PROFITABILITY PREDICTION USING MACROECONOMIC FORECASTS

THE INFORMATIVENESS OF GDP GROWTH EXPECTATIONS AND GEOGRAPHIC SEGMENT DISCLOSURES

GUSTAV HALL

VALTER LINDHAGEN

Master Thesis

Stockholm School of Economics

2021

Profitability Prediction Using Macroeconomic Forecasts: The Informativeness of GDP Growth Expectations and Geographic Segment Disclosures

Abstract:

Many firms today have an international footprint which means that they are exposed to different macroeconomic environments across the world. This Master Thesis investigates the usefulness of macroeconomic forecasts for the prediction of firm profitability. Recent research has shown that firm-level country exposures, determined based on geographic segment disclosures, can be combined with country-level predictions of real GDP growth to create a “MACRO” variable which has a significant relationship with future return on net operating assets, “RNOA”. Focusing on the time period 2000-2019, this thesis confirms the relationship on a sample of Swedish-listed manufacturing firms, for which global macroeconomic conditions play an important role. The study extends previous research by conducting a more comprehensive out-of-sample validation. Three different prediction models of one-year-ahead RNOA are estimated and compared. The results suggest that out-of-sample prediction accuracy is improved by including the MACRO variable, although not all tests yield significant results. In addition, a model containing only past RNOA and MACRO is shown to produce significantly lower out-of-sample forecast errors than a model which also contains additional accounting and financial market variables. The results shed light on the strong forecasting power that past RNOA has on its own, which has been documented several times in previous research and which can be attributed to its strong mean reversion properties.

Keywords:

Profitability prediction, Out-of-sample, Return on net operating assets, Macro to micro, Geographic segments

Authors:

Gustav Hall (41567)
Valter Lindhagen (23978)

Tutor:

Stina Skogsvik, Assistant Professor, Department of Accounting

Master Thesis
Master Program in Accounting, Valuation and Financial Management
Stockholm School of Economics

Acknowledgements

We would like to give our warmest thanks to our tutor Stina Skogsvik, Assistant Professor at the Department of Accounting, for her valuable guidance and support throughout the research process. We are also thankful to Antonio Vazquez, Assistant Professor at the Department of Accounting, for his engagement and valuable statistical advice.

Stockholm, May 2021

Gustav Hall Valter Lindhagen

Contents

1. INTRODUCTION 5
1.1. Background and research question 5
1.2. Disposition 6
2. LITERATURE REVIEW 7
2.1. Profitability as a value driver 7
2.2. Financial statements information and prediction of firm profitability 9
2.3. Macroeconomic information and prediction of firm profitability 11
2.3.1. Fundamental prediction models and macroeconomic contextual analysis 12
2.3.2. Profitability prediction with macroeconomic independent variables 12
2.4. Contribution 14
3. HYPOTHESES 16
3.1. Hypothesis A: Relationship between MACRO and future RNOA 16
3.2. Hypotheses B: Comparison of prediction accuracy across models 16
3.3. Hypothesis C: Comparison of prediction accuracy to analyst forecasts 17
4. RESEARCH DESIGN 19
4.1. The MACRO variable 19
4.1.1. Calculating the MACRO variable 19
4.1.2. Manual coding of country exposures 20
4.2. Estimation of prediction models 24
4.2.1. Timeline 26
4.3. Out-of-sample validation 28
4.3.1. Forecast errors 28
4.3.2. Proportion of correctly predicted increases and decreases in RNOA 28
4.4. Hypothesis testing 29
4.4.1. Hypothesis A: MACRO coefficient in estimation 29
4.4.2. Hypotheses B-C: Wilcoxon signed rank test 29
4.4.3. Hypotheses B-C: Proportion test 31
5. DATA 33
5.1. Industry focus and time period 33
5.2. Data collection and selection criteria 33
5.3. Descriptive statistics 37

6. RESULTS **41**
6.1. Estimating the prediction models **41**
6.2. Out-of-sample validation **44**
6.2.1. Forecast errors 44
6.2.2. Wilcoxon signed rank tests 47
6.2.3. Correctly predicted increases and decreases in RNOA 50
6.2.4. Proportion tests 52
7. DISCUSSION **54**
7.1. Summary of findings **54**
7.2. Predictive ability of the MACRO variable **54**
7.3. Predictive ability of past RNOA **56**
8. CONCLUSIONS AND FUTURE RESEARCH **58**
9. REFERENCES **60**
10. APPENDIX **63**

1. Introduction

1.1. Background and research question

Firms are becoming increasingly globalized and thereby exposed to different macroeconomic environments. Recognizing these exposures should be informative for future profitability. If firm A has 100 percent of its sales in Sweden and firm B has 50/50 sales split between Sweden and France, and Sweden is expected to outperform France in terms of economic growth, then holding all else equal, firm A could also be expected to show better future profitability development than firm B.

The objective of this Master Thesis is to investigate whether the inclusion of such macroeconomic information can improve predictions of firm profitability. Recent research by Li et al. (2014) has found that firm-specific geographic exposure in combination with country-level predictions of real GDP growth, in the form of a "MACRO" variable, is helpful for predicting one-year-ahead return on net operating assets ("RNOA"). They estimate a prediction model which contains the MACRO variable and find that a one percentage point increase in real GDP growth expectations translates, on average, into an increase of 27 basis points in one-year-ahead RNOA. However, the out-of-sample validation of the prediction model is limited and performed in unreported tests. To the best of our knowledge, no later articles have expanded on the topic of MACRO's predictive ability in out-of-sample forecasts. Thus, this thesis addresses the question of whether the inclusion of macroeconomic information in firm-level forecasting can improve the prediction of RNOA in an out-of-sample context.

In early profitability prediction research, a common conclusion was that the time-series behavior of earnings is best described as a "random walk" or a "martingale" process. Freeman et al. (1982) later showed that profitability, measured as return on equity ("ROE"), follows a mean-reverting process, meaning that it tends to revert to some average level over time. Following these findings, a number of studies have investigated the usefulness of accounting ratios in profitability prediction, but with diverse suggestions regarding which ratios to use. The common denominator has been the strong mean reversion characteristic of ROE and RNOA (e.g., Fairfield et al., 1996; Nissim & Penman, 2001; Skogsvik, 2002a). Limited attention has been given to the inclusion of external factors, such as macroeconomic information.

Among the few studies that examine the usefulness of macroeconomic information, many of them perform a contextual analysis. It means that the sample is sorted into groups based on macroeconomic characteristics that are found to alter the predictive content of fundamental factors for future profitability, proving that the macroeconomic context matters but not how to benefit from that knowledge. In contrast, recent research by Li et al. (2014) includes macroeconomic forecasts as independent variables in the prediction models. The results are promising with regards to explaining one-year-ahead profitability, but no comprehensive out-of-sample validation has yet been presented.

Drawing on previous research, this thesis estimates three different prediction models for one-year-ahead RNOA and compares the prediction accuracy of the models in out-of-sample forecasts. The first model contains past RNOA as the single independent variable, inspired by the findings in prior literature that past profitability has very high forecasting power on its own (e.g., Freeman et al., 1982; Fairfield et al., 1996; Skogsvik, 2002a). The second model adds the MACRO variable from Li et al. (2014), in order to evaluate the predictive ability of macroeconomic forecasts for future profitability. The third model is more comprehensive and contains additional accounting and financial market variables, similar to the original model used by Li et al. (2014).

In contrast to the study by Li et al. (2014), the research design of this thesis makes it possible to evaluate the incremental predictive content of MACRO over past RNOA in isolation, through the comparison of the first and second model described above. However, the primary contribution of the thesis is the more comprehensive validation of the estimated models in out-of-sample forecasts. Li et al. (2014) use absolute forecast errors from unreported tests to describe the predictive ability of MACRO out-of-sample. This thesis provides detailed reported tests and also use additional measures of prediction accuracy in the form of squared forecast errors as well as the proportion of correctly predicted increases and decreases in RNOA. Furthermore, the validation is extended by comparing the prediction accuracy of the models against that of analyst forecasts, in order to evaluate their usefulness against an external benchmark. Finally, this study focus on a more condensed set of firms for which the macroeconomic development is believed to be of particularly high importance, namely Swedish-listed manufacturing firms which generally have large global footprints and high dependency on the economic cycle.

1.2. Disposition

The disposition of the thesis is as follows. First, section 2 presents the theoretical background of the study and gives an overview of previous research on profitability prediction. The contribution of the study to the existing literature is also described further. Section 3 presents the hypotheses of the study. In section 4, the research design is described including a detailed presentation of the MACRO variable. This section also presents the statistical tests that are used to evaluate the hypotheses. Section 5 gives an overview of the data collection process and provides descriptive statistics. The results are presented in section 6, with the estimation results for the prediction models first, followed by the out-of-sample prediction results. Section 7 provides a discussion of the results in relation to prior literature. Section 8 summarizes the conclusions of the study and presents suggestions for future research.

2. Literature review

The following sections present the theoretical context of the study and gives an overview of prior literature on firm profitability prediction and related research fields. Research on the prediction of firm profitability is part of a broader research field on fundamental analysis. The concept of fundamental analysis is defined by Penman (2013, p.84) as "the method of analyzing information, forecasting payoffs from that information, and arriving at a valuation based on those forecasts". The research interest in this field has become very large, in part because of mounting evidence in the financial economics literature against the efficient market hypothesis (Kothari, 2001). Capital markets research in this area investigates whether fundamental analysis can be used to identify mispriced securities and earn excess returns by trading in those securities (Kothari, 2001).

The information that is used in fundamental analysis to "forecast payoffs" involves current and historical financial statements together with industry and macroeconomic data (Kothari, 2001). The majority of past research have focused on the former, financial statements data, without explicitly incorporating any information external to the firm itself (Li et al., 2014). This Master Thesis expands on the less explored area of forecasting firm profitability by also taking into account macroeconomic factors in conjunction with financial statements data.

In section 2.1, the RIV and VAV models are described briefly to demonstrate why forecasts of profitability are crucial for the assessment of firm value. Section 2.2 presents prior research on how financial statements information can be used to forecast profitability. Section 2.3 summarizes prior studies that have incorporated macroeconomic factors, in conjunction with accounting information, to predict profitability. Finally, section 2.4 highlights the contribution of this study to the existing literature. An overview of the most relevant literature is presented in Table A1 in Appendix.

2.1. Profitability as a value driver

Profitability prediction has a key role within the field of fundamental analysis, and when forecasting payoffs. The enterprise value of a firm can be divided into the value of equity and the value of debt (Penman, 2013). To value a firm's equity, either the enterprise value can be calculated followed by an adjustment for the value of debt, or the value of equity can be calculated directly (Penman, 2013). Under the same assumptions, although different underlying logics, the value of equity will be the same (Penman, 2013). In this section, with the purpose of highlighting profitability as a key value driver, two fundamental valuation models are presented, one valuing equity directly and one valuing equity via the calculation of enterprise value adjusted for the value of debt.

The residual income valuation (RIV) model values a firm's equity directly. It is algebraically equivalent to the well-known Dividend Discount Model (DDM) under the assumption that the clean-surplus relationship holds (Soliman, 2008). The RIV model is well known for displaying the relationship between accounting data and firm value, linking to a firm's value creation process, see Lee (1999) for an overview.

$$V(BV_0) = BV_0 + \sum_{t=1}^{\infty} \frac{BV_{t-1}(ROE_t - r_E)}{(1 + r_E)^t} \tag{2.1}$$

$V(BV_0) =$ Fundamental value of owners' equity at time t=0
$BV_0 =$ Book value of owners' equity at time t=0
$BV_{t-1} =$ Book value of owners' equity at time t-1
$ROE_t =$ Return on book value of owners' equity at time t
$r_E =$ Cost of equity capital

The value-added valuation (VAV) model, also referred to as the residual operating income model, values a firm's equity via the calculation of enterprise value (see e.g. Skogsvik, 2002b; Penman, 2013). It is equivalent to a free cash flow valuation model. The forecasted payoffs flow from the operating activities of a firm and is later adjusted for the value of the debt claims of those cash flows (Penman, 2013). The VAV model, similar in structure to the RIV model, also highlights the relationship between accounting data and firm value, linking to a firm's value creation process.

$$V(NOA_0) = NOA_0 + \sum_{t=1}^{\infty} \frac{NOA_{t-1}(RNOA_t - r_{Wacc})}{(1 + r_{Wacc})^t} \tag{2.2}$$

$V(NOA_0) =$ Fundamental value of net operating assets at time t=0
$NOA_0 =$ Book value of net operating assets at time t=0
$NOA_{t-1} =$ Book value of net operating assets at time t-1
$RNOA_t =$ Return on net operating assets at time t
$r_{Wacc} =$ Weighted average cost of capital

Both Equation 2.1 and 2.2 demonstrate that future profitability, in the form of ROE or RNOA, drives the value creation. In the RIV model, value is created if ROE exceeds the cost of equity and, similarly, value is created in the VAV model if RNOA exceeds the weighted average cost of capital. Because profitability drives value, much attention has been devoted to predicting it.

2.2. Financial statements information and prediction of firm profitability

Most of the literature concerning profitability prediction has focused on the prediction of earnings, that is, earnings per share or net income. A common conclusion in the early literature was that the time-series behavior of earnings is best described as a "random walk" or a "martingale" process (Freeman et al., 1982). Martingale means that, conditioned on past earnings, future changes in earnings are drawn from a distribution with mean zero. If the error terms are independent and identically distributed, the process is referred to as a random walk. Assuming that earnings follow a martingale or a random walk, the best prediction of next period's earnings is that they are equal to last period's earnings. Ball & Watts (1972) describes that earnings often follow a martingale or a random walk with a positive drift, meaning that last period's earnings is adjusted with a positive drift term to arrive at the best prediction of next period's earnings.

Freeman et al. (1982) challenge previous studies' inability to reject the random walk hypothesis and provide evidence that current ROE has predictive content with respect to future changes in EPS and ROE, at least when current ROE significantly deviates from its mean. The prediction results are validated out-of-sample. The authors use a logit approach and construct two different univariate prediction models, one generating probabilities for an increase in next year's ROE and one generating probabilities for an increase in next year's EPS. Both models use the prior year's ROE as the single independent variable. Just like the authors hypothesize, a negative coefficient is obtained for the independent variable in both models. The logic behind this is that ROE follows a mean-reverting process, meaning that it tends to revert to some average level over time. This also explains why the predictive content of ROE is highest when ROE deviates significantly from its mean. The mean reversion behavior had earlier been observed more informally by Beaver (1970) and Lookabill (1976). It is supported by standard economic arguments, saying that entrepreneurs will seek to leave less profitable industries for more profitable ones. Thus, in a competitive environment, profitability is mean-reverting both within and across industries (Fama & French, 2000).

Following the findings by Freeman et al. (1982) that ROE has predictive content with respect to earnings, a number of studies have investigated the usefulness of other accounting numbers and ratios in the prediction of earnings. Several authors, for example Bernard & Noel (1991) and Stober (1993), examine the predictive content of individual accounting numbers such as accounts receivables and inventory. A large number of authors have also designed multivariate prediction models containing several accounting variables. Ou & Penman (1989) pioneered the multivariate analysis by constructing a logit prediction model for one-year-ahead earnings containing a relatively large set of accounting descriptors: 16 in their first estimation period, 1965-1972, and 18 in their second estimation period, 1973-1977. The descriptors are selected based on statistical tests of 68 different accounting ratios and their individual forecasting power. Lev & Thiagarajan (1993) use an alternative approach for selecting the variables. They identify twelve "fundamental signals" based on a guided search of financial publications, with the purpose to identify variables that are used by financial analysts in practice. The identified variables include, for example, change in inventories and change in capital expenditures as well as non-financial statements information like order backlog and labor force. The variables are found to be highly useful for explaining both stock returns and earnings. The authors do not design an earnings prediction model but instead they group firms into high and low earnings quality, determined based on the fundamental signals, and show that future persistence and growth in earnings are significantly different between the groups.

Abarbanell & Bushee (1997) use a model containing nine of the twelve independent variables from Lev & Thiagarajan (1993) to predict one-year-ahead earnings changes and five-years-ahead average earnings growth. Two of the variables, capital expenditure and accounts receivables, are found to have opposite signs than documented by Lev & Thiagarajan (1993). This could represent different interpretations of the signals than proposed by Lev & Thiagarajan (1993). The prediction accuracy of the models is not validated out-of-sample. Dowen (2001) extends these studies by including additional signals inspired by finance literature, namely dividend yield, firm size and book-to-market ratio. Neither Dowen's models are tested out-of-sample. Both Dowen (2001), Lev & Thiagarajan (1993) and Abarbanell & Bushee (1997) also test how the information content of their fundamental signals are affected by different macroeconomic contexts, which is described further in section 2.3.1.

While most of the literature concerning profitability prediction has focused on the prediction of earnings, some research have also been conducted on the prediction of ROE and RNOA. Similar to Freeman et al. (1982), described above, Fairfield et al. (1996) also predict ROE in out-of-sample forecasts. They design prediction models for one-year-ahead ROE which disaggregate the prior year's ROE into different components, such as operating earnings, non-operating earnings and taxes, and special items. They show that such earnings disaggregation improves the prediction of ROE, compared to a benchmark model with the prior year's ROE as the single independent variable. The authors also confirm the mean reversion properties of ROE. Skogsvik (2002a) tests three different models for prediction of changes in medium-term (3-year) ROE. Interestingly, she finds that the most parsimonious model, where past average ROE is the only independent variable, performs better than the more elaborate models which include a large sets of accounting ratios. In other words, the inclusion of additional accounting ratios actually deteriorated the predictions. The predictions are performed in holdout samples and the study underlines the strong predictive power of past ROE documented in earlier studies (e.g., Freeman et al., 1982).

In contrast to ROE, RNOA captures a firm's operating profitability without the effects of capital structure. Nissim & Penman (2001) estimate multivariate models to forecast RNOA with good explanatory results in estimation, but with poor performance in prediction out-of-sample. As a result, their empirical analysis is more descriptive. They use DuPont analysis, decomposing RNOA into profit margin (PM) and asset turnover (ATO). They show that ATO is more persistent than PM and that changes in ATO are predictive of future changes in RNOA after controlling for RNOA. The study also presents evidence of mean reversion in RNOA, similar to what have earlier been observed for ROE. Soliman (2008) extends Nissim & Penman's (2001) work and show that DuPont components are still significant for explaining future changes in RNOA even after controlling for fundamental signals that have been used in other profitability prediction studies, including the signals used by Lev & Thiagarajan (1993) and Abarbanell & Bushee (1997). However, the prediction model is not validated out-of-sample. In summary, the studies by Nissim & Penman (2001) and Soliman (2008) suggest that decomposing RNOA enhances the ability to predict future RNOA compared to if the aggregated level of RNOA is used. Some additional prediction studies of RNOA which also take macroeconomic information into account are described in section 2.3.2.

In general, a few conclusions can be drawn based on the literature presented in this section. First, most of the literature concerning profitability prediction has focused on earnings rather than RNOA or ROE. Second, research on multivariate prediction models is relatively dispersed and not fully conclusive regarding what set of accounting ratios that should be used to forecast profitability. Third, not all studies perform out-of-sample validations, implying that it is uncertain how useful some of the estimated models are in practice. Fourth, past ROE and RNOA have particularly high forecasting power, because they i) contain much of the information captured in other accounting ratios, ii) have strong time-series behavior, meaning that historical profitability is very useful when predicting future profitability, and iii) follow a mean reversion process. Fifth, disaggregation of ROE and RNOA seem to improve forecasting power. In summary, the attractive properties of ROE and RNOA, together with their simplicity, make them particularly relevant for further research.

2.3. Macroeconomic information and prediction of firm profitability

The majority of prior research regarding the prediction of firm profitability, as presented in the previous section, does not take external information such as macroeconomic factors into account. In general terms, the line of research that examines the link between macroeconomic information, accounting data and stock returns can be considered to take a 'macro to micro' perspective (Doukakis et al., 2020). The perspective suggests that macroeconomic information is useful when predicting firm-level fundamentals and profitability. This section intends to demonstrate this usefulness by presenting previous research taking a 'macro to micro' perspective and to highlight the relatively uncharted area of using macroeconomic *forecasts* when predicting profitability.

2.3.1. Fundamental prediction models and macroeconomic contextual analysis

One way of taking external information such as macroeconomic factors into account when making predictions is by conducting a contextual analysis. Lev & Thiagarajan (1993), which was presented in section 2.2., relate the predictive content of the fundamental signals they identify to different macroeconomic contexts. The contexts applied are the change in GDP, inflation, and business activity. By dividing the years examined into groups, depending on the level of each factor, and running separate regressions for each regime, they find that the predictive content of the signals are highly dependent on the state of these macroeconomic factors.

Abarbanell & Bushee (1997) also conduct a contextual analysis. They sort data on two different macroeconomic factors: economic growth and inflation. In line with the findings in Lev & Thiagarajan (1993), the factors alter the relationship between the fundamental signals and future earnings. Prior research is not limited to the macroeconomic factors mentioned above. For example, Dowen (2001) introduces, apart from economic growth and inflation, a monetary policy factor. He conducts a contextual analysis by running annual regressions with dummy variables signaling whether any macroeconomic factor went above or below a certain threshold each year, with the purpose of better understanding the differences in the annually generated coefficients of the other variables. Dowen (2001) finds, in line with previous research, that the macroeconomic factors, including monetary policy, alter the coefficients.

The limitation of the contextual analysis is obvious. It points toward the fact that external factors matter in the context of predicting profitability, but it does not offer a solution for how to incorporate those effects in the predictions, thus, improving the accuracy.

2.3.2. Profitability prediction with macroeconomic independent variables

Instead of conducting a contextual analysis, prior research within the 'macro to micro' perspective have investigated whether the inclusion of macroeconomic factors as independent variables in prediction models can improve the predictability. A lot of attention has been devoted to the explanatory value of *current* macroeconomic information in prediction (see e.g., Nissim & Penman, 2003; Chordia & Shivakumar, 2005; Konchitchki, 2011). A less charted area in prior research instead incorporates macroeconomic *forecasts* when predicting firm profitability.

Early studies that include forecasts of external information apply "line-of-business" estimates (e.g., Kinney, 1971; Collins, 1976). The idea is that growth estimates could be applied to the different business segments of a company, improving the accuracy of the earnings prediction. The estimates are based on industry-specific growth predictions from external sources such as the U.S. Industrial Outlook and the Business and Defense Service Administration. The predictions are made on a year-to-year basis and applied by adjusting prior year earnings with a corresponding growth factor, as opposed to a "random walk" model.

Roberts (1989) goes further and applies a similar method for geographic segments, instead of business segments, when predicting earnings. She includes a firm-specific estimate of GNP[1] growth, based on both disclosed geographic segment sales and geographic segment earnings. The assumption behind such an inclusion is that firm performance depends on the performance of the economies which the firm is exposed to. She finds that the forecasts outperform that of a "random walk" model with and without a trend (drift component). The results indicate that macroeconomic predictions indeed improve predictability, however, the sample and time period is very limited. Balakrishnan et al. (1990) conducts similar research using firm-specific estimates of GNP growth based on geographic exposure, proxied by sales and earnings, and compares the predictability to a "random walk". The results support the fact that geographic segment data in combination with macroeconomic forecasts seem to enhance the predictive ability. An issue with both Roberts' (1989) and Balakrishnan et al.'s (1990) research, apart from the limited sample and time period, is that the comparisons do not take the mean reversion process of profitability into account, as shown by e.g. Freeman et al. (1982) and presented in section 2.2.

Later research by Li et al. (2014) develop these ideas by using a very large sample and a longer time period, with a total of 198,315 firm-year observations. They apply a multivariate model that acknowledges the mean-reverting property of profitability and other characteristics that is commonly known for explaining persistence in profitability (the model is inspired by e.g. Fama & French, 1995; Fama & French, 2006; Hou et al., 2012; So, 2013). In other words, they apply the idea of the earlier studies by Roberts (1989) and Balakrishnan et al. (1990) together with later findings regarding profitability characteristics. They construct a "MACRO" variable by combining geographic segment sales disclosure, for each firm and year, with country-level predictions of real GDP growth. The variable is found to improve forecasts of firm profitability. More specifically, they find that a one percentage point increase in real GDP growth expectations translates, on average, into an increase of 27 basis points in one-year-ahead RNOA. They also incorporate a variable corresponding to analyst forecasts of one-year-ahead RNOA with the purpose of investigating whether analysts take GDP growth expectations into account. For the full sample, the significance of the MACRO variable remains but decreases slightly when adding the analyst forecasts variable, indicating that analysts incorporate some of the expectations of real GDP growth but far from all. The authors also estimate a "naïve" version of their model. They create a non-domestic sample including only firms

[1] Gross National Product includes the market value of all goods and services produced by the citizens of a country, both abroad and domestically.

with over 50 percent in non-domestic sales and use a GDP growth estimate based on a weighted average across all non-domestic countries. This naïve version ignores how firms' non-domestic sales are actually distributed across countries and does therefore not require the same detailed use of geographic segments data. Interestingly, however, these predictions are found to have no explanatory value. The authors therefore conclude that the use of detailed geographic segment data is crucial for increasing the predictability of future RNOA. In unreported tests, Li et al. (2014) also compare the out-of-sample prediction accuracy of the estimated model with and without the MACRO variable. The comparison is made using absolute forecasts errors on an expanding window basis and they find that with the inclusion of the MACRO variable, the average and the median absolute forecast errors of RNOA is 2 basis points lower, and statistically significant.

Building on the research by Li et al. (2014), Doukakis et al. (2020) investigate whether the same MACRO variable improves the predictability even in times of economic crisis (when macroeconomic conditions change in a significant way). They confirm the findings of Li et al. (2014), that the MACRO variable significantly improves RNOA predictability in estimation. However, the variable is found to have no explanatory value in times of crisis. The findings are not validated out-of-sample.

2.4. Contribution

This study extends the research on how forecasts of macroeconomic factors can be used together with financial statements data to predict future profitability, a research area which is still relatively unexplored. Similar to Li et al. (2014), this study examines whether firm-level country exposures, determined based on geographic segment sales data, and forecasts of country-level GDP growth can be used jointly to improve predictions of profitability. In contrast to the study by Li et al. (2014), this thesis presents a more comprehensive validation of the estimated models in out-of-sample forecasts. Li et al. (2014) only briefly mention absolute forecast errors from unreported tests to describe the predictive ability of MACRO out-of-sample. This thesis provides detailed reported tests and additional measures of prediction accuracy in the form of squared forecast errors as well as the proportion of correctly predicted increases and decreases in RNOA, similar to a logit analysis. The out-of-sample validation is crucial for the practical usefulness of the prediction models. Given that the MACRO variable not only adds explanatory value but also improves out-of-sample predictions, promising trading strategies could be developed.

In addition, the research design of this thesis enables an isolation of the incremental predictive content of the MACRO variable over past RNOA. Three different prediction models for one-year-ahead RNOA are constructed. The first model contains past RNOA as the single independent variable, inspired by the high forecasting power of past profitability, as presented in the literature review. The second model adds the MACRO variable from Li et al. (2014), isolating the predictive content of the MACRO variable. The third model is more comprehensive and similar to the original model used by Li et al. (2014), enabling an out-of-sample validation of the additional variables' predictive content.

Furthermore, the out-of-sample prediction accuracy of the models is compared against that of analyst forecasts, something which has not been done in the aforementioned studies of the MACRO variable. This thesis also focus on a more condensed set of firms for which the macroeconomic development is believed to be of particularly high importance, namely Swedish-listed manufacturing firms. Finally, the study examines a later time period than Li et al. (2014) and all of the other studies mentioned above.

3. Hypotheses

This study examines three different hypotheses. Hypothesis A relates to the estimation of prediction models for one-year-ahead RNOA and tests whether the MACRO variable used in Li et al. (2014) and Doukakis et al. (2020) is useful for such predictions. Hypotheses B and C relates to the out-of-sample validation of the estimated models. Hypothesis B is divided into two and tests the out-of-sample prediction accuracy of the estimated models against each other. Hypothesis C compares the prediction accuracy to that of analyst forecasts. The hypotheses are explained in the sections below, and the statistical tests of the hypotheses are outlined in section 4.4.

3.1. Hypothesis A: Relationship between MACRO and future RNOA

As a first step, this study seeks to confirm the positive relationship between MACRO and one-year-ahead RNOA that has been documented by Li et al. (2014) and Doukakis et al. (2020). The relationship is tested by including the variable in the estimation of two prediction models of RNOA: SIMPLE_MACRO and ADVANCED_MACRO. Given the findings in prior research, the first hypothesis is one-sided and formulated as follows:

Hypothesis A

H_0: The MACRO variable does not have a significant positive association with one-year-ahead RNOA.

H_1: The MACRO variable has a significant positive association with one-year-ahead RNOA.

3.2. Hypotheses B: Comparison of prediction accuracy across models

Given that the estimated prediction models show that MACRO is significantly associated with one-year-ahead RNOA, it is also reasonable to expect that a prediction model which includes MACRO as an independent variable should produce more accurate forecasts of RNOA than a model with otherwise identical independent variables but which does not include MACRO. This hypothesis is tested by comparing the out-of-sample prediction accuracy of SIMPLE_MACRO and SIMPLE.[2] The latter model only contains past RNOA as independent variable, while the former also includes MACRO. The hypothesis is one-

[2] How "prediction accuracy" is measured is outlined in section 4.3.

sided, meaning that a rejection of the null hypothesis implies that SIMPLE_MACRO can be concluded to have higher prediction accuracy:

Hypothesis B.1

H_0: Prediction accuracy$_{\text{SIMPLE_MACRO}}$ ≤ Prediction accuracy$_{\text{SIMPLE}}$

H_1: Prediction accuracy$_{\text{SIMPLE_MACRO}}$ > Prediction accuracy$_{\text{SIMPLE}}$

Since SIMPLE_MACRO is still a simplistic model, one should expect prediction accuracy to improve even further by taking additional information into account that is not captured by MACRO or past RNOA. However, prior research suggest that measures of past profitability, such as RNOA and ROE, have very high forecasting power on their own.[3] In Skogsvik (2002a), the inclusion of additional accounting ratios as independent variables actually deteriorated the out-of-sample prediction accuracy. To examine the potential forecast improvements, or deteriorations, from using a more advanced model over a simplistic one, the SIMPLE_MACRO model will also be compared to the more comprehensive ADVANCED_MACRO. The variables included in this model are based on the variables used by Li et al. (2014). It primarily includes variables inspired by finance literature rather than only accounting ratios that were used by Skogsvik (2002a). Since the included variables have demonstrated an ability to predict RNOA in earlier studies the ADVANCED_MACRO model is expected to improve prediction accuracy relative to SIMPLE_MACRO. Thus, hypothesis B.2 is also one-sided:

Hypothesis B.2

H_0: Prediction accuracy$_{\text{ADVANCED_MACRO}}$ ≤ Prediction accuracy$_{\text{SIMPLE_MACRO}}$

H_1: Prediction accuracy$_{\text{ADVANCED_MACRO}}$ > Prediction accuracy$_{\text{SIMPLE_MACRO}}$

3.3. Hypothesis C: Comparison of prediction accuracy to analyst forecasts

Hypotheses B.1 and B.2 only concern the predictive ability of the models relative to each other. However, it is also of interest to examine whether the predictions are good or bad compared to some alternative benchmark. One such benchmark is analyst forecasts, which is often viewed as a proxy for financial market expectations. As a robustness check, a comparison is therefore made against analyst consensus expectations of one-year-ahead RNOA. The model which is found to have the highest prediction accuracy based on the testing of hypotheses B.1-B.2 is used for this comparison. The hypothesis is two-sided and tested out-of-sample:

[3] See e.g., Freeman et al., 1982; Fairfield et al., 1996; Skogsvik, 2002a; Soliman; 2008.

Hypothesis C

H_0: Prediction accuracy$_{\text{Best model}}$ = Prediction accuracy$_{\text{Analyst forecasts}}$

H_1: Prediction accuracy$_{\text{Best model}}$ $\neq$ Prediction accuracy$_{\text{Analyst forecasts}}$

4. Research Design

This section presents the methodology of the study. Section 4.1 describes the MACRO variable in detail. Section 4.2 describes the estimation of the prediction models and presents a timeline of the predictions. Section 4.3 describes the out-of-sample validation and section 4.4 presents the statistical tests that are used to evaluate the hypotheses.

4.1. The MACRO variable

The intuition behind the MACRO variable is straightforward. The future profitability of a firm should be related to the future economic performance of the countries that the firm is exposed to. Highlighting these cross-sectional and time-varying differences should be informative for future profitability and, in turn, improve the firm-specific predictability. In line with Li et al. (2014) and Doukakis et al. (2020), the MACRO factor proxies firm-level country exposures by using geographic segment sales data from the latest annual reports combined with country-specific forecasts of real GDP growth.

4.1.1. Calculating the MACRO variable

When calculating the MACRO variable, the first step is to manually code all segments within each firm-year with the corresponding individual countries (see next section, 4.1.2). After manually coding the specific countries for each segment, each country is linked to the corresponding real GDP growth forecast from IMF World Economic Outlook ("IMF") pertaining to the same year for which RNOA is predicted. If a segment consists of more than one country, each country-specific GDP forecast is value-weighted within that segment. The weights are determined, in line with Roberts (1989) and Li et al. (2014), based on actual GDP (in USD) from the previous year for each country. The value-weighted values are summed to a segment-specific MACRO variable. The segment-specific variable is then value-weighted once more with regards to the segment's share of total sales. Finally, the segment-weighted values are summed up to arrive at the firm-year-specific MACRO variable, which is later included in the predictions of RNOA. See Table 1 below for an overview of the calculations.

Table 1. Illustrative example of the calculation of MACRO

(A)	Segments	Segment A	Segment B	Segment C
(B)	Segment % of total sales	50%	30%	20%
(C)	Countries manually coded to segment	Country A Country B Country C	Country D Country E	Country F
(D)	Real GDP growth forecast per country	1.50% 2.00% -1.00%	3.00% 0.50%	2.00%
(E)	Within-segment country-weights based on actual GDP in the year before	40% 30% 30%	20% 80%	100%
(F) = (D)*(E)	Within-segment country-weighted real GDP growth forecast by country	1.5%*40%=0.6% 2.0%*30%=0.6% -1.0%*30%=-0.3%	3.0%*20%=0.6% 0.5%*80%=0.4%	2.0%*100%=2.0%
(G) = sum of (F)	Segment-specific MACRO variable	0.90%	1.00%	2.00%
(H) = (G)*(B)	Segment-specific MACRO variable value-weighted with regards to segment % of total sales	0.9%*50%=0.45%	1.0%*30%=0.30%	2.0%*20%=0.4%
(I) = sum of (H)	Firm-year MACRO variable		1.15%	

Note: Table 1 presents an overview of the calculations of the MACRO variable. The variable is calculated for a fictive firm-year with three segments.

If a country that have been coded to a segment lacks a GDP growth forecast from IMF, the country is excluded from the calculation of the MACRO variable, i.e. it does not receive any weight within the segment. This treatment, indirectly, increases the weights of the other countries in the segment for which GDP growth forecasts are available. If an entire segment (including one or several countries) lacks country-specific GDP forecasts from IMF for a certain year, that segment's share of total sales is instead multiplied with a world GDP growth forecast retrieved from IMF.

4.1.2. Manual coding of country exposures

As mentioned, the first step in creating the MACRO variable is to determine a firm's country exposures by manually coding specific countries to each of the geographic sales segments presented in the most recent annual report. To ensure consistency between firm-years, a number of rules and definitions are developed. First, all individual countries are sorted into a continent according to definitions from the UN (UN, n.d.). This, as an example, defines Russia as a European country and Turkey as an Asian country, removing the potential mistake of treating countries differently in different firm-years. Second, all countries are sorted into latitudes within each continent, where possible, in accordance with the UN definitions (UN, n.d.). As an example, Europe is divided into four latitudes (north, south, east and west) and Africa into northern and sub-Saharan. Third, every group of countries (not classifying as a continent) that is faced for the first time when coding gets a fixed definition. These constellations include for example "Mediterranean", "EU", "Asia-Pacific", "Middle East" etc. and are used for all firm-years with such segment names. See Table A2 and Table A3 in Appendix for a full disclosure of the definitions and defined constellations.

Given that the above has been defined, there are still some challenges with the manual coding. Particularly the fact that a specific firm-year could include individual countries, continents or constellations that overlap. To ensure consistency, the segments are ranked and coded in accordance with the rules in Table 2 below.

Table 2. Prioritization rules for manual coding of country exposures

1.	Individual countries Segments that consist of individual countries are coded first. This could for example be separate segments named "Sweden" and "Germany" but also segments named "Sweden and Germany".
2.	Constellations Segments that correspond to a defined constellation are coded next, excluding any individual countries already coded in the first step. There are instances where a firm-year consists of multiple, overlapping, constellations. In those cases, the following rules are applied: a) The constellation that does not make another constellation empty is prioritized, e.g., Scandinavia is prioritized over the Nordics. b) If a) cannot be applied, the priority is instead chosen based on the split that create the most even number of countries between the constellations, e.g., the Nordics is prioritized over the EU.
3.	Part of a continent Segments that correspond to parts of a continent based on the UN definitions are coded next, excluding any countries coded in the first and second steps.
4.	Continents Segments that correspond to an entire continent based on the UN definitions are coded next, excluding any countries coded in the first three steps.
5.	Rest of the world Segments that correspond to the rest of the world are coded last, excluding any countries coded in the first four steps.

Note: Table 2 presents an overview of the prioritization rules for the manual coding of country exposures.

There are a few exceptions to the ranking above. If prioritizing a constellation in a firm-year would result in that another segment, which corresponds to part of a continent or an entire continent, becomes empty, then this second segment is instead prioritized over the other segment that corresponds to a constellation. To illustrate, if a firm-year includes segments named “South America” (a continent) and “Latin America” (a constellation that includes all countries in South America and some other countries), then “South America” is prioritized even though it is a continent. If “South America” is not coded before “Latin America”, there would be no countries left to include in the segment “South America”. In other words, step 3 or 4 in the ranking above goes before step 2.

There is also an exception in the treatment of European segments which is worth highlighting. There are two European constellations that together include all countries in Europe (“Central & Eastern Europe” and “Western Europe”). Firms do not typically refer to the actual latitudes of Europe when they write “Western” or “Eastern”, as the UN definitions assume. Instead, the classifications are commonly used to divide Europe into two groups of countries with specific economic, political and cultural similarities.[4] If a firm-year consists of a Western and/or Eastern Europe segment, the constellations “Western Europe” and “Central & Eastern Europe” are prioritized, in line with step 2 in Table 2. However, if a firm-year consists of these two segments *and* additional European segments with other latitudes (north and/or south Europe), the UN definitions of north and/or south Europe are prioritized over the “Western Europe” and “Central & Eastern Europe” constellations. In other words, step 3 in Table 2 goes before step 2. This, in turn, reduces the number of countries that are coded to the Western and/or Eastern Europe segment. See Table A4 in Appendix for a detailed explanation of these rules.

Apart from the exceptions described above, many additional assumptions are made to ensure consistency. These additional assumptions mainly relate to differences in segment names between firms, such as “Pacific” instead of “Oceania” or “Holland” instead of the “Netherlands”, and situations where firms report more than one segment with the exact same meaning. For a full disclosure of these additional assumptions and solutions, see Table A5 in Appendix.

To better illustrate the manual coding and the ranking of different segments, the manual coding of Atlas Copco’s country exposures in 2018 is presented in Table 3 below.

[4] For example, the Nordic countries are considered to be part of “Western Europe” while the former Yugoslavian states are part of “Central & Eastern Europe”. According to UN’s classifications, which are based on the geographical latitudes within Europe, the Nordic countries are classified as “Northern” while the former Yugoslavian states are classified as “Southern”.

Table 3. Example of manual coding of country exposures: Atlas Copco in 2018

Segment name	Priority	Comment
Sweden	1	Individual country, coded first
United States	1	Individual country, coded first
China	1	Individual country, coded first
Other Europe	4	Other Europe represents a continent and has fourth priority. All European countries are coded to this segment, excluding those with a separate segment (Sweden, Germany, France, United Kingdom, Italy, Russia, and Belgium)
Germany	1	Individual country, coded first
Other Africa/Middle East	4/2	Middle East is a defined constellation and has second priority. All African countries are also coded to this segment, except those also included in Middle East, since they have already been coded to the segment, and except South Africa, since it is represented by an individual segment.
South Korea	1	Individual country, coded first
Other Asia/Australia[1]	4/4	Other Asia/Australia represents two continents and have fourth priority. All Asian and Oceanic countries are coded to this segment, excluding those with a separate segment (China, South Korea, India, Japan, Australia and all Asian countries that are included in Middle East).
France	1	Individual country, coded first
India	1	Individual country, coded first
Japan	1	Individual country, coded first
United Kingdom	1	Individual country, coded first
Italy	1	Individual country, coded first
Brazil	1	Individual country, coded first
Other North America	4	Other North America represents a continent and has fourth priority. All North American countries are coded to this segment, excluding those with a separate segment (United States and Canada).
Russia	1	Individual country, coded first
Canada	1	Individual country, coded first
Belgium	1	Individual country, coded first
Australia	1	Individual country, coded first
Other South America	4	Other South America represents a continent and has fourth priority. All South American countries are coded to this segment, excluding those with a separate segment (Brazil and Chile).
South Africa	1	Individual country, coded first
Chile	1	Individual country, coded first

Note: Table 3 presents an example of the manual coding for Atlas Copco in 2018. The priorities relate to the prioritization rules in Table 2. Continents are defined in Table A2 in Appendix and constellations are defined in Table A3 in Appendix.
[1] "Other Australia" is assumed to equal Oceania according to one of the defined rules, see Table A5 in Appendix.

4.2. Estimation of prediction models

Three different prediction models are estimated in this study. The models predict one-year-ahead RNOA and are validated in out-of-sample predictions. There are several reasons why this study focus on predicting RNOA instead of ROE. First, changes in GDP growth are expected to have a more direct and tangible effect on a firm's operations than on its financing or cost of financing. In other words, the effects are expected to be more concentrated to operating income and RNOA rather than earnings and ROE. Second, the country exposures in the MACRO variable are proxied using segment *sales*, which is closely related to a firm's operations. Third, the choice of RNOA is made to increase the comparability with previous studies of the MACRO variable, since both Li et al. (2014) and Doukakis et al. (2020) use RNOA as the dependent variable.

The sample is divided into five estimation periods of 15 years each which is followed by an out-of-sample prediction in the following year. See Table 4. The study applies a rolling scheme in order to control for possible prediction model instabilities over time.

Table 4. The estimation and validation periods

Estimation period	Validation period (out-of-sample)
2000-2014	2015
2001-2015	2016
2002-2016	2017
2003-2017	2018
2004-2018	2019

Note: Table 4 presents an overview of all five estimation periods and their corresponding out-of-sample validation period.

The first prediction model recognizes only the mean reversion process of RNOA, as observed by for example Nissim & Penman (2001). RNOA is calculated as operating income divided by average net operating assets. It is expected that a strong relationship between $RNOA_{t+1}$ and $RNOA_t$ exists. In all of the prediction models presented below, the subscript "i" (for each firm) is not disclosed for the case of brevity.

Prediction model 1: SIMPLE

$$RNOA_{t+1} = \alpha + \beta_1 RNOA_t + e_{t+1} \tag{4.1}$$

The second prediction model adds the MACRO variable to the mean reversion model. Given the findings by Li et al. (2014) and Doukakis et al. (2020), it is expected that a significant positive relationship between $MACRO_t$ and $RNOA_{t+1}$ exists.

Prediction model 2: SIMPLE_MACRO

$$RNOA_{t+1} = \alpha + \beta_1 RNOA_t + \beta_2 MACRO_t + e_{t+1} \tag{4.2}$$

The third prediction model is based on Li et al. (2014), and takes the mean reversion of RNOA and the MACRO variable into account as well as other independent variables that have shown evidence of isolating persistence in profitability.

Prediction model 3: ADVANCED_MACRO

$$RNOA_{t+1} = \alpha + \beta_1 RNOA_t + \beta_2 MACRO_t + \beta_3 BTM_t + \beta_4 Size_t + \beta_5 DNOA_t + \beta_6 D\_Loss_t + \beta_7 D\_Div_t + \beta_8 Div\_Yield_t + e_{t+1} \tag{4.3}$$

"BTM_t" is the book value of common equity divided by market capitalization of equity and "$Size_t$" is the natural logarithm of the market capitalization of equity (calculated in USD to ensure comparability). The variable "D_Loss_t" is a dummy variable that takes the value of 1 if the firm reported negative earnings in year *t* and 0 otherwise, and "D_Div_t" is also a dummy variable that takes the value of 1 if the firm paid a dividend in year *t* and 0 otherwise. "Div_Yield_t" is the dividend yield for year *t*, calculated by taking the dividend payment per share in year *t* divided by the year-end price per share in year *t-1*, and "$DNOA_t$" represents the change in net operating assets between the end of year *t* and *t-1*, scaled by total assets. For a comprehensive overview of the calculations of these independent variables, see Table A6 in Appendix.

One independent variable is excluded from ADVANCED_MACRO compared to the model used by Li et al. (2014), namely a variable representing analyst forecasts of $RNOA_{t+1}$. Instead of including analyst forecasts directly in the prediction models, the analyst forecasts are used as a robustness check as per hypothesis C. The prediction accuracy of the analysts is compared to that of the best model, as will be explained further in sections 4.3-4.4.

In all prediction models, it is expected that the β_1 coefficient will be less than one but non-negative, which is consistent with the mean reversion process. For the β_2 coefficient a positive relationship is expected, which is consistent with forecasts of higher future GDP resulting in predictions of higher $RNOA_{t+1}$, after controlling for current firm profitability (controlled for through the $RNOA_t$ variable). A higher β_2 coefficient is expected than in the studies by Li et al. (2014) and Doukakis et al. (2020), due to the relatively greater importance of macroeconomic conditions for the sample examined in this study, namely Swedish-listed manufacturing firms. The β_3 coefficient is expected to have a negative value since a lower "BTM_t" signals greater growth opportunities, and firms with greater growth opportunities are expected to have higher levels of future profitability after controlling for current profitability. In other words, the larger the "BTM_t", while holding current profitability constant, the less profitable the firm is expected to be in the future. Along the same line, the larger the market capitalization ("$Size_t$"), the higher the expectations of future profitability after controlling for current profitability, which translates to an expectation of a positive β_4 coefficient. The β_5 coefficient is expected to be negative due to the low persistence of accruals. For the β_6

coefficient a negative relationship is expected, as firms reporting a loss in terms of earnings should be less profitable. The coefficients β_7 and β_8 are expected to be positive because firms paying dividends should be more profitable.

In contrast to Li et al. (2014), only variables that are significant in estimation are included in the prediction models that are tested out-of-sample. If an independent variable (or several) is insignificant in one of the estimation periods (i.e. not significant at 10 percent significance level), the most insignificant variable is removed one at the time until all independent variables in that estimation period are significant. If an independent variable lacks significance in the estimation period, it cannot be argued to add explanatory value in the validation period.

As can be observed, none of the three models disaggregate $RNOA_t$ into components, although previous research suggest that such disaggregation improves forecasts of future RNOA (Nissim & Penman, 2001; Soliman, 2008). Disaggregating $RNOA_t$ would decrease the comparability of this study to that of Li et al. (2014) and it would also make it more difficult to investigate the mean reversion properties of RNOA.

All models are applied using pooled OLS regressions with clustered standard errors with regards to both firm and year effects. The reason for not adjusting for firm fixed effects and year fixed effects in the estimation period is because of the out-of-sample prediction. Year fixed effects, that control for variables constant between firms but time-varying, would not be useful when estimating models that will be applied out-of-sample. Firm fixed effects, which is the control for variables that are constant through time but varying cross-sectionally, would have been possible to adjust for in the estimation period but that would have limited the sample in the validation period to include only those companies that are included in the estimation sample. Since the aim is to develop a prediction model that can be applied regardless of what firms a period contains, no adjustments for fixed effects are made.

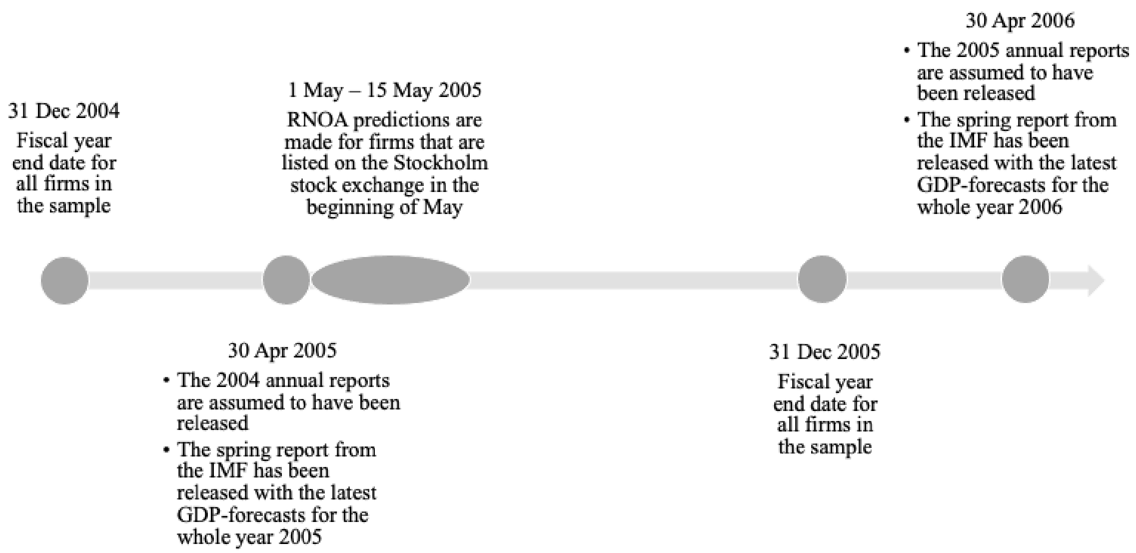
4.2.1. Timeline

To better illustrate when the information that is used becomes public and when the predictions are made, the timeline in Figure 1 has been constructed. The sample only contains firms with fiscal year-end in December. Under Swedish law (*Lag om börs- och clearingverksamhet*, SFS 1992:543) up until 2007, all firms with Sweden as home country that were listed on a stock exchange or an authorized marketplace were required to publish their annual report no later than four months after the fiscal year end. In 2007, this was reformulated (*Lag om värdepappersmarknaden,* SFS 2007:528) and the four-month rule applied to all firms with Sweden as a home country that are listed on a *regulated* stock exchange. For firms with a home country in the European Union that are listed on a Swedish exchange, there is a directive regarding the publication (first mentioned in the 2004 publication; EU, 2004) that states that the annual reports should be published no later than four months after the fiscal year-end. For non-EU members, other regulations apply.

Throughout the sample, there are firms listed on both regulated exchanges (Nasdaq Stockholm), and non-regulated exchanges (Nasdaq First North Growth Market). In addition, there are (albeit very few) firms registered in countries other than Sweden, both in the EU and outside the EU. This means that it cannot, with absolute certainty, be established that all annual reports would have been published by 30th of April for all firm-years in the sample period, but it is deemed very likely. Similarly, Li et al. (2014) also assumes that annual reports are made available four months after the fiscal period end.

The World Economic Outlook is published by the International Monetary Fund (IMF) two times each year, in April and in September/October. This study uses the report published in April each year to make predictions of RNOA for that same year.

Figure 1. Timeline



Note: Figure 1 presents the timeline for when information is made available and when the predictions are made. The time period illustrated in the figure constitutes an example.

As suggested by the timeline, when standing at the beginning of May 2005, the information from the 2004 annual reports (published on April 30, 2005 at the latest) is used together with the IMF forecasts for year 2005 (published in April 2005) when forecasting RNOA for 2005. Predictions are made for all firms whose shares are traded on the exchange at any time during the first 15 calendar days of May 2005. That the firms are tradeable at the prediction date is crucial for being able to apply the models when making trading decisions. In this example, 2004 is viewed as year *t* and 2005 as year *t+1*.

4.3. Out-of-sample validation

As presented in Table 4, the estimated prediction models are tested out-of-sample in the year following directly after each estimation period. The prediction accuracy of the models are evaluated through the tests of hypotheses B.1-B.2 and hypothesis C. Two different measures of prediction accuracy are used: 1) the forecast errors of the predictions, and 2) the proportion of correctly predicted changes in RNOA in terms of direction (i.e. whether RNOA increases or decreases between year *t* and *t+1*). The two measures are explained further below and the statistical tests are presented in section 4.4.

4.3.1. Forecast errors

The out-of-sample forecast error for each model is calculated as the difference between the actual outcome of $RNOA_{t+1}$ and predicted $RNOA_{t+1}$. The forecast error for analyst forecasts is calculated similarly, as the difference between actual $RNOA_{t+1}$ and analyst expectations of $RNOA_{t+1}$.[5] A forecast error close to zero indicates high predictive ability. Depending on whether actual $RNOA_{t+1}$ is higher or lower than predicted, the forecast error is either positive or negative. To avoid that positive and negative forecast errors balance out when comparing the models, the comparisons are based on the absolute forecast errors and squared forecast errors. By squaring the forecast errors, larger errors are assigned higher weights which is consistent with the general view that larger forecast errors are worse than smaller ones (Foster, 1986). The Wilcoxon signed rank test is used to test whether the differences in forecast errors are significant or not, which is explained further in section 4.4.2.

4.3.2. Proportion of correctly predicted increases and decreases in RNOA

To further evaluate the prediction accuracy of the models, the point estimates of $RNOA_{t+1}$ generated by each model are converted to a binary prediction, i.e. an expected increase or decrease in relation to $RNOA_t$. The binary predictions are then compared to the actual binary outcome, i.e. whether actual $RNOA_{t+1}$ is higher or lower than actual $RNOA_t$. This determines to what extent the estimated models, much like a logit regression, estimates the direction of change (an increase or decrease) correctly. The same method is applied to the analyst forecasts, which are then compared to the prediction model with the highest prediction accuracy, as outlined in hypothesis C. To test whether there are significant differences between the correct number of predicted increases and decreases across models and compared to analyst forecasts, a proportion z-test is applied, see section 4.4.3.

[5] Analyst expectations of $RNOA_{t+1}$ are computed based on analyst consensus estimates of EBIT. For a detailed description, see Table A6 in Appendix.

4.4. Hypothesis testing

This section describes how the hypotheses presented in section 3 are statistically tested. Hypothesis A, which relates to the association between $MACRO_t$ and $RNOA_{t+1}$ in estimation, is examined through the regression results. Hypotheses B-C, which concern the out-of-sample prediction accuracy across models and the comparison to analyst forecasts, are examined using two different tests: Wilcoxon signed rank tests for the differences in forecast errors and proportion tests for the differences in correctly predicted increases and decreases in RNOA.

4.4.1. Hypothesis A: MACRO coefficient in estimation

The regression results from the estimated prediction models SIMPLE_MACRO and ADVANCED_MACRO are used to test hypothesis A. Specifically, the estimated $MACRO_t$ coefficient, β_2, must be positive and statistically significant at the 10% level in order to reject the null hypothesis for hypothesis A. Thus, the hypothesis is reformulated as follows:

Statistical test for hypothesis A: Estimated $\boldsymbol{\beta_2}$ coefficient

H_0: $\beta_2 > 0$ with $p < 0.10$ are not both satisfied

H_1: $\beta_2 > 0$ with $p < 0.10$ are both satisfied

4.4.2. Hypotheses B-C: Wilcoxon signed rank test

The nonparametric Wilcoxon signed rank test is one of the two statistical tests that are used to evaluate hypotheses B-C, in addition to the proportion test which is explained in the next section. For hypotheses B.1-B.2, the Wilcoxon signed rank test is used to test whether the differences in forecast errors across models are statistically significant. The test is performed on the paired differences in absolute and squared forecast errors across models. For each observation (firm-year) in the validation period, the absolute forecast error of one model is subtracted from the absolute forecast error of the other model for the same observation. The same applies to squared forecast errors. The Wilcoxon signed rank test then tests whether the median of the paired differences is significantly different from zero. For hypothesis C, the same tests are performed on the paired differences in absolute and squared forecast errors between the best prediction model and analyst forecasts. All tests are performed on the paired differences across all validation periods collectively, similar to Fairfield et al. (1996).

Compared to a paired t-test, the Wilcoxon signed rank test does not require that the paired differences on which the test is performed are normally distributed. Using several common methods for assessing normality, it is concluded that the paired differences in this study cannot be assumed to follow a normal distribution.[6] The Wilcoxon signed rank test is therefore more appropriate than the paired t-test. In addition, the Wilcoxon signed rank test compares the median differences whereas the paired t-test compares the mean differences. When it comes to forecast errors, a comparison based on medians is viewed as more representative of the population, since medians are not affected by extreme values in the same way as means. The Wilcoxon signed rank test is used in several other prediction studies to compare forecast errors across models (e.g., Fairfield et al., 1996; Esplin et al., 2014). It is not clear what type of significance test that Li et al. (2014) use in their study, since the forecast errors are only mentioned very briefly and no tests are reported.

To test hypothesis B.1 with respect to forecast errors, the hypothesis is reformulated as per below. According to the hypothesis, SIMPLE is expected to have higher absolute and squared forecast errors than SIMPLE_MACRO:

Statistical test for hypothesis B.1: Wilcoxon signed rank test

H_0: Median of $(FE_{SIMPLE} - FE_{SIMPLE\_MACRO}) \leq 0$

H_1: Median of $(FE_{SIMPLE} - FE_{SIMPLE\_MACRO}) > 0$

where "FE" refers to both absolute and squared forecast errors.

To test hypothesis B.2 with respect to forecast errors, the hypothesis is reformulated as per below. According to the hypothesis, SIMPLE_MACRO is expected to have higher absolute and squared forecast errors than ADVANCED _MACRO:

Statistical test for hypothesis B.2: Wilcoxon signed rank test

H_0: Median of $(FE_{SIMPLE\_MACRO} - FE_{ADVANCED\_MACRO}) \leq 0$

H_1: Median of $(FE_{SIMPLE\_MACRO} - FE_{ADVANCED\_MACRO}) > 0$

where "FE" refers to both absolute and squared forecast errors.

To test hypothesis C with respect to forecast errors, the hypothesis is reformulated as per below. As mentioned in section 3, the hypothesis regards the robustness check against analyst forecasts and it is two-sided:

[6] Skewness-Kurtosis tests and Shapiro Wilk W tests are performed and histograms are studied visually. See Tables A7-A8 and Figures A1-A3 in Appendix.

Statistical test for hypothesis C: Wilcoxon signed rank test

H_0: Median of ($FE_{\text{Best model}} - FE_{\text{Analyst forecasts}}$) = 0

H_1: Median of ($FE_{\text{Best model}} - FE_{\text{Analyst forecasts}}$) ≠ 0

where "FE" refers to both absolute and squared forecast errors.

While the Wilcoxon signed rank test does not require a normally distributed population, it does require that the population has a symmetric distribution. The test is based on ranks[7], which also means that it is less sensitive to extreme values. Since the number of observations in the validation period is sufficiently large[8], the Wilcoxon test statistic *T* is assumed to be approximately normally distributed:

$$z = \frac{T - \mu_T}{\sigma_T} \ approx \sim N(0,1)$$

where $\mu_T = \frac{n(n+1)}{4}$ and $\sigma_T^2 = \frac{n(n+1)(2n+1)}{24}$

The z-score generated by the tests are converted to a one-tail or two-tailed p-value depending on the formulated hypothesis.

4.4.3. Hypotheses B-C: Proportion test

To determine if there are significant differences between the proportions of correctly predicted increases and decreases in RNOA, both across models and compared to analyst forecasts, a two-sample proportion z-test is applied. Similar to the Wilcoxon signed rank test, the two-sample proportion z-test uses a normally distributed test statistic, calculated as:

$$z = \frac{\hat{p}_1 - \hat{p}_2}{S_{d0}} \ approx \sim N(0,1)$$

where $\hat{p}_1$ and $\hat{p}_2$ are the observed proportion of correctly predicted increases and decreases in the two different samples and S_{d0} is the standard error of the difference in $\hat{p}_1$ and $\hat{p}_2$ calculated as:

[7] The paired differences are ranked in ascending order based on the absolute values of the differences. Each value gets a rank between 1 and *n*, where *n* is the total amount of pairs with non-zero differences. The ranks for positive and negative differences are summarized separately and the lower of the two sums is used as the test statistic *T* (Newbold et al., 2010).

[8] STATA, which is the statistical program used for this study, assumes normal approximation for samples where $n > 200$. In this study, $n = 728$ for the Wilcoxon tests of hypotheses B and $n = 495$ for the tests of hypothesis C (see Data section for further details).

$$S_{d0} = \sqrt{\frac{x_1 + x_2}{n_1 + n_2} * \left(1 - \frac{x_1 + x_2}{n_1 + n_2}\right) * \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

where x_1 and x_2 are the total number of correct predictions in the two samples respectively, and where n_1 and n_2 are the total number of observations in the two samples respectively. The z-score is converted to a one-tail or two-tailed p-value depending on the formulated hypothesis.

To test hypothesis B.1 with respect to the proportion of correctly predicted increases and decreases in RNOA, the hypothesis is reformulated as per below. According to the hypothesis, SIMPLE_MACRO is expected to have a higher proportion of correctly predicted increases and decreases compared to SIMPLE.

Statistical test for hypothesis B.1: Proportion test

H_0: $\hat{p}_{\text{SIMPLE\_MACRO}} \leq \hat{p}_{\text{SIMPLE}}$

H_1: $\hat{p}_{\text{SIMPLE\_MACRO}} > \hat{p}_{\text{SIMPLE}}$

To test hypothesis B.2 with respect to the proportion of correctly predicted increases and decreases in RNOA, the hypothesis is reformulated as per below. According to the hypothesis, ADVANCED_MACRO is expected to have a higher proportion of correctly predicted increases and decreases compared to SIMPLE_MACRO.

Statistical test for hypothesis B.2: Proportion test

H_0: $\hat{p}_{\text{ADVANCED\_MACRO}} \leq \hat{p}_{\text{SIMPLE\_MACRO}}$

H_1: $\hat{p}_{\text{ADVANCED\_MACRO}} > \hat{p}_{\text{SIMPLE\_MACRO}}$

To test hypothesis C with respect to the proportion of correctly predicted increases and decreases in RNOA, the hypothesis is reformulated as per below. As mentioned in section 3, the hypothesis regards the robustness check against analyst forecasts and it is two-sided:

Statistical test for hypothesis C: Proportion test

H_0: $\hat{p}_{\text{Best model}} = \hat{p}_{\text{Analyst forecasts}}$

H_1: $\hat{p}_{\text{Best model}} \neq \hat{p}_{\text{Analyst forecasts}}$

5. Data

5.1. Industry focus and time period

The sample covers manufacturing firms listed on the Stockholm stock exchange during the time period 2000-2019. An industry restriction is chosen in order to increase the homogeneity of the sample, since the level of accounting ratios typically vary between different industries (Gombola & Ketz, 1983). Including too many industries could weaken the predictive ability of the proposed models, as the level of accounting ratios would be less homogenous and the sample would include a wider range of business models and potentially industry-specific accounting techniques. The reason why the sample is restricted to manufacturing firms specifically is because such firms typically have large global footprints and thereby exposure to different macroeconomic environments in different countries. In addition, manufacturing firms include mostly industrial firms, for which sales and profitability are relatively dependent on the state of the economy. It is therefore believed that $MACRO_t$ has greater predictive ability for $RNOA_{t+1}$ in manufacturing firms compared to non-manufacturing firms. In addition, several previous studies of profitability prediction have focused on manufacturing firms specifically (e.g., Skogsvik, 2002a).

Swedish manufacturing firms in particular have a very large share of foreign sales, which is the main reason why the study is focused on firms listed on the Stockholm stock exchange.[9] Foreign sales account for more than half of total sales for 67% of all firm-years included in the sample. This can be compared to only 13% of the firm-years in Li et al. (2014), whose sample included firms from a range of different stock exchanges around the world. It would have been desirable to have a larger sample, for example with firms from all Nordic stock exchanges, but it has not been possible due to the very time-consuming process of manually coding the unique country exposures in each firm-year.

Each estimation period is 15 years long. A relatively long estimation period is chosen in order to ensure enough variation in the macroeconomic environment during the period. Sufficient variation in the level of $MACRO_t$ is required in order for it to add any explanatory value in the prediction of $RNOA_{t+1}$. Indeed, $MACRO_t$ shows large time-series variation during 2000-2019, as illustrated in Table 9.

5.2. Data collection and selection criteria

Data about listing status and industry classification of firms are sourced from Compustat, historical financial statements data, geographic segment data and analyst forecasts are sourced from FactSet Fundamentals, and macroeconomic data are sourced from IMF

[9] Both firms listed on Nasdaq Stockholm main list and Nasdaq First North Growth Market are included.

World Economic Outlook. All financial data obtained from FactSet Fundamentals are on a non-restated basis, since potentially restated information is not available at the prediction point in time.

Compustat's Security Daily file is used to identify all firms that were listed on the Stockholm stock exchange at the prediction point in time in each year during 2000-2019. Stock exchange code 256, corresponding to Stockholm stock exchange, is used for this purpose. All firms whose shares were traded on the exchange at any time during the first 15 calendar days of May in a certain year are initially included. SIC codes are sourced from Compustat's Fundamentals Annual file. Firms with SIC codes 2000-3999, which together make up "Division D: Manufacturing" according to the SIC structure, are included in the sample. Table 7 provides a breakdown of the firms-years in the final sample by major SIC group. The industry classification of firms can change over time and therefore historical SIC codes are used. The Fundamentals Annual file returns the SIC code prevailing at the end of each fiscal year for every firm, which in this study always corresponds to the calendar year since only firms with fiscal year-end in December are included. Given that the prediction point in time is in the beginning of May, it would have been optimal to find out the SIC code prevailing at exactly this point in time. However, an assumption has been made that a firm's SIC code at the prediction point in time in year *t+1* is equal to the SIC code at the end of year *t*. As illustrated in Table 5, the initial sample over the entire time period 2000-2019, comprising listed firms with SIC codes 2000-3999 and fiscal year-end in December, amounts to 2,933 firm-years and 364 firms.

Geographic segment data are derived from annual report segment disclosures collected by FactSet Fundamentals. Segment data have been collected for years 1999-2018, in order to construct the $MACRO_t$ variable that is used in the prediction models in years 2000-2019. Before choosing FactSet Fundamentals as the source, it was compared with S&P Capital IQ. Based on random sampling of a few firms, it was found that FactSet Fundamentals collected more granular segment information than S&P Capital IQ.[10] However, as illustrated in Table 5, 438 firm-years are dropped because segment data is unavailable in FactSet Fundamentals. Data could be unavailable either because a firm has not reported any segment data, for example if the firm has no sales, or because the data have not been collected by FactSet Fundamentals for some reason.

A control calculation is made to verify the reliability of the segment data obtained from FactSet Fundamentals. The sales per segment in each firm-year are added together and compared with reported total sales from the income statement, which is also obtained from FactSet Fundamentals. The difference between the two should be close to zero. However, relatively large differences are identified for some firm-years, indicating that

[10] A good example is an annual report by NIBE INDUSTRIER AB, where the company provided a table showing sales breakdown by geography as well as additional segment disclosures in a footnote below the table. The additional information had been collected by FactSet Fundamentals but not by S&P Capital IQ.

the segment data for those firm-years are probably incorrect in FactSet Fundamentals and not reflecting the numbers that were actually reported by those companies for those years.[11] To avoid erroneous data in the sample, all observations are dropped where the absolute difference between reported sales from the income statement and the sum of segment sales is larger than 1% of reported sales.[12] This reduces the sample with 171 firm-years.

An additional 86 firm-years are dropped because data for calculating $RNOA_t$ or $RNOA_{t+1}$ are unavailable in FactSet Fundamentals. Finally, 337 firm-years are dropped because $RNOA_t$ is greater than 100% or lower than –100%. This cut-off value is used in order to eliminate extreme values from the sample that would impair the regression results and thereby the prediction models. Similar methods for dealing with extreme values in profitability ratios have been used in prior studies. For example, Fairfield et al. (1996) use the same method and cut-off value for ROE_t in their prediction study.

78% of the 337 observations that are dropped due to the cut-off value in $RNOA_t$ are observations with negative $RNOA_t$, of which the vast majority are observations with negative operating income.[13] 67% of the eliminated observations belong to SIC groups 28 or 38, which primarily includes biotechnology and medical technology firms. Many of these firms are in early development stage and have negative profitability. $MACRO_t$ is expected to add very little value for predicting $RNOA_{t+1}$ in these firms. Thus, applying a cut-off value to $RNOA_t$ is also used with the purpose to increase homogenity across the sample towards more mature firms, for which $MACRO_t$ is believed to be more relevant.

Important to note is that the level of $RNOA_t$ is always known at the prediction point in time. The forecaster can screen out the extreme values and exclude those firms when applying the models out-of-sample. Observations with absolute values of $RNOA_t$ greater than 100% are therefore dropped both in the estimation periods and validation periods. In line with Fairfield et al. (1996), also observations with absolute values of one-year-ahead RNOA ($RNOA_{t+1}$) greater than 100% are dropped from the sample, but only in the estimation periods. Such observations are not dropped in the validation periods because the forecaster cannot screen out extreme future values that are unknown at the prediction point in time. The removal of observations with extreme values of $RNOA_{t+1}$ in estimation is not illustrated in Table 5. The reason is the rolling time periods, which means that certain years, e.g. 2016, are used both for estimation and validation (see Table 6). Because

[11] For some of the firm-years with large differences, manual checks were made in the annual reports to find out why the data could be incorrect in FactSet. For all of them, it was found that the segment reporting had some kind of special structure that would make it challenging for FactSet analysts or algorithms to collect the data, for example the segment reporting could be spread over several pages. Importantly, however, this means that the errors in FactSet Fundamentals should only be caused by missing data. There is no reason to believe that segment data that *have* been collected are incorrect.

[12] Absolute differences equal to or smaller than 1% are kept to account for the fact that some of the differences are caused by rounding and potentially FX effects.

[13] A very small number of the eliminated observations have negative $RNOA_t$ caused by positive operating income and negative average net operating assets for year *t*.

extreme values of $RNOA_{t+1}$ are dropped in estimation periods but not in validation periods, the number of observations that are included from 2016 is different depending on whether the year is used for estimation or validation. These dynamics cannot be illustrated in a simple way in Table 5. However, the number of observations that are dropped due to extreme values of $RNOA_{t+1}$ in estimation are relatively few.

Table 5. Final sample and selection criteria

Selection criteria	Firm-years	Firms
1. Listed on Stockholm stock exchange at the time of prediction		
2. SIC code 2000-3999 (manufacturing firms)		
3. Fiscal year-end in December		
Initial sample	**2,933**	**364**
4. Segment data for year *t* is unavailable in FactSet	-438	-49
5. Absolute difference between total reported sales and sum of segment sales is greater than 1% of total reported sales	-171	-5
6. Data to calculate $RNOA_t$ or $RNOA_{t+1}$ is unavailable in FactSet Fundamentals	-86	-11
7. $RNOA_t$>100% and $RNOA_t$<–100% are excluded	-337	-46
Final sample	**1,901**	**253**

Note: Table 5 shows how the final sample is formed based on a set of selection criteria. The final sample include observations from the period 2000-2019. The sample is divided into different estimation and validation periods as illustrated in Table 6.

In addition to the omitted observations that are illustrated in Table 5, some additional omissions are made within each combination of estimation and validation period. The reason is that the exact sample that is used in each estimation and validation period is governed by the regression results for ADVANCED_MACRO, since this is the model with the largest number of independent variables. If values for any of these variables are missing for any of the observations, then such observations are dropped from the sample so that each of the three models are estimated based on exactly the same sample in each estimation period.[14] Similarly, observations with missing values for any of the variables are also dropped from the corresponding validation period sample. However, as explained in section 4.2, the independent variables included in ADVANCED_MACRO are not the same across the five different periods since only variables that are found to be significant in each estimation period are included. Therefore, omissions are only made for observations where there are missing data for the variables that are actually included in the model in that particular period. For example, if it turns out that BTM_t and $Size_t$ are the only significant variables, besides $RNOA_t$ and $MACRO_t$, when running the regression for ADVANCED_MACRO in the first estimation period, then observations with missing data points for BTM_t and $Size_t$ are omitted. The omissions are made only for this particular estimation and validation period. Observations with missing data points for

[14] I.e. values could be missing for BTM_t, $Size_t$, $DNOA_t$, D_Loss_t, D_Div_t or Div_Yield_t because the data required to calculate these variables are unavailable in FactSet Fundamentals.

$DNOA_t$, D_Loss_t, D_Div_t or Div_Yield_t are not omitted from the sample because the variables were found to be insignificant and they are therefore not included in the estimated ADVANCED_MACRO model for the first period. However, if for example $DNOA_t$ is found to be significant in the second estimation period, then any observations with missing values for $DNOA_t$ are excluded from the second estimation period and the corresponding validation period.

Table 6. Sample for each estimation and validation period

Estimation period	Firm-years	Corresponding validation period (out-of-sample)	Firm-years
2000-2014	1,089	2015	122
2001-2015	1,167	2016	132
2002-2016	1,254	2017	143
2003-2017	1,273	2018	160
2004-2018	1,369	2019	171
		2015-2019	728

Note: Table 6 presents the number of observations included in each estimation and validation period. Compared to the total sample presented in Table 5, additional observations are omitted within each *estimation* period based on two criteria: 1) Observations with $RNOA_{t+1}$>100% and $RNOA_{t+1}$<–100% are dropped, 2) Observations with missing data for any of the independent variables included in the estimated prediction models for each period are dropped from that specific period. Within each *validation* period, additional omissions are made based on the second criteria only and not based on the first (observations cannot be dropped based on the first criteria since the value of $RNOA_{t+1}$ is unknown at the time of prediction in the validation periods).

5.3. Descriptive statistics

Tables 7, 8 and 9 provide descriptive statistics for the entire sample of 1,901 firm-years and 253 firms.[15] Table 7 gives a breakdown of the sample by major SIC group, including the median $RNOA_t$ for each group. The largest groups in terms of number of firm-years are groups 28, 35, 36 and 38, which together account for approximately 56% of the sample. The variation in $RNOA_t$ is fairly large, which is expected. However, the variation is expected to have been larger if also non-manufacturing firms were included.

[15] Note that this sample includes observations with absolute values of one-year-ahead RNOA ($RNOA_{t+1}$) greater than 100%, i.e. observations that are later dropped from the estimation periods (but not from the validation periods). However, the sample does not include observations with absolute values of past RNOA ($RNOA_t$) greater than 100%. See section 5.2 and Tables 5-6 for further details.

Table 7. Breakdown of sample and median $RNOA_t$ by major SIC group

SIC	Name of major SIC group	Firm-years	% Sample	$RNOA_t$ Median
20	Food And Kindred Products	54	2.84	11.11
21	Tobacco Products	19	1.00	47.08
22	Textile Mill Products	18	0.95	-0.80
23	Apparel And Other Finished Products Made From Fabrics And Similar Materials	43	2.26	8.55
24	Lumber And Wood Products, Except Furniture	71	3.73	15.18
25	Furniture And Fixtures	49	2.58	12.99
26	Paper And Allied Products	95	5.00	7.76
27	Printing, Publishing, And Allied Industries	63	3.31	6.26
28	Chemicals And Allied Products	228	11.99	-4.22
29	Petroleum Refining And Related Industries	3	0.16	16.78
30	Rubber And Miscellaneous Plastics Products	60	3.16	12.88
33	Primary Metal Industries	78	4.10	11.22
34	Fabricated Metal Products, Except Machinery And Transportation Equipment	123	6.47	14.05
35	Industrial And Commercial Machinery And Computer Equipment	305	16.04	12.44
36	Electronic And Other Electrical Equipment And Components, Except Computer Equipment	312	16.41	10.54
37	Transportation Equipment	143	7.52	13.35
38	Measuring, Analyzing, And Controlling Instruments; Photographic, Medical And Optical Goods; Watches And Clocks	217	11.42	1.28
39	Miscellaneous Manufacturing Industries	20	1.05	19.82
	Total	1,901	100.00	10.05

Note: Table 7 reports a breakdown of the final sample (all 1,901 firm-years across the entire sample period) by major SIC group (2-digit SIC code) including median $RNOA_t$ for each group (expressed in percentage points). Historical SIC classifications are used, implying that the same firm could appear in more than one group over the sample period if the SIC code for the firm has changed over time.

Table 8 provides summary statistics for all of the independent variables that are tested in the study. As described in section 4.2, not all of these variables will actually be included in the estimated prediction models, since it depends on their significance in estimation.

The average value of $MACRO_t$ across all firm-years is 2.18, expressed in percentage points, which is consistent with most countries experiencing real GDP growth during the time period. The standard deviation is 1.44, which is deemed to be a sufficiently large variation for $MACRO_t$ to help forecast $RNOA_{t+1}$. Except for time-series variation, which relate to a changing macroeconomic environment over time, the variation in $MACRO_t$ is also explained by cross-sectional variation since different firms are exposed to different countries with different GDP growth forecasts.

Average $RNOA_t$, which is also expressed in percentage points, is lower than the median, reflecting that the observations with negative $RNOA_t$ are relatively more extreme than the positive ones. The large variation in $RNOA_t$ is also reflected by the standard deviation of 32.16. BTM_t shows that the average firm in the sample has a book-to-market value of 0.60. The average value of D_Loss_t tells that firms had negative earnings in 31% of the firms-years. The average value of D_Div_t shows that firms paid dividend in 58% of the firm-years and Div_Yield_t shows that the average dividend yield across all firm-years was 2%.

The column in Table 8 showing the number of observations per variable illustrates that there are some missing values for most variables. As discussed in section 5.2, this impacts the final sample that is used in each estimation and validation period.

Table 8. Summary statistics for all independent variables included in the regressions

Independent variable	N	Mean	Std. dev.	Min	P25	Median	P75	Max
$RNOA_t$	1,901	3.38	32.16	-99.16	-4.37	10.05	19.40	99.40
$MACRO_t$	1,901	2.18	1.44	-4.66	1.80	2.44	2.89	5.78
BTM_t	1,863	0.60	0.55	-0.41	0.26	0.44	0.73	4.87
$Size_t$	1,863	5.20	2.30	-1.38	3.50	4.90	6.70	11.74
$DNOA_t$	1,901	0.06	0.21	-2.29	-0.03	0.03	0.11	1.53
D_Loss_t	1,826	0.31	0.46	0.00	0.00	0.00	1.00	1.00
D_Div_t	1,786	0.58	0.49	0.00	0.00	1.00	1.00	1.00
Div_Yield_t	1,775	0.02	0.03	0.00	0.00	0.02	0.03	0.26

Note: Table 8 presents summary statistics for all of the independent variables that are included in the regressions of the prediction models. All of the variables are not included in the final prediction models as this depends on the regression results. In line with Li et al. (2014), $RNOA_t$ and $MACRO_t$ are expressed in percentage points while the other variables are expressed in decimal form (e.g. average $RNOA_t$ is 3.38% and average Div_Yield_t is 2%). Variables starting with "D_" (D_Loss_t and D_Div_t) are dummy variables. Complete variable descriptions are available in Table A6 in Appendix. The summary statistics in the table pertains to the entire sample period (2000-2019). Some of the variables have missing values, which is why the number of observations is not the same for all variables.

Table 9 provides a breakdown of the sample by forecast year. The number of firms per year increases over time during the sample period. One reason is that the number of listed firms have generally increased over time. Another reason is that the number of observations that are dropped due to data unavailability in FactSet Fundamentals is relatively greater in the earlier part of the sample period compared to the later part.

Table 9 also shows the median values of $RNOA_{t+1}$ and $MACRO_t$ by forecast year. Both of the variables show clear variation over time, which generally is promising for the upcoming regressions and predictions. The lowest median value of $MACRO_t$ is –3.51 in 2009, in connection with the global financial crisis, and the highest median value is 3.61 in 2000. Interestingly, the lowest median value of $RNOA_{t+1}$ coincides with that of $MACRO_t$. Median $RNOA_{t+1}$ in forecast year 2009, i.e. the median actual outcome of RNOA in 2009, is only 5.10. This outcome is unknown at the prediction point in time in May 2009, since it is the value that is about to be predicted. Since the –3.51 value of $MACRO_t$ is known, the forecast of RNOA for 2009 should presumably improve if the information contained in $MACRO_t$ is taken into account. Overall, Table 9 indicates that there seem to be a positive time-series correlation between $MACRO_t$ and $RNOA_{t+1}$, where a high value of $MACRO_t$ in a certain year is typically associated with a high value of $RNOA_{t+1}$, and vice versa.

Table 9. Breakdown of sample and median values of $MACRO_t$ and $RNOA_{t+1}$ by forecast year

Forecast year (*t+1*)	Firms	% Sample	$MACRO_t$ Median	$RNOA_{t+1}$ Median
2000	39	2.05	3.61	14.82
2001	40	2.10	2.45	8.63
2002	58	3.05	1.69	6.14
2003	60	3.16	1.70	10.32
2004	61	3.21	2.78	13.00
2005	61	3.21	2.88	14.11
2006	69	3.63	3.14	14.98
2007	66	3.47	3.25	15.23
2008	79	4.16	2.16	9.80
2009	87	4.58	-3.51	5.10
2010	100	5.26	1.96	8.50
2011	90	4.73	3.39	8.17
2012	99	5.21	1.49	7.02
2013	109	5.73	1.42	5.39
2014	118	6.21	2.60	11.16
2015	126	6.63	2.38	11.31
2016	135	7.10	2.37	12.05
2017	150	7.89	2.39	10.25
2018	175	9.21	2.70	7.85
2019	179	9.42	1.78	8.05
2000-2019	1,901	100.00	2.44	10.06

Note: Table 9 presents a breakdown of the final sample (all 1,901 firm-years across the entire sample period) by forecast year including median $MACRO_t$ and median actual $RNOA_{t+1}$ for each year. Both variables are expressed in percentage points. For each forecast year, the value of $MACRO_t$ shown on the same row (which is known at the prediction point in time in that forecast year) could be used to forecast the value of $RNOA_{t+1}$ shown on the same row (which is unknown at the prediction point in time).

6. Results

Regression results for the prediction models are presented in section 6.1, resulting in the estimated models which are thereafter tested out-of-sample. The prediction results out-of-sample are presented in section 6.2.

6.1. Estimating the prediction models

The regression outputs for ADVANCED_MACRO for all five estimation periods are presented in Table 10. For each estimation period, Equation 4.3, as presented in section 4.2 and also at the top of Table 10, is applied and adjusted by removing insignificant variables one at the time. If more than one variable is insignificant, the most insignificant (highest p-value) is removed first. Two variables (D_Loss_t and Div_Yield_t) are not significant in any estimation period, therefore they are not disclosed in Table 10. Depending on the estimation period and which variables that turn out significant, the number of observations varies. The resulting sample for ADVANCED_MACRO in each estimation period thus governs the sample sizes for the regressions of SIMPLE_MACRO and SIMPLE. This is why the regression results for ADVANCED_MACRO are presented first. The regression outputs for SIMPLE_MACRO and SIMPLE for all five estimation periods are presented in Tables 11-12.

As expected, coefficient β_1 is significant for all estimation periods and models. The value is below one but non-negative which is consistent with the mean reversion process. For SIMPLE_MACRO and ADVANCED_MACRO, β_2 is significant and positive for all estimation periods, which is also expected and consistent with higher forecasts of future GDP growth resulting in predictions of higher $RNOA_{t+1}$. For ADVANCED_MACRO, β_3 is negative which is consistent with the expectation that the higher the book-to-market ratio, the less profitable the firm is expected to be in the future, after controlling for current profitability. Also in line with expectations, coefficient β_4 is positive, stemming from increased expectations of higher future profitability. Coefficient β_5 is also in line with expectations with a negative sign due to the low persistence of accruals and coefficient β_7 is positive, as expected, since firms paying dividends should be more profitable. As evident, two of the variables in ADVANCED_MACRO, D_Loss_t and Div_Yield_t, are not significant in any estimation period. The variables other than $RNOA_t$ and $MACRO_t$ that are significant in one or more periods does not show stability over time. This indicates that the relationship between the additional variables and $RNOA_{t+1}$ is not that robust.

For both SIMPLE_MACRO and ADVANCED_MACRO, the $MACRO_t$ variable is significant on, at least, a five percent significance level. Except for the first estimation period for ADVANCED_MACRO, it is significant on a one percent level. This means that hypothesis A is confirmed.

Table 10. Regression results for ADVANCED_MACRO for all estimation periods

$$RNOA_{t+1} = \alpha + \beta_1 RNOA_t + \beta_2 MACRO_t + \beta_3 BTM_t + \beta_4 Size_t + \beta_5 DNOA_t + \beta_6 D\_Loss_t + \beta_7 D\_Div_t + \beta_8 Div\_Yield_t + e_{t+1} \ ^{(A)}$$

Independent Variable	Estimation Period 1	Estimation Period 2	Estimation Period 3	Estimation Period 4	Estimation Period 5
$RNOA_t$	0.714***	0.698***	0.718***	0.698***	0.720***
	(19.6)	(17.3)	(16.8)	(18.52)	(17.52)
$MACRO_t$	0.827**	0.889***	0.883***	1.004***	0.956***
	(2.74)	(3.08)	(3.43)	(3.71)	(3.67)
BTM_t	-3.224***			-1.931**	-1.590*
	(-3.24)			(-2.33)	(-2.02)
$Size_t$		1.142**	1.097**		
		(2.64)	(2.63)		
$DNOA_t$	-9.201*	-9.016**	-8.527**		
	(-2.01)	(-2.60)	(-2.71)		
D_Div_t				5.415**	6.016***
				(2.66)	(3.19)
Intercept	3.937*	-4.629	-4.252	-0.835	-1.927
	(2.13)	(-1.69)	(-1.61)	(-0.35)	(-0.86)
N	1,089	1,167	1,254	1,273	1,369
adj. R-sq	0.577	0.593	0.613	0.614	0.632

Note: Table 10 presents the regression outputs for ADVANCED_MACRO for all five estimation periods after removing insignificant variables, one at the time. Depending on the estimation period and which variables turn out significant, the number of observations in each estimation period varies. $RNOA_t$ and $MACRO_t$ are expressed in percentage form while the other independent variables are expressed in decimal form. To illustrate, in estimation period 1, the coefficient for $MACRO_t$ is 0.827 which indicates that a one percentage point increase in expectation of real GDP growth corresponds to an 82.7 basis points increase of $RNOA_{t+1}$. For the other variables, such as BTM_t, the coefficient in estimation period 1 (-3.224) indicates that a BTM_t increase of 1 translates into a 3.224 percentage point decrease of $RNOA_{t+1}$. Complete variable descriptions are available in Table A6 in Appendix. All regressions are pooled OLS regressions controlling for clustered standard errors at firm and year level. Any insignificant variables are not presented in Table 10 (blanks). For significant variables, the t-statistics are presented within the brackets and the stars represents the significance level corresponding to the following p-values *p<0.10, **p<0.05, ***p<0.01.
$^{(A)}$ Only independent variables that turn out significant are included in the estimated model for ADVANCED_MACRO in each period. Thus, the estimated models are not exactly equal to Equation 4.3.

Table 11. Regression results for SIMPLE_MACRO for all estimation periods

$RNOA_{t+1} = \alpha + \beta_1 RNOA_t + \beta_2 MACRO_t + e_{t+1}$

Independent Variable	Estimation Period 1	Estimation Period 2	Estimation Period 3	Estimation Period 4	Estimation Period 5
$RNOA_t$	0.715***	0.736***	0.754***	0.751***	0.779***
	(19.28)	(18.27)	(18.26)	(23.3)	(20.77)
$MACRO_t$	1.204***	1.165***	1.121***	1.186***	1.098***
	(3.70)	(3.39)	(3.67)	(3.86)	(3.67)
Intercept	0.628	0.297	0.525	0.478	0.004
	(0.57)	(0.28)	(0.53)	(0.49)	(0.00)
N	1,089	1,167	1,254	1,273	1,369
adj. R-sq	0.570	0.583	0.604	0.606	0.624

Note: Table 11 presents the regression outputs for SIMPLE_MACRO for all five estimation periods. The regressions are run on the same samples as ADVANCED_MACRO. All variables are expressed in percentage form. To illustrate, in estimation period 1, the coefficient for $MACRO_t$ is 1.204 which indicates that a one percentage point increase in expectation of real GDP growth corresponds to a 120.4 basis points increase of $RNOA_{t+1}$. Complete variable descriptions are available in Table A6 in Appendix. All regressions are pooled OLS regressions controlling for clustered standard errors at firm and year level. The t-statistics are presented in the brackets and the stars represents the significance level corresponding to the following p-values *p<0.10, **p<0.05, ***p<0.01.

Table 12. Regression results for SIMPLE for all estimation periods

$RNOA_{t+1} = \alpha + \beta_1 RNOA_t + e_{t+1}$

Independent Variable	Estimation Period 1	Estimation Period 2	Estimation Period 3	Estimation Period 4	Estimation Period 5
$RNOA_t$	0.716***	0.736***	0.754***	0.751***	0.779***
	(19.27)	(18.29)	(18.38)	(23.46)	(20.81)
Intercept	3.073**	2.645**	2.829**	2.962**	2.398**
	(2.62)	(2.23)	(2.60)	(2.71)	(2.19)
N	1,089	1,167	1,254	1,273	1,369
adj. R-sq	0.564	0.578	0.599	0.601	0.620

Note: Table 12 presents the regression outputs for SIMPLE for all five estimation periods. The regressions are run on the same samples as ADVANCED_MACRO. $RNOA_t$ is expressed in percentage form. To illustrate, in estimation period 1, the coefficient for $RNOA_t$ is 0.716 which indicates that a one percentage point increase in $RNOA_t$ corresponds to a 71.6 basis points increase of $RNOA_{t+1}$. Complete variable descriptions are available in Table A6 in Appendix. All regressions are pooled OLS regressions controlling for clustered standard errors at firm and year level. The t-statistics are presented in the brackets and the stars represents the significance level corresponding to the following p-values *p<0.10, **p<0.05, ***p<0.01.

6.2. Out-of-sample validation

The three estimated prediction models are validated out-of-sample in the year following directly after each estimation period. As an example, the ADVANCED_MACRO model that is used to predict RNOA in the first validation period (2015) is the following, based on the regression results presented in Table 10:

$$RNOA_{t+1} = 3.937 + 0.714RNOA_t + 0.827MACRO_t - 3.224BTM_t - 9.201DNOA_t$$

where t = 2014

The next section, 6.2.1, presents the out-of-sample prediction accuracy of the models with respect to forecast errors, followed by Wilcoxon signed rank tests in section 6.2.2. Section 6.2.3 describes the out-of-sample prediction accuracy with respect to correctly predicted increases and decreases in RNOA, followed by proportion tests in section 6.2.4.

6.2.1. Forecast errors

Table 13 provides descriptive statistics on forecast errors, absolute forecast errors and squared forecast errors for the three models across all five validation periods (2015-2019). Based on previous profitability prediction studies with out-of-sample validation (e.g., Fairfield et al., 1996), large forecast errors are expected. The statistics in Table 13 confirm this. As illustrated in panel B, the average absolute forecast error for all three models is close to 30, meaning that predicted $RNOA_{t+1}$ is on average almost 30 percentage points off compared to the actual outcome. The standard deviations of approximately 377 are also very large and arise from the presence of some extreme observations.[16] As a comparison, the ROE prediction models tested by Fairfield et al. (1996) shows similar absolute forecast errors of around 27-30 percentage points on average, but with standard deviations of more than 600 percentage points.

The median forecast error for all three models is close to zero, as illustrated in panel A of Table 13. This means that the prediction models seem to overestimate $RNOA_{t+1}$ approximately as many times as they are underestimating it.

As mentioned in the research design section (4.3.1), comparisons across models should primarily be made based on the absolute and squared forecast errors in panel B-C rather than the forecast errors in panel A, since positive and negative forecast errors balance out in panel A. As expected, according to hypothesis B.1, SIMPLE_MACRO has lower absolute and squared forecast error than SIMPLE, as shown by the average and median values in panel B-C, although the differences are small. The standard deviations are also very similar for SIMPLE and SIMPLE_MACRO. In terms of interquartile range,

[16] As mentioned in the data section (5.2), observations with extreme values of $RNOA_{t+1}$ in the validation periods are not excluded from the sample, since these values are unknown at the prediction point in time. The presence of extreme values of $RNOA_{t+1}$ leads to very large forecast errors for some observations.

however, SIMPLE is actually slightly better than SIMPLE_MACRO, but with very small differences.

Opposite to the expectations in hypothesis B.2, ADVANCED_MACRO has higher absolute forecast error than SIMPLE_MACRO, as shown by the average and median values in panel B. In fact, its absolute forecast errors are also higher than those of SIMPLE. For squared forecast errors, ADVANCED_MACRO is better than both of the other models in terms of mean error, but worse in terms of median error. The standard deviation of ADVANCED_MACRO is slightly lower than for the other models for both absolute and squared forecast errors and its interquartile range is also better.

Table 13. Descriptive statistics on forecast errors for all models across all validation periods (2015-2019)

Panel A: Forecast errors			
	SIMPLE	SIMPLE_MACRO	ADVANCED_MACRO
N	728	728	728
Mean	11.03	10.67	10.96
Std. dev.	378.35	378.36	378.05
Median	0.01	-0.35	-0.57
Interquartile range	10.86	10.97	10.91
Panel B: Absolute forecast errors			
	SIMPLE	SIMPLE_MACRO	ADVANCED_MACRO
N	728	728	728
Mean	29.75	29.71	29.92
Std. dev.	377.34	377.34	377.02
Median	5.49	5.23	5.61
Interquartile range	13.86	13.96	12.90
Panel C: Squared forecast errors			
	SIMPLE	SIMPLE_MACRO	ADVANCED_MACRO
N	728	728	728
Mean	143,071.90	143,071.70	142,846.60
Std. dev.	3,814,243.00	3,814,231.00	3,807,864.00
Median	30.15	27.40	31.46
Interquartile range	242.37	245.10	222.30

Note: Table 13 presents descriptive statistics on forecast errors, absolute forecast errors and squared forecast errors for all three prediction models. The sample includes all firm-years across all of the five validation periods (2015-2019), except for firm-years that have been dropped due to missing values for any of the independent variables included in ADVANCED_MACRO. Forecast errors are calculated as actual $RNOA_{t+1}$ less predicted $RNOA_{t+1}$, meaning that the errors are expressed in percentage form (e.g. average forecast error for SIMPLE is 11.03 percentage points, as shown in panel A). Absolute forecast errors are the absolute values of the forecast errors and squared forecast errors are the squared values. Interquartile range refers to the difference between 75th and 25th percentiles.

Table 14 provides the same statistics as in Table 13, but also includes analyst forecasts as a benchmark. Since not all firms have analyst coverage, the sample is smaller and amounts to 495 firm-years compared to 728 for the statistics in Table 13. An interesting observation is that the absolute and squared forecast errors are clearly lower for the sample of firms that have analyst coverage. For example, the average absolute forecast errors shown in panel B of Table 14 are almost 18 percentage points lower for each model compared to Table 13. The median absolute forecast error is about 1 percentage point lower for each model. Also the standard deviations are substantially lower. For absolute forecast errors, the standard deviations are around 21 per model in Table 14 compared to 377 in Table 13. There are several potential reasons for why the models produce better predictions for firms with analyst coverage. One is that firms with analyst coverage are likely to be more mature on average compared to firms without analyst coverage, meaning that they have lower time-series variation in RNOA and thereby fewer extreme values for the actual outcome of $RNOA_{t+1}$. This results in lower forecast errors. Another explanation could be that firms with analyst coverage potentially have higher quality of their financial reporting, which in turn leads to more accurate calculations of the "true" RNOA. For example, firms that are scrutinized by analysts may be more prone to disclose unusual items in the income statement, which affect the calculation of RNOA.[17]

Regarding the differences in prediction accuracy across models, the observations that could be made based on Table 13 are largely the same also for the sample of firms with analyst coverage in Table 14. SIMPLE_MACRO shows the best accuracy with regards to most parameters, especially with regards to the median errors in panel B-C, which are considered the most important measures.[18] SIMPLE shows slightly better prediction accuracy than ADVANCED_MACRO, just like in Table 13.

However, the most interesting take-aways from Table 14 relates to the prediction accuracy of the models compared to analyst forecasts. Based on the median values in panel B-C, the prediction models are outperformed by analyst forecasts. The same conclusion holds for interquartile range. Analyst forecasts have higher standard deviations and average errors in panel B-C, but this is caused by some extreme values.

[17] As described in the variable descriptions (Table A6 in Appendix), RNOA in this study uses operating income before unusual items as the numerator (similar to the study by Li et al. (2014)).
[18] The medians are not affected by extreme values in forecast errors in the same way as the means are.

Table 14. Descriptive statistics on forecast errors for all models and for analyst forecasts across all validation periods (2015-2019), using only firm-years with analyst coverage

Panel A: Forecast errors				
	SIMPLE	SIMPLE_MACRO	ADVANCED_MACRO	Analyst forecasts
N	495	495	495	495
Mean	-1.22	-1.57	-1.70	1.19
Std. dev.	24.08	24.12	24.20	36.58
Median	0.56	0.22	-0.37	-0.36
Interquartile range	8.58	8.28	9.07	7.70
Panel B: Absolute forecast errors				
	SIMPLE	SIMPLE_MACRO	ADVANCED_MACRO	Analyst forecasts
N	495	495	495	495
Mean	12.00	11.94	12.14	14.82
Std. dev.	20.91	21.00	20.99	33.46
Median	4.47	4.14	4.60	3.83
Interquartile range	11.62	11.31	11.29	10.33
Panel C: Squared forecast errors				
	SIMPLE	SIMPLE_MACRO	ADVANCED_MACRO	Analyst forecasts
N	495	495	495	495
Mean	580.37	582.91	587.29	1,337.16
Std. dev.	2,703.97	2,740.32	2,764.71	9,749.61
Median	20.02	17.12	21.17	14.68
Interquartile range	169.21	161.26	168.96	137.67

Note: Table 14 presents descriptive statistics on forecast errors, absolute forecast errors and squared forecast errors for all three prediction models and for analyst forecasts. The sample includes all firm-years across all of the five validation periods (2015-2019) that have analyst coverage, except for firm-years that have been dropped due to missing values for any of the independent variables included in ADVANCED_MACRO. Forecast errors are calculated as actual $RNOA_{t+1}$ less predicted $RNOA_{t+1}$, meaning that the errors are expressed in percentage form (e.g. average forecast error for SIMPLE is –1.22 percentage points, as shown in panel A). Absolute forecast errors are the absolute values of the forecast errors and squared forecast errors are the squared values. Interquartile range refers to the difference between 75th and 25th percentiles.

6.2.2. Wilcoxon signed rank tests

Wilcoxon signed rank tests are used to test the statistical significance of the differences in absolute and squared forecast errors across models and compared to analyst forecasts. Table 15 provides the test results for hypothesis B.1. The tests are performed on the medians of the paired differences in absolute and squared forecast errors, which in Table 15 is calculated as the errors for SIMPLE less the errors for SIMPLE_MACRO. A positive difference indicates that SIMPLE_MACRO has lower absolute and squared forecast errors, and thereby higher prediction accuracy. Panel A of Table 15 shows that the median difference in absolute forecast errors is positive 1.88 basis points.[19]

[19] Note that this value is not equal to the median absolute forecast error for SIMPLE less the median absolute forecast error for SIMPLE_MACRO from Table 13. Instead, it is the median of the *paired differences* in absolute forecast errors across the two models. For each of the 728 observations, the

Furthermore, the Wilcoxon signed rank test confirms that the median is larger than zero with 10% significance level. The null hypothesis of hypothesis B.1 is thereby rejected and it is concluded that SIMPLE_MACRO has significantly higher prediction accuracy than SIMPLE in terms of absolute forecast errors. However, the null hypothesis is not rejected for the difference in squared forecast errors, as illustrated in panel B.

Table 15. Wilcoxon signed rank test for hypothesis B.1: Paired differences in forecast errors between SIMPLE and SIMPLE_MACRO

Panel A: Paired differences in absolute forecast errors			
Median	0.0188		
Sign	N	Sum ranks	Expected
Positive	371	140,032	132,678
Negative	357	125,324	132,678
Zero	0	0	0
Total	728	265,356	265,356
z	1.296		
Prob > z	0.0976		
H_0	Rejected at 10% sig. level		
Panel B: Paired differences in squared forecast errors			
Median	0.0753		
Sign	N	Sum ranks	Expected
Positive	371	135,966	132,678
Negative	357	129,390	132,678
Zero	0	0	0
Total	728	265,356	265,356
z	0.579		
Prob > z	0.2812		
H_0	Not rejected		

Note: Table 15 presents the test results of the Wilcoxon signed rank test for hypothesis B.1. The paired differences in forecast errors are calculated as $FE_{SIMPLE} - FE_{SIMPLE\_MACRO}$, where "FE" refers to both absolute forecast errors (reported in panel A) and squared forecast errors (reported in panel B). A positive difference means that SIMPLE_MACRO has higher prediction accuracy than SIMPLE. Positive (negative) signs refer to the number of observations with positive (negative) differences. The test statistic *T* is the smaller of the positive and negative rank sums and is converted to a z-score using the formula presented in section 4.4.2. H_0: Median of ($FE_{SIMPLE} - FE_{SIMPLE\_MACRO}$) ≤ 0. The medians are expressed in percentage points. Reported p-values are one-tailed as the hypothesis is one-sided. The tests are performed on paired differences across all validation periods (2015-2019).

Table 16 provides the test results for hypothesis B.2. As suggested by the descriptive statistics in Tables 13-14 in the previous section, the tests in Table 16 confirm that SIMPLE_MACRO produce better forecasts than ADVANCED_MACRO, opposite to what was hypothesized. The medians of the paired differences reported in panel A and B of Table 16 are negative, which is against hypothesis B.2. Both of the Wilcoxon signed rank tests generate p-values above 0.90, effectively meaning that the opposite to

absolute forecast error of SIMPLE_MACRO has been subtracted from that of SIMPLE. 1.88 basis points is the median value of all these 728 differences.

hypothesis B.2 is confirmed: SIMPLE_MACRO has significantly better prediction accuracy than ADVANCED_MACRO in terms of absolute and squared forecast errors. The significance level for this finding is 5% for absolute forecast errors and 10% for squared forecast errors.

Table 16. Wilcoxon signed rank test for hypothesis B.2: Paired differences in forecast errors between SIMPLE_MACRO and ADVANCED_MACRO

Panel A: Paired differences in absolute forecast errors			
Median	-0.1007		
Sign	N	Sum ranks	Expected
Positive	350	121,634	132,678
Negative	378	143,722	132,678
Zero	0	0	0
Total	728	265,356	265,356
z	-1.946		
Prob > z	0.9742		
H_0	Not rejected		
Panel B: Paired differences in squared forecast errors			
Median	-0.4796		
Sign	N	Sum ranks	Expected
Positive	350	124,880	132,678
Negative	378	140,476	132,678
Zero	0	0	0
Total	728	265,356	265,356
z	-1.374		
Prob > z	0.9153		
H_0	Not rejected		

Note: Table 16 presents the test results of the Wilcoxon signed rank test for hypothesis B.2. The paired differences in forecast errors are calculated as $FE_{SIMPLE\_MACRO} - FE_{ADVANCED\_MACRO}$, where "FE" refers to both absolute forecast errors (reported in panel A) and squared forecast errors (reported in panel B). A positive difference means that ADVANCED_MACRO has higher prediction accuracy than SIMPLE_MACRO. Positive (negative) signs refer to the number of observations with positive (negative) differences. The test statistic *T* is the smaller of the positive and negative rank sums and is converted to a z-score using the formula presented in section 4.4.2. H_0: Median of ($FE_{SIMPLE\_MACRO} - FE_{ADVANCED\_MACRO}$) ≤ 0. The medians are expressed in percentage points. Reported p-values are one-tailed as the hypothesis is one-sided. The tests are performed on paired differences across all validation periods (2015-2019).

Table 17 reports the test results for hypothesis C, which concerns the prediction accuracy of the best model compared to analyst forecasts. Based on the aforementioned tests for hypotheses B.1 and B.2, SIMPLE_MACRO is considered to be the best model and therefore constitute the model that is tested against analyst forecasts. The sample in this test is smaller than in the previous ones since only firms with analyst coverage are included.

Hypothesis C is two-sided as opposed to the other hypotheses which are one-sided. Table 17 shows that the median difference in forecast errors between SIMPLE_MACRO and analyst forecasts is positive for both absolute and squared forecast errors. This indicates that analyst forecasts have higher prediction accuracy, similar to what could be observed based on the descriptive statistics in Table 14. However, the test results show that the null hypothesis of a median that is equal to zero cannot be rejected. In other words, it is not possible to conclude with statistical significance that SIMPLE_MACRO is either better or worse than analyst forecasts.

Table 17. Wilcoxon signed rank test for hypothesis C: Paired differences in forecast errors between SIMPLE_MACRO and analyst forecasts

Panel A: Paired differences in absolute forecast errors			
Median	0.1381		
Sign	N	Sum ranks	Expected
Positive	256	62,055	61,380
Negative	239	60,705	61,380
Zero	0	0	0
Total	495	122,760	122,760
z	0.212		
Prob > |z|	0.8321		
H_0	Not rejected		
Panel B: Paired differences in squared forecast errors			
Median	0.2272		
Sign	N	Sum ranks	Expected
Positive	256	63,075	61,380
Negative	239	59,685	61,380
Zero	0	0	0
Total	495	122,760	122,760
z	0.532		
Prob > |z|	0.5945		
H_0	Not rejected		

Note: Table 17 presents the test results of the Wilcoxon signed rank test for hypothesis C. The paired differences are calculated as $FE_{SIMPLE\_MACRO} - FE_{Analyst\ forecasts}$, where "FE" refers to both absolute forecast errors (reported in panel A) and squared forecast errors (reported in panel B). A positive difference means that analyst forecasts have higher prediction accuracy than SIMPLE_MACRO. Positive (negative) signs refer to the number of observations with positive (negative) differences. The test statistic *T* is the smaller of the positive and negative rank sums and is converted to a z-score using the formula presented in section 4.4.2. H_0: Median of ($FE_{SIMPLE\_MACRO} - FE_{Analyst\ forecasts}$) = 0. The medians are expressed in percentage points. Reported p-values are two-tailed as the hypothesis is two-sided. The tests are performed on paired differences across all validation periods (2015-2019).

6.2.3. Correctly predicted increases and decreases in RNOA

Table 18 provides descriptive statistics on correctly predicted increases and decreases in RNOA for the three models across all five validation periods as well as the three models compared to analyst forecasts. The means represent the proportion of correctly estimated increases and decreases, disclosed in decimal form. As visible in panel A, the means for all three models exceed 50 percent. This means that, on average, the three models predict the direction of increases and decreases better than what would be expected from a coin toss. In line with hypothesis B.1, the mean for SIMPLE_MACRO exceeds that of SIMPLE, indicating that SIMPLE_MACRO better predicts, on average, the increases and decreases in RNOA. In contrast, ADVANCED_MACRO has a lower mean than SIMPLE_MACRO which contradicts the expectations of hypothesis B.2. It also has slightly lower mean than SIMPLE. This is similar to the findings regarding forecast errors. Even though the means differ between the three models, the differences are not that large. As visible in panel B, the same relationship exists between the three models for the sample that only includes firm-years with analyst coverage. Interestingly, the mean for analyst forecasts is almost 10 percentage points higher than for SIMPLE_MACRO, the model with the highest mean of the three prediction models. This indicates that analysts better predict the increases and decreases in RNOA.

As mentioned in section 4.4.3, to determine whether the differences in means between the models and the analyst forecasts are statistically significant, proportion z-tests have been conducted. These tests are presented in the next section.

Table 18. Descriptive statistics on correctly predicted increases and decreases in RNOA for all models across all validation periods (2015-2019)

Panel A: All prediction models

	SIMPLE	SIMPLE_MACRO	ADVANCED_MACRO
N	728	728	728
Mean	0.5371	0.5536	0.5316
Std. dev.	0.4990	0.4975	0.4993

Panel B: All prediction models and analyst forecasts

	SIMPLE	SIMPLE_MACRO	ADVANCED_MACRO	Analyst forecasts
N	494	494	494	494
Mean	0.5304	0.5486	0.5344	0.6457
Std. dev.	0.4996	0.4981	0.4993	0.4788

Note: Table 18 presents descriptive statistics on correctly predicted increases and decreases in RNOA. The means represent the proportion of correctly estimated increases and decreases, disclosed in decimal form. In panel A, the sample includes all firm-years across all of the five validation periods (2015-2019), except for firm-years that have been dropped due to missing values for any of the independent variables included in ADVANCED_MACRO. In panel B, the sample includes all firm-years across all of the five validation periods (2015-2019) that have analyst coverage, except for firm-years that have been dropped due to missing values for any of the independent variables included in ADVANCED_MACRO. One of the 495 observations with analyst forecasts (see Table 14) did not have a $RNOA_t$ value, therefore there are only 494 observations in panel B.

6.2.4. Proportion tests

Table 19 displays the proportion z-tests for hypotheses B.1, B.2 and C. In panel A, the mean value, representing the proportion of correctly predicted increases and decreases in RNOA, is hypothesized to be higher for SIMPLE_MACRO than for SIMPLE. As panel A shows, the z-value is relatively small which, converted to a p-value (for this one-tailed test), means that the null hypothesis cannot be rejected. Therefore, the proportion of correctly predicted increases and decreases by SIMPLE_MACRO cannot be said to significantly outperform SIMPLE. In panel B, the mean value for ADVANCED_MACRO is hypothesized to be higher than the mean for SIMPLE_MACRO. As panel B shows, the z-value is negative and relatively small which, converted to a p-value (for this one-tailed test), means that the null hypothesis cannot be rejected. Therefore, the proportion of correctly predicted increases and decreases by ADVANCED_MACRO cannot be said to outperform SIMPLE_MACRO. In fact, the opposite hypothesis (that SIMPLE_MACRO has a higher mean than ADVANCED_MACRO) would have a much smaller p-value (0.2000). Although not significant, it indicates that ADVANCED_MACRO has lower rather than higher ability to predict the direction of change in RNOA compared to SIMPLE_MACRO. In panel C, the proportion of correctly predicted increases and decreases by SIMPLE_MACRO (the prediction model with the highest mean) is compared to that of analyst forecasts. As panel C shows, the z-value is negative and relatively large and converted to a p-value (for this two-tailed test) the null hypothesis can be rejected at a one percent significance level. Therefore, the proportion of correctly predicted increases and decreases by analyst forecasts can be said to be significantly different from SIMPLE_MACRO. Since the mean for the analyst forecasts is higher, it can also be said that the analysts significantly outperform the best prediction model with regards to estimating the direction of RNOA.

Table 19. Proportion z-tests for hypotheses B.1, B.2 and C

	Mean	Std. Err.	x	N
Panel A: Hypothesis B.1				
SIMPLE_MACRO	0.5536	0.0184	403	728
SIMPLE	0.5371	0.0185	391	728
$\hat{p}_1 - \hat{p}_2$	0.0165			
S_{d0}	0.0261			
z	0.6316			
Prob > z	0.2638			
H_0	Not rejected			
Panel B: Hypothesis B.2				
	Mean	Std. Err.	x	N
ADVANCED_MACRO	0.5316	0.0185	387	728
SIMPLE_MACRO	0.5536	0.0184	403	728
$\hat{p}_1 - \hat{p}_2$	-0.0220			
S_{d0}	0.0261			
z	-0.8417			
Prob > z	0.8000			
H_0	Not rejected			
Panel C: Hypothesis C				
	Mean	Std. Err.	x	N
SIMPLE_MACRO	0.5486	0.0224	271	494
Analyst forecasts	0.6457	0.0215	319	494
$\hat{p}_1 - \hat{p}_2$	-0.0972			
S_{d0}	0.0312			
z	-3.1135			
Prob > |z|	0.0018			
H_0	Rejected at 1% sig. level			

Note: Table 19 presents the test results of the proportion z-tests for hypotheses B.1, B.2 and C. The means represent the proportion of correctly predicted increases and decreases in RNOA, in decimal form. The expression $\hat{p}_1 - \hat{p}_2$ represents the difference in mean between the models compared and S_{d0} represents the difference in standard error between the models, which takes the number of observations (N) and number of correct prediction (x) into account. These inputs can be converted to a z-score applying the formula presented in section 4.4.3. The z-score is then converted to a corresponding p-value. Panel A shows a one-tailed test if SIMPLE_MACRO significantly better predicts the direction of change in RNOA compared to SIMPLE. Panel B shows a one-tailed test if ADVANCED_MACRO significantly better predicts the direction of change in RNOA compared to SIMPLE_MACRO. Panel C shows a two-tailed test of whether SIMPLE_MACRO and analyst forecasts are significantly different from each other in terms of correctly predicting the direction of change in RNOA.

7. Discussion

7.1. Summary of findings

The main findings of this study can be summarized in three bullets:

- The MACRO variable has a significantly positive relationship with one-year-ahead RNOA in estimation, thereby confirming hypothesis A. The variable also improves predictions out-of-sample, based on the comparison of SIMPLE_MACRO and SIMPLE, but this finding is significant only with respect to absolute forecast errors. If prediction accuracy is instead measured in terms of squared forecast errors or in terms of correctly predicted increases and decreases in RNOA, the results point in the same direction but are not statistically significant. In summary, hypothesis B.1 is only weakly supported.
- Prediction accuracy of one-year-ahead RNOA does not improve by using a more elaborate model which includes additional variables other than past RNOA and MACRO. In fact, such a model produce forecasts with significantly lower accuracy measured in terms of absolute and squared forecast errors, as illustrated by the comparison of ADVANCED_MACRO and SIMPLE_MACRO. In terms of correctly predicting the direction of change in RNOA, the accuracy of the models is not significantly different from each other. In summary, hypothesis B.2 is not confirmed.
- As a robustness check, the fairly simplistic model containing only past RNOA and MACRO as independent variables (i.e. SIMPLE_MACRO) is shown to be neither significantly worse nor better than analyst forecasts at predicting one-year-ahead RNOA in terms of forecast errors. However, analysts are significantly better at predicting the direction of change in RNOA. In summary, the conclusions for hypothesis C are different depending on how prediction accuracy is measured.

The following sections discuss the findings in more detail, in particular the first two bullets, by connecting to prior literature.

7.2. Predictive ability of the MACRO variable

The findings of this study confirm those of Li et al. (2014) and Doukakis et al. (2020). Taking firm-specific geographic exposure into account, proxied by segment sales, in combination with country-level predictions of real GDP growth, i.e. the MACRO variable, is useful for explaining changes in one-year-ahead RNOA. In line with the previous studies, the MACRO variable is shown to have a positive relationship with future RNOA, meaning that an increase in expected GDP growth corresponds to an increase in expected one-year-ahead RNOA. As shown in Table 10 and Table 11, the MACRO coefficients during the estimation periods vary from a low of 0.827 to a high of

1.204. In Li et al. (2014) and Doukakis et al. (2020) the coefficients are substantially lower, 0.27 and 0.012, respectively. There could be several reasons for this. One contributing reason could be the differences in time period and sample. Li et al. (2014) apply an earlier time period (1998-2010) on a large global sample and Doukakis et al. (2020) a slightly more narrow time period (2005-2015) on a European sample. Neither of the studies have a particular industry focus. This study focuses on Swedish-listed manufacturing firms between 2000 and 2019, which could have several effects. First, as expressed in the data section (section 5.1), manufacturing firms include mostly industrial firms for which sales and profitability are relatively dependent on the state of the economy. For a sample of such firms, the MACRO coefficient is expected to be more positive than for a sample of firms without industry focus. Second, as shown by Doukakis et al. (2020), the MACRO variable has no explanatory value in times of crisis. This implies that a sample with less crisis exposure should have a stronger relationship to the MACRO variable. The estimation periods in this study are a few years longer compared to the earlier studies, and the additional years included are likely to be classified as non-crisis years. In Doukakis et al. (2020), they use a time period and geographic focus with relatively large concentration to the global financial crisis and the European debt crisis. Thus, the number of crisis years in relation to non-crisis years should be higher in their study, which could explain the lower MACRO coefficient documented. In comparison with Li et al. (2014), the sample in this study is also likely to have less crisis exposure since more non-crisis years following the global financial crisis are included. In summary, the focus on manufacturing firms and the longer time period including less crisis exposure could be contributing reasons for the substantially higher MACRO coefficients.

In unreported out-of-sample tests, Li et al. (2014) find that with the inclusion of the MACRO variable, both the mean and median of absolute forecast errors are significantly reduced by 2 basis points compared to when they exclude the variable from their model. In line with Li et al. (2014), this thesis shows significant results for the difference in *absolute* forecasts errors between SIMPLE and SIMPLE_MACRO (significant on ten percent level). Interestingly, the median of the absolute forecast errors decrease by 26 basis points with the inclusion of the MACRO variable in this study (see Table 13), compared to the 2 basis points in Li et al. (2014). It is recognized that these numbers are not directly comparable because of the additional independent variables in Li et al.'s (2014) model compared to SIMPLE_MACRO. Nevertheless, it indicates a potentially higher predictive ability for the MACRO variable than previously documented. The differences in squared forecast errors and the proportion test also point in the direction of increased prediction accuracy, with lower forecast errors and a higher proportion of correctly estimated increases and decreases in RNOA when the MACRO variable is included. However, these results are not statistically significant. Although the predictive ability of MACRO is only weakly supported in the out-of-sample validation, the variable shows stability over time in estimation, as opposed to the additional variables included in

ADVANCED_MACRO. The MACRO variable is significant on, at least, a five percent confidence level throughout all estimation periods.

7.3. Predictive ability of past RNOA

As summarized in the second bullet in section 7.1, this study has also shown that using a more comprehensive model for predicting RNOA does not improve prediction accuracy compared to a more simplistic model. Instead, prediction accuracy actually deteriorates. This is true despite the fact that the additional variables included in the comprehensive model have a significant relationship with one-year-ahead RNOA in estimation.

The findings are similar to the out-of-sample prediction results for ROE in Skogsvik (2002a), where a parsimonious model with past ROE as the only independent variable is shown to outperform a more elaborate model. Besides the fact that this study includes a MACRO variable and predicts RNOA instead of ROE, one of the main differences compared to Skogsvik (2002a) is that most of the additional variables included in ADVANCED_MACRO are not accounting ratios. Instead, most of them are financial market variables. In line with the formulation of hypothesis B.2, one could expect that such variables should improve predictions of one-year-ahead RNOA as they contain useful non-accounting information that is not captured by past RNOA, but this study have shown the opposite. The study thereby adds to Skogsvik's (2002a) findings by showing that the addition of financial market variables to an otherwise fairly simplistic model, as opposed to adding a set of accounting ratios, does not improve prediction accuracy either. Even in estimation, the proportion of the variance in $RNOA_{t+1}$ that can be explained by the independent variables barely change with the addition of the accounting and financial market ratios, as illustrated by the difference in adjusted R-squared between SIMPLE_MACRO and ADVANCED_MACRO (see Table 10 and 11).

The findings underlines the high forecasting power that past RNOA has on its own, which has been documented several times in prior research (e.g., Soliman, 2008) and which can be explained by its strong mean reversion properties (documented in e.g., Nissim & Penman, 2001). The mean reversion properties of RNOA are confirmed in this study, since all of the estimated coefficients of past RNOA are positive and lower than one.

Another observation concerns the lack of stability over time with regard to the significant set of independent variables in ADVANCED_MACRO (except $RNOA_t$ and $MACRO_t$). The instability partly explains why the model did not produce better predictions than SIMPLE_MACRO. Even more importantly, it highlights something for practitioners to bear in mind when forecasting future firm profitability: factors that historically have demonstrated a strong relationship with future profitability in estimation, do not necessarily improve predictions of profitability in the future. Since advanced models with many variables are also more costly to produce, it is likely a better option in many cases to use a more simplistic model that acknowledges the mean reversion in profitability.

Relying on analyst forecasts is another alternative, at least for predicting the direction of change in RNOA, for which analysts have shown to outperform a model containing past RNOA and MACRO.

8. Conclusions and future research

This Master Thesis has extended the research on how forecasts of macroeconomic factors can be used to predict future profitability, a research area which is still relatively unexplored. The thesis investigates whether firm-specific geographic exposure in combination with country-level predictions of real GDP growth is helpful for explaining one-year-ahead RNOA. Compared to previous research, a comprehensive out-of-sample validation of the estimated models has been conducted. In addition, the study has been designed to enable an isolation of the incremental predictive content of the MACRO variable over past RNOA. The thesis also focuses on a later, industry-specific sample, consisting of Swedish-listed manufacturing firms with large global footprints and high dependency on macroeconomic conditions.

The findings show, in line with previous research, that the MACRO variable adds explanatory value for one-year-ahead RNOA in estimation. As expected, the relationship between MACRO and one-year-ahead RNOA is greater in magnitude for a sample of Swedish-listed manufacturing firms compared to the samples studied by Li et al. (2014) and Doukakis et al. (2020). Depending on prediction model and estimation period, a one percentage point increase in expectation of real GDP growth corresponds to an increase of between 82.7 and 120.4 basis points in one-year-ahead RNOA. The out-of-sample prediction accuracy is improved by including the MACRO variable. However, whether this improvement is statistically significant or not depends on the measure of prediction accuracy. Interestingly, a more elaborate model, which includes additional variables other than past RNOA and MACRO, has significantly higher forecast errors out-of-sample, demonstrating that a strong relationship with future profitability in estimation not necessarily translates to improved out-of-sample predictions. This also underlines the high forecasting power of past RNOA, which has been documented several times in prior research. Analysts forecasts significantly outperform the best prediction model with regards to the direction of change in RNOA, indicating that analysts have skill. However, the results are inconclusive with regards to differences in forecast errors.

A limitation of this study is the relatively small sample size, reducing the reliability and the generalizability of the results. The sample size was limited by the very time-consuming process of manually coding the unique country exposures in each firm-year. In addition, there is some nosiness associated with the MACRO variable. Firms disclose their segment reporting in a subjective manner and the manual coding of country exposures is based on certain necessary assumptions about how sales within reported segments are distributed across countries.

This thesis has found that the out-of-sample prediction accuracy is improved by including the MACRO variable, with statistically significant results for one measure of prediction accuracy. Even though not all test turned out significant, the variable shows promise.

Therefore, it is worth exploring further. It would be of interest to investigate *when* or for which other *type of firms* the MACRO variable could have higher relevance. Doukakis et al. (2020) have found evidence that the MACRO variable lacks explanatory value in times of economic crisis, but there should be other conditions to identify and also test out-of-sample. It would also be of interest to explore the predictive ability of other macroeconomic variables. For example, as presented in section 2.3.1, Abarbanell & Bushee (1997) and Dowen (2001) have shown that profitability is affected by inflation, which is worth exploring further. Given strong out-of-sample statistical support in future research, promising trading strategies could be developed.

9. References

Academic sources

Abarbanell, J. S., & Bushee, B. J. (1997). Fundamental Analysis, Future Earnings, and Stock Prices. *Journal of Accounting Research, 35*(1), 1-24.

Balakrishnan, R., Harris, T. S., & Sen, P. K. (1990). The predictive ability of geographic segment disclosures. *Journal of Accounting research, 28*(2), 305-325.

Ball, R., & Watts, R. (1972). Some time series properties of accounting income. *Journal of Finance, 27*(3), 663-681.

Beaver, W. H. (1970). The Time Series Behavior of Earnings. *Journal of Accounting Research, 8,* 62-99.

Bernard, V., & Noel, J. (1991). Do Inventory Disclosures Predict Sales and Earnings? *Journal of Accounting, Auditing & Finance, 6*(2), 145-181.

Chordia, T., & Shivakumar, L. (2005). Inflation illusion and post-earnings-announcement drift. *Journal of Accounting Research*, *43*(4), 521-556.

Collins, D. W. (1976). Predicting earnings with sub-entity data: Some further evidence. *Journal of Accounting Research*, 163-177.

Doukakis, L., Ghicas, D. C., Siougle, G., & Sougiannis, T. (2020). The Informativeness of Micro and Macro Information During Economic Crisis and Non-Crisis Periods: Evidence from Europe. *European Accounting Review, 29*(3), 467-492.

Dowen, R. J. (2001). Fundamental information and monetary policy: The implications for Earnings and Earnings Forecasts. *Journal of Business Finance & Accounting, 28*(3-4), 481-501.

Esplin, A., Hewitt, M., Plumlee, M., & Yohn, T. (2014). Disaggregating Operating and Financial Activities: Implications for Forecasts of Profitability. *Review of Accounting Studies, 19*(1), 328-362.

Fairfield, P., Sweeney, R., & Yohn, T. (1996). Accounting Classification and the Predictive Content of Earnings. *The Accounting Review, 71*(3), 337-355.

Fama, E. F., & French, K. R. (1995). Size and Book-to-Market Factors in Earnings and Returns. *The Journal of Finance, 50*(1), 131-155.

Fama, E. F., & French, K. R. (2000). Forecasting Profitability and Earnings. *The Journal of Business, 73*(2), 161-175.

Fama, E. F., & French, K. R. (2006). Profitability, investment and average returns. *Journal of Financial Economics, 82*(3), 491-518.

Freeman, R. N., Ohlson, J. A., & Penman, S. H. (1982). Book Rate-of-Return and Prediction of Earnings Changes: An Empirical Investigation. *Journal of Accounting Research, 20*(2), 639-653.

Gombola, M., & Ketz, J. E. (1983). Financial Ratio Patterns in Retail and Manufacturing Organizations. *Financial Management, 12*(2), 45-56.

Hou, K., van Dijk, M. A., & Zhang, Y. (2012). The implied cost of capital: A new approach. *Journal of Accounting and Economics, 53,* 504-526.

Kinney, W. R. (1971). Predicting earnings: entity versus subentity data. *Journal of Accounting Research*, 127-136.

Konchitchki, Y. (2011). Inflation and Nominal Financial Reporting: Implications for Performance and Stock Prices. *The Accounting Review, 86*(3), 1045-1085.

Kothari, S. P. (2001). Capital markets research in accounting. *Journal of Accounting and Economics*, *31*(1-3), 105-231.

Lee, C. M. (1999). Accounting-Based Valuation: Impact on Business Practices and Research. *American Accounting Association, 13*(4), 413-425.

Lev, B., & Thiagarajan, S. R. (1993). Fundamental Information Analysis. *Journal of Accounting Research, 31*(2), 190-215.

Li, N., Richardson, S., & Tuna, I. (2014). Macro to micro: Country exposures, firm fundamentals and stock returns. *Journal of Accounting and Economics, 58*(1), 1-20.

Lookabill, L. (1976). Some Additional Evidence on the Time Series Properties of Accounting Earnings. *The Accounting Review, 51*(4), 724-738.

Nissim, D., & Penman, S. H. (2001). Ratio Analysis and Equity Valuation: From Research to Practice. *Review of Accounting Studies, 6*(1), 109-154.

Nissim, D., & Penman, S. H. (2003). The association between changes in interest rates, earnings, and equity values. *Contemporary Accounting Research, 20*(4), 775-804.

Ou, J. A., & Penman, S. H. (1989). Financial Statement Analysis and the Prediction of Stock Returns. *Journal of Accounting and Economics, 11,* 295-329.

Richardson, S. A., Sloan, R. G., Soliman, M. T., & Tuna, I. (2005). Accrual Reliability, Earnings Persistence and Stock Prices. *Journal of Accounting and Economics, 39*(3), 437-485.

Roberts, C. B. (1989). Forecasting earnings using geographical segment data: Some UK evidence. *Journal of International Financial Management & Accounting, 1*(2), 130-151.

Skogsvik, K. (2002b). *A Tutorial on Residual Income Valuation and Value Added Valuation.* Stockholm School of Economics, EFI, Stockholm.

Skogsvik, S. (2002a). *Redovisningsmått, värderelevans och informationseffektivitet.* PhD Thesis, Stockholm School of Economics, EFI, Stockholm.

So, E. C. (2013). A new approach to predicting analyst forecast errors: Do investors overweight analyst forecasts? *Journal of Financial Economics, 108*, 615-640.

Soliman, M. (2008). The Use of DuPont Analysis by Market Participants. *The Accounting Review, 83*(3), 823-853.

Stober, T. L. (1993). The Incremental Information Content of Receivables in Predicting Sales, Earnings, and Profit Margins. *Journal of Accounting, Auditing & Finance, 8*(4), 447-473.

Books

Foster, G. (1986). *Financial Statement Analysis* (2nd Edition ed.). Englewood Cliffs, New Jersey: Prentice-Hall.

Newbold, P., Carlson, W. L., & Thorne, B. (2010). *Statistics for Business and Economics* (7th Edition ed.). Upper Saddle River, New Jersey: Pearson.

Penman, S. H. (2013). *Financial Statement Analysis and Security Valuation.* New York: McGraw-Hill.

Other sources

EU. (2004). DIRECTIVE 2004/109/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. *Official Journal of the European Union*.

SFS 1992:543. *Lag om börs- och clearingverksamhet.* Ministry of Finance. https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/lag-1992543-om-bors--och-clearingverksamhet_sfs-1992-543

SFS 2007:528. *Lag om värdepappersmarknaden*. Ministry of Finance. https://www.riksdagen.se/sv/dokument-lagar/drokument/svensk-forfattningssamling/lag-2007528-om-vardepappersmarknaden_sfs-2007-528

UN. (n.d.). *Standard country or area codes for statistical use (M49)*. Retrieved from United Nations: https://unstats.un.org/unsd/methodology/m49/

10. Appendix

Table A1. Overview of relevant literature

Author(s)	Data	Forecast variable	Independent variables	Statistical method
Studies of profitability prediction models that do not take macroeconomic factors into account				
Freeman et al. (1982)	30 industrial firms in the US, 1946-1977	Earnings at time t+1	ROE at time t	Logit
	30 industrial firms in the US, 1946-1977	ROE at time t+1	ROE at time t	Logit
Ou & Penman (1989)	19,579 firm-years, US-listed firms, 1973-1983	Earnings at time t+1	A set of accounting ratios at time t	Logit
Fairfield et al. (1996)	33,334 firm-years, industrial firms in the US, 1973-1990	ROE at time t+1	ROE at time t	Regression
	33,334 firm-years, industrial firms in the US, 1973-1990	ROE at time t+1	Disaggregated components of ROE at time t	Regression
Nissim & Penman (2001)[1]	A large set of US-listed firms, 1963-1999	RNOA and growth in NOA	A set of accounting ratios	Unknown
Skogsvik (2008)	~65 Swedish-listed manufacturing firms per estimation period, 1972-1994	Change in medium-term ROE (past 3-years average) between year t and t+3	Average ROE from year t-2 to t	Logit
	~65 Swedish-listed manufacturing firms per estimation period, 1972-1994	Change in medium-term ROE (past 3-years average) between year t and t+3	A set of accounting ratios in year t	Logit
	~65 Swedish-listed manufacturing firms per estimation period, 1972-1994	Change in medium-term ROE (past 3-years average) between year t and t+3	Average ROE from year t-2 to t, and a set of accounting ratios in year t	Logit

Soliman (2008)	38,716 firm-years, no information about geography, 1984-2002	Change in RNOA in year t+1	DuPont components of RNOA in year t, and control variables[2]	Regression
	38,716 firm-years, no information about geography, 1984-2002	Change in RNOA in year t+1	Changes in DuPont components of RNOA in year t, and control variables[2]	Regression
Studies of profitability prediction models including macroeconomic contextual analysis				
Abarbanell & Bushee (1997)[3]	4,180 firm-years, US-listed firms, 1983-1990	Change in earnings in year t+1	Nine fundamental accounting signals in year t	Regression
Dowen (2001)	4,533 US firm years, 1985-1995	Change in EPS at time t+1	Nine fundamental accounting signals, three financial market signals and three macroeconomic dummy variables (economic growth, inflation, and monetary policy) at time t	Regression
Studies of profitability prediction models including macroeconomic independent variables				
Roberts (1989)	78 UK industrial companies, 1981-1983	Earnings at time t+1	Earnings and a firm-specific GNP growth estimate at time t	Growth adjusted model
Balakrishnan et al. (1990)	89 US companies, 1979-1983	Earnings at time t+1	Earnings and a firm-specific GNP growth estimate at time t	Growth adjusted model
Li et al. (2014)	198,315 firm-years, global sample, 1998-2010	RNOA in year t+1	A set of accounting and financial market variables, and a firm-specific GDP growth estimate, at time t	Regression
Doukakis et al. (2020)	15,343 firm-years, European sample, 2005-2015	RNOA in year t+1	A set of accounting and financial market variables, a firm-specific GDP growth estimate, and an economic crisis indicator variable, at time t	Regression

Note: Table A1 provides an overview of the most central articles in the literature review. Time "t" corresponds to the prediction point in time. [1]Information about the prediction models is very scarce as the models are only mentioned very briefly in the light of the poor prediction results out of sample. [2]The control variables include earnings predictors used in prior literature, namely the fundamental signals used in Abarbanell & Bushee (1997) and the accruals used in Richardson et al. (2005). [3]Applies the prediction model in two different macroeconomic contexts; economic growth, and inflation.

Table A2. List of IMF countries and the UN M49 classifications

Country (IMF)	Continent (UN M49)	Latitude (UN M49)	Country (IMF)	Continent (UN M49)	Latitude (UN M49)
Algeria	Africa	Northern	Malawi	Africa	Sub-Saharan
Egypt	Africa	Northern	Mali	Africa	Sub-Saharan
Libya	Africa	Northern	Mauritania	Africa	Sub-Saharan
Morocco	Africa	Northern	Mauritius	Africa	Sub-Saharan
Sudan	Africa	Northern	Mozambique	Africa	Sub-Saharan
Tunisia	Africa	Northern	Namibia	Africa	Sub-Saharan
Angola	Africa	Sub-Saharan	Niger	Africa	Sub-Saharan
Benin	Africa	Sub-Saharan	Nigeria	Africa	Sub-Saharan
Botswana	Africa	Sub-Saharan	Republic of Congo	Africa	Sub-Saharan
Burkina Faso	Africa	Sub-Saharan	Rwanda	Africa	Sub-Saharan
Burundi	Africa	Sub-Saharan	São Tomé and Príncipe	Africa	Sub-Saharan
Cabo Verde	Africa	Sub-Saharan	Senegal	Africa	Sub-Saharan
Cameroon	Africa	Sub-Saharan	Seychelles	Africa	Sub-Saharan
Central African Republic	Africa	Sub-Saharan	Sierra Leone	Africa	Sub-Saharan
Chad	Africa	Sub-Saharan	Somalia	Africa	Sub-Saharan
Comoros	Africa	Sub-Saharan	South Africa	Africa	Sub-Saharan
CÙte d'Ivoire	Africa	Sub-Saharan	South Sudan	Africa	Sub-Saharan
Democratic Republic of the Congo	Africa	Sub-Saharan	Tanzania	Africa	Sub-Saharan
Djibouti	Africa	Sub-Saharan	The Gambia	Africa	Sub-Saharan
Equatorial Guinea	Africa	Sub-Saharan	Togo	Africa	Sub-Saharan
Eritrea	Africa	Sub-Saharan	Uganda	Africa	Sub-Saharan
Eswatini	Africa	Sub-Saharan	Zambia	Africa	Sub-Saharan
Ethiopia	Africa	Sub-Saharan	Zimbabwe	Africa	Sub-Saharan
Gabon	Africa	Sub-Saharan			
Ghana	Africa	Sub-Saharan			
Guinea	Africa	Sub-Saharan			
Guinea-Bissau	Africa	Sub-Saharan			
Kenya	Africa	Sub-Saharan			
Lesotho	Africa	Sub-Saharan			
Liberia	Africa	Sub-Saharan			
Madagascar	Africa	Sub-Saharan			

Country (IMF)	Continent (UN M49)	Latitude (UN M49)	Country (IMF)	Continent (UN M49)	Latitude (UN M49)
Kazakhstan	Asia	Central	Armenia	Asia	West
Kyrgyz Republic	Asia	Central	Azerbaijan	Asia	West
Tajikistan	Asia	Central	Bahrain	Asia	West
Turkmenistan	Asia	Central	Cyprus	Asia	West
Uzbekistan	Asia	Central	Georgia	Asia	West
China	Asia	East	Iraq	Asia	West
Hong Kong SAR	Asia	East	Israel	Asia	West
Japan	Asia	East	Jordan	Asia	West
Korea	Asia	East	Kuwait	Asia	West
Macao SAR	Asia	East	Lebanon	Asia	West
Mongolia	Asia	East	Oman	Asia	West
Taiwan Province of China	Asia	East	Qatar	Asia	West
Afghanistan	Asia	South	Saudi Arabia	Asia	West
Bangladesh	Asia	South	Syria	Asia	West
Bhutan	Asia	South	Turkey	Asia	West
India	Asia	South	United Arab Emirates	Asia	West
Islamic Republic of Iran	Asia	South	West Bank and Gaza	Asia	West
Maldives	Asia	South	Yemen	Asia	West
Nepal	Asia	South	Australia	Oceania	
Pakistan	Asia	South	Fiji	Oceania	
Sri Lanka	Asia	South	Kiribati	Oceania	
Brunei Darussalam	Asia	Southeast	Marshall Islands	Oceania	
Cambodia	Asia	Southeast	Micronesia	Oceania	
Indonesia	Asia	Southeast	Nauru	Oceania	
Lao P.D.R.	Asia	Southeast	New Zealand	Oceania	
Malaysia	Asia	Southeast	Palau	Oceania	
Myanmar	Asia	Southeast	Papua New Guinea	Oceania	
Philippines	Asia	Southeast	Samoa	Oceania	
Singapore	Asia	Southeast	Solomon Islands	Oceania	
Thailand	Asia	Southeast	Tonga	Oceania	
Timor-Leste	Asia	Southeast	Tuvalu	Oceania	
Vietnam	Asia	Southeast	Vanuatu	Oceania	

Country (IMF)	Continent (UN M49)	Latitude (UN M49)	Country (IMF)	Continent (UN M49)	Latitude (UN M49)
Belarus	Europe	East	Netherlands	Europe	West
Bulgaria	Europe	East	Switzerland	Europe	West
Czech Republic	Europe	East	Antigua and Barbuda	North America	Caribbean
Hungary	Europe	East	Aruba	North America	Caribbean
Moldova	Europe	East	Barbados	North America	Caribbean
Poland	Europe	East	Dominica	North America	Caribbean
Romania	Europe	East	Dominican Republic	North America	Caribbean
Russia	Europe	East	Grenada	North America	Caribbean
Slovak Republic	Europe	East	Haiti	North America	Caribbean
Ukraine	Europe	East	Jamaica	North America	Caribbean
Denmark	Europe	North	Puerto Rico	North America	Caribbean
Estonia	Europe	North	St. Kitts and Nevis	North America	Caribbean
Finland	Europe	North	St. Lucia	North America	Caribbean
Iceland	Europe	North	St. Vincent and the Grenadines	North America	Caribbean
Ireland	Europe	North	The Bahamas	North America	Caribbean
Latvia	Europe	North	Trinidad and Tobago	North America	Caribbean
Lithuania	Europe	North	Belize	North America	Central America
Norway	Europe	North	Costa Rica	North America	Central America
Sweden	Europe	North	El Salvador	North America	Central America
United Kingdom	Europe	North	Guatemala	North America	Central America
Albania	Europe	South	Honduras	North America	Central America
Bosnia and Herzegovina	Europe	South	Mexico	North America	Central America
Croatia	Europe	South	Nicaragua	North America	Central America
Greece	Europe	South	Panama	North America	Central America
Italy	Europe	South	Canada	North America	Northern America
Kosovo	Europe	South	United States	North America	Northern America
Malta	Europe	South	Argentina	South America	
Montenegro	Europe	South	Bolivia	South America	
North Macedonia	Europe	South	Brazil	South America	
Portugal	Europe	South	Chile	South America	
San Marino	Europe	South	Colombia	South America	
Serbia	Europe	South	Ecuador	South America	
Slovenia	Europe	South	Guyana	South America	
Spain	Europe	South	Paraguay	South America	
Austria	Europe	West	Peru	South America	
Belgium	Europe	West	Suriname	South America	
France	Europe	West	Uruguay	South America	
Germany	Europe	West	Venezuela	South America	
Luxembourg	Europe	West			

Note: Table A2 presents a list of all IMF countries and the UN M49 classifications. UN M49 is the standard country and area codes prepared by the Statistics Division of the United Nations Secretariat.

Table A3. List of constellations

European Union	Central Europe/Middle Europe	Central & Eastern Europe (vs. Western)	Western Europe (vs. Central & Eastern)
Austria	Germany	Belarus	Denmark
Belgium	Switzerland	Bulgaria	Finland
Bulgaria	Austria	Czech Republic	Iceland
Croatia	Poland	Hungary	Ireland
Cyprus	Czech Republic	Moldova	Norway
Czech Republic	Slovak Republic	Poland	Sweden
Denmark	Hungary	Romania	United Kingdom
Estonia	Slovenia	Russia	Greece
Finland		Slovak Republic	Italy
France		Ukraine	Malta
Germany		Estonia	Portugal
Greece		Latvia	San Marino
Hungary		Lithuania	Spain
Ireland		Albania	Austria
Italy		Bosnia and Herzegovina	Belgium
Latvia		Croatia	France
Lithuania		Kosovo	Germany
Luxembourg		Montenegro	Luxembourg
Malta		North Macedonia	Netherlands
Netherlands		Serbia	Switzerland
Poland		Slovenia	
Portugal			
Romania			
Slovak Republic			
Slovenia			
Spain			
Sweden			
United Kingdom			

North East Asia	NAFTA	DACH
China	United States	Germany
Hong Kong SAR	Mexico	Austria
Japan	Canada	Switzerland
Korea		
Taiwan Province of China		

Nordics	**Benelux**	**Mediterranean**	**Eurasia**
Denmark	Belgium	Portugal	Afghanistan
Finland	Luxembourg	Spain	Armenia
Iceland	Netherlands	France	Azerbaijan
Norway		Italy	Belarus
Sweden		Malta	Georgia
		Slovenia	Kazakhstan
		Croatia	Kyrgyz Republic
		Bosnia and Herzegovina	Mongolia
		Montenegro	Moldova
		Albania	Tajikistan
		Greece	Turkmenistan
		Turkey	Ukraine
		Syria	Uzbekistan
		Cyprus	
		Lebanon	
		Israel	
		Jordan	
		Egypt	
		Libya	
		Tunisia	
		Algeria	
		Morocco	
		San Marino	

Scandinavia	**Commonwealth of Independence States (CIS)**	**Baltics**	**Far East**
Sweden	Armenia	Estonia	China
Norway	Azerbaijan	Latvia	Hong Kong SAR
Denmark	Kazakhstan	Lithuania	Japan
	Kyrgyz Republic		Korea
	Tajikistan		
	Uzbekistan		
	Russia		
	Belarus		
	Moldova		
	Ukraine		
	Turkmenistan		

EMEA	**EMEA (cont.)**	**EMEA (cont.)**	**Emerging Markets**
Belarus	Jordan	Mozambique	Argentina
Bulgaria	Kuwait	Namibia	Brazil
Czech Republic	Lebanon	Niger	Chile
Hungary	Libya	Nigeria	China
Moldova	Mauritania	Republic of Congo	Colombia
Poland	Morocco	Rwanda	Czech Republic
Romania	Oman	São Tomé and Príncipe	Egypt
Russia	West Bank and Gaza	Senegal	Greece
Slovak Republic	Qatar	Seychelles	Hungary
Ukraine	Saudi Arabia	Sierra Leone	India
Denmark	Sudan	Somalia	Indonesia
Estonia	Syria	South Africa	Korea
Finland	Tunisia	South Sudan	Kuwait
Iceland	United Arab Emirates	Tanzania	Malaysia
Ireland	Yemen	The Gambia	Mexico
Latvia	Algeria	Togo	Pakistan
Lithuania	Angola	Uganda	Peru
Norway	Benin	Zambia	Philippines
Sweden	Botswana	Zimbabwe	Poland
United Kingdom	Burkina Faso		Qatar
Albania	Burundi		Russia
Bosnia and Herzegovina	Cabo Verde		Saudi Arabia
Croatia	Cameroon		South Africa
Greece	Central African Republic		Taiwan Province of China
Italy	Chad		Thailand
Kosovo	Comoros		Turkey
Malta	CÙte d'Ivoire		United Arab Emirates
Montenegro	Democratic Republic of the Congo		
North Macedonia	Djibouti		
Portugal	Equatorial Guinea		
San Marino	Eritrea		
Serbia	Eswatini		
Slovenia	Ethiopia		
Spain	Gabon		
Austria	Ghana		
Belgium	Guinea		
France	Guinea-Bissau		
Germany	Kenya		
Luxembourg	Lesotho		
Netherlands	Liberia		
Switzerland	Madagascar		
Bahrain	Malawi		
Egypt	Mali		
Iraq	Mauritius		

Latin America	**Middle East**	**Continental Europe**	**Asia-Pacific**
Argentina	Bahrain	Belarus	China
Belize	Egypt	Bulgaria	Hong Kong SAR
Bolivia	Iraq	Czech Republic	Japan
Brazil	Jordan	Hungary	Korea
Chile	Kuwait	Moldova	Macao SAR
Colombia	Lebanon	Poland	Mongolia
Costa Rica	Libya	Romania	Taiwan Province of China
Ecuador	Mauritania	Russia	Afghanistan
El Salvador	Morocco	Slovak Republic	Bangladesh
Guatemala	Oman	Ukraine	Bhutan
Guyana	West Bank and Gaza	Denmark	India
Honduras	Qatar	Estonia	Islamic Republic of Iran
Mexico	Saudi Arabia	Finland	Maldives
Nicaragua	Sudan	Latvia	Nepal
Paraguay	Syria	Lithuania	Pakistan
Peru	Tunisia	Norway	Sri Lanka
Panama	United Arab Emirates	Sweden	Brunei Darussalam
Suriname	Yemen	Albania	Cambodia
Venezuela		Bosnia and Herzegovina	Indonesia
Uruguay		Croatia	Lao P.D.R.
		Greece	Malaysia
		Italy	Myanmar
		Kosovo	Philippines
		Montenegro	Singapore
		North Macedonia	Thailand
		Portugal	Timor-Leste
		San Marino	Vietnam
		Serbia	Australia
		Slovenia	Fiji
		Spain	Kiribati
		Austria	Marshall Islands
		Belgium	Micronesia
		France	Nauru
		Germany	New Zealand
		Luxembourg	Palau
		Netherlands	Papua New Guinea
		Switzerland	Samoa
			Solomon Islands
			Tonga
			Tuvalu
			Vanuatu

Note: Table A3 presents a list of all constellations. It is recognized that these constellations might have changed over time, however, the constellations by the end of the sample period have been applied for all firm-years. For example, Croatia is included in the “EU” constellation although the country has not been a member state during the entire sample period.

Table A4. Additional rules regarding the manual coding of European segments

<table>
<tr><td>Table 3. still governs the overall priority of the manual coding. If there are European segments, after accounting for individual countries (step 1) and potentially emptying smaller constellations (step 2a), the following ranking applies:</td></tr>
<tr><td>Segments that contain a latitude in the segment name and there is only one such segment in a firm-year:
- If it is West: Use the Western countries in the constellation Western vs. Central & Eastern
- If it is East: Use the Central & Eastern countries in the constellation Western vs. Central & Eastern
- If it is South: Use the countries classified as South according to the UN M49 classification
- If it is North: Use the countries classified as North according to the UN M49 classification</td></tr>
<tr><td>Segments that contain a latitude in the segment name and there are two such segments in a firm-year (also applies if there is only one segment but it covers two latitudes, e.g. one segment called “Northern & Western Europe”):
- If the latitudes are West vs. East, use the constellations Western vs. Central & Eastern
- If the latitudes are either West or East, plus either North or South: 1) allocate the countries to the North or South segment according to UN M49 classification, 2) use the Western vs. Central & Eastern constellation for the West or East segment, and exclude the countries that have already been allocated to North or South
- If the latitudes are North and South: use UN M49 classification</td></tr>
<tr><td>Segments that contain a latitude in the segment name and there are three such segments in a firm-year (also applies if there are less than three segments but they cover three latitudes, e.g. one segment called “Northern & Western Europe” and one segment called “Southern Europe”):
- If the latitudes are both West and East, plus either North or South: 1) allocate the countries to the North or South segment according to UN M49 classification, 2) use the Western vs. Central & Eastern constellation for the rest of the European countries
- If the latitudes are both North and South, plus either West or East: 1) allocate the countries to all segments according to UN M49 classification</td></tr>
<tr><td>Segments that contain a latitude in the segment name and there are four such segments (also applies if there are less than four segments but they cover four latitudes, e.g. one segment called “Northern & Western Europe” and one segment called “Southern & Eastern Europe”):
- Allocate the countries to all segments according to UN M49 classification</td></tr>
</table>

Note: Table A4 presents additional rules regarding the manual coding of European segments.

Table A5. Additional general rules regarding the manual coding

Segments starting with "Other…", "Rest of…" and "Miscellaneous..." followed by a region is treated the same as a segment with only the name of the region, e.g., "Rest of Asia" is treated the same way as "Asia".
If there are two (or more) segments within the same firm-year with the same meaning, e.g., "Rest of Europe" and "Europe", the segment containing data will be prioritized. If both segments contain data, they are coded with the exact same countries which, in effect, merges the two segments.
If there are segments called something unrelated to geographic exposure such as "items effecting comparability", "Investments of disposal group" etc. and they have no segment sales, they are simply excluded from that firm-year. If they do have segment sales, the countries manually coded for the rest of that firm-year across all other segments are all coded to those segments. Which means that those (very few) segments will get a segment specific MACRO factor in-line with the total geographic exposure of that firm-year.
"Australia" is coded as the individual country "Australia".
"Other Australia", "Rest of Australia" and "Australia, etc." are treated as the continent "Oceania".
A segment that states multiple continents where one of them is "Australia" will be treated as "Oceania", e.g. a segment called "Africa, Asia & Australia" will be treated as "Africa, Asia & Oceania". Following the same type of reasoning, "AustralAsia" is treated as the continents "Asia" and "Oceania", and "Asia, Australia" is also treated as the continents "Asia" and "Oceania".
If a segment is called "South East Asia", the countries specified by the UN will be applied, and it should not include the countries defined as only "South Asia" or "East Asia" by the UN and vice-versa.
In a segment that starts broad and then narrows down, e.g., "Asia (mainly China)", the more specific part will be prioritized, which for this example means "China".
"South Korea" is treated as "Korea" fom the IMF list.
"Great Britain" and "England" are treated as the "United Kingdom".
"Asia/Far East" is treated as the continent "Asia".
"Pacific" and "South Pacific" are treated as the continent "Oceania".
"Holland" is treated as the "Netherlands".
"Asia-Pacific" is treated as the constellation "Asia-Pacific" but "Asia/Pacific" is treated as the continents "Asia" and "Oceania".
"Americas" and "America" are treated as the continents "North America" and "South America" taken together.

Note: Table A5 presents additional general rules regarding the manual coding of segments.

Table A6. Descriptions of independent variables and other data items used in the study

Variable or data item	**Description**
Analyst forecasts of $RNOA_{t+1}$ *	Analyst forecasts of EBIT in year *t+1* divided by estimated average NOA in year *t+1*. The EBIT estimate is the median analyst consensus estimate as of 30 April in year *t+1* (the last date before the prediction point in time), collected by FactSet Fundamentals. Estimated average NOA is computed as NOA at the end of year *t* plus half of the EBIT estimate.
BTM_t	Book-to-market ratio, computed as the ratio of common equity to equity market capitalization, both measured at the end of year *t*.
D_Div_t	Indicator variable equal to one for firms that paid dividend in year *t* (dividend pertaining to year *t-1*) and zero otherwise. Includes normal and extra dividends, but not special dividends.
Div_Yield_t	Dividend per share paid in year *t* (dividend pertaining to year *t-1*) divided by share price at the end of year *t-1*. Includes normal and extra dividends, but not special dividends.
D_Loss_t	Indicator variable equal to one for firms that have negative earnings before extraordinary items in year *t* and zero otherwise.
$DNOA_t$	Change in net operating assets in year *t* divided by average total assets in year *t*.
$MACRO_t$ *	The sum product of a firm's geographic sales exposure to a country (based on the segment reporting in the annual report for year *t*) and the GDP growth forecast of the country for year *t+1* (from IMF World Economic Outlook published in April in year *t+1*).
NOA	Net operating assets, computed as the difference between operating assets and operating liabilities. Operating assets is total assets less cash and short-term investments. Operating liabilities is total liabilities less the sum of short- and long-term debt.
$RNOA_t$ *	Return on net operating assets in year *t*, computed as the ratio of operating income before unusual items to average net operating assets.
$Size_t$	Natural logarithm of equity market capitalization, measured in USDm at the end of year *t*.

Note: Table A6 provides a detailed description of the independent variables and other data items used in the study. Variables denoted with * are expressed in percentage form.

Table A7. Skewness-Kurtosis test for normality in the distribution of paired differences in absolute and squared forecast errors

Panel A: Paired differences in forecast errors between SIMPLE and SIMPLE_MACRO					
				Joint test	
	N	Pr(skewness)	Pr(kurtosis)	Adj. chi2(2)	Prob > chi2
Differences in absolute forecast error	728	0.0455	0.0000	18.7700	0.0001
Differences in squared forecast error	728	0.0000	0.0000	.	.
Panel B: Paired differences in forecast errors between SIMPLE_MACRO and ADVANCED_MACRO					
				Joint test	
	N	Pr(skewness)	Pr(kurtosis)	Adj. chi2(2)	Prob > chi2
Differences in absolute forecast error	728	0.0063	0.0002	18.4000	0.0001
Differences in squared forecast error	728	0.0000	0.0000	.	.
Panel C: Paired differences in forecast errors between SIMPLE_MACRO and analyst forecasts					
				Joint test	
	N	Pr(skewness)	Pr(kurtosis)	Adj. chi2(2)	Prob > chi2
Differences in absolute forecast error	495	0.0000	0.0000	482.2700	0.0000
Differences in squared forecast error	495	0.0000	0.0000	788.5700	0.0000

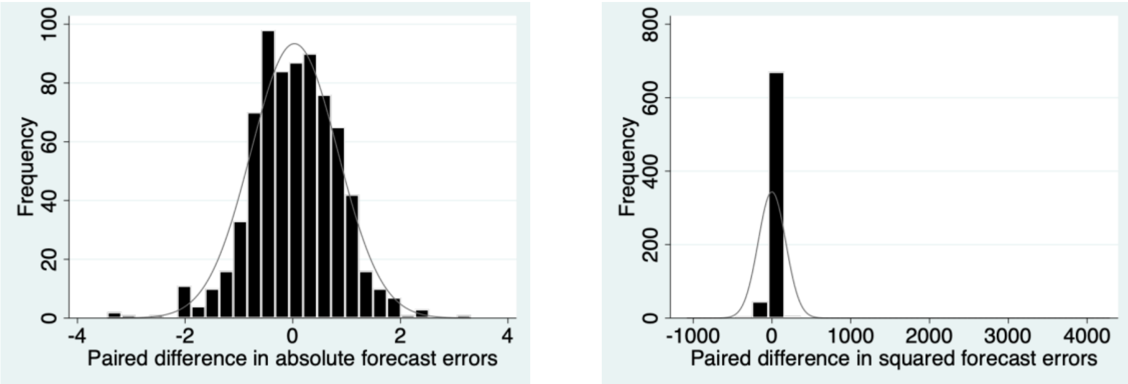
Note: Table A7 presents a Skewness-Kurtosis test for normality in the distribution of paired differences in absolute and squared forecast errors. The p-values are very low, meaning that the null hypothesis of normal distribution can be rejected. The paired differences in panel A are calculated as the absolute forecast error for SIMPLE less the equivalent for SIMPLE_MACRO, and the squared forecast error for SIMPLE less the equivalent for SIMPLE_MACRO. The paired differences in panel B are calculated as the absolute forecast error for SIMPLE_MACRO less the equivalent for ADVANCED_MACRO, and the squared forecast error for SIMPLE_MACRO less the equivalent for ADVANCED_MACRO. The paired differences in anel C are calculated as the absolute forecast error for SIMPLE_MACRO (the best model) less the equivalent for analyst forecasts, and the squared forecast error for SIMPLE_MACRO less the equivalent for analyst forecasts.

Table A8. Shapiro-Wilk W test for normality in the distribution of paired differences in absolute and squared forecast errors

Panel A: Paired differences in forecast errors between SIMPLE and SIMPLE_MACRO					
	N	W	V	z	Prob > z
Differences in absolute forecast error	728	0.9894	5.0320	3.9490	0.0000
Differences in squared forecast error	728	0.1426	405.1660	14.6730	0.0000
Panel B: Paired differences in forecast errors between SIMPLE_MACRO and ADVANCED_MACRO					
	N	W	V	z	Prob > z
Differences in absolute forecast error	728	0.9896	4.9240	3.8960	0.0001
Differences in squared forecast error	728	0.0202	462.9820	14.9990	0.0000
Panel C: Paired differences in forecast errors between SIMPLE_MACRO and analyst forecasts					
	N	W	V	z	Prob > z
Differences in absolute forecast error	495	0.4365	187.8390	12.5800	0.0000
Differences in squared forecast error	495	0.1055	298.2010	13.6900	0.0000

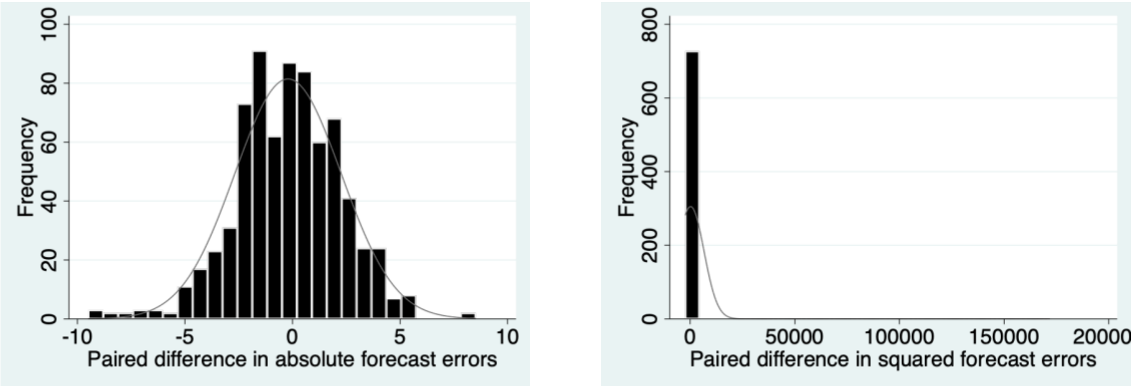
Note: Table A8 presents a Shapiro-Wilk W test for normality in the distribution of paired differences in absolute and squared forecast errors. The p-values are very low, meaning that the null hypothesis of normal distribution can be rejected. The paired differences in panel A are calculated as the absolute forecast error for SIMPLE less the equivalent for SIMPLE_MACRO, and the squared forecast error for SIMPLE less the equivalent for SIMPLE_MACRO. The paired differences in panel B are calculated as the absolute forecast error for SIMPLE_MACRO less the equivalent for ADVANCED_MACRO, and the squared forecast error for SIMPLE_MACRO less the equivalent for ADVANCED_MACRO. The paired differences in panel C are calculated as the absolute forecast error for SIMPLE_MACRO (the best model) less the equivalent for analyst forecasts, and the squared forecast error for SIMPLE_MACRO less the equivalent for analyst forecasts.

Figure A1. Histogram of the frequency distribution of paired differences in absolute and squared forecast errors between SIMPLE and SIMPLE_MACRO



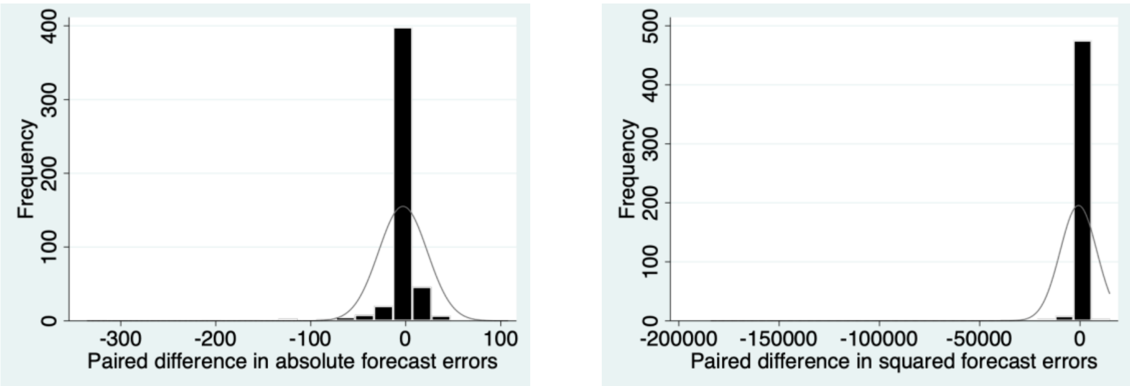
Note: Figure A1 presents histograms of the frequency distribution of paired differences in absolute and squared forecast errors between SIMPLE and SIMPLE_MACRO. The left chart shows tendencies of normal distribution but the evidence is deemed too weak when also considering the test results in Table A7 and Table A8. The distribution in right chart has a too high peak to be considered normally distributed. The paired differences presented in the charts are calculated as the absolute forecast error for SIMPLE less the equivalent for SIMPLE_MACRO, and the squared forecast error for SIMPLE less the equivalent for SIMPLE_MACRO.

Figure A2. Histogram of the frequency distribution of paired differences in absolute and squared forecast errors between SIMPLE_MACRO and ADVANCED_MACRO



Note: Figure A2 presents histograms of the frequency distribution of paired differences in absolute and squared forecast errors between SIMPLE_MACRO and ADVANCED_MACRO. The left chart shows tendencies of normal distribution but the evidence is deemed too weak when also considering the test results in Table A7 and Table A8. The distribution in right chart has a too high peak and the observations are also too centered to be considered normally distributed. The paired differences presented in the charts are calculated as the absolute forecast error for SIMPLE_MACRO less the equivalent for ADVANCED_MACRO, and the squared forecast error for SIMPLE_MACRO less the equivalent for ADVANCED_MACRO.

Figure A3. Histogram of the frequency distribution of paired differences in absolute and squared forecast errors between SIMPLE_MACRO and analyst forecasts



Note: Figure A3 presents histograms of the frequency distribution of paired differences in absolute and squared forecast errors between SIMPLE_MACRO and analyst forecasts. The distributions in both charts have too high peaks to be considered normally distributed. The paired differences presented in the charts are calculated as the absolute forecast error for SIMPLE_MACRO less the equivalent for analyst forecasts, and the squared forecast error for SIMPLE_MACRO less the equivalent for analyst forecasts.