Johnson, 1991; Jacoby et al., 2001; Hershey & Walsh, 2001). In Hershey and Walsh (2001) the participants were asked to solve six different complex pension investment problems. The authors let some of their participants take a six-hour training program. The goal of the training was to teach them basic financial planning and various problem-solving techniques for common pension investment issues. The participants who took the training were able to significantly improve their performance. In the study by Gaeth and Shanteau (1984) two different training methods were used. The first consisted of a lecture where the participants were made aware of some of the difficulties involved in forecasting. The second training method contained an interactive component. The authors found no evidence for improvements in the accuracy from the first method. However, the second method led to a significant reduction in the usage of irrelevant information, as well as an overall improvement in accuracy.

Other research has instead shown little or no improvements from training (Staël von Holstein, 1972; Fischer, 1982). Staël von Holstein (1972) investigated the stock price forecasting ability of his participants. The experiment was divided up into ten different sessions. After each session the respondents received training in the form of feedback about the true values for the stock price movements and their accuracy score. His findings showed that the training was ineffective and yielded virtually no improvement in accuracy over the 20 weeks the experiment lasted.

Training has also been used with the purpose of affecting overconfidence. Several researchers have successfully reduced overconfidence through different training schemes (Lichtenstein & Fischhoff, 1980; Koriat, Lichtenstein & Fischhoff, 1980; Arkes, Christensen, Lai and Blumer, 1987; Kunda, 1990). For example people tend to justify their judgements by focusing on the aspects that support their stance. By forcing the person who forms the judgement to also list the factors that speak against their choice, Koriat et al. (1980), as well as Kunda (1990) managed to reduce their participants' overconfidence. However, the evidence is far from conclusive as other researchers argue that overconfidence is difficult to reduce (Block & Harper, 1991; Trafimow & Sniezek, 1994).

As we can see from Table 4, the results from training are ambiguous. However, some of the more explicit training sessions have been successful and improved the overall forecasting accuracy as well as the overconfidence. Thus, we believe that we will answer our final research question and show that one can change the participant's answers through training.

2.6 Hypotheses

In this section we formulate the hypotheses generated from the literature review. Finally, we present a figure (Figure 1) that sums up the hypothesis in relation to our research questions.

A*re stock market professionals better than novices in forecasting stock prices?* As stated in section 2.3, the literature concludes that there exists an inverse experience effect (Yates et al., 1991; Muradoglu & Önkal, 1994). Therefore, we believe that there will be a negative correlation between experience and the accuracy of the forecasts. Thus, we form the following hypothesis:

H1. Financial experts will perform worse than novices in forecasting the stock prices

Are financial experts more confident than novices and has confidence a deteriorating effect on the forecasting accuracy? As we have shown in section 2.4 much research concludes that financial experts are more confident than novices (Törngren & Montgomery, 2004; Bhandari & Deaves, 2006; Glaser et al., 2005). Moreover, a negative relationship appears to exist between overconfidence and the quality of the forecasts (Sniezek et al., 1990). Based on the literature review, the following hypotheses are formulated:

H2. Financial experts will be more confident than novices of their ability to forecast the stock prices
H3. Confidence will show a negative relationship to the ability to forecast the stock prices

Is it possible to change the participant's answers through training? Although we have a simple form of training i.e. educational information, we are still optimistic about receiving positive correlation between the information and the accuracy of the forecast. Muradoglu and Önkal (1994) showed that prior knowledge of subjective probability concepts had a significant impact on the respondents' answers. Our educational information aims to inform the participants about the difficulties involved in making stock price forecast and that people tend to provide too optimistic and narrow forecasts. If the information indeed has an effect, we believe that it will make the participants spread out the probability estimates over a larger number of intervals. Accordingly, we formulate the following hypotheses:

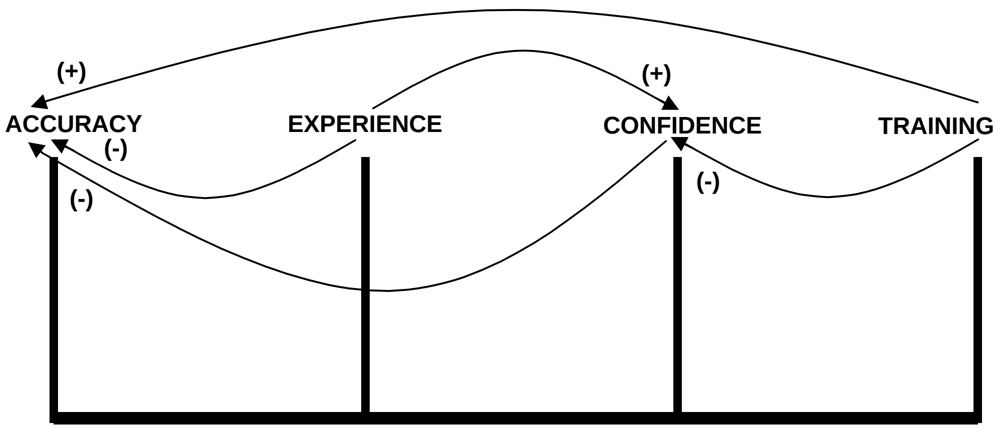
H4. The participants who receive the information will make more accurate stock price forecasts than the participants who do not receive this information

Finally, based on earlier successful attempts to reduce overconfidence presented in section 2.5 (Koriat et al., 1980; Kunda, 1990), we form the following hypothesis:

H5. The participants who receive the information will be less confident in their ability to forecast the stock prices than the participants who do not receive the information

Figure 1 summarises the patterns we expect to see between the dependent variables accuracy and confidence and the independent variables; experience, confidence and training. Note that overconfidence is an independent variable when it affects the accuracy of the forecasts and a dependent variable when we try to change the level of confidence through training.

Figure 1
Roadmap of the correlations between the variables.



Our first research question asks if stock market professionals are better than novices in forecasting stock prices. As depicted in Figure 1 we believe experience will have a negative effect on accuracy. Thus, stock market professionals will be less accurate than the novices. Further on we ask if financial experts are more confident than novices and if confidence has a deteriorating effect on the forecasting accuracy? As Figure 1 shows we believe that experience will increase the level of confidence and that it will have a negative effect on the forecasting accuracy. Finally we ask if it is possible to change the participant's answers through training?

Accordingly, our last two hypotheses states that training will improve the forecasting accuracy and decrease the level of confidence.

3. METHOD AND DATA

3.1 Participants

Two groups of participants were used for the experiment. *Financial experts* (35 participants) consisted of people who had professional experience related to the stock market. Roughly 50% of the participating financial experts were members of The Alumni Association of the Stockholm School of Economics (Handelshögskolans i Stockholm Kamratförening). The other half comprised a mixture of graduated students at Stockholm School of Economics, currently working within the finance industry, and a small fraction of teachers at the Stockholm School of Economics within the finance department. The second group, the *novices* (40 participants) consisted of undergraduate students at the Stockholm School of Economics.

This setup allowed us to test how experience affects the forecasting accuracy and confidence level and thereby provided a basis for answering Hypotheses 1-3.

The average participation rate of the contacted financial experts was approximately 13%. Their average age was 31.18 and experience from working in the stock market was 6.30 years. Of the 35 participants only two were female. The average participation rate of the novices was approximately 43%. Of the 40 respondents, 32 were 4th-year students, while the other eight were 2nd and 3rd year students. Their average age was 24.67, while the experience of working in the stock market was 0 years. Out of the novice group, seven were female, and 28 were male.

3.2 The Design of the Questionnaire

An initial e-mail was sent out to 264 financial experts and 92 students asking whether they wanted to participate in our study or not. On the 27th of March 2007 we sent out the questionnaire to those who agreed to participate. The participants were asked to respond within one week. This choice of time reflects the trade-off between giving the participants sufficient time to answer and including information that is as updated as possible.

The distributed e-mail included the questionnaire and instructions in an Excel-file. The first worksheet consisted of instructions of how the test was designed and how the participants were supposed to fill in their answers. The second worksheet included the questionnaire and an example of how to fill in the answers. There were two versions of this worksheet; one including information about the difficulties involved in stock price forecasting and one without this information. The questionnaire consisted of graphs of the stock price development for 28 of the 30 stocks listed on the OMXS30-index, over the period 25th of September 2006 to 23rd of March 2007. We also presented key financial ratios for each company (for further instructions see

Appendix 1). As Atlas Copco had two shares on the OMXS30-index we excluded their A-share and only included their B-share. We also excluded Alfa Laval, in order to use it as an example.

The stocks on the OMXS30-index represent the largest and most liquid companies in Sweden. We chose these stocks in order to minimize the risk that the participants would not recognise the companies. The stocks where grouped according to industry in the same fashion as that chosen by the OMX-group.

The participants were asked to assess the development of the buying price for each of the different stocks from the 2nd of April to the 2nd of May. The one-month forecast horizon has been used in earlier research (Törngren & Montgomery, 2004) and reflects practical aspects, such as our limited time to complete the thesis. For each stock the participants were asked to assess the probabilities that the stock price movement would fall in any of the six defined intervals. Six intervals where used which is the same structure used in earlier research (Yates et al., 1991; Muradoglu & Önkal, 2004). The size of the intervals (>+5%, +5-3%, +3-0%, -0-3%, -3-5%, <-5%) were also chosen to reflect earlier research but the intervals were adjusted to a one-month forecast horizon (Yates et al., 1991; Staël von Holstein, 1972).

Our interval was symmetrical around zero with the same percentage change on both the positive and the negative side. The alternative would have been to use the historical stock market development over a specific time period and then choose the intervals so that each interval would be equally probable. The rationale for the symmetric interval is that it is simpler for the participants to work with intervals consisting of symmetric and rounded off numbers. The probabilities for the different intervals, for each stock, should add up to 100% and the participants were allowed to spread their probabilities in any way they wanted.

To facilitate the interpretation of the graphs, we chose to present the stock price movements as the percentage movement instead of the absolute stock price movement. In order to further simplify for the participants, we included a graph of the development for the OMXS30-index over the same period. The participants were told that they were allowed to use all sources of information except for discussing with the other participants.

3.3 Incentive Scheme

To ensure that the participants provided accurate answers that reflected their true expectations, we incorporated an incentive scheme. An incentive scheme is important in order to reduce bias in the calibration (Johnson & Alistair, 2001). Our choice of incentives includes the following items which largely reflect the incentive structure in Yates et al. (1991):

- Feedback on the correctness of the participants' forecasts which will allow them to learn more about the various errors in forecasting the stock price
- A ranking table which increases competition and the participants' willingness to invest time in answering the questions thoughtfully

- All participants were explicitly guaranteed complete anonymity. This was especially important since they could otherwise have felt reluctant to participate. In order for the participants to track their results in the ranking table we asked them to chose a username. Hence, this setup also gave them the option not to participate in the ranking table
- In our initial letter we explicitly stated that the survey should take approximately 45 minutes to complete. By doing so we mitigated the risk that some participants would not provide well thought answers due to time constraint

It could be argued that informing the participants of how the scoring method was constructed could provide a suitable incentive measure. However, this would increase the risk for obtaining answers based on the scoring method rather than answers based on the participant's actual expectations of the stock price development. Thus, we chose not to explain the scoring method to the participants.

3.4 Independent Variables

The Participants' Experience: We used the participants' level of experience to sort them into the group of financial experts or the group of novices. People who were currently employed within the finance industry or had prior experience of working in the finance industry were selected to the group of financial experts. Furthermore, we measured how many years the participants had been active as financial professionals.

The Participants' level of Confidence: In order to test the aggregated level of confidence among the financial experts and novices groups respectively, we asked them to indicate, for each stock, how certain they felt about their forecasts. This choice of method does not provide us with a direct measure of possible overconfidence but rather their subjective confidence level. The participants were given five different options for their answers stretching from very uncertain to very certain. These different options were converted in the analysis to the numerical scale 1-5; one being very uncertain and five very certain.

Information: All the participants were randomly assigned to two different groups. The participants in the first group were informed about the difficulties involved in forecasting stock prices. These participants will from now on be referred to as the informed financial experts (18 participants) and the informed novices (19 participants). The other group of financial experts and novices, who did not receive this training, will from now on be referred to as the uninformed financial experts (17 participants) and uninformed novices (21 participants). Hence, this setup represents our experimental manipulation and enables us to answer Hypothesis 4 and 5.

The participants received the information when they opened the questionnaire. The information content was emphasising the following aspects:

- According to the efficient market hypothesis it should not be possible to predict the future stock price movements since all relevant information is already reflected in today's stock prices

- People are in general overconfident about the precision of their knowledge. This effect is larger with financial experts than with novices
- People who are overconfident tend to assign higher probabilities to events that in real life occur fairly infrequently and vice versa.

The purpose with this experimental manipulation was to test if information would make the participants alter their responses according to our hypotheses. We regard this information as a simple form of training. Similar simple training arrangements have been tested in other papers e.g. Block and Harper (1991). Their training was however much more explicit in that they urged the participants to spread out the distributions. We have purposely chosen a much less direct approach in our training. The rationale for doing so is that we believe that the underlying intention with the text would have become too obvious, had we used very explicit statements.

Use of Information Cues: The participants were asked to indicate which factors that had been most important for forming their judgements. They could choose as many as they liked from eight different information cues. These were; the stocks volatility, volatility changes during the period, large stock price movements, stock price trends, key ratios, information from the companies, external analysis and information from colleagues and friends (see Appendix). The participants could also state if any other sources of information had affected their judgements. 16 participants stated that for example dividends, other key ratios and the overall macro economy were important information cues when they formed their judgements.

3.5 Performance Measures

Throughout our thesis the Brier score is used for the scoring rule. This is the most commonly used and accepted measure of probability judgement accuracy and it is also applicable for multiple events, which is used in our thesis (Brier 1950; Yates, 1988). The first measure that we looked at is the overall accuracy of the forecasts, the *Probability Score for Multiple events* (PSM). In order to gain a greater understanding of the overall accuracy, we decompose the PSM-score into three different smaller components which are explained further below. The PSM-score for a participant is constructed by creating a vector f (f = 1, 2,…, 6) for each stock, where f_k denotes the probability forecast that the stock's price change will fall into interval k (k = 1, 2,…, 6). We also define an outcome index vector d_k (d = 1, 2,…,6) where d_k assumes the value 1 if the actual price change falls within interval k. If the actual price change falls within any of the other intervals, d_k assumes the value 0. The PSM for a single stock is then calculated in the following way:

$$PSM(f,d)=(f_1-d_1)^2+(f_2-d_2)^2+...+(f_6-d_6)^2=\sum_{k=1}^{K}(f_k-d_k)^2 \qquad (1)$$

where K=6 in this thesis. PSM ranges from 0 to 2. The lower the value the better is the accuracy of the participant for this particular stock. The average *PSM* over all stock forecasts is denoted as: $\overline{PSM}$. This is also the measure that is used for the overall accuracy.

Decomposition of the Brier-score: No matter how good a forecaster is, he/she will still not be able to make perfect forecasts due to the simple fact that there is an inherent uncontrollable element in forecasting. The stock price development is both exposed to uncontrollable factors, such as the overall economy, and controllable factors, such as company performance (Fama, 1970; 1991). It is therefore useful to decompose the accuracy measure in controllable and uncontrollable aspects. This will allow us to filter out the uncontrollable effects and thereby achieve a more correct measurement of the participant's forecasting skills. Yates (1988) developed the following model:

PSM = *f*(Uncontrollable: Base Rate; Controllable: Calibration, Covariation and Scatter) (2)

We will not hypothesize on how our different performance measures relate to our hypotheses. For example, it is expected that the financial experts will perform worse than the novices, but we do not know how this inferior performance will express itself in terms of our different performance measures. Looking at previous research, Yates et al (1991) for example found that financial experts had a higher scatter index than novices.

Calibration shows how well the forecaster managed to assign appropriate probabilities to the various outcomes. In other words, assume that the average probability for a 3-5% rise would be 25%. The participant's forecast is perfectly calibrated if in fact the stock prices rose by 3-5% for seven of the 28 examined stocks, i.e. 25% of the stocks. Hence, calibration does not concern whether the forecasts were correct for the individual stocks, but only for the aggregate result. The calibration score can hence be understood as a measure of how well the participant forecasts the average stock market development. The lower the calibration score, the better calibrated is the forecaster. The calibration score over all intervals is expressed in the following way:

$$CI = \sum_{k=1}^{K} \left(\overline{f}_k - \overline{d}_k \right)^2 \qquad (3)$$

where $\overline{f}_k$ and $\overline{d}_k$ denotes the individual participant's mean forecast and base rate for interval *k*, where $K = k_1, k_2, \ldots, k_6$.

Historical calibration is calculated in the same way as calibration but we use the average monthly development for the OMXS30-index during the last two years as the base rate.

Covariation is the second measure that we used. It looks at how well the participant assigns high probabilities to the correct intervals, and low probabilities to the incorrect intervals. In an ideal case, the participant would indicate a 100% probability in those instances where the event actually did occur and 0%

when it did not occur. The covariation score ranges from minus one to plus one. The higher the score, the better the forecaster is at discriminating between the instances when the actual price change will and will not fall within the specified intervals (Muradoglu & Önkal, 1994). Covariation for a given interval *k* is expressed in the following way:

$$\mathrm{Cov}(f_k, d_k) = \mathrm{Slope}_k \mathrm{Var}(d_k) \qquad (4)$$

where

$$\mathrm{Slope}_k = \left(\bar{f}_{1k} - \bar{f}_{0k}\right) \qquad (5)$$

Hence $\left(\bar{f}_{1k}\right)$ represents the mean forecast for the forecasts in interval *k* that actually occurred and $\left(\bar{f}_{0k}\right)$ represents the mean forecast for the forecasts in the interval *k* that did not occur. Also,

$$\mathrm{Var}(d_k) = \bar{d}_k\left(1 - \bar{d}_k\right) \qquad (6)$$

represents the variance of the outcome index d_k for interval *k*. The slope is under the forecaster's control whereas the variance is not. Since we are more concerned about the controllable factors, the covariance skills of the forecaster, over all six intervals, is defined as:

$$\text{Mean Slope} = (1/K)\sum_{k=1}^{K} \mathrm{Slope}_k \qquad (7)$$

Scatter is the third measure that we used. A good forecaster should be able to vary his/her forecasts with the actual occurrences. He/she should also be able to avoid varying his/her predictions independently of those occurrences. Independent variation is referred to as scatter and is usually caused by two factors. Firstly, the forecaster might base his/her predictions on pieces of information that he/she expects are reliably related to the stock price but which in fact are not. Secondly, the forecaster might be inconsistent in what effect he/she assigns to a specific information cue. It is clearly ideal to have as low scatter index as possible. The scatter index for interval *k* is measured in the following way:

$$SI_k = (1/N)\left[N_{1k}\mathrm{Var}(f_{1k}) + N_{0k}\mathrm{Var}(f_{0k})\right] \qquad (8)$$

where $\mathrm{Var}(f_{1k})$ represents the conditional variance of the forecasts in interval *k* when this interval actually did occur. Additionally, $\mathrm{Var}(f_{0k})$ is the conditional variance for those instances when the actual outcome did not fall into interval *k*. N_{1k} is then the number of forecasts that were assigned to interval *k* when that was the actual outcome and N_{0k} is the number of forecasts that were assigned to interval *k* when that was not the actual outcome. Consequently $N = N_{1k} + N_{0k}$ and the scatter index over all intervals are calculated in the following way:

$$SI = \sum_{k=1}^{K} SI_k \qquad (9)$$

Forecast profile variance was used to compare the forecasts of a participant to the forecasts of a uniform forecaster. As the uniform forecaster provides a constant forecast for every stock, this measure allowed us to examine how much the forecasts from a participant varied from stock to stock. The lower the forecast profile variance, the more similar are the participant's forecasts to a uniform forecaster. We calculated the forecast profile variance in the following way:

$$\text{Forecast profile variance} = \frac{1}{N}\sum\left[\frac{1}{6}\sum_{k}\left(f_k - 0.167\right)^2\right] \qquad (10)$$

The performance measures are summarized in Table 5.

Table 5
Performance measures used.

Performance measure	
Overall Accuracy	How accurate the participant's stock price forecasts are. According to Hypothesis 1, we expect a lower overall accuracy for financial experts than novices.
Calibration	How well the participant forecasts the average stock market movement, irrelevant of the specific stocks.
Historical calibration	Identical to calibration but using the monthly stock price change over the last two years as the base rate instead of the actual outcome during the period 2nd of April to 2nd of May.
Covariation	How well the participant assigned high probabilities to true events and low probabilities to false events.
Scatter	How well the participant managed to avoid varying his/her forecasts independently from false events.
Forecast profile variance	How closely the forecasts of the participant resemble a uniform forecaster, i.e. someone who assigns the same probability to all the different intervals.

3.6 Benchmarking Profiles

Similar to Yates et al. (1991), we will compare the participants' forecasts to four different static benchmarking profiles. This setup allows us to test whether the participants' active choices in the forecasts are superior to a simple mechanical decision model.

- The uniform forecaster assigns the same probability to each interval. Our thesis utilizes six different intervals and accordingly he/she assigns the probability 1/6 to each interval. Hence, he/she will always earn $\overline{PSM} \approx 0.83$
- The historical forecaster looks at the historical monthly performance of the stock market. He/she then calculates the average probability that the stock price change falls in each of the intervals and assigns these probabilities to his/her forecasts. As suggested by Staël von Holstein (1972) we use a two year historical horizon for the OMXS30-index. The historical relative frequency for the period 2005-04-01 to 2007-04-02 for the six different intervals was: 26%, 23%, 14%, 23%, 9% and 5% (falling order of intervals, starting with >+5%)
- The base rate forecaster is a fictional forecaster who can anticipate the average actual relative frequencies for the six intervals. He/she then assigns these probabilities to all stocks. The base rate used is hence the ex-post actual outcome for OMXS-30. This gives us the following probability distribution for the six different intervals: 42%, 4%, 14%, 11%, 11% and 18% (falling order of intervals, starting with >+5%). Although this profile is hypothetical, a skilled forecaster could approximate the base rate by using the historical outcome. He/she then adjusts this to his/her perception of how the present period will systematically differ from the historical outcome.

4. EMPIRICAL FINDINGS – HYPOTHESES TESTING

The overall accuracy of the forecasts is measured with the $\overline{PSM}$ -score as described in section 3.5. Firstly, we will look at the descriptive statistics and see how the respondent groups performed in comparison to each other and to the three benchmark profiles used. Secondly, we will test our five hypotheses. Thirdly, we will test the robustness of our results. Fourthly, we will discuss the implications of our findings and finally we will discuss possible limitations with our choice of method.

4.1 Descriptive Statistics

Table 6 presents the median values for the forecasting accuracy and the confidence level of the different respondent groups together with the results for the benchmarking profiles (see section 3.6). Obviously, the forecaster's goal is to obtain a forecasting score as close to 0 as possible. The worst possible forecast would result in a score of 2.

Table 6
Median values for the forecasting accuracy and the confidence level for all respondent groups and the three benchmarking profiles.

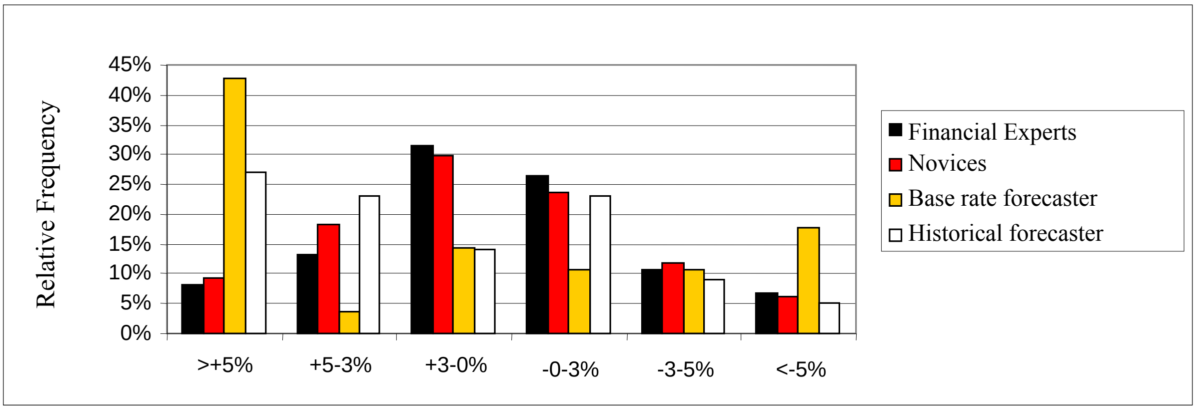
Participant group	Forecasting accuracy	Confidence
Informed financial experts	1.08	2.50
Uninformed financial experts	1.15	2.63
Aggregated financial experts	1.09	2.61
Informed novices	0.97	2.61
Uninformed novices	0.99	2.45
Aggregated novices	0.98	2.52
Uniform forecaster	0.87	-
Historical forecaster	0.83	-
Base Rate forecaster	0.74	-

The data in Table 6 indicates that all groups are inaccurate in their forecasts as all groups score worse than the benchmarking profiles. The $\overline{PSM}$ -scores for the three benchmarking profiles are the following; the uniform forecaster: 0.87, the historical forecaster: 0.83 and the base rate forecaster: 0.74. Only seven participants manage to outperform the uniform forecaster scheme. Three of these respondents are financial experts and four are novices. The top-three forecasters score 0.82, 0.83 and 0.84, respectively. The first two are novices and the third best participant is a financial expert. Only one participant manages to outperform the historical forecaster. The three forecasters with the worst performance score 1.26, 1.26 and 1.44, respectively. These three were all financial experts. The data indicates a slightly higher $\overline{PSM}$ -score for the financial experts which suggest that their forecasts were less accurate than those of the novices.

After analysing the average values it shows that there is very little difference between the median and the average values, so a decision was made to not include the average values as well.

Figure 2 displays the average relative frequencies for the different groups and the corresponding numbers for the historical forecaster and the base rate forecaster. The base rate forecaster is obviously, identical to the actual outcome, hence a different way of looking at the calibration. Clearly, the participants underestimate the likelihood for large stock price increases as the calibration is the worst for the interval for a growth of more than five percent. Furthermore, the participants are highly unsuccessful in all intervals except for the interval for a 3-5% decline where both respondent groups are very close to the actual base rate. It also appears as if the financial experts are slightly worse calibrated than the novices.

Figure 2
Relative frequencies for the forecasts for financial experts, novices, the base rate forecaster (the actual outcome) and the historical forecaster (the base rate for the last two years).



It should be stressed that the actual outcome for the examined time period deviates quite substantially from the historical outcome over the last two years. For example, as shown in Figure 2, the relative frequency for a stock price increase of more than five percent was 16 percentage points higher in the actual period than the average over the last two years. If we instead compare the calibration of the participants' forecasts with the relative frequencies over the last two years (the historical forecaster), both groups are better calibrated. In three of the six intervals the participant groups were less than ten percentage points away from the true outcome. This could be interpreted as that both groups had some knowledge of the historical base rate. Still, the participants were far from well calibrated, especially for the intervals for >+5% and +3-0%.

The low numbers for the confidence level in Table 6 suggest that the participant groups are fairly uncertain about their forecasts. Recall that a score of 1 indicates that the participant is very uncertain and a score of 5 indicates that the participant is very certain about his/her answers. There is also little difference in the confidence levels between the different groups.

4.1 Hypothesis 1

The first hypothesis predicts that the financial experts will perform worse than the novices in predicting stock price changes. Table 7 presents the median forecasting accuracy for the financial experts and the novices. As a higher $\overline{PSM}$ -score *indicates* worse forecasting accuracy, Hypothesis 1 suggests that the financial experts should have a higher $\overline{PSM}$ -score than the novices. There appears to be some difference in the forecasting accuracy, in that the financial experts are less accurate.

Table 7
Median values for the performance measures for financial experts and novices.

Performance Measures	Financial experts	Novices
Forecasting accuracy	1.09	0.98
Calibration	0.24	0.21
Historical calibration	0.17	0.09
Covariation	0.00	0.00
Scatter	0.06	0.03
Forecast profile variance	0.03	0.02

In order to examine whether the financial experts' forecasting accuracy is significantly different from the novices we use the Mann Whitney U-test. This method is also used in earlier research (Yates et al., 1991; Muradoglu & Önkal, 1994). Due to the sample being fairly small and the unknown sampling distribution of variables, the U-test was selected. The Mann-Whitney U-test is a non-parametric test that examines whether two samples of observations come from the same distribution. The Z-value indicates whether there is a significant difference between the two samples. In order to clarify this, we also indicate the significance level, where * means that there is a significant difference at the 5% level and ** that there is a significant difference at the 1% level. The Mann Whitney U-test will be used frequently throughout our hypotheses testing. Table 8 shows that the financial experts' forecasting accuracy was significantly worse (higher PSM) than the accuracy of the novices (Z-value: -2.51*). Hence, Hypothesis 1 is not rejected.

In order to verify our results, we look at the differences between the financial experts and the novices in the four different decomposed performance measures (calibration, covariation, scatter and forecast profile variance). Table 8 shows that the financial experts had a significantly higher forecast profile variance than the novices (Z-value: -2.01*). This suggests that the forecasts of the novices are more similar to a uniform forecaster and that the financial experts are more prone to make forecasts that vary from one stock to the next. The difference between the two groups in absolute value is however small.

Table 8
Mann Whitney U-test data for the differences in the performance measures between the financial experts and the novices.

Mann Whitney U-test data			
Differences in forecasting accuracy		**Differences in calibration**	
Mean rank (financial experts)	44.74	Mean rank (financial experts)	43.06
Mean rank (novices)	32.10	Mean rank (novices)	33.58
Z-value	-2.51*	Z-value	-1.88
Differences in Covariation		**Differences in scatter**	
Mean rank (financial experts)	35.20	Mean rank (financial experts)	41.20
Mean rank (novices)	34.68	Mean rank (novices)	35.20
Z-value	-1.41	Z-value	-1.19
Differences in forecast profile variance			
Mean rank (financial experts)	43.41		
Mean rank (novices)	33.26		
Z-value	-2.01*		

* $p < 0.05$ (2-tailed)

We also look at the correlations between the different variables. We use Spearman correlations as the sampling distribution is unknown. For Hypothesis 1 a dummy variable that assumes the variable 1 for the financial experts and 0 for the novices is used. The correlation numbers are presented in Table 9. The data tells us that the dummy variable is significantly correlated with the forecasting accuracy (ρ: 0.29*). As the correlation is positive, this suggests that the forecasting accuracy of the financial experts is lower than the novices. Similar to the Mann Whitney U-test, we also find that the forecast profile variance is higher with the financial experts, as indicated by the positive correlation (ρ: 0.23*).

Table 9
Spearman correlation numbers.

Measure	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dummy for informed or uniformed (1)	-								
Dummy for financial experts or novices (2)	0.04	-							
Number of information cues used (3)	0.03	0.05	-						
Number of years of experience (4)	-0.06	-	-0.31	-					
PSM (5)	-0.11	0.29*	0.14	0.19	-				
Calibration (6)	-0.14	0.22	0.05	0.09	0.84**	-			
Covariation (7)	-0.23*	0.16	0.09	-0.18	-0.02	0.08	-		
Scatter (8)	-0.15	0.14	0.12	-0.16	0.54**	0.42**	0.13	-	
Confidence (9)	-0.03	0.04	0.30**	-0.12	0.36**	0.35**	0.03	0.17	-
Forecast profile variance (10)	-0.16	0.23*	0.04	0.24*	0.81**	0.81**	0.08	0.70**	0.29*

* p < 0.05 (2-tailed)
** p < 0.01 (2-tailed)

4.2 Hypothesis 2

Our second hypothesis assumes that financial experts, on average, are more confident than novices. The median values in Table 6 show that the confidence levels of the financial experts and the novices are both fairly low and almost equal between the two groups.

In order to compare the confidence level of the financial experts with the confidence level of the novices, the Mann Whitney U-test is used. The results, presented in Table 10, indicate that the confidence level is not significantly different between the financial experts and the novices (Z-value: -0.30). Thus, Hypothesis 2 is rejected.

Table 10
Mann Whitney U-test data for the differences in the confidence level between the financial experts and novices.

Mann Whitney U-test data	
Difference in the confidence level	
Mean rank (1)	38.81
Mean rank (2)	37.29
Z-value	-0.30

4.2 Hypothesis 3

The third hypothesis suggests that an increased level of confidence will decrease the forecasting accuracy. As a higher $\overline{PSM}$ -score means worse forecasting accuracy, Hypothesis 3 suggests that there should be a positive correlation between the confidence level and the $\overline{PSM}$ -score. To test this we compare the correlation between the participants' confidence level and their overall accuracy, as well as the decomposed performance measures. Table 9 shows that there is positive and significant correlation between confidence and forecasting accuracy (ρ: 0.36), suggesting that as the confidence level increases, the forecasting accuracy decreases. Hence, Hypothesis 3 is not rejected.

The correlation coefficients in Table 9 provide further evidence for why increased levels of confidence deteriorate the forecasting accuracy. Firstly, we find a positive and significant correlation between the confidence level and the calibration index (ρ: 0.35**). As a higher calibration score equals worse calibration, the positive correlation suggests that calibration is worsened when the confidence level increases. Moreover, a higher calibration score implies worse forecasting accuracy. Consequently, we can conclude that increased confidence has a deteriorating effect on forecasting accuracy. Secondly, the positive correlation to forecast profile variance (ρ: 0.29*) suggests that the more confident the participant is, the less similar his/her forecasts are to the forecasts of a uniform forecaster.

4.4 Hypothesis 4

The fourth hypothesis states that the informed participants will have a higher forecasting accuracy than the uniformed participants. In order to examine this we decompose the participants into three different groups: informed vs. uniformed participants, informed financial experts vs. uninformed financial experts and informed novices vs. uninformed novices. By splitting the sample in this way we can isolate the effect of the information and thereby minimize the risk of having the results distorted by other factors. The median and the average values for all the performance measures are presented in Table 11. There appears to be little difference between the informed and the uninformed participants in any of the groups.

Table 11
Median/average values for the performance measures for the informed and uniformed participants.

Performance measure	Informed participants	Uninformed participants	Informed financial experts	Uninformed financial experts	Informed novices	Uninformed novices
Forecasting accuracy	0.99	1.02	1.08	1.15	0.97	0.99
Calibration	0.21	0.23	0.23	0.24	0.20	0.22
Calibration using 2-year historical data for base rate	0.12	0.11	0.16	0.17	0.08	0.09
Covariation	0.00	0.01	0.00	0.01	-0.01	0.00
Scatter	0.03	0.05	0.06	0.06	0.03	0.04
Forecast profile variance	0.02	0.02	0.03	0.02	0.01	0.02

We also test if there are any significant differences in the performance measures between the two samples in all the different groups using the Mann Whitney U-test. The results from these tests are presented in Table 12. The table shows that there is no significant difference in the forecasting accuracy between the informed and the uniformed financial experts (Z-value: -1.19). The results also show no significant difference in the forecasting accuracy between the informed and the uniformed novices (Z-value: -1.15). Moreover, Table 12 shows that there is no significant difference in the forecasting accuracy between the aggregated informed and the aggregated uninformed participants (Z-value: -0.94). Hence, there are no indications of that the informed participants' forecasting accuracy should be higher than the forecasting accuracy of the uninformed participants.

Table 12
Mann Whitney U-tests for differences in the performance measures between the different groups of informed and uniformed participants.

Performance measure		Informed participants (1) vs. uninformed participants (2)	Informed financial experts (1) vs. uninformed financial experts (2)	Informed novices (1) vs. uninformed novices (2)
Accuracy	Mean rank (1)	35.59	19.68	18.26
	Mean rank (2)	40.34	16.42	22.52
	Z-value	-0.94	-1.19	-1.15
Calib. Index	Mean rank (1)	34.95	16.83	18.21
	Mean rank (2)	40.97	19.24	22.57
	Z-value	-1.20	-0.69	-1.18
Calibration using 2- year	Mean rank (1)	37.49	17.28	20.26
hist. data for base rate	Mean rank (2)	38.50	18.76	20.71
	Z-value	-0.20	-0.43	-0.12
Covariation	Mean rank (1)	32.95	15.28	17.92
	Mean rank (2)	42.92	20.88	22.83
	Z-value	-1.98*	-1.62	-1.33
Scat. Index	Mean rank (1)	34.78	19.11	15.79
	Mean rank (2)	41.13	16.82	24.76
	Z-value	-1.26	-0.66	-2.43*
Forecast profile var.	Mean rank (1)	34.54	17.92	15.95
	Mean rank (2)	41.37	18.09	24.62
	Z-value	-1.36	-0.05	-2.34*

* $p < 0.05$ (2-tailed)

These findings are also supported by the correlation data in Table 9. Since a lower $\overline{PSM}$ -score means better forecasting accuracy, Hypothesis 4 suggests that there should be a negative correlation between the dummy variable for the financial experts and the $\overline{PSM}$ -score. There is no significant correlation between this dummy variable and the forecasting accuracy. Thus, Hypothesis 4 is rejected.

Table 12 also shows that there is little difference in the decomposed performance measures between the informed and uninformed participants in either of the samples. Our results do show that the informed novices have a significantly lower scatter index than the uninformed novices (Z-value: -2.43*). While the informed novices have a significantly lower forecast profile variance than the uninformed novices (Z-value: -2.34*). These two findings suggest that the forecasts of the informed novices are more similar to a uniform forecaster with less independent variation in the forecasts between the different stocks. However, the difference in size, in absolute value, is rather small in both the scatter and forecast profile variance. Finally, we find that the aggregated informed participants have a significantly lower (worse) covariation than the aggregated uniformed participants.

4.5 Hypothesis 5

The fifth hypothesis suggests that the informed participants should have a lower confidence level than the uniformed participants. In order to test this we perform a similar decomposition of the participants into three different test groups, as in Hypothesis 4. Table 13 shows that there is little difference in the median and average confidence levels between the different groups. It is also apparent that the confidence level is fairly low in all the different participant groups.

Table 13
Median/average confidence levels for the different groups of informed and uniformed participants.

Informed participants	Uninformed participants	Informed financial experts	Uninformed financial experts	Informed novices	Uninformed novices
2.51/2.52	2.61/2.59	2.50/2.48	2.63/2.67	2.61/2.55	2.45/2.53

We use the Mann Whitney U-test to examine whether there are significant differences in the confidence level between the different groups. The test data is presented in Table 14. The table shows that there is no significant difference in the confidence level between the informed and the uniformed financial experts (Z-value: -0.27). Though there is no significant difference between the informed novices and the uniformed novices (Z-value: -0.94). Finally, we could not find any significant difference between the aggregate informed and uninformed participants (Z-value: -0.26).

Table 14
Mann Whitney U-tests for differences in the confidence level between the different groups of informed and uniformed participants.

Mann Whitney U-test data

Informed financial experts vs. uninformed financial experts		**Informed novices vs. uninformed novices**	
Mean rank (informed financial experts)	37.32	Mean rank (informed novices)	16.42
Mean rank (uninformed financial experts)	38.66	Mean rank (uninformed novices)	19.68
Z-value	-0.27	Z-value	-0.94
Informed vs. uninformed participants			
Mean rank (informed participants)	21.00		
Mean rank (uniformed participants)	20.05		
Z-value	-0.26		

* $p < 0.05$ (2-tailed)

These findings are also supported by the correlation data in Table 9. The information dummy, which assumes the value 1 for the informed participants and 0 for the uniformed participants, does not have a significant correlation with the level of confidence. This means that the confidence level of the informed participants does not differ from the confidence level of the uniformed participants. Hence, Hypothesis 5 is rejected.

4.6 Robustness Tests

In order to further test the conclusions reached in the hypothesis testing we will perform additional robustness tests. This part can hence be viewed as an explorative examination to identify potential limitations in our findings. Firstly, we will run linear regressions. Secondly, we will examine whether the fact that the stocks in our questionnaire belonged to different industries had any impact on the results reached.

Regression analysis: We use linear regressions for our regression analysis. Table 15 presents the data from the linear regression using the full sample of participants. In accordance with our findings in Hypothesis 1 we find that the forecast accuracy of the financial experts is, on average, worse than the accuracy of the novices as the coefficient for the dummy variable is positive and significant. In Table 16 and 17 we run the regression separately for the financial experts and the novices. The difference in forecasting accuracy is now instead reflected in the constant. The constant is higher, but only marginally, for the financial experts which is in line with our previous findings.

The positive and significant coefficient for the confidence level suggests that the forecasting accuracy decreases as the confidence increases. This finding is in line with our conclusions in Hypothesis 3, showing that this effect seems to be more pronounced with the financial experts than with the novices.

The lack of a significant coefficient for the dummy variable for the informed participants supports our findings in Hypothesis 4. Consequently, the regression analysis supports our findings in the hypothesis testing.

Table 15
Linear regression using the full sample of participants – Dependent variable: PSM-score.

Variable	Coefficient	t	Std error
Constant	0.86**	25.03	0.03
Dummy that assumes value 1 for financial experts and 0 for novices	0.09**	3.78	0.02
Confidence level	0.09**	4.54	0.02
Number of information cues used	0.00	-0.23	0.01
Dummy that assumes value 1 for participants who indicated that they used additional information cues	-0.04	-1.30	0.03
Dummy that assumes value 1 for informed participants and 0 for uninformed	-0.03	-1.31	0.02

** $p < 0.01$

Table 16
Linear regression using only the financial experts - Dependent variable: PSM-score.

Variable	Coefficient	t	Std error
Constant	0.87**	16.69	0.05
Confidence level	0.08**	5.51	0.01
Number of information cues used	-0.03	-3.09	0.01
Dummy that assumes value 1 for participants who indicated that they used additional information cues	0.06	1.91	0.03
Dummy that assumes value 1 for informed participants and 0 for uninformed	0.03	0.94	0.03
Number of years of experience	0.00	-0.72	0.00
Dummy that assumes value 1 for participants working as stock traders	0.03	0.95	0.04

** $p < 0.01$

Table 17
Linear regression using only the novices - Dependent variable: PSM-score.

Variable	Coefficient	t	Std error
Constant	0.86**	25.03	0.03
Confidence level	0.04**	3.15	0.01
Number of information cues used	0.02	2.40	0.01
Dummy that assumes value 1 for participants who indicated that they used additional information cues	-0.01	-0.27	0.03
Dummy that assumes value 1 for informed participants and 0 for uninformed	-0.02	0.86	0.02

** $p < 0.01$

Industry analysis: Our analysis of the possible impact from the industry aspect is divided into two components. Firstly, we examine whether the forecasting accuracy differs between the different industries. Secondly, we examine whether the results reached in our hypothesis testing are in fact simply caused by variations between the different industries.

The 28 stocks that the participants were asked to forecast are unevenly spread out over nine different industries. The questionnaire presents the stocks industry by industry. One can speculate that especially the financial experts should have better fundamental knowledge in some industries than other. If this is true, we can expect to see differences in the forecasting accuracy between the industries. Another factor to consider is that the participants might have lost focus the further into the questionnaire they progressed, which can result in lower forecasting accuracy for the industries presented at the end of the questionnaire. To test these two effects we group the stocks in accordance with their respective industry.

Table 18

Presentation order of the industries in the questionnaire. Ranking order of the industries according to the median $\overline{PSM}$ -score for the financial experts, the novices and the three benchmarking profiles.

Presentation order in the questionnaire	Ranking order for financial experts	Forecast accuracy for financial experts	Ranking order for novices	Forecast accuracy for novices	Uniform forecaster	Historical forecaster	Base rate forecaster
Energy (1 stock)	9th	1.16	8th	1.11	0.87	1.11	0.90
Materials (3 stocks)	7th	1.10	7th	1.03	0.87	0.84	0.83
Industrials (9 stocks)	5th	1.04	5th	1.00	0.87	0.82	0.64
Consumer Discretionary (4 stocks)	4th	1.03	4th	0.98	0.87	0.78	0.72
Consumer Staples (1 stock)	1st	0.71	1st	0.71	0.87	0.75	1.05
Health Care (1 stock)	6th	1.10	6th	1.00	0.87	1.03	1.06
Financials (5 stocks)	3rd	0.97	3rd	0.97	0.87	0.74	0.65
Information Technology (2 stocks)	2nd	0.94	2nd	0.92	0.87	0.80	0.69
Telecom. Services (2 stocks)	8th	1.11	9th	1.11	0.87	1.07	0.97
OMXS30-index		1.09		0.98	0.87	0.83	0.74

The first column in Table 18 presents the order in which the industries were presented in the questionnaire. The third and fifth columns list the industries according to how well the financial experts and novices, respectively, manage to forecast the stocks in the specific industry. There appears to be a clear difference in the forecasting accuracy between the different industries. Both groups are very successful in forecasting the consumer staples industry. However, this industry only comprises one stock. Furthermore, the ranking of the industries, according to the participants' forecasting accuracy within that industry, is almost similar between the financial experts and the novices.

To test whether the differences in forecasting accuracy between the different industries are significant we use the Wilcoxon test. This test examines two industries at a time and indicates whether there is a significant difference in the forecasting accuracy between them by splitting the data into financial experts and novices. Furthermore, industries that only have one stock have been excluded in order to avoid insignificant comparisons. Table 19 and 20 indicate that there are indeed significant differences in the forecasting accuracy between the two industries. There does not appear to be any pattern-like differences between the financial experts and the novices. It appears as if materials and telecommunication are the two most difficult industries to forecast for both groups.

Table 19
Wilcoxon tests for differences in the financial experts' forecasting accuracy between the different industries.

Industry	(1)	(2)	(3)	(4)	(5)
Materials (1)	-				
Industrials (2)	-2.98**	-			
Consumer discretionary (3)	-2.16*	-0.44	-		
Financials (4)	-3.64**	-3.16**	-3.93**	-	
Information technology (5)	-4.13**	-3.39**	-3.52**	-2.52*	-
Telecommunication services (6)	-1.09	-2.92**	-2.97**	-3.75**	-4.01**

* $p < 0.05$
** $p < 0.01$

Table 20
Wilcoxon tests for differences in the novices' forecasting accuracy between the different industries.

Industry	(1)	(2)	(3)	(4)	(5)
Materials (1)	-				
Industrials (2)	-1.79	-			
Consumer discretionary (3)	-2.330*	-0.954	-		
Financials (4)	-3.28**	-2.39*	-0.91	-	
Information technology (5)	-3.95**	-3.79**	-2.74**	-2.55*	-
Telecommunication services (6)	-3.11**	-4.01**	-4.01**	-4.33**	-4.59**

* $p < 0.05$
** $p < 0.01$

If we compare the order in which the industries were presented and the ranking order of the participants' we find no particular pattern. Consequently, there is no indication that the participants' forecasting accuracy is affected by lack of focus in the later part of the questionnaire. Instead, the differences in the forecasting accuracy between the industries seem to derive from the industry itself.

We now proceed to test whether the industries affected our conclusions reached in the hypothesis testing. For this purpose we run the regressions again, but this time using the data for each individual stock price forecast rather than using each individual's average value. This allows us to assign each stock forecast to its corresponding industry by using a dummy variable for each industry. Furthermore, we create additional interactive variables by multiplying the confidence variable with the industry dummy variables. The rationale for this setup is that if the relationships that we previously found to be significant now turn out to be incorrect, it is likely that the relationships were instead caused by simple variations between the industries.

We run the regressions in the same order as above. As the regressions include a large number of variables we present the data in the Appendix. We received similar results to our hypothesis testing when running the regression on the full sample of participants as is evident from Table 22. However, when we split the sample into financial experts and novices, the confidence variable is no longer significant as is displayed in Table 23 and 24. This suggests that our findings in Hypothesis 1 and 3 could in fact be the result of variations caused by the specific industry that the stocks belong to.

5. DISCUSSION AND SUMMARY

We summarize our findings from our five different hypotheses in Table 21 below.

Table 21
Summary of findings in our hypotheses.

Hypotheses	Outcome	Findings
1) Financial experts will perform worse than novices in forecasting the stock prices	Not rejected	We find that the financial experts, on average, received a lower forecasting accuracy than the novices
2) Financial experts will be more confident than novices of their ability to forecast the stock prices	Rejected	We find no significant difference between financial experts and novices in terms of confidence level
3) Confidence will show a negative relationship to the ability to forecast the stock prices	Not rejected	We find that the more confident the participant is, the lower is his/her average forecasting accuracy
4) The participants who receive the information will make more accurate stock price forecasts than the participants who do not receive this information	Rejected	We find no significant difference in the forecasting accuracy between the informed and the uninformed participants
5) The participants who receive the information will be less confident in their ability to forecast the stock prices than the participants who do not receive the information	Rejected	We find no significant difference in the confidence level between the informed and the uninformed participants

Our results show that the stock prices are difficult to forecast over a one-month time period. Similar to both Staël von Holstein (1972) and Yates et al (1991) we find that the overall forecasting accuracy for all participants is poor and that the less experienced novices make more accurate forecasts than the financial experts, which is displayed in Hypothesis 1. The difference between the two groups is also large in absolute value. The overall accuracy for both the financial experts and the novices is also worse than all the three different benchmarking profiles with the exception of a few participants. The reasons to why the novices perform better than the financial experts can be found in their lower forecast profile variance. In other words, the novices make forecasts that are more similar to a uniform forecaster than the financial experts. The merits of such a forecasting strategy can be questioned, as the more similar a forecast is to a uniform forecaster, the more similar it is to a simple mechanical forecasting scheme. It is hence not surprising that the financial experts, who own their living on actively forecasting stock prices, chose a more active forecasting strategy. However, with a one-month time horizon and a wide array of stocks to forecast, choosing a forecasting strategy more similar to a uniform forecaster can be a way to minimize the risks and thus increase the average accuracy.

The data shows that overall calibration is poor as the participants are unable to successfully assess the relative frequencies for the different intervals for the stock price development. For example they underestimate the probability for stock price increases by more than five percent. This can partly be attributed to the fact that the actual development during the period 2nd of April to 2nd of May differs quite substantially from the monthly pattern over the previous two years. There is little difference in the calibration between the financial experts and the novices and hence calibration cannot explain the financial experts' inferior performance. From our results we find it surprising that the financial experts' calibration differs so much from the average monthly stock price development over the last two years, as displayed in Figure 2. Consequently, the past stock price movements seem to have little effect on the financial experts' forecasts for the future which we can only speculate as to why this is so. The financial experts might be inconsiderate about the historical stock market development. This can also be an indication for that the financial experts believe in the efficient market hypothesis (Fama, 1970; 1991). Under such a scenario they might in fact be right to deviate from the historical outcome, as the stock market is assumed to follow a random walk.

Our data also suggests two other noteworthy relationships. Firstly, in Hypothesis 2 we find that the forecaster's confidence level has a significant impact on his/her forecasting accuracy. It seems as the more confident the forecaster is in his/her ability, the less accurate are his/her forecasts. This can mainly be attributed to two factors. Firstly, we find a positive correlation between confidence and calibration. Increased confidence with the forecaster should hence result in that he/she makes forecasts that are either more, or less, extreme than the average outcome for the stock market. One can hypothesise that increased confidence results in a larger ignorance for indicative signs of the stock price development. According to Daniel et al. (2001), overconfident investors overreact to private information. Thus, the overconfident forecaster might be inclined to dismiss the public historical information cues, in favour for other information, which he/she perceives as superior. Secondly, we find a positive correlation between confidence and the forecast profile variance. Hence, the less confident the individual is, the more his/her forecasts resemble a uniform forecaster. One can speculate that this is caused by that the less confident forecasters better know the limits of their knowledge and the difficulty of the forecasting problem than the more confident forecasters. Hence, the less confident forecasters are likelier to make more cautious forecasts which are more similar to a uniform forecaster.

The participants, on average, display fairly low confidence in their forecasts. Hypothesis 2 confirms that there was also little difference between the financial experts and the novices in their level of confidence. One can only speculate on the causes to the low level of confidence. The participants might of course be underconfident or the financial experts might be more accustomed to focus on a few specific companies or industries. Hence, they might feel uncomfortable with the fixed, and large, sample of stocks used in our questionnaire.

Section 4.6 shows that there are large differences in the forecasting accuracy between the different industries. The telecommunications, and the financials industries seem to be the most difficult to forecast, whereas the consumer staples and the information technology industries are the easiest. This pattern is also very similar between the financial experts and the novices. We do not believe that these industries share any characteristics that would make them either particularly easy or difficult to forecast which could explain the observed pattern. Instead we regard this as a pure chance effect. From this we can draw two conclusions. Firstly, the financial experts' high forecasting accuracy in the consumer staples industry can not be attributed to their higher experience and knowledge. Secondly, the differences in forecasting accuracy between the financial experts and the novices do not derive from any large differences between them in their forecasting ability for the different industries.

In our robustness test we also find indications that our findings may be affected by variations between the different industries. It is hard to tell which picture is correct as several of our tests also support our findings in the hypothesis testing.

As the results from the informed participants differ very little from the uniformed participants, we do not find any significant effects from our educational information on the forecasting accuracy and the confidence level. These results are also fairly consistent when comparing informed and uniformed participants within both the financial experts and the novices groups. We do however observe that the informed novices tend to make forecasts that are more constant from stock to stock than the uninformed novices. We can again only speculate as to why we see this pattern. Firstly, the novices might have had less prior knowledge about the topics in the information and hence the informed participants learnt something new. Secondly, the novices might be more susceptible to the information than the financial experts. One possible explanation to why the confidence levels of the informed participants do not differ from the confidence level of the uninformed is the low average confidence level in the whole sample. Consequently, there is very little leeway to decrease confidence.

6. CRITICISM

Our thesis uses a one-month time period for the forecasts. It might be argued that the time span was too short. Stock market professionals often work with either much shorter or longer time horizons. Moreover, one month might not be enough time to allow for any greater fundamental changes in the examined companies.

However, there are financial experts who provide one-month forecasts. From a practical point of view it was also necessary to complete the research in time in order to finish the thesis.

Experts and novices stated probabilities that were clearly concentrated to the two middle intervals with fairly low weight given to the extreme intervals. At the same time, the base rate for the actual outcome was largely biased towards the extreme intervals. This observation led us to speculate whether the choice of interval structure was a major contributing factor to the results observed. Had we used a non-symmetrical interval structure or changed the size of the intervals we might have gotten significantly different results. This aspect is not discussed in either Staël von Holstein (1972), Yates et al. (1991) or Muradoglu and Önkal (1994). As we did not have access to their data we could not verify whether it displayed a similar pattern. It would hence be of interest for future research to see if the results would differ if one uses a different setup for the intervals.

In our paper the financial experts were asked to forecast 28 of the 30 stocks on OMXS30-index. These stocks covered nine different industries. It could be questioned whether it is reasonable to assume that financial experts can have a qualified opinion on all the 28 stocks. Moreover, one can hypothesise that the forecasting accuracy would have improved if we had narrowed the number of stocks and let the participants choose the stocks in accordance with their expertise. Still, it seems fair to assume that the financial experts should have had excellent knowledge about some of the stocks in the sample. Also, for the stocks where they had little knowledge one could expect that they would have provided forecasts that were similar to a uniform forecaster. Had they acted in such a way they would likely have outperformed the novices.

7. CONCLUSIONS

Investors put a lot of faith into financial experts. Before the companies publish their annual reports or IPO-prospects the experts are asked to predict the outcome of those events. Their forecasts are often the subject of great interest by the media and the general public who generally believe that financial experts are good at forecasting. Consequently, the quality of these forecasts is of great importance to many parties.

Our thesis examines three different topics. Firstly, we test whether professionals in the stock market are able to make better stock market forecasts than novices. Secondly, we test whether stock market professionals are more confident than novices and also if confidence has a deteriorating effect on the forecasting accuracy. Thirdly, we examine whether it is possible to change a participant's answers through training.

We let 75 people answer a questionnaire in which we asked them to assess the stock prices for 28 stocks listed on the Stockholm Stock Exchange. Of the participants 35 were stock market professionals and 40 were undergraduate students from the Stockholm School of Economics. Added to the questionnaire was a confidence evaluation, where the participants were asked to indicate how confident they felt about their stock market forecasts.

Our data shows that people have limited stock price forecasting abilities as the overall forecasting accuracy was very low. For example, only seven participants were able to beat a uniform forecaster, i.e.

someone who would set the same probability to all intervals. Only one participant was able to outperform a forecaster who follows historical data. Furthermore, we find that the stock market professionals made less accurate forecasts than the students. This was mainly caused by that the students' forecasts changed less from stock to stock than the forecasts of the stock market professionals.

The overall reported confidence level with the participants was low and there were no notable differences between the stock market professionals and the students. However, it appears as if forecasting accuracy decreases the more confident the individual is.

We also tested whether we could affect the forecasting accuracy and the confidence level by informing some of the participants about the difficulties involved in making stock price forecasts. Overall, we find no support in our data for that this simple form of training had any significant effect on the participants.

REFERENCES

Articles and Books

Andersson, P. (2001). Does Experience Matter in Lending? A Process-Tracing Study on Experienced Loan Officers' and Novices' Decision Behvaior. *Journal of Economic Psychology 25*, 471-492.

Andersson, P. (2004). How Well Do Financial Experts Perform? A Review of Empirical Research on Performance of Analysts, Day-Traders, Forecasters, Fund Managers, Investors, and Stockbrokers. *SSE/EFI Working Paper Series in Business Administration No. 2004:9*, 1-29.

Andersson, P., Edman, J., & Ekman, M. (2005). Predicting the World Cup 2002 in Soccer: Performance and Confidence of Experts and Non-Experts. *International Journal of Forecasting 21*, 565-576.

Andersson, P. (2007). Gör Oddssättare Realistiska Bedömningar? En Ekonomisk-Psykologisk Analys av Realismen i Odds. *SSE/EFI Working Paper in Business Administration No 2007:02*, 1-24.

Barber, B., & Odean T. (2000). Trading is Hazardous to Your Wealth: the Common Stock Investment Performance of Individual Investors. *The Journal of Finance 55,* 773-806.

Benos, A. (1998). Aggressiveness and Survival of Overconfident Traders – Heuristics and Biases. *Journal of Financial Markets 1*, 353-383.

Bhandari, G., & Deaves, R. (2006). The Demographics of Overconfidence. *The Journal of Behavioral Finance 7*, 5-11.

Block, R., & Harper, D. (1991). Overconfidence in Estimation: Testing the Anchoring-and-Adjustment Hypothesis. *Organizational Behaviour and Human Decision Processes 49*, 188-207.

Bloomfield, R., Libby, R., & Nelson, M. (2000). Underreactions, Overreactions and Moderated Confidence. *Journal of Financial Markets 3*, 113-137.

Brenner, L., Koehler, D., Liberman, V., & Tversky, A. (1996). Overconfidence in Probability and Frequency Judgements: a Critical Examination. *Organizational Behavior and Human Decision Processes 65*, 212-220.

Brier, W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review 78.*

Camerer, C., & Johnson, E. (1991). The process-performance paradox in expert judgement: How can experts know so much and predict so badly?, In Ericsson, K., & Smith, J. (Eds.), *Toward a general theory of expertise: Prospects and limits*, (pp. 195-217). New York: Cambridge Press.

Cowles, A. (1933). Can Stock Market Forecasters Forecast? *Econometrica 1*, 309-324.

Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (2001). Overconfidence, Arbitrage, and Equilibrium Asset Pricing. *The Journal of Finance 56*, 921-965.

Daniel, K., Hirshleifer, D., & Teoh, S.H. (2002). Investor Psychology in Capital Markets: Evidence and Policy Implications. *Journal of Monetary Economics 56*, 139-209.

De Bondt, W. (1991). What Do Economists Know About the Stock Market? *Journal of Portfolio Management 17*, 84-91.

Dittrich, D., Güth, W., & Maciejovsky, B. (2005). Overconfidence in Investment Decisions: an Experimental Approach. *The European Journal of Finance 11*, 471-491.

Fama, E. (1970). Efficient Capital Markets: a Review of Theory and Empirical Work. *Journal of Finance 25*, 383-417.

Fama, E. (1991) Efficient Capital Markets: II. *Journal of Finance 46*, 1575-1617.

Fischer, G. (1982). Scoring-Rule Feedback and the Overconfidence Syndrom in Subjective Probability Forecasting. *Organizational Behavior and Human Performance 29*, 352-369.

Fischhoff, B., & McGregor, D. (1982). Subjective Confidence in Forecasts. *Journal of Forecasting 1*, 155-172.

Gaeth, G., & Shanteau, J., (1984). Reducing the Influence of Irrelevant Information on Experienced Decision Makers. *Organizational Behaviour and Human Performance 33*, 263-283.

Gervais, S., & Odean, T. (2001). Learning to be overconfident. *The Review of Financial Studies 14*, 1-27.

Gigerenzer, G., Hoffrage, U., & Kleinböltin, H., (1991). Probabilistic Mental Models: a Brunswikian Theory of Confidence. *Psychological Review 98*, 506-528.

Gigerenzer, G., & Todd, P. (2000). Précis of Simple Heuristics That Make Us Smart. *Behavioral and Brain Sciences 69*, 97-124.

Glaser, M., Nöth, M., & Weber, M. (2004). Behavioral Finance. In Koehler, D., & Harvey, N. (Eds.), *The Blackwell Handbook of Judgement and Decision Making*, (pp. 527-548). Blackwell.

Glaser, M., Langer, T., & Weber, M. (2005). Overconfidence of Professionals and Lay Men: Individual Differences Within and Between Tasks? *Working Paper from the University of Mannheim 05:25*, 1-40.

Griffin, D., & Tversky, A. (1992). The Weighing of Evidence and The Determinants of Confidence. *Cognitive Psychology 24*, 411-435.

Griffin, D., & Brenner, L. (2004). Perspectives on Probability Judgement Calibration, In Koehler, D., & Harvey, N. (Eds.), *The Blackwell Handbook of Judgement and Decision Making*, (pp. 177-199). Blackwell.

Hershey, D., & Walsh, D. (2000). Knowledge Versus Experiences in Financial Problem Solving Performance. *Current Psychology 19*, 261-292.

Hirshleifer, D. (2001). Investor Psychology and Asset Pricing. *Journal of Finance 56*, 1533-1597.

Hirshleifer, D., & Shumway, T. (2001). Good Day Sunshine: Stock Returns and the Weather. *Journal of Finance 58*, 1009-1032.

Jacoby, J., Morrin, M., Johar, G., Gürhan, Z., Küss, A., & Mazursky, D. (2001). Training novice investors to become more expert: The role of information assessing strategies. *The Journal of Psychology and Financial Markets 2*, 69-76.

Johnson, J., & Alistair, B. (2001). Calibration of Subjective Probability Judgements in a Naturalistic Setting. *Organizational Behavior and Human Decision Processes 85*, 265-290.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review 80*, 237-251.

Keren, G. (1991). Calibration and Probability Judgements: Conceptual and Methodological Issues. *Acta Psychologica 77*, 217-273.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for overconfidence. *Journal of Experimental Psychology: Human Learning and Memory 6*, 107-118.

Kunda, Z. (1990). The Case for Motivated Reasoning. *Psychological Bulletin 108*, 480-498.

Langer, E., & Roth, J. (1975). Heads I Win, Tail It's Chance: The Illusion of Control as a Function of the Sequence of Outcomes in a Purely Chance Task. *Journal of Personality and Social Psychology 32*, 951-955.

Lichtenstein, S., & Fischoff, B. (1980). Training for Calibration. *Organizational Behaviour and Human Performance 26*, 149-171.

Lichtenstein, S., Fischhoff, B., & Phillips, L. (1982). Calibration of Probabilities: The State of the Art to 1980. In Kahneman, D., Slovic, P., & Tversky, A. (Eds.), *Judgement Under Uncertainty: Heuristics and Biases,* Cambridge Cambridge University Press.

Menkhoff, L., Schmidt, U., & Brozynski, T. (2006) The Impact of Experience on Risk Taking, Overconfidence, and Herding of Fund Managers: Complementary Survey Evidence. *European Economic Review 50*, 1753-1766

Muradoglu, G., & Önkal, D. (1994). An Exploratory Analysis of Managers' Probabilistic Forecast of Stock Prices. *Journal of Forecasting 13*, 565-578.

Murphy, H., & Brown, G. (1985). A Comparative Evaluation of Objective and Subjective Weather Forecasts in the United States. *Journal of Forecasting 3*, 369-393.

Odean, T. (1998). Volume, Volatility, Price and Profit When All Traders Are Above Average. *The Journal of Finance 53*, 1887-1934.

Russo, E., & Schoemaker, P. (1992). Managing Overconfidence. *Sloan Management Review 33*, 7-17.

Shanteau, J. (1992). Compentence in Experts: The Role of Task Characteristics. *Organizational Behavior and Human Decision Processes 53*, 252-266.

Shefrin, H., & Statman, M. (1994). Behavioral Capital Asset Pricing Theory. *The Journal of Financial and Quantitative Analysis 29*, 323-349.

Sniezek, J., Paese, P., Switzer III, F. (1990). The Effect of Choosing on Confidence in Choice. *Organizational Behaviour and Human Decision Processes 46*, 264-282.

Staël von Holstein, C-A. (1972). Probabilistic Forecasting: an Experiment Related to the Stock Market. *Organizational Behaviour and Human Performance 8*, 139-158.

Svenson, O. (1981). Are We All Less Risky and More Skilful Than Our Fellow Drivers? *Acta Psychologica 47*, 143-148.

Taylor, S., & Brown, J. (1988). Illusion and Well-Being: a Social Psychological Perspective on Mental Health. *Psychological Bulletin 103*, 193-210.

Trafimow, D., & Sniezek, J., (1994). Perceived Expertise and Its Effect on Confidence. *Organizational Behavior and Human Decision Processes 57*, 290-302.

Törngren, G., & Montgomery, H. (2004). Worse Than Chance? Performance and Confidence Among Professionals and Laypeople in the Stock Market. *The Journal of Behavioural Finance 5*, 148-153.

Weinstein, N. (1980). Unrealistic Optimism About Future Life Events. *Journal of Personality and Social Psychology 39*, 806-820

Wärneryd, K-E. (2001). *Stock-market Psychology – How People Value and Trade Stocks.* Cheltenham/Northampton: Edward Elgar Publishing Limited.

Yates, J. F. (1988). Analyzing the Accuracy of Probability Judgements for Multiple Events: An Extension of the Covariance Decomposition. *Organizational Behaviour and Human Decision Processes 41*, 281-299.

Yates, J. F., McDaniel, L., Brown, E. (1991). Probabilistic Forecasts of Stock Prices and Earnings: The Hazards of Nascent Experience. *Organizational Behaviour and Human Decision Processes 49*, 60-79.

Ågerup, K. (2002). *Sagan om Adcore* (Stock Letter, Sweden)

Databases

Key financial ratios, Dagens Industri - Stockwatch, URL: http://di.se/Nyheter/?page=/Avdelningar/Indikatorer.aspx%3FMainIndexOptions%3DOMXS30%3BSSE%26indexview%3D1%26EcovID%3DO2 (2007-03-25)

Share prices, OMX official database, URL: http://www.omxgroup.com/nordicexchange/Abouttrading/Indexes/OMX_index_family/OMXS_Local_Index/?languageId=1

APPENDIX – INDUSTRY REGRESSIONS

Table 22
Regression results using the full sample of participants – Dependent variable: PSM-score.

Variable	Coefficient	t	Std error
Constant	1.01***	16.23	0.06
Confidence level	0.04**	2.46	0.02
Dummy that assumes value 1 for financial experts and 0 for novices	0.04*	1.67	0.02
Number of information cues used	0.00	-0.89	0.01
Dummy that assumes value 1 for participants who indicated that they used additional information cues	0.01	0.73	0.02
Dummy that assumes value 1 for informed participants and 0 for uninformed	-0.01	-0.48	0.02
Dummy for stocks belonging to the materials industry	-0.07	-1.34	0.05
Dummy for stocks belonging to the industrials industry	-0.13**	-2.22	0.06
Dummy for stocks belonging to the consumer discretionary industry	-0.16**	-2.43	0.07
Dummy for stocks belonging to the consumer staples industry	0.40***	-4.92	0.08
Dummy for stocks belonging to the health care industry	-0.06	-0.67	0.08
Dummy for stocks belonging to the finance industry	-0.19***	-2.61	0.07
Dummy for stocks belonging to the IT industry	-0.23***	-3.08	0.07
Dummy for stocks belonging to the telecommunications industry	-0.05	-0.67	0.07
The confidence variable multiplied with the materials dummy	0.00	0.48	0.02
The confidence variable multiplied with the industrials dummy	0.02	0.81	0.02
The confidence variable multiplied with the consumer discretionary dummy	0.02	0.75	0.02
The confidence variable multiplied with the consumer staples dummy	-0.03	-1.05	0.03
The confidence variable multiplied with the health care dummy	0.02	0.70	0.02
The confidence variable multiplied with the finance dummy	0.01	0.25	0.02
The confidence variable multiplied with the IT dummy	0.02	0.80	0.03
The confidence variable multiplied with the telecommunications dummy	0.01	0.18	0.03

* $p < 0.05$
** $p < 0.01$
*** $p < 0.001$

Table 23
Regression results using only the financial experts – Dependent variable: PSM-score.

Variable	Coefficient	t	Std error
Constant	0.92***	4.40	0.21
Confidence level	0.11	1.57	0.07
Dummy that assumes value 1 for informed participants and 0 for uninformed	0.03	0.85	0.03
Number of years of professional experience	0.00	-0.81	0.00
Number of information cues used	-0.03***	-3.15	0.01
Dummy that assumes value 1 for participants who indicated that they used additional information cues	0.06*	1.89	0.03
Dummy that assumes value 1 for participants who worked as stock traders	0.03	1.01	0.03
Dummy for stocks belonging to the materials industry	0.06	0.25	0.23
Dummy for stocks belonging to the industrials industry	-0.03	-0.14	0.22
Dummy for stocks belonging to the consumer discretionary industry	-0.07	-0.30	0.23
Dummy for stocks belonging to the consumer staples industry	0.17	0.59	0.29
Dummy for stocks belonging to the health care industry	-0.22	-0.79	0.28
Dummy for stocks belonging to the finance industry	-0.04	-0.17	0.23
Dummy for stocks belonging to the IT industry	-0.22	-0.89	0.24
Dummy for stocks belonging to the telecommunications industry	0.09	0.38	0.25
The confidence variable multiplied with the materials dummy	-0.04	-0.51	0.08
The confidence variable multiplied with the industrials dummy	-0.03	-0.37	0.07
The confidence variable multiplied with the consumer discretionary dummy	-0.02	-0.26	0.08
The confidence variable multiplied with the consumer staples dummy	0.25**	-2.36	0.11
The confidence variable multiplied with the health care dummy	0.07	0.67	0.10
The confidence variable multiplied with the finance dummy	-0.06	-0.76	0.08
The confidence variable multiplied with the IT dummy	0.00	0.04	0.09
The confidence variable multiplied with the telecommunications dummy	-0.06	-0.64	0.09

* $p < 0.05$
** $p < 0.01$
*** $p < 0.001$

Table 24
Regression results using only the novices – Dependent variable: PSM-score.

Variable	Coefficient	t	Std error
Constant	0.91***	6.44	0.14
Confidence level	0.06	1.18	0.05
Dummy that assumes value 1 for informed participants and 0 for uninformed	-0.02	-0.80	0.02
Number of information cues used	0.02**	2.51	0.01
Dummy that assumes value 1 for participants who indicated that they used additional information cues	-0.01	-0.25	0.02
Dummy for stocks belonging to the materials industry	-0.06	-0.37	0.16
Dummy for stocks belonging to the industrials industry	-0.02	-0.17	0.15
Dummy for stocks belonging to the consumer discretionary industry	-0.02	-0.16	0.16
Dummy for stocks belonging to the consumer staples industry	-0.21	-0.92	0.22
Dummy for stocks belonging to the health care industry	0.05	0.26	0.19
Dummy for stocks belonging to the finance industry	-0.02	-0.14	0.16
Dummy for stocks belonging to the IT industry	-0.19	-1.08	0.18
Dummy for stocks belonging to the telecommunications industry	-0.04	-0.21	0.18
The confidence variable multiplied with the materials dummy	-0.01	-0.12	0.06
The confidence variable multiplied with the industrials dummy	-0.03	-0.58	0.05
The confidence variable multiplied with the consumer discretionary dummy	-0.03	-0.61	0.06
The confidence variable multiplied with the consumer staples dummy	-0.07	-0.80	0.08
The confidence variable multiplied with the health care dummy	-0.06	-0.088	0.07
The confidence variable multiplied with the finance dummy	-0.04	-0.73	0.06
The confidence variable multiplied with the IT dummy	0.00	0.08	0.06
The confidence variable multiplied with the telecommunications dummy	0.02	0.34	0.06

* $p < 0.05$
** $p < 0.01$
*** $p < 0.001$

APPENDIX – THE QUESTIONNAIRE

This document was sent to the participants who were chosen to receive the training information. The only difference between this document and the one being sent to the participants who did not receive information is the request to read the information.

Thank you for taking the time to fill out our questionnaire!

With this questionnaire we want to examine your skills in assessing future stock price developments. You are guaranteed complete anonymity and you will be presented with the results after the thesis has been submitted.

Below you are asked to assess 30 different stocks from the A-list on the Stockholm Stock Exchange. For each stock there are six different intervals indicated for potential stock price developments during the period 2nd of April till 2nd of May 2007. The intervals are the following:

Increase >5%, Increase 5-3%, Increase 3-0%, Decrease 0-3%, Decrease 3-5%, Decrease >5%

You are asked to assess the probability that the actual stock price change will end up in each of the different intervals. The probabilities are to be indicated on a scale from 0 to 100% and the sum of the six intervals should sum up to 100%. See example below:

Change:	**>+5%**	**+5-3%**	**+3-0%**	**-0-3%**	**-3-5%**	**>-5%**	**Total:**	Verify below:
Probability:	0%	5%	20%	25%	30%	20%	100%	Correct Summation!

In the example above, the participant believes that there is a 20% probability that the stock price will decrease with more than 5% during the period 2nd of April to 2nd of May. Furthermore, she believes that there is a 30% probability that the stock price will decrease with 3-5% etc. If the sum of the probabilities in the six different intervals sum up to 100%, the box *Verify below* will indicate *Correct Summation!*.

You have complete discretion in your choice of tools and information to use, although we ask you to complete the questionnaire on your own. We ask you to submit the questionnaire no later than 1st of April.

Good luck!

(The questionnaire can be found under the second sheet)

Andreas Malmquist
20727@student.hhs.se
Phone: +46 709 40 79 74

André Rhodin
20749@student.hhs.se
Phone: +46 707 92 14 65

Please note!

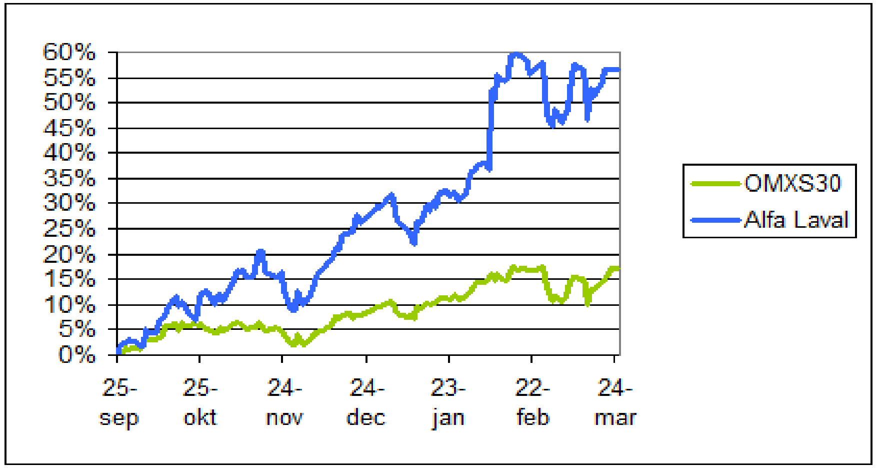
According to common financial theory, all public information is already reflected in the stock price. Hence, it is very difficult to correctly forecast the future stock price development. Furthermore, several research papers have shown that it is difficult for individuals to assess probabilities, the reason for this often being that the individual provides to narrow intervals.

Even though some individuals have the capacity to assess probabilities, their assessments tend to be distorted by systematic errors, such as an exaggerated confidence in oneself. Researchers have concluded that this exaggerated confidence in oneself is often stronger with “professionals” than with novices, even though this belief is often not matched by a higher performance.

It has been shown that individuals differ in their forecast accuracy depending on, for example the difficulty level of the task and the time span between the assessment and the outcome, the so called feedback-horizon. Also, individuals with an exaggerated confidence in oneself indicate too high probabilities in events that occur fairly infrequently, and vice versa for events that occur more often.

Practise example - Alfa Laval

Stock price development September 2006 - Mars 2007



Facts 2006 - SEK

Turnover/stock	177,32
Profit/stock	15,31
Equtiy/stock	60,11
Assets/stock	167,90
Dividends/stock	6,25
Earnings/stock	16,09

Indicate probabilities that the stock price change for the period 2 April till 2 May will fall within the following intervals.

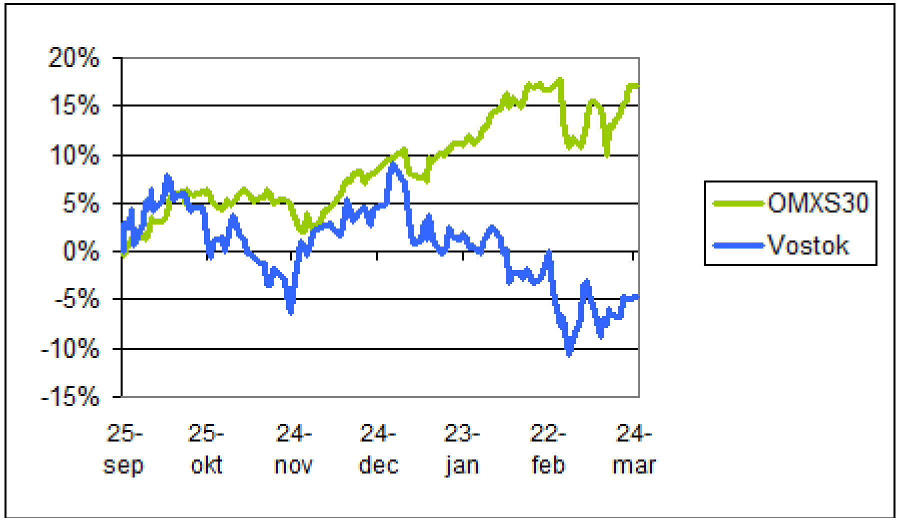
Change:	>+5%	+5-3%	+3-0%	-0-3%	-3-5%	>-5%	Total:	Verify below:
Probability:	10%	50%	40%	0%	0%	0%	100%	Correct Summation!

How confident are you in your assessments? Indicate your answer with an X:

Very uncertain	Uncertain	Somewhat certain	Certain	Very certain
			X	

Vostok Nafta SDB

Stock price development September 2006 - Mars 2007 **Facts 2006 - SEK**



Turnover/stock	234,26
Profit/stock	167,00
Equtiy/stock	526,43
Assets/stock	602,76
Dividends/stock	
Earnings/stock	-5,57

Indicate probabilities that the stock price change for the period 2 April till 2 May will fall within the following intervals.

Change:	**>+5%**	**+5-3%**	**+3-0%**	**-0-3%**	**-3-5%**	**>-5%**	**Totalt:**	Verify below:
Probability:							0%	Correct Summation!

How confident are you in your assessments? Indicate your answer with an X:

Very uncertain	Uncertain	Somewhat certain	Certain	Very certain

Indicate the aspects that you used the most with an X

Volatility level during the period:

Volatility change during the period:

Large single movements:

Trend during the period:

Financial key numbers:

Official information from the companies:

External analysis:

Information from friends or colleagues:

Other information, for example:

Background questions:

Graduation year:

Age:

Profession:

Industry:

Years of professional experience:

Percentage of salary comprised of bonus:

Do you used timeseries/graphs to collect information in your daily work?

User name (the user name will be indicated in the ranking list):

Own comments: