The Stockholm School of Economics
Department of Accounting and Financial Management
Bachelor Thesis
May 2021

Earnings Forecasts and Stock Price Data

How stock prices can be used to forecast less biased and more accurate earnings

Erik Bergmark
24335@student.hhs.se

Gustav Möller
24376@student.hhs.se

Abstract

In this study, we investigate the effect of incorporating stock price data in model-based earnings forecasts. Previous research shows that stock price data can be used to forecast less bias and more accurate earnings. We expand the literature of earnings forecasts on Swedish data and our results are suggestive that stock price can be used to forecast less biased, but not more accurate, earnings. Our findings also indicate that the stock price contains incremental information about future earnings in the short-term, compared to pure-accounting based models and forecasts by financial analysts. Furthermore, we also look into the effect that firm characteristics have on forecast performance. The results are suggestive that the forecast accuracy for model-based forecasts is dependent on firm characteristics, such as industry membership, size, and earnings-to-price ratio.

Tutor: Peter Alexander Aleksziev
Keywords: Forecasts, earnings, stock price, bias, accuracy
Acknowledgements: We would like to thank Peter Alexander Aleksziev for always taking the time to help us and for his much-appreciated feedback and assistance.

Table of Content

1. Introduction

Earnings forecasts are a key component of modern-day finance. They are, among other things, used when making investment decisions, estimating implied cost of capital, and in corporate valuation. There are several different methods used to forecast earnings. Common forecasting methods include financial analysts' forecasts, time-series models, random walk models, and cross-sectional models. Throughout the last 60 years, these models and methods have been compared and tested to conclude how to generate the most optimal forecast.

With the purpose of finding a less biased and more accurate forecast model, Harris and Wang, (2019) develop a cross-sectional model that forecasts earnings based on stock price and accounting data. They present their model in the article "*Model-based earnings forecasts vs. financial analysts' earnings forecasts*" in which they compare the quality of their model to other model-based forecasts and with forecasts made by financial analysts. They conducted their study on US listed firms, and we intend to replicate this on Swedish listed companies. To the best of our knowledge, this has not been done before. In our study, we also include the HDZ model by Hou et al., (2012), a cross-sectional model based on accounting and dividend data that is considered a benchmark for model-based forecasts (Harris and Wang, 2019). To further expand the study, we have also added an additional model, the earnings persistence (EP) model. This model was developed by Li and Mohanram, (2014) who found that it outperformed the HDZ model. Apart from these models, we also include a first order autoregressive and a random walk model. These models are also often used as a benchmark for model-based forecasts (Harris and Wang, 2019).

To measure the quality of the forecasts we calculate the mean bias and accuracy, in line with previous literature (Harris and Wang, 2019). In addition to Harris and Wang, (2019), we also conduct pairwise t-tests on the forecasts to determine if they are statistically significantly different. Forecast bias measures if the forecast over- or underestimates earnings, while accuracy is the mean forecast error. Previous literature shows that analysts' consensus is more accurate than models over short time horizons while models perform better on longer horizons and are less biased (Harris and Wang, 2019; Bradshaw et al., 2012; Azevedo et al., 2020). The models also have the benefit of covering more companies since they only base their forecast on quantitative data while the analysts' forecasts are more time consuming as their analyses are deeper and incorporate several different information sources in their predictions (Harris and Wang, 2019).

To forecast future earnings, we follow the method used by Harris and Wang, (2019) and estimate the models using a pooled cross-section ordinary-least-square (OLS) regression. In line with Harris and Wang, (2019), we also use a rolling window of ten years to generate an out-of-sample forecast so that each forecast is only based on information available at that time. Thereafter, we calculate the mean bias and accuracy for these forecasts and compare the result to the analyst consensus forecasts and conduct efficiency and encompassing tests. In line with Harris and Wang, (2019), we find support for analyst forecasts being more accurate than model-based forecasts for the one- and two-year ahead forecasts, while also being upwardly biased. Our results are suggestive, but not conclusive, that stock price can be used to forecast less biased earnings, similar to Harris and Wang, (2019). However, we do not find any indication that this yields more accurate forecasts compared to pure-accounting based models for Swedish companies, which contradicts previous research (Harris and Wang, 2019). The efficiency test further indicates that there is incremental information contained within the stock price, and the encompassing test also suggests that each model-based forecast contains information not found in the I/B/E/S consensus, at the one- and two-year ahead forecasts. For the five-year ahead forecasts, the results were not statistically significant. Furthermore, we find suggestive support for dividends' predictive power of future earnings, in line with previous literature (Hou et al., 2012; Zhou and Ruland, 2006). We also conduct an analysis based on firm characteristics where we find that forecast performance is dependent on industry membership, size, and earnings-to-price ratio (E/P).

1.1. Purpose

The purpose of this study is to investigate *whether stock price data can be used to forecast more accurate and less biased earnings for Swedish listed companies*. Earnings forecasts are an important concept of modern-day finance, however, most of the research has been focused on the US market. Therefore, expanding the literature for countries such as Sweden, where the research is more limited, is highly relevant. To further investigate the applicability of model-based forecasts, we also look into the possible effects of firm characteristics on forecast performance.

1.2. Contribution

This study contributes to existing literature in three ways. Firstly, it expands the research of earnings forecasts on Swedish data. Secondly, we compare the earnings persistence (EP) model developed by Li and Mohanram, (2014) to the HDZ model developed by Hou et al., (2012), and the PW model developed by Harris and Wang, (2019). The PW and EP model have both been found to outperform the HDZ model, but to the best of our knowledge, these three models have not been compared using the same dataset. Thirdly, we expand the analysis of how firm characteristics affect forecast performance (Harris and Wang, 2019) by conducting t-tests between the model-based and the analyst consensus forecasts.

1.3. Scope

We replicate the study of Harris and Wang, (2019) on Swedish data and therefore follow their methodology as closely as possible. Harris and Wang, (2019) used a sample of approximately 2 500 firms between July 1976 and June 2015 to estimate their models. Due to limited data for Swedish companies, our sample is smaller and contains 375 firms between January 1988 and December 2019. We evaluate forecast performance using the same measurements as Harris and Wang, (2019), by calculating forecast bias and accuracy. We also conduct efficiency and encompassing tests, and break the sample down into groups based on firm characteristics to test the effects of industry membership, size, and E/P ratio on forecast performance. Furthermore, we expand the benchmark study (Harris and Wang, 2019) by conducting t-tests between the model-based and analyst consensus forecasts for certain firm characteristics.

1.4. Disposition

This study is structured into 8 sections. Section 2 consists of a literature review and a description of the theories behind different forecasting methods. Section 3 describes the methodology and introduces our hypotheses while section 4 describes our data collection and sample selection process. Section 5 presents descriptive statistics, a Pearson correlation matrix, the average estimated coefficients for the models, the calculated bias and accuracy for each forecast, the efficiency and encompassing tests, and the forecast performance based on firm characteristics. Section 6 contains the analysis of the results and a discussion of their implications, as well as reflections on our research method. In section 7 we provide suggestions for future research within the topic and section 8 presents our conclusions. After section 8, references and the appendix can be found.

2. Literature review and theory

In this section, we review some of the most central findings related to earnings forecasts methods and models. We prioritize more recent findings and older key-articles since a large quantity of research has been conducted on the subject in the past 60 years. The review starts with a general overview of previous research, including the background to the field and some central findings. This is followed by expanding on the predictive power of different financial metrics included in the models we test. Afterwards, we elaborate on the theoretical framework of the model-based forecasts, and we present the theoretical framework and function of the model developed by Harris and Wang, (2019). Lastly, we present previous findings on Swedish data and our hypotheses development.

2.1. Previous research

A great deal of modern-day earnings forecasts research derives from the early findings of Ball and Brown, (1968). They argue that earnings information in accounting statements can affect investors' valuation of companies and that abnormal profits can be made given superior information of future earnings. This resulted in a large quantity of research being conducted comparing the forecasts from financial analysts to time-series based models to determine which method yielded superior information of future earnings. According to Bradshaw et al., (2012), the literature culminated in the study of Brown et al., (1987). They concluded that forecasts by financial analysts were superior. Their argument was based on the timing and information advantage of analysts over time-series models. The timing advantage refers to analysts' abilities to better update their forecasts as new information becomes available, before the forecast is published. The information advantage refers to their ability to better use the currently available information in their forecasts.

Analyst consensus forecasts refers to the aggregation of numerus forecasts from financial analysts covering a specific company retrieved from databases such as the Institutional Brokers Estimate System (I/B/E/S). The findings of Brown et al., (1987) still holds up to some extent as analyst forecasts are still more accurate on shorter forecast horizons (one- and two-year ahead) compared to more recent models (Azevedo et al., 2020). However, the superiority of analysts' forecasts has been questioned and challenged in later research (Bradshaw et al., 2012).

The main downsides of analyst's earnings forecasts are that they are upwardly biased (Harris and Wang, 2019), their sluggishness and poor long-term estimates (Azevedo et al., 2020). One of the more prominent explanations for the bias is that it results from a conflict of interest (Dugar and Nathan, 1995; Francis and Philbrick, 1993; Lin and McNichols, 1998; Ramnath et al., 2008, cited in Harris and Wang, 2019, p. 425). Becker, (2001) expanded further on the conflict of interest explanation and argues that analysts at investment banks may generate positively biased forecasts in exchange for investment banking business. The conflict of interest could also potentially be linked to the high short-term accuracy of analysts' forecasts. Previous research indicates that analysts incorporate different types of private information and other market information (Alford and Berger, 1999; Fried and Givoly, 1982; Kross et al., 1990; Sougiannis and Yaekura, 2001, cited in Harris and Wang, 2019, p. 426). The sluggishness issue is a result of analysts not updating their forecasts when new information becomes available (Guay et al., 2011). The sluggishness issue somewhat contradicts the argument of Brown et al., (1987) regarding the timing advantage of financial analysts. Guay et al., (2011) proceed to explain the problem of sluggish forecasts by showing that it is problematic to use them for calculating the implied cost of capital. The drawback of poor long-term estimates is in line with the findings of Bradshaw et al., (2012). They show that for time periods longer than two years, a random walk model provides more accurate forecasts than analysts. Mitigating these issues has been the goal for many of the model-based forecast methods (Harris and Wang, 2019).

Model-based forecasts are not a new concept. Ball and Watts, (1972) modelled future earnings using a submartingale approach, which is closely related to the random walk process, while Ball and Brown, (1968) used a time-series model that used historical values to forecast future earnings. However, after the findings of Brown et al., (1987) that concluded analysts being superior compared to model-based forecasts, the amount of research in this area decreased heavily (Bradshaw et al., 2012). More recently, the literature has seen a proliferation of more information intensive models compared to the RW and autoregressive models. Richardson et al., (2010) is a central piece of research within model-based earnings forecast literature. They created a formalized forecasting framework that forecast next period's earnings as a function of book value, changes in book value, current earnings, and some non-accounting-based metrics. Several of the more recent models can be viewed as different variations of this model (Harris and Wang, 2019). Hou et al., (2012) developed the HDZ model, a cross-sectional forecast model based on accounting and dividend data. The structure of this model reduces the bias and increases firm coverage and has become a benchmark for cross-sectional earnings forecast models (Harris and Wang, 2019). However, the HDZ model has received criticism by Li and Mohanram, (2014) who showed that cross-sectional models in general perform worse for companies without I/B/E/S analyst coverage, weakening the argument of increased firm coverage.

To develop the research on cross-sectional forecast models, Li and Mohanram, (2014) introduced two models of their own. One being the earnings persistence model (EP model), which forecasts earnings based on past earnings and the persistence of profits and losses, and the residual income model (RI model), which is based on the concept residual income valuation developed by Ohlson, (1995) and Feltham and Ohlson, (1995; 1996). Both of these models are relatively simple and are less information intensive compared to the HDZ model. Harris and Wang, (2019) advanced the field further by creating the PW model, a cross-sectional earnings forecast model including accounting and stock price data. This is the benchmark study we are replicating. They deepened their analysis by investigating how firm characteristics affect the quality of earnings forecasts. The more recent model by Azevedo et al., (2020) takes a different approach to forecasting earnings. Their model combines analyst forecasts with time-series models. Using this combination, they take advantage of the high short-term accuracy of financial analysts while mitigating the poor long-term performance of analysts by also using a time-series forecasts component in their model.

2.2. Predictive power of financial metrics

Central to all of the mentioned model-based forecasts is the selection of independent variables. In this section, we will explain the predictive power of different financial metrics that are used to forecast earnings. We have divided them into accounting data, dividend data, and stock price data. This makes it easier to address the explanatory power of each different financial metric.

2.2.1. Accounting data

The idea behind the predictive power of accounting metrics also partially derives from the findings of Ball and Brown, (1968). They showed that accounting data affect investors' valuations of stocks. The connection between accounting numbers and profitability is further explained by the cross-sectional profitability models by Fama and French, (2000; 2006). They also present findings arguing that earnings are highly persistent. Expanding on the persistence argument, it is probable that historic earnings can be used to predict future earnings. However, findings from Chan et al., (2003) have criticised the persistence argument as they find no significant connection between historic and future earnings for long-term horizons. Explaining on the arguments of Chan et al., (2003) earnings are not optimal to use when making long-term forecasts. However, following the arguments of Fildes, (1991), forecasts based on multiple non-perfectly correlated information sources provide better forecasts and therefore weakens the potential argument of excluding earnings as a variable, even in the long-term. Support for the general predictive power of accounting variables can also be found in Richardson et al., (2010). They show that the information in financial statements can be used to forecast earnings.

Commonly used accounting-based financial metrics are operating accruals, assets, and current earnings. These are metrics included in the models developed by Harris and Wang, (2019) and Hou et al., (2012) which both found a relationship between them and future earnings. Findings from Freeman et al., (1982) further strengthen the support for predictive power in accounting data as they show that an estimate using book-rate-of-return yields more accurate forecasts than the previously mentioned submartingale process of Ball and Watts, (1972).

2.2.2. Dividend data

Another common variable is dividends. Using the widely known Gordon Growth Model (Gordon, 1962), the relationship between dividends, earnings and dividend growth rate, and constant cost of equity, can be explained from the following equation:

$$PV = \frac{Div_1}{r - g}$$

Under the assumption of a constant payout ratio and that expected returns are not affected by dividend policy, this function can be rewritten to indicate that a lower payout ratio yields a higher growth rate. Higher dividends should then be negatively correlated with earnings growth. However, findings from Zhou and Ruland, (2006) and Arnott and Asness, (2003) show the opposite. They find that high-dividend paying companies see stronger earnings growth. A possible explanation for their findings is the dividend smoothing model by Lintner, (1956). He argues that corporate executives are reluctant to lower dividends due to the negative signal it sends to the markets and therefore only increase the dividends when they are sure that they can maintain a higher level of dividends in the long-term. This theory would imply that only companies who are certain that they will see a positive earnings growth and can sustain higher dividends will increase them, hence explaining the predictive. This reasoning is supported by

Watts, (1973) who found a positive relationship between future earnings and unexpected dividend changes. The predictive power is also displayed in Hou et al., (2012) as they find a positive relationship between dividends and future earnings.

2.2.3. Stock price data

The previously mentioned inclusion of stock price data by Harris and Wang, (2019) is motivated by the efficient market hypothesis, i.e., that all currently available information is reflected in the stock price (Fama, 1970). The notion of stock price having predictive power for future earnings is supported by previous literature as shown by Azevedo et al., (2020) and Gao and Wu, (2014), who developed forecasting models including stock prices. Azevedo et al., (2020) motivate their inclusion of stock prices on the findings of Richardson et al., (2010) and Ashton and Wang, (2013), while Gao and Wu, (2014) based their inclusion of stock price on the findings of Elgers and Murray, (1992), Nekrasov and Ogneva, (2011), and Weiss et al., (2008). Gao and Wu, (2014) primarily find that incorporation of past returns improves long-term growth estimates. Guay et al., (2011) find that the information content of stock prices can help reduce the sluggishness of analyst forecasts by allowing the analysts more time to incorporate the recent price movements of stocks in their own forecasts. A potential downside of using stock price is the proliferation of tech companies where stock prices soar for companies that are still unprofitable (Financial Times, 2019), which could potentially weaken the predictive power in the future.

2.3. Theoretical framework of forecasting models

2.3.1. Random walk models

Random walk models forecast future earnings as the current period's earnings and find support from Ball and Watts, (1972). They find that earnings follow a submartingale process, which means future values are equal to the current value, but with a positive error term. The random walk method indirectly yields results pointing towards persistent earnings in line with the findings of Fama and French, (2006).

2.3.2. Cross-sectional models

The other models included in this study are cross-sectional forecasting models that use a specified set of independent variables to create a linear function to forecast earnings. These cross-sectional earnings forecast models are based upon the cross-sectional profitability models of Fama and French, (2000; 2006). Using cross-sectional models has the advantage of not requiring time-series data. As long as data is available for two or more consecutive years, the data can be used to estimate the models. Using time-series data, successful companies become overrepresented and non-successful companies are removed from the sample, which increases survivorship bias and lowers the quality of the forecasts (Hou et al., 2012).

The cross-sectional models that are included in our study are the HDZ model, the PW model, and the EP model. We also include the first order autoregressive model (AR(1) model) as used by Harris and Wang, (2019). Cross-sectional models can be univariate or multivariate. One advantage of multivariate models is that they build on the idea that several non-perfectly correlated sources will lead to better predictions of future earnings, compared to any variable in isolation Fildes, (1991). The HDZ and PW models are multivariate as they forecast earnings based on several independent variables, while the EP and AR(1) models are univariate as they forecast earnings solely on historical values of past earnings.

2.3.3. Forecasting with the PW model

Of the previously mentioned categories, the financial metrics used in the multivariate cross-sectional PW model by Harris and Wang, (2019) are accounting and stock price metrics. The cross-sectional structure derives from the profitability models of Fama and French, (2000; 2006) while the model builds on the PW framework developed by Pope and Wang, (2005), hence the name “the PW model”. Their model could potentially also be seen as one of many variations of the previously mentioned Richardson et al., (2010) framework. Harris and Wang, (2019) expands on the existing theoretical forecasting model by Ashton and Wang, (2013) by adding operating accruals as a variable. They also use the model to calculate future earnings and not only implied cost of capital and long-term growth, which was the original function of the model by Ashton and Wang, (2013). The PW model forecasts earnings based on stock price, operating accruals, book value, and current earnings.

2.4. Swedish findings

Skogsvik, (2008) investigates if information in financial statements can be used to predict future profitability using a model including dividends, new share issues, book value of owner's equity, earnings and return on equity. She finds that a model solely based on return on equity yields more accurate forecasts compared to the more information rich accounting-based model. The method of using return on equity to predict future profitability has also been used by Runsten, (1998). Skogsvik, (2008) does not include stock prices as an independent variable in their model but presents results which suggest that a possible relationship exists between stock prices and changes in return on equity. However, to the best of our knowledge, a potential relationship between stock prices and the quality of forecasts for firms with specific characteristics has not been explored previously for Swedish data. Other earnings forecast models such as the HDZ model have been tested on Swedish data but solely on a bachelor thesis level.

2.5. Hypotheses development

As stated in the previous sections, the relationship between earnings and different types of financial metrics, including stock price, has received a lot of attention in previous research. The quality of model-based forecasts, however, can be largely affected by firm characteristics (Harris and Wang, 2019), and to the best of our knowledge, this research is also limited for Swedish data. Increasing the literature within this area is of relevance when determining the practical applicability of model-based forecasts.

Firstly, characteristics of the financial statements for certain industries, such as financial services or telecom, are fundamentally different from those of other industries (Harris and Wang, 2019). Some industries are also heavily affected by regulatory requirements, such as financial services, which further affects their financial statements (Harris and Wang, 2019). It is therefore reasonable to assume that models that base the earnings forecasts on accounting data will perform worse for these industries (Harris and Wang, 2019), and hence will be outperformed by analyst consensus forecasts. Secondly, Harris and Wang, (2019), based on the finds of Lee and So, (2017), suggests that analyst coverage and forecast accuracy is related to the size of companies. Large companies generally have more analysts covering them. Therefore, for short-term forecast horizons we can expect analyst consensus to be more accurate than model-based forecasts (Harris and Wang, 2019). Firm size is defined as market capitalization in line with previous literature (Harris and Wang, 2019). One argument for this definition is that market capitalization is forward looking, as it is based on the stock price (Dang et al., 2018). Thirdly, analysts usually use a high E/P ratio as a screening tool to find undervalued companies and these companies should therefore be covered by more analysts, and just as for larger firms, the analyst consensus should outperform model-based forecasts (Harris and Wang, 2019).

Both the PW and EP models outperformed the HDZ model in their original articles. However, these models have, to the best of our knowledge, not been compared directly to each other, and therefore testing these models on the same dataset could yield interesting results. Since the PW model developed by Harris and Wang, (2019) includes more information, we argue it should outperform the EP model, based on the arguments of Fildes, (1991).

3. Method

In this section we explain the methodology behind each model and which variables are included. We also introduce our hypotheses and explain the motives behind them.

3.1. Hypotheses

Harris and Wang, (2019) expand their analysis by looking at how firm characteristics affect forecast quality. These characteristics are industry membership, company size, and E/P value, as explained in section 2.6. Following Harris and Wang, (2019), we define firm size based on market capitalization, where the number of shares outstanding is multiplied with the stock price. When determining whether one forecast outperforms another, this is based on the forecast bias and accuracy measurement. The hypotheses that we test are:

Hypothesis 1: Analyst consensus outperforms model-based forecasts for industries with discrete financial statements.

Hypothesis 2: Analyst consensus outperforms model-based forecasts for larger companies.

Hypothesis 3: Analyst consensus outperforms model-based forecasts for companies with a high E/P value.

We have also decided to include the earnings persistence model by Li and Mohanram, (2014), which they found outperformed the HDZ model developed by Hou et al., (2012). Based on the reasoning in section 2.6., we will also test the hypothesis that the PW model outperforms the EP model.

Hypothesis 4: The PW model outperforms the EP model.

3.2. Forecast models

To investigate our research questions and to test the hypotheses in section 3.1., we use the AR(1), HDZ, PW, EP, and RW models to forecast earnings. Thereafter, we calculate the forecast bias and accuracy for each of the model-based forecasts and for the analysts' consensus forecasts retrieved from I/B/E/S. The following section presents each model in detail and explains how they are used to forecast earnings. When estimating the models, realized earnings per share is used as the dependent variable. The independent variables use the same definition across all models.

3.2.1. Random walk model

The random walk model (RW model) uses the current earnings as the forecast for next year's earnings. It is included for replication purposes and as a benchmark following previous literature (Hou et al., 2014; Harris and Wang, 2019). The random walk is not used as extensively for forecasting purposes, one reason being that it is impractical for calculating implied cost of capital (Li and Mohanram, 2014). However, findings from Gerakos, and Gramacy, (2013) show that the model is still relevant as they found that the RW model under some circumstances outperforms the HDZ model. The RW model predicts one-period ahead earnings as:

$$EPS_{j,t+k} = EPS_{j,t}$$

where $EPS_{j,t}$ is the earnings per share for firm j at time t. Earnings per share is measured as net income before extraordinary items in accordance with Harris and Wang, (2019).

3.2.2. Autoregressive model

The autoregressive model of the first order (AR(1) model) used by Harris and Wang, (2019) forecast earnings based on current earnings and an indicator variable determining whether the earnings were negative or not. The AR(1) model is often used in combination with the RW model as benchmarks when comparing model-based forecasts (Harris and Wang, 2019), and uses the following function to forecast one-period ahead earnings:

$$EPS_{j,t+k} = \beta_{0jk} + \beta_{1jk} * EPS_{j,t} + \beta_{2jk} * NEGE_{j,t} + \varepsilon_{j,t+k}$$

where $EPS_{j,t}$ is the earnings per share for firm j at time t, $NEGE_{j,t}$is an indicator variable that equals one if the earnings for firm j at time t is negative and zero otherwise. The indicator variable $NEGE_{j,t}$ is included because negative earnings are less persistent compared to positive earnings (Harris and Wang, 2019) and is used consistently in the other models, except the RW model.

3.2.3. HDZ model

In the original study by Hou et al., (2012), the HDZ model outperforms analysts in terms of lower bias and larger firm coverage. The following function is used to forecast one-period ahead forecasts for the HDZ model:

$$EPS_{j;t+k} = \beta_{0jk} + \beta_{1jk} * ASSETS_{j,t} + \beta_{2jk} * DIVIDEND_{j,t} + \beta_{3jk} * DD_{j,t} + \beta_{4jk} * EPS_{j,t} + \beta_{5jk} * NEGE_{j,t} + \beta_{6jk} * ACCRUALS_{j,t} + \varepsilon_{j,t+k}$$

where $EPS_{j,t}$ is the earnings per share for firm j at time t. $ASSETS_{j,t}$, $DIVIDEND_{j,t}$ and $ACCRUALS_{j,t}$ are, respectively, total assets at the end of the period, total common dividend paid during the year, and total operating accruals, all on a per share basis, for firm j at time t. $NEGE_{j,t}$ is an indicator variable that equals one if the earnings for firm j at time t is negative and zero otherwise. $DD_{j,t}$ is an indicator variable that equals one if firm *j* paid a dividend during the year *t*, and zero otherwise. The accounting and dividend variables are divided by the number of shares outstanding to increase comparability between firms, and to decrease heteroscedasticity (Harris and Wang, 2019).

Operating accruals are defined as the difference between earnings and cash flow from operating activities (Harris and Wang, 2019). In their original study, Harris and Wang, (2019) used two different methods to calculate accruals, one before 1988, and one 1988 and onwards. Before 1988, operating accruals were defined as “the change in non-cash current assets less the change in current liabilities, excluding short-term debt and taxes payable, minus depreciation and amortization expense” (Harris and Wang, 2019). From 1988 and onwards, operating accruals were redefined as “the difference between earnings and cash flows from operations” (Harris and Wang, 2019). Our dataset only includes one year where the first definition would be used, which only includes one observation. Therefore, we have decided to exclude the first definition since it would not have a measurable effect on the result.

3.2.4. PW model

This is the model developed by Harris and Wang, (2019) that we covered the background to in section 2.3.3. Harris and Wang, (2019) found this model outperformed financial analysts and the HDZ model in terms of mean bias and accuracy. This model incorporates accounting data and stock price data to forecast earnings. The following function is used to forecast one-period ahead earnings:

$$EPS_{j;t+k} = \beta_{0jk} + \beta_{1jk} * PRICE_{j,t} + \beta_{2jk} * EPS_{j,t} + \beta_{3jk} * NEGE_{j,t} + \beta_{4jk} * BOOK_{j,t} + \beta_{5jk} * BOOK_{j,t-1} + \beta_{6jk} * PRICE_{j,t-1} + \beta_{7jk} * ACCRUALS_{j,t} + \varepsilon_{j,t+k}$$

where $EPS_{j,t}$ is the earnings per share for firm j at time t. $PRICE_{j,t}$, $BOOK_{j,t}$and $ACCRUALS_{j,t}$ is, respectively, the stock price three months after the end of the fiscal year, book value of equity, and operating accruals on a per share basis, for firm j at time t. The definition for operating accruals is the same as in used for the HDZ model above. $NEGE_{j,t}$is an indicator variable that equals one if the earnings for firm j at time t is negative and zero otherwise.

3.2.5. Earnings persistence model

The earnings persistence model by Li and Mohanram, (2014) is structurally similar to the autoregressive. It forecasts earnings based on historical values and allows for different persistence in profits and losses and is referred to as the EP model. The EP model uses the following regression to forecast one-period ahead earnings:

$$EPS_{j;t+k} = \beta_{0jk} + \beta_{1jk} * EPS_{j,t} + \beta_{2jk} * NEGE_{j,t} + \beta_{3jk} * NEGE_{j,t} * EPS_{j,t} + \varepsilon_{j,t+k}$$

where $EPS_{j,t}$ is the earnings per share for firm j at time t and $NEGE_{j,t}$ is an indicator variable that equals zero if the earnings for firm j at time t is negative and zero otherwise. This model also includes a term combining the $EPS_{j,t}$ and $NEGE_{j,t}$ variables ($NEGE_{j,t} * EPS_{j,t}$).

3.3. Forecast performance

3.3.1. Bias and accuracy

We measure the quality of the forecasts from both the models and the analysts primarily by using forecast bias and accuracy. Following previous literature, forecast bias is defined as “the mean difference between realized earnings and forecast earnings, scaled by price” and forecast accuracy is defined as “the mean absolute value of the difference between realized and forecasted earnings, scaled by price” (Harris and Wang, 2019). Accuracy is the absolute forecast error, i.e., the closer to zero the value, the more accurate the forecast. A negative bias value means the forecast is higher than realized earnings, i.e. upwardly biased, while a positive bias value means the forecast is lower than realized earnings, i.e. downwardly biased.

3.3.2. Efficiency test

We evaluate the efficiency of the forecasts from the models and the analysts to examine the incremental information each forecast contains about realized earnings. Forecast efficiency is a measurement of how well the forecast reflects the information that is currently available. An efficient forecast is associated with a forecast error that is unpredictable. If a forecast error is predictable, the forecast will consistently over- or underestimate earnings. This is tested using the following regression creased by Mincer and Zaronwitz, (1969):

$$EPS_{j,t+1} = \beta_0 + \beta_1 * \widehat{EPS}_{j,t} + \varepsilon_{j,t+1}$$

For the Mincer and Zarowitz regression, the slope (β_1) coefficient for the forecast should be close to one for the forecast error to be unpredictable. If the slope coefficient is statistically significantly different from one, the forecast error is predictable. A predictable forecast error is associated with a lower quality forecast. The R-squared statistic from this regression is a measurement of the information contained within each forecast, irrespective of their bias and inefficiency (Harris and Wang, 2019).

3.3.3. Encompassing test

The Mincer and Zarowitz, (1969) regression can also be used to measure the incremental information of competing forecasts, irrespective of their bias and accuracy, and therefore determine if one forecast encompasses another:

$$EPS_{j,t+1} = \beta_0 + \beta_1 * \widehat{EPS}^1{}_{j,t} + \cdots + \beta_N * \widehat{EPS}^N{}_{j,t} + \varepsilon_{j,t+1}$$

If the coefficient of forecast N is zero ($\beta_N = 0$), that model does not contain any information beyond what is contained in other forecasts, and therefore the forecast N is encompassed by the other models. The scale and significance of each coefficient ($\beta_1, \dots, \beta_N$) is an indicator of the relative information contained in the series of forecasts included in the regression. If a forecast's coefficient is not statistically different from zero, the model does not contain any information not contained in the other models. Therefore, it is considered encompassed by the forecasts where the coefficient is statistically significantly different from zero. Similar to the efficiency test, the R-squared statistic is a measurement of the information contained in the combination of forecasts. Given that more than one coefficient is statistically significantly different from zero in one combination, the optimal combination for forecasting earnings can be determined. The optimal combination can be provided by using the estimated coefficients from the encompassing regression as weights for each forecast (Harris and Wang, 2019). If one forecast coefficient is statistically significant, and higher than the other coefficients in a combination, that forecast is considered to dominate the other forecasts (Harris and Wang, 2019).

3.4. Firm characteristics

To test the hypotheses whether the quality of earnings forecasts are dependent on industry membership, firm size, or E/P ratio, the sample is divided into groups based on these characteristics and the models are re-estimated for each group. When testing the difference between the industries, the companies were divided into twelve groups based on Ken French's industry classifications (French, K.R., 2021). To further test whether analyst forecasts outperform model-based forecasts for larger companies and for firms with high E/P firms, the sample was divided into quantiles based on either market capitalization or E/P ratio. Accuracy and bias were then calculated again for each of the different groups (Harris and Wang, 2019).

3.5. Differences to the replicated study

The main difference between our study and the original study by Harris and Wang, (2019) is the use of Swedish data instead of US data, which results in a smaller sample. Apart from this, the databases used to acquire the data are also US based and are less complete for Swedish companies, which further limits the numbers of observations.

To conduct the analysis on a firm characteristics basis, the sample is divided into groups based on each characteristic and the models are re-estimated for each group. Harris and Wang, (2019) divided their sample into deciles. Due to our limited dataset, we instead divided our sample into quantiles for size and E/P ratio, to increase the statistical strengths of the results. Furthermore, we also expand the analysis on firm characteristics by conducting t-tests on the difference in performance between the model-based and analyst forecasts.

As previously mentioned, we also complement the replication by including an additional model, the earnings persistence model, which was not used by Harris and Wang, (2019). This model was added due to both Harris and Wang, (2019) and Li and Mohanram, (2014) concluding that their respective models outperformed the HDZ model. To the best of our knowledge, the PW and EP model have not been directly compared in any existing research.

4. Empirics

In this section we explain our data collection, sample selection, and estimation methodology.

4.1. Data Collection

The models described in section 3.2. use accounting data, dividends, and stock price as independent variables to forecast earnings for one-period ahead forecasts. The accounting and dividend variables are acquired from WRDS Compustat (Global - Daily, Fundamentals Annual file) and the stock price data from FinBas (Stockholm Stock Exchange). From Compustat we also acquired the Standard Industry Classification (SIC) codes for each company that was used to divide the sample based on industry membership, when testing Hypothesis 1. To compare the model-based forecasts with financial analyst forecasts, we need the analyst consensus for the same period which was acquired from the Institutional Brokers Estimate System (I/B/E/S) database.

We acquired the stock price data for the Nasdaq Stockholm Stock Exchange between June 1987 and December 2019 from FinBas. This generates the International Standard Identification Number (ISIN) for all stocks listed on the Nasdaq Stockholm Stock Exchange. The dataset initially includes all the shares, both A and B (sometimes C) shares as well as preferred shares. When using the stock price to estimate earnings forecasts, we use the stock price of the shares with the least amount of voting rights, which is also often the most liquid and therefore better reflect available information (Chordia et al., 2008). The ISIN codes from the FinBas dataset were then used to search for the accounting data on Compustat. Compustat identifies companies using the ISIN code for one security connected to the company, and therefore by using the ISIN codes for all shares we minimize the risk that any companies are missed. This method results in our dataset including companies that are now delisted, either due to being taken private or because of bankruptcy, and therefore also minimizes the survivorship bias (Hou et al., 2012).

In the original article, Harris and Wang, (2019) used the Center for Research in Security Prices (CRSP) database for the stock price, dividends and shares outstanding. However, we instead use FinBas for the stock price, which does not contain the total number of shares outstanding. Therefore, we instead collect this from Compustat by taking the number of shares used to calculate the earnings per share as reported. To acquire the five-year ahead analyst consensus forecasts, Harris and Wang, (2019) used the four-year ahead forecasted earnings combined with the long-term growth rate from I/B/E/S. We followed their method, however, the I/B/E/S database is not as extensive for Swedish companies. Therefore, to increase the number of observations, we also added the five-year ahead analyst consensus forecasts from I/B/E/S which complemented the four-year forecasts with the growth rate.

4.2. Sample Selection

The original study by Harris and Wang, (2019), the period between July 1976 and June 2015. Due to limitations in Compustat's international file, data further back than June 1987 was not available. We added an additional four years until 2019 in our sample to increase the quality of the dataset. Compustat provided an initial dataset of 441 Swedish companies, with a total of 7 463 observations. These observations were then reduced by a combination of missing data and extreme observations which we will cover in detail below. After these adjustments, the sample covers the period between January 1988 and December 2019.

Firstly, there is a lot of missing data in the earlier period of our time interval. Since we are using a per share basis for the models, every observation with the number of shares outstanding missing was removed. This reduced our sample to 5 947 observations (434 firms).

Secondly, when matching the accounting data from Compustat with the stock price data from FinBas, there is some discrepancy between the two datasets. After matching the accounting data with the stock price data, the number of observations was reduced to 4 872 (428 firms).

Thirdly, following Harris and Wang, (2019) we removed observations in the extreme percentiles of earnings, book values, assets, stock prices, and one period ahead earnings forecasts. Harris and Wang, (2019) do not provide any definitions for the extreme percentile, therefore we have chosen to define it as all values below the 1st percentile and all values above the 99th percentile. This leaves us with 4 439 observations (422 firms).

Fourthly, we removed observations with a share price of less than 8 SEK. This is also following Harris and Wang, (2019), who removed all observations with a share price of less than $1 based on Khan and Watts, (2009). As 8 SEK is approximately $1, we excluded all share prices below that line. This leaves us with 3 623 observations (411 firms).

Fifthly, some models require lagged variables (for book value and stock price) and all model's dependent variable is realized earnings per share in the forecast period. For the lagged variables (book value and stock price), the first observation for every firm in the period is therefore missing as the year before that is not included in the dataset. For the forward variable (earnings per share), the last observation of every period is missing as the next year is not included in the dataset. For one-year ahead forecasts, this gives a total of 3 048 observations (375 firms). For two-year ahead forecasts, the total number of observations is 2 750 (351 firms) and for five-year ahead forecasts, the total number of observations is 1 977 (272 firms). These datasets are used to estimate the models for each forecast horizon.

Table 1
Sample development

	Observations	Firms
Initial Dataset	7463	441
Accounting data missing	-1516	434
Stock price data missing	-1075	428
Elimination of extreme observations	-433	422
Share price of less than 8 SEK	-816	411
Lagged variables	-575	375
Sample for estimating one-year ahead forecast	**3048**	**375**
Sample for estimating two-year ahead forecast	**2750**	**351**
Sample for estimating five-year ahead forecast	**1925**	**272**

The table shows how the sample was adjusted and the effect of each adjustment. The final sample for each year ahead forecast is the sample used to estimate the models for the one-, two-, and five-year ahead forecasts.

After estimating the future earnings for the one-, two-, and five-year ahead forecasts, we compared the model-based forecasts with each other and to the I/B/E/S consensus using the bias and accuracy measurements. However, I/B/E/S reports an adjusted earnings per share which cannot be compared directly with the reported earnings per share and instead must be compared with the I/B/E/S reported actual values. Therefore, when calculating the bias and accuracy, we limit the dataset to observations where the I/B/E/S actual data is available. This further reduces our sample to 2 083, 1 811, and 130 observations respectively for the one-, two, and five-year forecasts, when determining forecast performance.

4.3. Estimation methodology

The PW, HDZ, AR(1), and EP models are estimated using a pooled cross-section ordinary-least-square (OLS) regression with a rolling window of ten years. The estimated coefficients are then used with the independent variables to forecast out-of-sample earnings for the next period *k* (where *k* = 1, 2, 5). The dataset used to estimate the models is between January 1988 and December 2019, and due to the rolling window of ten years, the first forecast occurs in 1997.

5. Results

In this section, we present the results of estimating the models and calculating the forecast bias and accuracy. Afterwards, we present the efficiency and encompassing tests and the result of re-estimating the models based on firm characteristics.

Table 2
Descriptive statistics

Variable	N	Mean	Std. Dev.	Min	Max
EPS	3048	5.6000	8.2160	-19.9900	54.9860
$REPS_1$	3048	5.4150	8.1590	-19.8430	54.9860
$AFEPS_1$	2107	5.1290	5.3340	-26.4600	70.3700
$IBES\_ACTUAL_1$	2107	4.4760	11.5020	-173.5460	403.7000
$PRICE$	3048	61.6500	59.4620	8.0500	597.7370
$BOOK$	3048	45.4990	45.2480	0.0780	541.1530
$DIVIDEND$	3048	2.6030	4.0320	0.0000	125.4870
$ASSETS$	3048	133.9640	237.3250	0.3520	2364.3350
$ACCRUALS$	3048	-0.5700	9.8100	-90.7850	72.4030

The table reports the descriptive statistics for the current earnings per share *(EPS)*, realized earnings per share *($REPS_1$)*, analyst forecasts for one-year ahead *($AFEPS_1$)*, I/B/E/S realized earnings *($IBES_1$)*, stock price *($PRICE$)*, book value per share *($BOOK$)*, dividend per share *($DIVIDEND$)*, assets per share *($ASSETS$)*, and operating accruals per share *($ACCRUALS$)*.

5.1. Descriptive statistics

Table 2 shows descriptive statistics for key variables. The table shows higher earnings for the I/B/E/S consensus forecast compared to the realized earnings (analyst one-year ahead forecasts *afeps1* is compared to the I/B/E/S actual earnings *ibes1*). From the t-test in appendix 6, the I/B/E/S forecast is statistically significantly upwardly biased, which is in line with previous research (Harris and Wang, 2019).

5.2. Correlation matrix

Appendix 7 shows the Pearson correlations between the dependent and key independent variables as well as their significance level. Analyst forecasts and realized future earnings are statistically significantly positively correlated with stock price, book value, dividends, and assets. The correlation between analyst forecasts and operating accruals is not statistically significant.

5.3. Model estimations

Appendix 3, 4, and 5 shows the average estimated coefficients for the PW, HDZ, and EP models for the period. The number of observations per forecast horizon are lower than what is presented in table 1 due to the use of a rolling window of ten years, which results in the first forecast occurring in 1997. The R-squared statistic decreases over the forecast period for all models, and the coefficient for current earnings is also positive and statistically significantly different from zero for the one- and two-year ahead forecasts, at the 1 % level. For the HDZ model, the current earnings coefficient is positive and statistically significant at the 1 % level for the one-, two-, and five-year ahead forecasts. For the EP model, the current earnings coefficient is positive and statistically significant at the 1 % level for all time horizons.

5.4. Forecast performance

In this section, we will cover the results of the forecast bias and accuracy tests as well as the efficiency and encompassing tests.

5.4.1. Bias and accuracy

Table 3 shows the mean bias and accuracy for the one-, two-, and five-year ahead forecasts as well as the significance level for testing the null hypothesis that the value is equal to zero. Appendix 1 and 2 shows the pairwise mean difference for each pair of forecasts for bias and accuracy and the significance level for testing the null hypothesis that the value for each pair is equal. For the one- and two-year ahead forecasts, the I/B/E/S consensus forecast is statistically significantly upwardly biased at the 5 % and 1 % levels respectively. This is in line with previous literature that analyst consensus is upwardly biased (Harris and Wang, 2019). The mean bias for the PW model is not statistically significantly different from zero for any time horizon. For the two-year ahead forecast, the PW model is statistically significantly less biased than all other model-based forecasts, at the 1 % level. For the five-year ahead forecast, the PW model is statistically significantly less biased than all forecasts except for the RW model, at least at the 5 % level. For the one-year ahead forecast however, the EP model is statistically significantly less biased than all other forecasts, at least at the 5 % level.

Table 3

Forecast bias and accuracy

Panel A: One-year ahead forecast: bias and accuracy						
	AR(1)	HDZ	PW	RW	EP	IBES
Bias	-0.0021	0.0127***	-0.0047	0.0056	0.0001	-0.0098**
Std. Dev.	0.2349	0.2230	0.2249	0.2592	0.2316	0.2273
Accuracy	0.1150***	0.1053***	0.1064***	0.1091***	0.1125***	0.0500***
Std. Dev.	0.2048	0.1970	0.1981	0.2352	0.2025	0.2219
N	2083	2083	2083	2083	2083	2083
Panel B: Two-year ahead forecasts: bias and accuracy						
	AR(1)	HDZ	PW	RW	EP	IBES
Bias	0.0099*	0.0254***	0.0028	0.0136*	0.0115*	-0.0185***
Std. Dev.	0.2608	0.2490	0.2399	0.2966	0.2579	0.2568
Accuracy	0.1326***	0.1224***	0.1233***	0.1369***	0.1302***	0.0712***
Std. Dev.	0.2248	0.2183	0.2057	0.2634	0.2229	0.2474
N	1811	1811	1811	1811	1811	1811
Panel C: Five-year ahead forecasts: bias and accuracy						
	AR(1)	HDZ	PW	RW	EP	IBES
Bias	0.0840**	0.0556*	0.0363	0.0590	0.0845**	-0.0312
Std. Dev.	0.3815	0.3367	0.3586	0.4653	0.3824	0.3143
Accuracy	0.2164***	0.1848***	0.2167***	0.2137***	0.2533***	0.1569***
Std. Dev.	0.3248	0.2865	0.2875	0.3277	0.3941	0.2737
N	130	130	130	130	130	130

The table shows the average and standard deviation for bias and accuracy for the AR(1), HDZ, PW, RW, EP and IBES forecasts for the one-, two-, and five-year ahead forecasts. A negative value means the forecast is upwardly biased, while a positive value means it is downwardly biased. The accuracy is the forecast error, and therefore the lower the value, the more accurate the forecast. The significance level is the result of testing the null hypotheses that each value is equal to zero. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

All forecast accuracy is statistically significantly different from zero, with the I/B/E/S consensus having the lowest mean forecast error. This forecast error was statistically significantly lower than all model-based forecasts at the 1 % level for the one- and two-year ahead forecasts. For the five-year ahead forecasts, the I/B/E/S consensus was statistically significantly more accurate than all model-based forecasts except for the HDZ model, at least at the 10 % level. Of the model-based forecasts, the HDZ model had the lowest mean forecast error for all periods. However, this was not statistically significantly lower than the PW and RW models for the one-year ahead forecasts and the PW model for the two-year ahead forecasts. For the other models, the HDZ model mean forecast error was statistically significantly lower at the 1 % level for the one- and two-year ahead forecasts. For the five-year ahead forecast, the HDZ model was statistically significantly more accurate for all model-based forecasts, and not statistically significantly less accurate than the I/B/E/S consensus.

Table 4

Efficiency test (Mincer-Zarnowitz regression)

Panel A: One-year ahead forecasts				
	Constant	Slope	$Adj-R^2$	N
AR(1)	0.0057	1.0113	0.3607	2083
HDZ	0.2985	1.1009***	0.3977	2083
PW	0.2978	0.9382**	0.3881	2083
RW	2.1001***	0.6175***	0.3682	2083
EP	0.0719	1.0003	0.3688	2083
IBES	0.1719	0.8382***	0.1495	2083
Panel B: Two-year ahead forecasts				
	Constant	Slope	$Adj-R^2$	N
AR(1)	0.0311	1.0542***	0.2534	1 811
HDZ	0.3305	1.1253***	0.2994	1 811
PW	0.5044	0.9527***	0.3009	1 811
RW	2.7968***	0.5262***	0.2664	1 811
EP	0.1499	1.0338***	0.2616	1 811
IBES	2.4896***	0.3618***	0.0350	1 811
Panel C: Five-year ahead forecasts				
	Constant	Slope	$Adj-R^2$	N
AR(1)	3.6236***	0.7082***	0.0575	130
HDZ	1.5534	1.0491***	0.1854	130
PW	3.7244	0.5844***	0.0796	130
RW	5.5045***	0.2467***	0.0599	130
EP	3.8655***	0.6675***	0.0558	130
IBES	3.5207***	0.2764***	0.1253	130

The table shows the result of the efficiency test, estimating the Mincer Zaronwitz regression for the AR(1), HDZ, PW, RW, EP and IBES forecast. The closer to one the slope coefficient is, the more efficient the forecast. The significance level for the slope coefficient is the result of testing the null hypotheses that each value is equal to one. *** p<0.01, ** p<0.05, * p<0.1.

5.4.2. Efficiency test

Table 4 shows the result of the efficiency test. For all the one-year ahead forecasts, the slope coefficient is close to one, except for the RW model. Therefore, the model-based and the I/B/E/S consensus forecasts have similar efficiency, and a forecast error that is unpredictable. From the R-squared statistic, the HDZ model is the most informative in the one-year ahead forecast, followed by the PW model. As the time horizon increases, the information contained within each model-based forecast falls, indicated by the R-squared statistic. For the two-year ahead forecast, the PW model is the most informative, and in the five-year ahead forecast, the HDZ model is the most informative. For the five-year ahead forecasts, the information contained in the I/B/E/S consensus forecast increased.

5.4.3. Encompassing test

Table 5, 6, and 7 shows the result of the encompassing tests for the one-, two-, and five-year ahead forecasts. The results from the efficiency test shows that the RW model was less informative compared to the other models. Therefore, we have omitted the RW forecasts from the encompassing test, following Harris and Wang, (2019).

For the one-year ahead forecasts (Table 5), the HDZ, PW, and I/B/E/S consensus forecasts had a coefficient that was statistically significantly different from zero for all combinations where they were included, at the 1 % level. When comparing all these three forecasts (Model 15), the HDZ forecast dominates, and when comparing the HDZ to only the PW (Model 5) or to the I/B/E/S consensus (Model 7), the HDZ forecast dominates in both. When comparing the PW and I/B/E/S, the PW forecast dominates. Model 15 also gives the second highest R-squared statistic. The highest R-squared is obtained when combining all the forecasts. For this combination, every coefficient is statistically significantly different from zero, at the 1 % level (Model 22).

Table 5

One-year encompassing test (N = 2083)

	AR(1)	HDZ	PW	EP	IBES	$Adj - R^2$
Model 1	0.126	0.979***				0.398
Model 2	0.362***		0.648***			0.397
Model 3	0.117			0.888***		0.369
Model 4	0.855***				0.291***	0.391
Model 5		0.644***	0.437***			0.413
Model 6		0.849***		0.256***		0.401
Model 7		0.954***			0.263***	0.423
Model 8			0.604***	0.407***		0.400
Model 9			0.808***		0.269***	0.414
Model 10				0.850***	0.288***	0.399
Model 11	-0.081	0.702***	0.456***			0.413
Model 12	-0.585***	0.947***		0.733***		0.403
Model 13	0.059	0.898***			0.261***	0.423
Model 14		0.605***	0.420***	0.057		0.413
Model 15		0.573***	0.375***		0.241***	0.434
Model 16			0.532***	0.345***	0.253***	0.422
Model 17	-0.574***	0.702***	0.418***	0.527***		0.415
Model 18	-0.119	0.659***	0.403***		0.243***	0.434
Model 19	-0.632***	0.867***		0.714***	0.259***	0.428
Model 20	-0.209		0.542***	0.537***	0.254***	0.423
Model 21		0.555***	0.367***	0.028	0.241***	0.433
Model 22	-0.620***	0.659***	0.365***	0.535***	0.244***	0.437

The table reports the results for estimating the encompassing regression for the AR(1), HDZ, PW, EP, and IBES forecasts for the one-year ahead forecasts. Each model to the left shows each combination of forecasts, and the significance level is the result of testing the null hypotheses that the slope of each forecast is equal to zero.*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

The results are similar for the two-year ahead forecast (Table 6). However, one key difference exists, the PW forecast coefficient is not statistically significantly different from zero when compared with the EP and I/B/E/S consensus (Model 20), at the 1 % level. The HDZ and I/B/E/S forecasts coefficients, however, are still statistically significantly different from zero for all models, also at the 1 % level. Looking at model 5, 7, and 15, the result is similar to the one-year ahead forecasts, with the HDZ forecast dominating. The highest R-squared is, just as for the one-year forecast, achieved when combining all forecast models. All coefficients are also still statistically significantly different from zero in this case, at the 1 % level (Model 22), except for the EP forecast which is at the 5 % level.

Table 6

Two-year encompassing test (N = 1811)

	AR(1)	HDZ	PW	EP	IBES	$Adj-R^2$
Model 1	0.229***	0.929***				0.302
Model 2	0.323***		0.730***			0.308
Model 3	0.041			0.995***		0.261
Model 4	0.994***				0.114***	0.261
Model 5		0.607***	0.527***			0.327
Model 6		0.859***		0.302***		0.305
Model 7		1.078***			0.088***	0.304
Model 8			0.699***	0.354***		0.310
Model 9			0.906***		0.128***	0.311
Model 10				0.976***	0.115***	0.269
Model 11	-0.058	0.640***	0.543***			0.327
Model 12	0.498**	0.898***		0.741***		0.306
Model 13	0.212**	0.899***			0.084***	0.306
Model 14		0.596***	0.520***	0.019		0.327
Model 15		0.556***	0.529***		0.090***	0.332
Model 16			0.695***	0.302***	0.111***	0.317
Model 17	-0.083	0.603***	0.552***		0.092***	0.332
Model 18	-0.530**	0.867***		0.755***	0.086***	0.310
Model 19	-0.221	0.699***		0.507**	0.113***	0.317
Model 20	-0.221		0.699	0.507**	0.113***	0.317
Model 21		0.556***	0.529***	0.000	0.090***	0.332
Model 22	-0.519**	0.596***	0.528***	0.458**	0.093***	0.333

The table reports the results for estimating the encompassing regression for the AR(1), HDZ, PW, EP, and IBES forecasts for the two-year ahead forecasts. Each model to the left shows each combination of forecasts, and the significance level is the result of testing the null hypotheses that the slope of each forecast is equal to zero.*** p<0.01, ** p<0.05, * p<0.1.

Lastly, for the five-year ahead forecasts (Table 7), the HDZ forecast is the only coefficient that is statistically significantly different from zero in all combinations at the 1 % level. The other model-based forecast coefficients are only statistically significantly different from zero when comparing them to the I/B/E/S consensus forecast. However, the I/B/E/S consensus forecasts' coefficient is not statistically significantly different from zero except in Model 18, Model 21 and Model 22, at the 10 % level. There is no combination where all coefficients are statistically significantly different from zero.

Table 7

Five-year encompassing test (N = 130)

	AR(1)	HDZ	PW	EP	IBES	$Adj-R^2$
Model 1	-0.386	1.418***				0.210
Model 2	0.277		0.498**			0.082
Model 3	0.594			0.180	0.053	0.053
Model 4	0.693***				0.146	0.064
Model 5		1.334***	-0.164			0.204
Model 6		1.454***		-0.418		0.212
Model 7		1.147***			0.131	0.211
Model 8			0.520**	0.217		0.080
Model 9			0.584***		0.141	0.089
Model 10				0.636***	0.152	0.062
Model 11	-0.352	1.458***	-0.067			0.204
Model 12	0.065	1.451***		-0.474		0.206
Model 13	-0.487	1.430***			0.160	0.218
Model 14		1.481***	-0.046	-0.395		0.206
Model 15		1.338***	-0.217		0.145	0.210
Model 16			0.497**	0.171	0.135	0.083
Model 17	0.074	1.478***	-0.048	-0.458		0.200
Model 18	-0.435	1.492***	-0.104		0.164*	0.213
Model 19	-0.034	1.463***		-0.476	0.161	0.215
Model 20	0.320		0.486*	-0.103	0.131	0.077
Model 21		1.511***	-0.085	-0.463	0.163*	0.215
Model 22	-0.019	1.512***	-0.085	-0.447	0.164*	0.209

The table reports the results for estimating the encompassing regression for the AR(1), HDZ, PW, EP, and IBES forecasts for the five-year ahead forecasts. Each model to the left shows each combination of forecasts, and the significance level is the result of testing the null hypotheses that the slope of each forecast is equal to zero.*** p<0.01, ** p<0.05, * p<0.1.

5.5. Firm characteristics

We have omitted the five-year ahead forecasts for the firm characteristics. The reason being that the number of observations are too few to estimate the models for the different groups. Instead we only show the result for re-estimating each model for the one- and two-year ahead forecasts per sector, size, and E/P ratio. There is also no significant difference in bias for the forecasts with different firm characteristics and therefore we have omitted these results.

5.5.1. Industry membership

Compared to the original article (Harris and Wang, 2019) we have a smaller sample. This resulted in certain sectors containing too few observations to re-estimate the models and therefore all sectors with observations below 150 were reclassified as sector 12 (Others). Table

8 shows the accuracy for the one- and two-year ahead forecasts. The remaining sectors are Manufacturing (3), Business Equipment (6), Shops (9), Healthcare (10), Finance (11) and Others (12). The I/B/E/S consensus mean forecast error was lower for all sectors compared to the model-based forecasts. For the financial services sector (Sector 11), the mean forecast error is significantly higher for the model-based forecasts compared to the other sectors.

Table 8

Forecast accuracy by sector

Panel A: One-year ahead forecasts, N = 2083						
	3	6	9	10	11	12
AR(1)	0.121***	0.071***	0.095***	0.081***	0.161***	0.138***
HDZ	0.118***	0.075***	0.110***	0.076***	0.167***	0.136***
PW	0.117***	0.073***	0.101***	0.083***	0.141***	0.135***
RW	0.106***	0.072***	0.074***	0.062***	0.162***	0.129***
EP	0.117***	0.072***	0.092***	0.081***	0.189***	0.137***
IBES	0.033***	0.047***	0.083*	0.035***	0.063***	0.052***
N	408	327	184	189	285	690
Panel B: Two-year ahead forecasts, N = 1811						
	3	6	9	10	11	12
AR(1)	0.141***	0.076***	0.102***	0.084***	0.179***	0.161***
HDZ	0.141***	0.077***	0.116***	0.083***	0.165***	0.154***
PW	0.138***	0.078***	0.109***	0.084***	0.172***	0.154***
RW	0.142***	0.084***	0.099***	0.077***	0.194***	0.161***
EP	0.141***	0.075***	0.101***	0.083***	0.217***	0.161***
IBES	0.054***	0.065***	0.106*	0.052***	0.084***	0.075***
N	363	278	162	156	247	605

The table reports the average accuracy for the AR(1), HDZ, PW, RW, EP, and IBES forecast for the one- and two-year ahead forecasts based on industry classification. The industries are 3 Manufacturing, 6 Business Equipment, 9 Shops, 11 Finance, 12 Others. Industries that contained fewer than 150 observations were re-classified as "Others". The significance level is the result of testing the null hypotheses that the accuracy of each forecast is equal to zero. *** p<0.01, ** p<0.05, * p<0.1.

To test whether analyst consensus forecasts outperform model-based forecasts for sectors with discrete financial statements (Hypothesis 1), the model-based forecast errors were pooled together. The pooled forecast error was then compared to the analyst consensus using a paired t-test, testing the null hypothesis that the difference in accuracy between the two forecasts was zero (Appendix 9). The t-test shows that the forecast error was statistically significantly lower, at the 1 % level, for the analyst forecasts compared to the pooled-model-based forecasts. This was the case for both the one- and two-year ahead forecasts for the financial services sector.

5.5.2. Size

The mean forecast error is lower for analyst forecasts compared to model-based forecasts for companies with larger market capitalization, shown in table 9. For both types of forecasts, the mean forecast error decreases for larger companies. To test whether the analyst consensus outperforms the model-based forecasts for larger companies (Hypotheses 2), the mean forecast error in the 5th quantile was pooled. The pooled forecast error and the analyst consensus was then compared using a paired t-test testing the null hypothesis that the difference in accuracy between the two forecasts was zero (Appendix 10). The result showed that the forecast error was statistically significantly lower for the analyst consensus compared to the pooled-model-based forecasts for both the one- and two-year ahead forecasts for companies in the 5th quantile.

Table 9

Forecast accuracy by size

Panel A: One-year ahead forecasts, N = 2083					
	1	2	3	4	5
AR(1)	0.220***	0.161***	0.150***	0.088***	0.056***
HDZ	0.214***	0.157***	0.135***	0.084***	0.061***
PW	0.210***	0.148***	0.135***	0.088***	0.059***
RW	0.216***	0.145***	0.123***	0.078***	0.051***
EP	0.218***	0.156***	0.151***	0.086***	0.054***
IBES	0.073***	0.062***	0.066***	0.042***	0.025***
N	249	377	434	490	533
Panel B: Two-year ahead forecasts, N = 1811					
	1	2	3	4	5
AR(1)	0.258***	0.186***	0.154***	0.114***	0.069***
HDZ	0.252***	0.205***	0.151***	0.105***	0.068***
PW	0.246***	0.186***	0.147***	0.105***	0.071***
RW	0.270***	0.184***	0.142***	0.115***	0.067***
EP	0.254***	0.183***	0.153***	0.112***	0.068***
IBES	0.108***	0.089***	0.095***	0.055***	0.039***
N	199	322	379	430	481

The table reports the average accuracy for the AR(1), HDZ, PW, RW, EP, and IBES forecast for the one- and two-year ahead forecasts for each size quantile. The significance level is the result of testing the null hypotheses that the accuracy of each forecast is equal to zero. *** p<0.01, ** p<0.05, * p<0.1.

5.5.3. E/P ratio

Table 10 shows that as the E/P ratio increases, the mean forecast error decreases, except for the similar to the results for firm size, except for the 5th quantile. To test whether the analyst consensus outperforms the model-based forecasts for the larger companies (Hypotheses 3), the mean forecast error in the 5th quantile was pooled. Just as for firm size, the pooled forecast error and the analyst consensus was then compared using a paired t-test testing the null hypothesis that the difference in accuracy between the two forecasts was zero (Appendix 11). The results showed that the forecast error was statistically significantly lower for the analyst consensus compared to the pooled-model-based forecasts for both the one- and two-year ahead forecasts for companies in the 5th quantile.

Table 10

Forecast accuracy by E/P ratio

Panel A: One-year ahead forecasts, N = 2083					
	1	2	3	4	5
AR(1)	0.123***	0.043***	0.043***	0.074***	0.289***
HDZ	0.115***	0.048***	0.050***	0.083***	0.292***
PW	0.125***	0.044***	0.046***	0.077***	0.289***
RW	0.153***	0.042***	0.039***	0.069***	0.259***
EP	0.124***	0.043***	0.043***	0.074***	0.289***
IBES	0.089***	0.035***	0.028***	0.052**	0.055***
N	349	436	439	441	418
Panel B: Two-year ahead forecasts, N = 1811					
	1	2	3	4	5
AR(1)	0.128***	0.055***	0.052***	0.095***	0.325***
HDZ	0.127***	0.062***	0.061***	0.097***	0.328***
PW	0.123***	0.057***	0.054***	0.094***	0.324***
RW	0.178***	0.053***	0.049***	0.090***	0.326***
EP	0.128***	0.055***	0.052***	0.095***	0.325***
IBES	0.095***	0.050***	0.043***	0.078***	0.095***
N	295	370	378	393	375

The table reports the average accuracy for the AR(1), HDZ, PW, RW, EP, and IBES forecast for the one- and two-year ahead forecasts for each E/P quantile. The significance level is the result of testing the null hypotheses that the accuracy of each forecast is equal to zero. *** p<0.01, ** p<0.05, * p<0.1.

6. Analysis

In this section we analyze our results and discuss their implications. We begin by addressing the forecast performance and analyzing the results to draw conclusions about the explanatory power of stock price and other financial metrics. Afterwards, we continue with an analysis of our hypotheses and how forecast performance is affected by firm characteristics. Lastly, we discuss the validity, reliability, and comparability of our research method.

6.1. Forecast performance

6.1.1. Bias and accuracy

As seen in table 3, the I/B/E/S consensus is both statistically significantly upwardly biased, and more accurate than the model-based forecasts for the one- and two-year ahead forecasts. This is in line with previous literature that finds analyst consensus being more accurate when compared to model-based forecasts in the short-term (Li and Mohanram, 2014; Harris and Wang, 2019; Azevedo et al., 2020). For the five-year ahead forecast, the I/B/E/S was statistically significantly more accurate than all model-based forecasts except the HDZ model, where the results were inconclusive. That the analyst consensus is more accurate than the PW model for the five-year ahead forecasts contradicts the findings of Harris and Wang, (2019). However, the results were only statistically significant at the 10 % level, which means that it is suggestive, and not conclusive, that the I/B/E/S consensus is more accurate than the PW model. The limited number of comparable observations also makes our result difficult to compare to Harris and Wang, (2019). In terms of bias, the results indicate that the PW model is less biased than the HDZ and EP models as well as the I/B/E/S consensus for the two- and five-year ahead forecasts. When comparing accuracy across the model-based forecast, we find no indication that the PW model yields more accurate results compared to pure-accounting based models.

6.1.2. Efficiency

As the time horizon increases, the slope of each forecast moves further from one. For the two-year ahead forecasts the results show that all coefficients are statistically significantly different from one. This indicates that as the forecast horizon increases, the forecast error becomes more predictable, which decreases the quality of the forecasts. Generally, the information contained in each forecast (R-squared statistic) decreases as the time horizon increases. The one exception is the I/B/E/S consensus forecast where the amount of information contained in the forecast has increased, while the forecast error has become more predictable at the five-year ahead forecast horizon. This contradicts the findings of Harris and Wang, (2019) which found that analyst forecasts were more negatively affected by an increased time horizon, compared to model-based forecasts. One explanation for this contradictory result could be our limited dataset, especially for the five-year ahead forecasts, which is further discussed in section 6.3.2.

6.1.3. Encompassing test

When combining the I/B/E/S consensus with any model-based forecast, the slope of the analysts forecast decreases compared to the efficiency test. This indicates that the importance of analyst forecasts decreases when combined with model-based forecasts. Therefore, the results imply that the model-based forecasts incorporate much of the information contained in the analyst forecasts. For both the one- and two-year ahead forecasts, the encompassing test also showed that when combining all forecasts, all coefficients are statistically significantly different from zero at the 1 % level. This further indicates that there is incremental information in all forecasts in the short-term and suggests that the optimal forecast method of future earnings is a combination of model-based and analyst forecasts. These findings are similar to those of Harris and Wang, (2019), who also reached the conclusion that a combination of model-based and analyst forecasts is the most optimal method. This finding supports the reasoning of Azevedo et al., (2020) who combines analyst forecasts with a time-series model to forecast earnings. For the five-year ahead forecast, there is no combination where all coefficients are statistically significantly different from zero, even at the 10 % level. Therefore, this encompassing test does not provide any indication of what combination of models is the most optimal for this time horizon.

For the one-year ahead forecast, the combination including all forecasts (Model 22) were the most informative, with all coefficients statistically significantly different from zero, at the 1 % level. Therefore, the optimal combination of forecasts gives weights to the AR(1), HDZ, PW, EP, and I/B/E/S consensus forecasts of approximately 62.0 %, 65.9 %, 36.5 %, 53.5 %, and 24.4 %, respectively. For the two-year ahead forecast, the combination with the highest information content with all coefficients being statistically significantly different from zero at the 1 % level, was the HDZ, PW, and I/B/E/S consensus (Model 15). This combination gives weights of approximately 55.6 %, 52.9 %, and 9.0 %, respectively.

6.1.4. Firm characteristics

The analyst consensus forecasts from I/B/E/S statistically significantly outperformed the model-based forecasts in terms of accuracy for all forecasts in the one- and two-year horizons, at the 1 % level. However, the firm characteristics also had a significant effect on the quality of the forecasts, both for the I/B/E/S consensus and the model-based forecasts. Since there was in general no significant difference in bias between the model-based and I/B/E/S consensus forecasts depending on firm characteristics, determining whether one forecast outperforms another is solely based on forecast accuracy.

When comparing the mean forecast error for the pooled-model-based forecasts and the I/B/E/S consensus in the financial services sector, the mean forecast error was statistically significantly lower for the I/B/E/S forecast, at the 1 % level (Appendix 9). This supports Hypothesis 1, that analyst consensus outperforms model-based forecasts for firms with discrete financial statements. Similar results were found when comparing the model-based forecasts for the 5th quantile for size and E/P ratio. The I/B/E/S consensus mean forecast error was statistically significantly lower compared to the model-based forecasts, at the 1 % level (Appendix 10 and 11). This supports Hypotheses 2 and 3, that analyst consensus outperforms model-based forecasts for larger firms and for firms with a high E/P ratio. The findings are in line with Harris and Wang, (2019), who also found support for the same hypotheses when testing them on US firms.

6.1.5. Earnings persistence model

When comparing the HDZ, PW, and EP models, the HDZ model has the lowest mean forecast error. This is statistically significantly lower than the EP model for all forecast horizons, at the 1 % level, and statistically significantly lower than the PW model for the five-year ahead forecast, at the 1 % level. The PW model is statistically significantly more accurate than the EP model. This indicates that both the HDZ and PW models outperform the EP model in terms of forecast accuracy. When comparing bias, the EP model is statistically significantly less biased than the HDZ and PW models for the one-year ahead forecast, and compared to the HDZ model, also for the two-year ahead forecast. Except for the one-year ahead forecast, the PW model is statistically significantly less biased than the EP model at the 1 % level, and always statistically significantly less biased than the HDZ model, at the 1 % or 5 % level. This indicates that the EP model outperforms the HDZ model, and that the PW model outperforms the EP model, in terms of forecast bias.

From the efficiency test, the R-squared statistic is the highest for the HDZ model, followed by the PW model and then the EP model. However, the slope of the forecast error is closer to one for the EP model, and for the one-year ahead forecast not statistically significantly different from one, meaning that the EP model had a less predictable forecast error. When looking at the encompassing test for the one- and two-year ahead forecasts, the PW and HDZ forecasts dominate the EP forecast when compared separately (Model 6 and Model 8). When comparing the HDZ, PW, and EP together (Model 14), the EP forecast coefficient is not statistically significantly different from zero, meaning the EP model is encompassed by the HDZ and PW models. This further indicates that the HDZ and PW model outperform the EP model. For the five-year ahead forecast, the HDZ encompasses both the PW and EP models, as both their coefficients are not statistically significantly different from zero while the HDZ is. Therefore, the results provide suggestive evidence that the HDZ and PW model outperforms the EP model which supports Hypothesis 4.

6.2. Explanatory power of financial metrics

6.2.1. Accounting data

In appendix 2, 3, and 4 we present the average coefficients for the PW, HDZ, and EP models and their respective statistical significance for testing the null hypothesis that the coefficient is zero. The values of the coefficients indicate support for the predictive power of accounting metrics as there are numerous statistically significant coefficients, especially for the one- and two-year ahead forecasts. All models show a positive statistically significant coefficient for current earnings at the 1 % level (the HDZ and PW model for the one- and two-year ahead forecasts and the EP model for all time horizons). This indicates that current earnings have predictive power for future earnings, which is in line with previous literature (Harris and Wang, 2019; Hou et al., 2012). For operating accruals, the coefficients are negative and statistically significantly different from zero for both the HDZ and PW models, at the 1 % and 5 % level, for the one- and two-year ahead forecast. This is also in line with Harris and Wang, (2019) and indicates that earnings are persistent for the one- and two-year ahead forecasts. Similar to current earnings, the coefficient for assets is positive and statistically significantly different from zero, at the 1 % level for all forecast horizons in the HDZ model. This indicates that they have a positive correlation with earnings, for all time horizons. Lastly, the book value variable in the PW model was also positive and statistically significant for all forecast horizons (except four-years), at least at the 5 % level, which indicates that they have a positive correlation with earnings, similar to assets.

6.2.2. Dividend data

Out of the models tested, the HDZ model was the only one to incorporate dividends. This model had the lowest mean forecast error, however, when compared to the PW and RW models, the difference was not statistically significant. Therefore, the performance of the HDZ model cannot be taken as an indicator that dividend has predictive power for future earnings. In appendix 5, however, we see that the coefficients for both dividend related variables are positive and statistically significant, which indicates that dividends have a strong predictive power for future earnings. Furthermore, from appendix 7 we also see that dividends are positively and statistically significantly correlated with realized earnings. This is in line with the findings of Zhou and Ruland, (2006) and Arnott and Asness, (2003).

6.2.3. Stock price data

The PW model had a mean bias that was not statistically significantly different from zero for all forecast horizons. For the two- and five-year ahead forecasts, the PW model was also statistically significantly less biased than the other models, except for the RW model for the five-year ahead forecasts. This indicates that stock prices can be used to generate less biased forecasts. In terms of accuracy, the PW model had a higher mean forecast error compared to the HDZ model, however, this was also not statistically significant. Therefore, we find no indication that stock price can be used to forecast more accurate earnings, when compared to a pure-accounting based model. From the efficiency test, the information contained in the PW model is the second highest and the highest for the one- and two-year ahead respectively. The slope for the PW model was also close to one, indicating that the forecast error is unpredictable.

The encompassing test indicates that the PW model contains some incremental information when compared to certain models. For the one- and two-year ahead forecasts, the PW model's forecast coefficients are always statistically significantly different from zero at the 1 % level for all combinations, except one (Model 20) in the two-year ahead forecast. Since the main difference between the PW model and the HDZ models is that the PW model incorporates stock price data, this suggests that there is some incremental information in stock price data. This result is in line with Harris and Wang, (2019), which also found that stock price contains incremental information in their encompassing test. For the five-year ahead encompassing test, however, the PW model's forecast coefficient is not statistically significantly different from zero in any combination involving the HDZ model. The PW model was therefore encompassed by the HDZ model. This indicates that for the five-year ahead forecasts, the PW and HDZ model contributes only with the accounting variables, and that the HDZ model's combination of variables contains more incremental information. In the long-term, it is therefore likely that the PW model's strong performance in the encompassing test is due to its accounting variables and not due to the incorporation of stock price data. This is further supported by the efficiency test where the PW model's R-squared statistic decreases significantly for the five-year ahead forecasts. Harris and Wang, (2019) finds similar results, indicating that current and lagging stock prices do not have a statistically significant relationship with future earnings long-term.

6.3. Analysis of research method

In this section we present an analysis of our research method. We begin with the validity of our study, which affects our ability to draw conclusions. This is followed by a discussion regarding how reliable our study is and then concluded by a discussion regarding the comparability with existing literature, such as Harris and Wang, (2019), Hou et al., (2012), and Li and Mohanram, (2014).

6.3.1. Validity

When conducting the analysis, we use a cross-sectional OLS regression with a ten-year rolling widow to estimate the models, following our benchmark study (Harris and Wang, 2019). However, there is a lot of missing data for the earlier periods, primarily before 2004. Furthermore, the I/B/E/S database also has a substantial amount of missing data for Swedish companies in general, especially for the five-year ahead forecasts. This resulted in an issue with analysing the forecasts performance for the five-year time horizon. According to previous literature (Harris and Wang, 2019), model-based forecasts perform better in the long-term, and therefore this issue decreases the validity of our study.

When testing our hypotheses, our results indicate that analyst forecasts outperform model-based forecasts for the one- and two-year ahead forecasts, for companies with certain firm characteristics. However, previous literature suggests that model-based forecasts are more accurate than analysts in the long-term (Harris and Wang, 2019). We were not able to test this since our five-year ahead sample did not contain enough observations. This drawback decreases the validity of our study regarding how firm characteristics can affect forecast performance. Another issue related to the validity of our study is the use of Ken French's 12 industry divisions. This classification system was chosen to make our results more comparable with those of Harris and Wang, (2019). However, this is arguably not optimal for Swedish companies as six of the sectors contained too few observations. This prohibited us from re-estimating the models for these sectors, and they were instead reclassified as "Others".

6.3.2. Reliability

After analysing our research method, we have reached the conclusion that our thesis should be viewed as relatively reliable. The reliability of our method is strengthened by our extensive literature review including numerous well-known references. We also used the same established data sources as Harris and Wang, (2019) with the exception of FinBas instead of CRSP for the stock price data. Using FinBas to acquire the ISIN codes of all companies listed on Nasdaq Stockholm between 1987 and 2019 also minimizes the risk of survivorship bias, as delisted companies are still included in our dataset. One of the major drawbacks of our process is the limited dataset which is mainly caused by the lack of I/B/E/S data for Swedish companies. Another issue related to the dataset arises when matching the accounting and stock price datasets. For the FinBas dataset, the ISIN codes change every time a company does a stock split and FinBas keeps these historical variations. Compustat on the other hand uses the latest code for one security connected to the company. The codes in the FinBas data were then manually changed to be constant so both datasets had one common variable, and therefore could be matched. This introduces the risk of human error, which decreases the reliability of the study.

6.3.3. Comparability

Since this is a replication study, we focus on the largest deviations from our benchmark study (Harris and Wang, 2019). The main deviation is the limited dataset, especially for the five-year ahead forecasts, and therefore comparing the results to those of Harris and Wang, (2019) becomes somewhat problematic. Our sample also covers a shorter time frame, which further limits the comparability. Another critical factor that also impacts comparability is differences in accounting regulations. US companies apply the General Accepted Accounting Principles (GAAP) while Swedish companies use the International Financial Reporting Standard (IFRS). One key difference between the two regulations is that IFRS is principle-based while GAAP is rules-based, meaning IFRS leaves more room for interpretation which could potentially affect the numbers (Staff of the U.S. Securities and Exchange Commission, 2019).

7. Suggestions for future research

Conducting this study has provided two primary insights into how the research within the area could be expanded. One is related to the breakdown of firm characteristics and the other to how the limited data sample could be handled.

Firstly, an area for future research that could be of interest for further analysis is how firm characteristics affect the quality of forecasts. In our study, we focused on the firm characteristics presented by Harris and Wang, (2019), being industry membership, size and E/P ratio. Size was defined as market capitalization following Harris and Wang, (2019). Market capitalization is forward looking (Dang et al., 2018), and can therefore be argued to be the most appropriate for earnings forecasts. However, other definitions of firm size, such as revenue or total assets. can also be used and are therefore relevant to consider. Revenue is a measurement of the product market competition while total assets focus on the amount of resources that is available to the firm (Dang et al., 2018). This would remove the potential issue with inflated stock prices for companies with low profits (Financial Times, 2019). While industry members also can provide interesting insight, the classification system we used is not suitable for Swedish data as six of the twelve industries contained too few observations and were reclassified as “Others”. Therefore, utilizing another industry classification system could yield stronger results.

Secondly, the Compustat database from which we acquired the accounting data has a lot of missing data for Swedish companies before 2004, when the number of observations per year was below 100 (Appendix 8). The dataset after 2004 is much more complete. Therefore, testing the same methodology again in e.g. 10 years would give an additional 1000-2000 observations and allow for more comprehensive conclusions. Another way of solving the problem of limited observations would be to expand the sample. The original study by Harris and Wang, (2019) covered the US market while we focused on Swedish data. To make the samples more comparable, a study could be conducted on the Nordic or European region, since this would make the sample size closer to that of Harris and Wang, (2019).

8. Conclusion

This study investigates how stock price can be used to forecast less biased and more accurate earnings using model-based forecasts for Swedish listed companies between 1988 and 2019. In line with previous research, we find that stock price can be used to forecast less biased earnings. We consider this result suggestive, but not conclusive, since the level of significance varies between 1 % and 5 %. For the five-year ahead forecasts, the bias is also not statistically significantly different from the random walk model. We find no indication that this yields more accurate forecasts compared to pure-accounting based models, which is contradicting previous research. Furthermore, we find indications that stock price does contain incremental information about future earnings in the one- and two-year ahead forecast, which is supported by previous research (Harris and Wang, 2019; Azevedo et al., 2020). However, due to the limited number of observations, the comparability of our results for the five-year ahead forecast is somewhat problematic.

Firm characteristics also had a significant effect on the quality of the forecasts. For the financial services sector, the analyst consensus was statistically significantly more accurate than the model-based forecasts, for the one- and two-year ahead forecasts, at the 1 % level. This was also the case for the larger companies and those with a high E/P ratio. We therefore find suggestive support for analyst consensus outperforming model-based forecasts for companies with these firm characteristics. Due to limited number of observations, we were not able to test our hypotheses for the five-year ahead forecasts. This decreases the validity of our study since previous research indicates that model-based forecasts are more accurate than analyst consensus in the long-term (Harris and Wang, 2019). When comparing the PW and EP models, we found that the PW outperformed the EP model. This is also suggestive, since the EP model is statistically significantly less biased, at the 1 % level, for the one-year ahead forecast and was associated with a forecast error that is less predictable.

For future research, we recommend further analysis of firm characteristics based on a broader definition of firm size. We also suggest that another industry classification system should be used when analysing Swedish data. Another interesting area for future research is conducting a similar study on a broader dataset, such as the Nordic or European region. This would increase the sample size and therefore likely lead to statistically stronger results and allow for more comprehensive conclusions.

References

Alford AW, Berger PG., 1999, “A Simultaneous Equations Analysis of Forecast Accuracy, Analyst Following, and Trading Volume”, *Journal of Accounting, Auditing & Finance,* vol. 14 no. 3, pp. 219-240.

Ashton D, Wang P 2013, “Terminal valuations, growth rates and the implied cost of capital” *Review of Accounting Studies* vol. 18 no. 1, pp. 261–290.

Ball, R., Kothari, S. P., & Robin, A. 2000, “The effect of international institutional factors on properties of accounting earnings”, *Journal of Accounting and Economics*, vol. 29, no. 1, pp. 1–51.

Brown, L., Griffin, P., Hagerman, R., & Zmijewski, M. 1987, “Security analyst superiority relative to univariate time series models in forecasting quarterly earnings”, *Journal of Accounting and Economics*, vol. 9, no. 1, pp. 61–87.

Dang, Chongyu & (Frank) Li, Zhichuan & Yang, Chen, 2018, "Measuring firm size in empirical corporate finance," *Journal of Banking & Finance, Elsevier,* vol. 86, no. C, pp. 159-176.

Dugar, Amitabh and Nathan, Siva, 1995, “The Effect of Investment Banking Relationships on Financial Analysts' Earnings Forecasts and Investment Recommendations*”, *Contemporary Accounting Research*, vol. 12, no. 1, pp. 131-160.

Elgers, P., & Murray, D. 1992, “The relative and complementary performance of analyst and security-price-based measures of expected earnings”, *Journal of Accounting Research,* vol. 15, no. 2–3, pp. 303–316.

Eugene F. Fama 1970, “Efficient Capital Markets: A Review of Theory and Empirical Work”, *The Journal of finance,* vol. 25, no. 2, pp. 383-417.

Fama, E., & French, K. 2000, “Forecasting profitability and earnings”, *Journal of Business,* vol. 73, no. 2, pp. 161–175.

Fama, E., & French, K. 2006, “Profitability, investment and average returns”, *Journal of Financial Economics*, vol. 82, no. 3, 491–518.

Feltham, G., & Ohlson, J. 1995, “Valuation and clean surplus accounting for operating and financial activities” *Contemporary Accounting Research*, vol. 11, no. 2, pp. 689–732.

Feltham, Ga and Ohlson, Ja, 1996, “Uncertainty resolution and the theory of depreciation measurement”, *Journal of Accounting Research*, vol. 34, no. 2, pp. 209-234.

Fildes, 1991, “Efficient use of information in the formation of subjective industry forecasts”, *Journal of Forecasting*, vol. 10 no. 6, pp. 597-617.

Francis, Jennifer, and Donna Philbrick, 1993, “Analysts' Decisions As Products of a Multi-Task Environment”, *Journal of Accounting Research*, vol. 31, no. 2, pp. 216–230.

Freeman, R.; J. A. Ohlson; and S. H. Penman, 1982, “Book Rate-of-Return and Prediction of Earnings Changes: An Empirical Investigation”, *Journal of Accounting Research*, vol. 20, no. 2, pp. 639-653.

Gordon, M. J. 1962, “The Investment Financing and Valuation of the Corporation”. Homewood, IL: Richard D. Irwin. Web

Guay W, Kothari SP, Shu S, 2011, “Properties of implied cost of capital using analysts' forecasts”, *Australian Journal of Management* vol. 36, no. 2, pp. 125–149.

Joseph D. Piotroski 2000, “Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers”, *Journal of Accounting Research*, vol. 38, pp. 1-41.

Kevin K. Li, Partha Mohanram, 2014, “Evaluating cross-sectional forecasting models for implied cost of capital”, *Review of accounting studies*, vol. 19, pp. 1152-1185.

Kewei Hou . Mathijs A. van Dijk. Yinglei Zhang. 2012, "The implied cost of capital: A new approach", *Journal of accounting and economics* vol. 53, no. 3, 2012 pp. 504-526.

Khan, M., & Watts, R. 2009, "Estimation and empirical properties of a firm-year measure of conservatism", *Journal of Accounting and Economics*, vol. 48, no. 2-3, pp. 132–150.

Kross, W., Ro, B., & Schroeder, D. 1990, "Earnings Expectations: The analysts´ information advantage", *The Accounting Review*, vol. 65, no. 2, pp. 461-476.

Lee, Charles M.C, So, Eric C 2017, "Uncovering expected returns: Information in analyst coverage proxies", *Journal of financial economics*, vol. 124, no. 2, pp. 331-348.

Lin, Hsiou-wei and McNichols, Maureen F., 1998, "Underwriting relationships, analysts' earnings forecasts and investment recommendations", *Journal of Accounting and Economics*, 25, no. 1, pp. 101-127.

Lintner, J. 1956, "Distribution of Incomes of Corporations among Dividends, Retained Earnings, and Taxes", *The American Economic Review*, vol. 46, no. 2, pp. 97-113.

Louis K. C. Chan, Jason Karceski and Josef Lakonishok. 2003, "The Level and Persistence of Growth Rates", *The Journal of Finance*, vol. 58, no. 2 , pp. 643-684.

Mark T. Bradshaw, Michael S. Drake, James N. Myers, Linda A. Myers. 2012, "A re-examination of analysts' superiority over time series forecasts of annual earnings", *Review of accounting studies*, 2012 vol. 17, no. 4, pp. 944-968.

Mincer, J., & Zarnowitz, V. 1969, "The evaluation of economic forecasts." In J. Mince (Ed.), *Economic forecasts and expectations*, New York: National Bureau of Research.

Nekrasov, A., & Ogneva, M. 2011, "Using earnings forecasts to simultaneously estimate firm-specific cost of equity and long-term growth", *Review of Accounting Studies*, vol. 16, no. 3, pp. 414–457.

Ohlson, J. 1995, “Earnings, book values, and dividends in equity valuation”, *Contemporary Accounting Research*, vol. 11 no. 2, 661–687.

Ping Zhou and William Ruland, 2006, “Dividend Payout and Future Earnings Growth” *Financial Analysts Journa*l vol. 62, no. 3, pp. 58-69.

Pope, P., & Wang, P., 2005,. “On the relevance of earnings components: Valuation and forecasting links”, *Review of Accounting Studies*, vol. 10 no. 4, pp. 387–407.

Ramnath, Sundaresh & Rock, Steve & Shane, Philip, 2008, "The financial analyst forecasting literature: A taxonomy with suggestions for further research," *International Journal of Forecasting, Elsevier*, vol. 24, no.1, pages 34-75.

Ray Ball and Philip Brown. 1968, “An Empirical Evaluation of Accounting Income Numbers”, *Journal of Accounting Research*, vol. 6, no. 2, pp. 159-178.

Ray Ball and Ross Watts. 1972, “Some Time Series Properties of Accounting Income” *The Journal of Finance*, vol. 27, no. 3, pp. 663-681.

Richard D.F. Harris, Pengguo Wang. 2019, “Model-based earnings forecasts vs. financial analysts' earnings forecasts”, *The British Accounting Review*, vol. 51, no. 4, pp. 424-437.

Richardson S, Tuna I, Wysocki P 2010, “Accounting anomalies and fundamental analysis: a review of recent research advances”, *Journal of Accounting and Economics*, vol. 50, no. 2–3, pp. 410–454.

Robert D. Arnott and Clifford S. Asness 2003, “Surprise! Higher Dividends = Higher Earnings Growth”, *Financial analysts journal*, vol. 59, pp.70-87.

Ross Watts, 1973, “The Information Content of Dividends”, *The Journal of Business*, vol. 46, no. 2, pp. 191-211.

Runsten, M. 1998, "The association between accounting information and stock prices: model development and empirical tests based on Swedish data"

Skogsvik, S. 2008, "Financial statement information, the prediction of book return on owners' equity and market efficiency: the Swedish case", *Journal of Business Finance & Accounting*, vol. 35, no. 7-8, pp. 795-817.

Sougiannis, Theodore & Yaekura, Takashi., 2001, "The Accuracy and Bias of Equity Values Inferred from Analysts' Earnings Forecasts", *Journal of Accounting, Auditing & Finance,* vol. 16 no. 4. pp. 331-362.

Tarun Chordia, Richard Roll, Avanidhar Subrahmanyam 2008, "Liquidity and market efficiency", *Journal of Financial Economics,* vol. 87, no. 2, pp. 249-268.

Vitor Azevedo, Patrick Bielstein, Manuel Gerhart, 2020, "Earnings forecasts: the case for combining analysts' estimates with a cross-sectional model", *Review of Quantitative Finance and Accounting,* vol. 56, no. 2, pp. 545–579.

Weiss, D., Naik, P., & Tsai, C. 2008, "Extracting forward-looking information from security prices: A new approach", *The Accounting Review,* vol. 83, no. 4, pp. 1101–1124.

Zhan Gao, Wan-Ting Wu 2014, "Predicting long-term earnings growth from multiple information sources", *International Review of Financial Analysis,* vol. 32, pp. 71-84.

Electronic Sources

Becker, D. 2001, "Analyzing analysts. Remarks before the committee on federal regulation of securities of the American bar association" Available: http://www.sec.gov/news/speech/spch510.htm [2021-05-12]

Financial Times, Vincent, Mathew, 2019 "Loss-making tech companies are floating like it's 1999" Available: https://www.ft.com/content/c5705ca2-605c-11e9-9300-0becfc937c37 [2021-05-13]

French, K.R . 2021. *“Detail for 12 Industry Portfolios”* Available: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_12_ind_port.html [2021-03-26]

Gerakos, Joseph and Gramacy, Robert, Regression-Based Earnings Forecasts (July 31, 2013). Chicago Booth Research Paper no. 12-26, Available at SSRN: https://ssrn.com/abstract=2112137 or http://dx.doi.org/10.2139/ssrn.2112137

Staff of the U.S. Securities and Exchange Commission, 2019, “A Comparison of U.S. GAAP and IFRS” Available: ttps://www.sec.gov/spotlight/globalaccountingstandards/ifrs-work-plan-paper-111611-gaap.pdf [2021-05-13]

Appendix

Appendix 1

Forecast bias pairwise mean difference

Panel A: One-year ahead forecast, mean difference and t-statistic					
	HDZ	PW	RW	EP	IBES
AR(1)	-0.0148***	0.0026	-0.0077***	-0.0022***	0.0077
HDZ		0.0174***	0.0071**	0.0126***	0.0225***
PW			-0.0103***	-0.0048**	0.0050
RW				0.0055**	0.0154**
EP					0.0098
Panel B: Two-year ahead forecast, mean difference and t-statistic					
	HDZ	PW	RW	EP	IBES
AR(1)	-0.0155***	0.0071***	-0.0037	-0.0016*	0.0285***
HDZ		0.0226***	0.0118***	0.0139***	0.0439***
PW			-0.0108**	-0.0087***	0.0213***
RW				0.0021	0.0321***
EP					0.0301***
Panel C: Five -year ahead forecast, mean difference and t-statistic					
	HDZ	PW	RW	EP	IBES
AR(1)	0.0269**	0.0479***	0.0190	-0.0005	0.1092***
HDZ		0.0210**	-0.0079	-0.0274***	0.0823**
PW			-0.0288	-0.0484***	0.0613
RW				-0.0195	-0.8737*
EP					0.1097***

The table shows the mean bias difference for each pair of forecasts for the one-, two-, and five-year ahead forecasts. The AR(1) model is first compared to the HDZ model, then to the PW model, and so on. The significance level is the result of testing the null hypotheses that the bias of each pair is equal. *** p<0.01, ** p<0.05, * p<0.1.

Appendix 2

Forecast accuracy pairwise mean difference

Panel A: One-year ahead forecast, mean difference and t-statistic					
	HDZ	PW	RW	EP	IBES
AR(1)	0.0098***	0.0086***	0.0059**	0.0025***	0.0650***
HDZ		-0.0011	-0.0039	-0.0072***	0.0552***
PW			-0.0027	-0.0061***	0.0564***
RW				-0.0033	0.0591***
EP					0.0624***

Appendix 2, cont.

Panel B: Two-year ahead forecast, mean difference and t-statistic

	HDZ	PW	RW	EP	IBES
AR(1)	0.0101***	0.0093***	-0.0043	0.0024***	0.0614***
HDZ		-0.0009	-0.0145***	-0.0078***	0.0513***
PW			-0.0136***	-0.0069***	0.0521***
RW				0.0067*	0.0658***
EP					0.0591***

Panel C: Five-year ahead forecast, mean difference and t-statistic

	HDZ	PW	RW	EP	IBES
AR(1)	0.0323***	0.0000	-0.0376*	0.0027	0.0589*
HDZ		-0.0323***	-0.0699***	-0.0296***	0.0266
PW			-0.0376*	0.0027	0.0588*
RW				0.0403*	0.0965**
EP					0.0562*

The table shows the mean accuracy difference for each pair of forecasts for the one-, two-, and five-year ahead forecasts. The AR(1) model is first compared to the HDZ model, then to the PW model, and so on. The significance level is the result of testing the null hypotheses that the accuracy of each pair is equal. *** p<0.01, ** p<0.05, * p<0.1.

Appendix 3

PW model average estimated coefficients

	t=1	t=2	t=3	t=4	t=5
$PRICE_t$	0.0010	-0.0101	-0.0108	-0.0159	-0.0105
EPS_t	0.4000***	0.1532***	0.0734	0.1011	0.0439
$NEGE_t$	-2.0252***	-2.5316***	-2.5867***	-3.3562***	-4.3626***
$BOOK_t$	0.0270**	0.0576***	0.0806***	0.0304	0.0272***
$BOOK_{t-1}$	0.0243*	0.0193	-0.0017	0.0157	0.0095
$PRICE_{t-1}$	-0.0021	0.0004	0.0073	0.0138	0.0112
$ACCRUALS_t$	-0.0676***	-0.0359**	0.0103	0.0143	-0.0053
$CONSTANT$	1.1736***	1.8449***	1.8762***	3.1690***	4.1136***
N	3 005	2 707	2 429	2 174	1 934
R^2	0.405	0.247	0.219	0.141	0.109

The table reports the coefficients for the PW model. The constant ($CONSTANT$), stock price ($PRICE$), earnings per share in the current period (EPS), whether earnings where negative in the current period, one if they were and zero otherwise ($NEGE$), book value per share ($BOOK$), and operating accruals ($ACCRUALS$). The significance level is the result of testing the null hypotheses that the coefficient is zero. *** p<0.01, ** p<0.05, * p<0.1.

Appendix 4

HDZ model average estimated coefficients

	t=1	t=2	t=3	t=4	t=5
$DIVIDEND_t$	0.2712***	0.2685***	0.1967	0.1897	0.1792
EPS_t	0.4510***	0.2809***	0.2504***	0.1249	0.0453
$NEGE_t$	-1.1421	-0.2824	-0.1751	-0.8297	-2.0577**
DD_t	1.1289**	2.8361***	2.6447***	3.7263***	3.4562***
$ASSETS_t$	0.0042***	0.0055***	0.0062***	0.0048***	0.0044***
$ACCRUALS_t$	-0.0826***	-0.0726***	-0.0413	0.0023	-0.0131
$CONSTANT$	0.8196	0.1573	0.4162	0.5421	1.6115**
N	3 005	2 707	2 429	2 174	1 934
R^2	0.4050	0.2272	0.1939	0.1639	0.1323

The table reports the coefficients for the PW model. The constant ($CONSTANT$), dividend per share (DPS), earnings per share in the current period (EPS), whether earnings where negative in the current period, one if they were and zero otherwise ($NEGE$), whether the firm paid a dividend, one if they did and zero otherwise (DD), assets per share ($ASSETS$) and operating accruals ($ACCRUALS$). The significance level is the result of testing the null hypotheses that the coefficient is zero. *** p<0.01, ** p<0.05, * p<0.1.

Appendix 5

EP model average estimated coefficients

	t=1	t=2	t=3	t=4	t=5
EPS_t	0.5683***	0.4427***	0.3926***	0.2845***	0.1969***
$NEGE_t$	-3.3770***	-3.1986***	-2.7590**	-3.8934***	-5.4044***
$NEGE * EPS_t$	-0.5311	-0.5240	-0.3940	-0.3308	-0.4655
$CONSTANT$	2.5672***	3.1707***	3.3120***	3.9794***	4.7607***
N	3 005	2 707	2 429	2 174	1 934
R^2	0.3613	0.1747	0.1426	0.1106	0.0911

The table reports the coefficients for the EP model. The constant ($CONSTANT$), earnings per share in the current period (EPS), whether earnings where negative in the current period, one if they were and zero otherwise ($NEGE$), and a variable multiplying $NEGE$ with EPS ($NEGE * EPS$). The significance level is the result of testing the null hypotheses that the coefficient is zero.*** p<0.01, ** p<0.05, * p<0.1.

Appendix 6

Paired t-test I/B/E/S forecast vs. I/B/E/S actual

Variable	Mean	Std.Err.	Std.Dev.	95 % Conf. Interval	
IBES Forecast	5.1294	0.1162	5.3336	4.9016	5.3573
IBES Actual	4.4757	0.2506	11.5017	3.9843	4.9671
Difference	0.6538	0.2316	10.6319	0.1995	1.1080
t-statistics	2.8226				
df	2106				

The table reports the result of the paired t-test testing the null hypotheses that the I/B/E/S consensus and the I/B/E/S actual data is equal.

Appendix 7

Pairwise correlations (N = 3048)

	EPS	$AFEPS_1$	$REPS_1$	$PRICE$	$BOOK$	$DIVIDEND$	$ASSETS$	$ACCRUAL$
EPS	1.000							
$AFEPS_1$	0.439***	1.000						
$REPS_1$	0.628***	0.431***	1.000					
$PRICE$	0.234***	0.547***	0.197***	1.000				
$BOOK$	0.678***	0.375***	0.552***	0.224***	1.000			
$DIVIDEND$	0.523***	0.324***	0.446***	0.121***	0.450***	1.000		
$ASSETS$	0.345***	0.218***	0.331***	0.066***	0.466***	0.297***	1.000	
$ACCRUALS$	0.475***	0.011	0.208***	0.059***	0.245***	0.104***	0.207***	1.000

The table reports the correlation matrix (Pearson) for current earnings per share *(EPS)*, analyst forecast for one-year ahead *($AFEPS_1$)*, realized earnings in one year *($REPS_1$)*, stock price *($PRICE$)*, book value per share *($BOOK$)*, dividend per share *($DIVIDEND$)*, assets per share *($DIVIDEND$)*, and operating accruals per share *($ACCRUALS$)*. The significance level is the result of testing the null hypothesis that the value is equal to zero. *** p<0.01, ** p<0.05, * p<0.1.

Appendix 8

Observations per year

Year	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Obs.	1	2	3	2	2	1	1	4	7	20	48	55	60	73	92	96
Year	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Obs.	103	134	149	155	154	141	144	158	156	152	155	170	191	209	212	198

The table reports the number of observations per year for the adjusted sample. The total number of observations is 3048. The total number of observations is not equal to the number of observations for estimating the coefficients (Appendix 3, 4, and 5) because the models are estimated using a rolling window of ten years to generate out-of-sample forecasts, and since our sample beings in 1988, the first forecast is made in 1997.

Appendix 9

Paired t-test (pooled model-based vs. analyst consensus forecasts) for Sector 11

Variable	Mean	Std.Err.	Std.Dev.	95 % Conf. Interval		t	df
One-year ahead forecast	0.1014	0.0155	0.2621	0.0708	0.1319	6.5292	284
Two-year ahead forecast	0.1016	0.0192	0.3021	0.0637	0.1394	5.2836	246

The table shows the result of the paired t-test testing the null hypotheses that the forecast error for the pooled model-based forecast is equal to the analyst consensus forecast, for the financial services sector.

Appendix 10

Paired t-test (pooled model-based vs. analyst consensus forecasts) for the 5th size quantile

Variable	Mean	Std.Err.	Std.Dev.	95 % Conf. Interval		t	df
One-year ahead forecast	0.0311	0.0033	0.0755	0.0247	0.0376	9.5205	532
Two-year ahead forecast	0.0291	0.0038	0.0834	0.0216	0.0366	7.6593	480

The table shows the result of the paired t-test testing the null hypotheses that the forecast error for the pooled model-based forecast is equal to the analyst consensus forecast, for the 5th quantile for firm size.

Appendix 11

Paired t-test (pooled model-based vs. analyst consensus forecasts) for the 5th E/P ratio quantile

Variable	Mean	Std.Err.	Std.Dev.	95 % Conf. Interval		t	df
One-year ahead forecast	0.2290	0.0161	0.3301	0.1972	0.2607	14.1813	417
Two-year ahead forecast	0.3247	0.0206	0.3987	0.2842	0.3652	15.7715	374

The table shows the result of the paired t-test testing the null hypotheses that the forecast error for the pooled model-based forecast is equal to the analyst consensus forecast, for the 5th quantile for the E/P ratio.