The Relationship Between Earnings Management and Model-Based Earnings Forecast Accuracy*

Luca Brunke[a] and Tim Vater[b]

Cologne, March 2022

Abstract

The aim of this study is to analyze the relationship between earnings management and model-based earnings forecast accuracy. We provide evidence that firms with higher level of earnings management tend to exhibit larger forecast errors, i.e., earnings forecast accuracy for these firms is lower. Further, we show that accounting for the level of earnings management in earnings forecast models increases forecast accuracy, which then translates to more reliable ICC estimates.

Keywords: Earnings Prediction, Earnings Management, Earnings Quality
JEL Classification: G17, M41

[a]University of Cologne, Germany, Albertus-Magnus-Platz, 50923 Köln, Germany; E-mail: luca.brunke@wiso.uni-koeln.de

[b]University of Cologne, Germany, Albertus-Magnus-Platz, 50923 Köln, Germany; E-mail: vater@wiso.uni-koeln.de

*For their insightful discussions and suggestions, we are grateful to Dieter Hess, Simon Wolf, and Frederik Simon.

Contents

List of Tables

List of Figures

1 Introduction

Earnings are a central measure of a firm's performance. Hence, it is of special interest for investors, analysts, and firms themselves to obtain accurate information about future earnings (Tian, Yim and Newton (2021)). For practitioners and academics alike, earnings forecasts are an important input for firm valuation, asset allocation, or cost of capital calculation (Azevedo, Bielstein and Gerhart (2021)). In recent years, research on cross-sectional model forecasts as an alternative to analysts' earnings forecasts emerged, focusing particularly on using these forecasts for the computation of the implied cost of capital (ICC). The ICC is an expected return proxy and it is computed as the discount rate that equates expected future cash flows to current stock price. Several studies provide evidence that model-based ICCs are more reliable expected return proxies than analyst-based ICCs (e.g., Hou, Van Dijk and Zhang (2012) and Hess, Meuter and Kaul (2019)).

A common denominator in most earnings forecast models is that last period's reported earnings are a key explanatory variable for future earnings (e.g., Hou, Van Dijk and Zhang (2012) and Li and Mohanram (2014)). This is unsurprising, as previous literature finds that earnings are highly persistent (e.g., Fama and French (2006) and Hou and Van Dijk (2019)). Thus, the reliability of the reported earnings figure is likely related to the predictive ability of the forecast models. However, among others, one factor affecting reported earnings, and thus earnings forecasts, has not been covered by research on model-based earnings forecasts, yet. This factor is the extent of a firm's earnings management (EM). A widely accepted definition of EM is the adjustment of financial reports in order to deceive certain stakeholders about a firm's economic performance or to affect contractual obligations that are based on reported financial numbers (e.g., Healy and Wahlen (1999), Dechow and Skinner (2000) and Lo (2008)). Hence, the occurrence of EM, i.e., intentionally modification earnings, should intuitively compromise the reliability of reported earnings. This assumption is further supported when looking at managers' incentives to manipulate earnings (Dechow, Sloan and Sweeney (1996) and Dechow and Schrand (2004)). For instance, managers use EM to increase stock prices before initial public offerings, to meet analysts' earnings targets or to maximize bonuses that are based on the respective earnings. Literature provides evidence for the occurrence of EM as a response to these incentives (e.g., Healy (1985), Perry and Williams (1994), Teoh, Welch and Wong (1998) and Doyle, Jennings and Soliman (2013)). Teoh and Wong (2002) provide evidence that discretionary accounting accruals are an important determinant for earnings surprises. These studies provide a first indication that EM reduces the reliability of reported earnings and in turn possibly negatively affect the accuracy of model-based earnings forecasts.

Consequently, with our paper, we aim to examine the relationship between EM and the predictability of future earnings. Higashikawa (2020) studies a similar relationship in his work, whereas he investigates the relationship between earnings quality measures, e.g. smoothness or persistence, and the respective earnings forecast accuracy. He finds that a higher earnings quality is associated with a better forecast accuracy, i.e., a lower earnings forecast error. However, there are two important differences between the study by Higashikawa (2020) and our study: First, compared to Higashikawa (2020) who uses the earnings forecast model by Hou, Van Dijk and Zhang (HVZ), we rely on the residual income (RI) earnings forecast by Li and Mohanram (2014) model in our study. To us, the use of the HVZ model appears somewhat counter-intuitive since former studies found the RI model to perform better in terms of forecasting accuracy (e.g., Li and Mohanram (2014) and Hess, Meuter and Kaul (2019)). Second, we do not incorporate descriptive earnings quality measures which Higashikawa (2020) applies, but EM figures which aim to detect consequences of managers' manipulations.[1] In other words, behind the concept of earnings quality all choices made inside a firm are hidden, while EM specifically focuses on the choices made and actions taken by managers to modify earnings. Further, we seek to use the relation between a firm's extent of EM and the respective earnings forecast accuracy to improve the predictive ability of earnings forecasts models. That is, we incorporate information about firms' EM in the earnings forecast approach and furthermore evaluate if this results in more accurate forecasts and more finally in more reliable ICC estimates.

For our primary analysis, we require measures of (i) earnings forecast accuracy and (ii) the extent of a firm's EM. To evaluate forecast accuracy, we first generate earnings forecasts for up to three years ahead using the RI model by Li and Mohanram (2014). Then, we calculate the price-scaled absolute forecast error (PAFE). Firms manage earnings either through the manipulation of cash flows or accruals (Dechow and Schrand (2004)). In line with the bigger part of previous literature, we focus on the accruals component and use absolute discretionary accruals to measure the degree of firms' EM (e.g., Frankel, Johnson and Nelson (2002), Klein (2002), Bergstresser and Philippon (2006), among others). Discretionary accruals are de-

[1]Note that manipulation in this case is neither connotated in a good nor a bad way.

fined as the residuals from the estimation of an accruals model. We use the modified Jones (1991) model by Dechow, Sloan and Sweeney (1995) for the estimation.[2]

The results of the empirical analysis support our assumption, i.e., we provide evidence for a negative relationship between the extent of a firm's EM and the ability to accurately forecast its respective earnings. When running annual cross-sectional regressions of the firm-year specific PAFE on the respective EM measure for one-, two-, and three-year ahead earnings forecasts, we find significantly positive average parameter estimates of 0.0204, 0.0189, and 0.0182, respectively. In other words, we provide empirical evidence that a higher level of EM corresponds to less accurate model-based earnings forecasts. Subsequently, we capitalize on this finding and use the relation between EM and the predictability of future earnings to improve forecast accuracy. We annually rank firms into quintiles based on the extent of a firm's EM and create five dummy variables that indicate a firm's respective quintile. We then interact the earnings forecast model with the EM quintile dummy variables. Again, we generate earnings forecasts for up to three years ahead and find that the forecasts of the interacted model show significantly lower PAFEs compared to the initial RI model. For instance, for one-year (two-year, three-year) ahead forecasts, the median PAFE of the RI model is 3.72% (4.88%, 6.41%), whereas the PAFE of the interacted model is 3.18% (4.58%, 5.64%). Further, analogous to the methodology used in previous studies (e.g., Hou, Van Dijk and Zhang (2012) and Li and Mohanram (2014)), we implement an ICC analysis and provide evidence that ICCs based on the interacted model are more reliable expected return proxies in comparison to ICCs based on the RI model. For the cross-section of firms, we annually regress realized future returns on ICCs. We show that ICCs based on the interacted model exhibit higher correlations to realized future returns. For example, for one-year ahead forecasts, the RI model shows an average parameter estimate of 0.1904 and an R^2 of 10.70%, while the interacted model shows values of 0.2176 and 12.80%, respectively. Moreover, we annually rank firms into deciles based on the ICCs and implement a long-short-strategy, i.e., we compute the spread between the highest and lowest decile. We find that this portfolio approach yields higher returns for holding periods of up to three years when using ICCs based on the interacted model compared to using ICC estimates based on the RI model (e.g., 12.32% vs. 10.63% for a one year holding period). Lastly, we ensure that our findings are robust to

[2]We further elaborate on the selection of the earnings forecast model as well as on the selection of this specific accruals model in section 2.

alternative underlying earnings forecast models. We rerun the previous tests and provide evidence that the tenor of results is unchanged when using the earnings persistence (EP) model by Li and Mohanram (2014) and the model by HVZ (2012).

Our findings contribute to the literature as follows. First, to our knowledge, we are the first to examine the relationship between the extent of a firm's EM and the possibility to accurately forecast its respective earnings figure. In line with our expectations, we provide evidence for a significantly negative relationship. That is, when the level of a firm's EM increases, the PAFE seems to increase as well. Our results suggest that a firm's extent of EM should be considered when generating model-based earnings forecasts. It improves forecast accuracy and results in more reliable ICCs that yield higher investment strategy returns. This is important as it supports previous research (e.g., Hou, Van Dijk and Zhang (2012) and Hess, Meuter and Kaul (2019)) that identifies model-based earnings forecasts as a viable alternative to analysts' earnings forecasts. Finally, our findings add to the debate on managers' incentives for EM. Beneish (2001) points out that there are potentially two perspectives on EM. On the one hand, the "informative perspective" suggests that managerial discretion is used to reveal private expectations about future cash flows to stakeholders. For our analysis, this could improve the information content of reported earnings and lead to more accurate earnings forecasts. However, there is no empirical evidence for this perspective, and our results do not support it either. On the other hand, the "opportunistic perspective" states that managers manipulate earnings to mislead investors with the intention of obtaining personal gain. This should impair the reliability of reported earnings, resulting in less accurate earnings forecasts. Our findings match this perspective, and therefore support the results of previous studies focusing on opportunistic managers' actions (e.g., Perry and Williams (1994), Teoh, Welch and Wong (1998), and Bergstresser and Philippon (2006)).

The remainder of this paper is structured as follows. Section 2 provides a brief overview of related literature. Section 3 outlines the methodology and section 4 describes the data we use for our empirical analysis. Section 5 covers the empirical results and section 6 concludes.

2 Related Literature

This section provides an overview regarding the literature related to our study. First, we present studies focusing on cross-sectional earnings forecasts and their relation to ICCs.[3] Second, we briefly discuss studies that implement models to estimate discretionary accruals as a measure for the extent of a firm's EM.

Model-Based Earnings Forecasts and Implied Cost of Capital

Information about the expected rate of return is crucial in various economic settings, e.g., to ensure an efficient allocation of scarce resources or capital budgeting (e.g., Botosan and Plumlee (2005) and Lee, So and Wang (2021)). There exists a vast amount of literature on different approaches for deriving an estimate of a firm's expected rate of return. It is well documented that using realized returns to proxy for expected returns bears a range of problems and leads to noisy and biased estimates (e.g., Fama and French (1997) and Easton and Monahan (2016)). Thus, in recent years, a stream of literature that approximates the expected rate of return with the ICC emerged (e.g., Gebhardt, Lee and Swaminathan (2001), Claus and Thomas (2001), Ohlson and Juettner-Nauroth (2005), and Easton (2004)). An advantage of the ICC estimation is that it does not rely on noisy realized returns to derive a proxy for expected returns (Lee, So and Wang (2010)). Although it is an important source of information for researchers and practitioners alike, the ICC of a firm itself is unobservable. As mentioned before, it is defined as the internal rate of return that results from equating the current stock price to the present value of expected future cash flows. Whereas the current stock price is directly observable, information about future cash flows has to be approximated. In order to derive a reliable ICC estimate, this approximation relies heavily on the accuracy of the respective input factors, especially unobservable future cash flows (Botosan and Plumlee (2005)). While future cash flows are usually proxied by future earnings, future earnings itself are unobservable as well.

Literature provides two popular options to derive estimates of a firm's future earnings. On the one hand, for a subsample of firms, analyst forecasts of the respective firm's earnings are available. Easton and Monahan (2005) show that more reliable ICCs are the result of more accurate analysts' forecasts. Thus, they provide evidence for the necessity of accurate input factors for the ICC estimation.

[3]Throughout this paper, we will use the terms "cross-sectional" and "model-based" earnings forecasts interchangeably.

However, since analysts mainly cover larger firms or more generally for a smaller subsample of companies, various models to forecast future earnings emerged (e.g., Hou, Van Dijk and Zhang (2012) and Li and Mohanram (2014)). These models allow to cover all firms with available financial statement data. The majority of recent studies on model-based earnings forecasts implements a cross-sectional estimation approach. While the model-based forecasts show lower forecast accuracy, Hou, Van Dijk and Zhang (2012) find that these forecasts beat analysts' earnings forecasts in terms of coverage, forecast bias and earnings response coefficient. Further, ICCs based on cross-sectional earnings forecasts are more reliable expected return proxies than analyst-based ICC estimates. Thus, Hou, Van Dijk and Zhang (2012) provide evidence that suggests deriving ICC estimates from model-based earnings forecasts rather than from analyst forecasts. Whereas Easton and Monahan (2016) question these results, Hess, Meuter and Kaul (2019) confirm the superiority of model-based earnings forecast in terms of the reliability of ICC estimates. However, the puzzle why mechanical earnings forecast models result in less accurate forecasts compared to analyst earnings forecasts, but in more reliable ICC estimates, remains unanswered at this point (Hess, Meuter and Kaul (2019)). Additionally, Gerakos and Gramacy (2013) as well as Li and Mohanram (2014) note that the forecast errors resulting from the Hou, Van Dijk and Zhang (2012) model are quite similar to or even worse than those derived from a random walk model. They express doubt whether the forecasts from that model should be used at all. Thus, Li and Mohanram (2014) propose two new models to improve the approach of Hou, Van Dijk and Zhang (2012) by differentiating between the earnings persistence of profit and loss firms, adjusting the earnings metric for special items, and estimating earnings per share instead of firm-level earnings. They provide evidence that their adjusted model outperforms the model by Hou, Van Dijk and Zhang (2012) regarding forecast bias, accuracy, earnings response coefficient, and ICC reliability.

Evans, Njoroge and Yong (2017) and Tian, Yim and Newton (2021) show that using the least absolute deviation method, i.e., median regressions, further improves forecast performance. However, since our analysis is mainly concerned with the relation between EM and mean earnings forecast accuracy, we follow Hou, Van Dijk and Zhang (2012) and Li and Mohanram (2014) and employ the ordinary least square (OLS) method.[4]

[4]Untabulated tests show that our results remain unchanged when median regressions are used.

In addition to forecasting mean or median earnings, Konstantinidi and Pope (2016) and Chang, Monahan, Ouazad and Vasvari (2021) use quantile regressions to estimate the distribution of expected earnings. Using these estimates, they compute the higher moments of future earnings. They argue that these moments are measures of risk in future earnings and provide evidence that they are related to common risk measures such as credit risk ratings or corporate bond spreads. Although they also develop models to forecast future earnings, their work is mainly concerned with forecasting higher moments of future earnings and not with a mean forecast of earnings. Thus, in our study, we will not cover the models suggested by Konstantinidi and Pope (2016) and Chang, Monahan, Ouazad and Vasvari (2021) due to a deviating research focus. Additionally, both studies do not provide evidence that their models outperform established mean earnings forecast models in terms of forecasting accuracy.

Throughout our empirical analysis, we focus on the RI model introduced by Li and Mohanram (2014), since previous studies find that it performs best in terms of forecast accuracy. However, we will disclose the results based on the EP model by Li and Mohanram (2014) and the HVZ model by Hou, Van Dijk and Zhang (2012) in section A.3 and A.4, respectively. Our main findings are robust to changes in the underlying earnings forecast model.

Estimation of the Earnings Management Measure

A widely accepted definition of EM in previous studies is the adjustment of financial reports in order to deceive certain stakeholders about a firm's economic performance or to affect contractual obligations that are based on reported financial numbers (e.g., Healy and Wahlen (1999) and Dechow and Skinner (2000)). However, this concept is difficult to measure directly, as it focuses on unobservable managerial intent (Dechow and Skinner (2000)). The most common approach to measure EM is isolating the discretionary part of accruals (Dechow, Hutton, Kim and Sloan (2012)). This part of accruals reflects distortions due to active EM, while the non-discretionary part captures adjustments based on fundamental performance (Dechow, Ge and Schrand (2010)). Estimates of discretionary accruals are obtained by directly modeling the accruals process. Widely used accruals models are developed by Jones (1991), Dechow, Sloan and Sweeney (1995), Dechow and Dichev (2002), McNichols (2002) and Dechow, Hutton, Kim and Sloan (2012)).

Jones (1991) analyzes whether firms use EM to decrease earnings during import relief investigations. Her model includes total accruals as dependent variable and change in revenues and property, plants and equipment as independent variables. The fitted value of the regression represents non-discretionary accruals and the residual represents discretionary accruals. Jones (1991) finds that managers actively decrease earnings to profit from import reliefs. Dechow, Sloan and Sweeney (1995) point out that the model by Jones (1991) implicitly assumes that revenues are non-discretionary. In consequence, if EM occurs through discretionary revenues, it is not accounted for in the discretionary accruals estimate. Dechow, Sloan and Sweeney (1995) propose a solution by modifying the model by Jones (1991). They use cash revenue instead of reported revenue, i.e., the change in revenues is adjusted for change in receivables. They provide empirical evidence that the modified model better detects EM compared to the initial model by Jones (1991). Dechow and Dichev (2002) suggest a new measure for accruals and earnings quality. While they do not explicitly intent to measure EM, their measure is based on the standard deviation of the residuals, i.e., discretionary accruals. Their model includes change in working capital as dependent variable and past, current and future cash flows as independent variables. They find that a larger standard deviation of discretionary accruals results in less persistent earnings, longer operating cycles and more volatile cash flows, accruals and earnings (Dechow, Ge and Schrand (2010)). McNichols (2002) links the approach of Jones (1991) to Dechow and Dichev (2002). She adds the variables of Jones (1991) to the model by Dechow and Dichev (2002) and shows that the explanatory power regarding working capital accruals increases. Moreover, Francis, LaFond, Olsson and Schipper (2005) use this model to compute the accruals quality measure proposed by Dechow and Dichev (2002). However, they further differentiate between accruals quality due to economic fundamentals and due to management choices. They find that lower accruals quality yields higher cost of debt, smaller price multiples on earnings and larger equity betas. Yet, they conclude that accruals quality driven by economic factors has a larger effect on cost of capital than accruals quality driven by management choices.

In this study, we use the modified Jones (1991) model by Dechow, Sloan and Sweeney (1995) to compute the EM measure for the following reasons: First, we exclude the original Jones (1991) model from the set of possibly applicable accruals models, because, as stated before, Dechow, Sloan and Sweeney (1995) show that their modified model better detects EM. Second, the accruals models by Dechow and Dichev (2002) and McNichols (2002) appear to be neither suitable for our specific research design since both models contain cash flows from period $t + 1$ as an explanatory variable for the discretionary accruals in period t. In other words, those two models incorporate information from a future period in order to model accruals in the current period. This induces a timing problem, because we aim to investigate the relationship between the firm's earnings management and the resulting earnings forecast error in the following period. A conceptual mismatch follows if we on the one hand calculate an earnings management measure for period t with information from period $t+1$ and at the same time pretend to not have information for period $t+1$ when forecasting earnings for that period. To prevent such a look-ahead-bias in our analysis, we would have to relate the forecast error of current period's earnings forecasts to last period's EM measure. However, we want to avoid such a timing lag between both measures. Additionally, Dechow and Dichev (2002) point out that their model is not specifically intended to estimate firms' EM. Based on those two arguments, we decided to also exclude the two accruals models by Dechow and Dichev (2002) and McNichols (2002) from our analysis. Third, the accruals quality measure of Francis, LaFond, Olsson and Schipper (2005) that is driven by management choices requires a seven-year time-series of firm-specific data. This potentially induces a survivorship bias that we intent to avoid. Finally, Dechow, Hutton, Kim and Sloan (2012) suggest caution when using their performance matching approach for detecting earnings management. They claim that this approach is in general only effect if knowledge about correlated omitted variables can be used to identify appropriate matched pairs. Since we do not have information about such, we refrain from implementing that approach. Additionally, according to Dechow, Hutton, Kim and Sloan (2012) performance matching entails a significant reduction in test power. This selection process leaves us with the modified Jones (1991) model as the best suited accruals model, which we thus in the following base our empirical analysis on.

To the best of our knowledge, we are the first to analyze the relationship between the extent of a firm's EM and model-based earnings forecast accuracy. Only the study by Higashikawa (2020) has a similar setup, since it investigates the relationship between earnings quality measures and earnings forecast accuracy. As elaborated before there are multiple important differences between the two studies. First, we study the influence of EM on earnings forecast accuracy not of earnings quality measures. This differentiation is important since both concepts cover different information. Whereas EM aims to specifically detect managers' earnings manipulation, earnings quality measures mainly describe earnings characteristics which are the result of all choices made within a firm. Second, somewhat counter-intuitive, Higashikawa (2020) uses the HVZ model instead of the in terms of forecast accuracy better performing RI model. Finally, Higashikawa (2020) does neither study the possibility of improving earnings forecast by incorporating insights about the studied relationship into earnings forecast models nor the implications of such improvements for the model-based ICC computation.

As noted in the previous section, model-based earnings forecasts are an important measure in practice as well as in academic studies. Thus, understanding the factors influencing their accuracy is worth investigating further. In the following, we present the methodology we applied to study such relationship.

3 Methodology

This section outlines the methodology we employ in this study. First, it shows how we generate earnings forecasts and the corresponding PAFEs. Second, it presents how we compute the EM measure, i.e., absolute discretionary accruals. Third, it depicts how we (i) examine the relation between the extent of a firm's EM and earnings forecast accuracy, (ii) use information about firms' EM to improve the predictive ability of earnings forecast models, and (iii) test if this information enhances ICC reliability.

Model-Based Earnings Forecasts

To forecast earnings, we use the RI model introduced by Li and Mohanram (2014). The model is defined as follows:

$$\begin{aligned} Earn_{i,t+\tau} = \beta_0 + \beta_1 Earn_{i,t} + \beta_2 NegE_{i,t} + \beta_3 NegExE_{i,t} \\ + \beta_4 BkEq_{i,t} + \beta_5 TACC_{i,t} + \epsilon_{i,t+\tau}, \end{aligned} \tag{1}$$

where $Earn$ reflects earnings, $NegE$ is a dummy variable that equals 1 for firms with negative earnings and 0 otherwise and $NegExE$ is an interaction term of the dummy variable and earnings. Further, $BkEq$ is the book value of equity, $TACC$ reflects total accruals, t represents the time index and τ is a time constant. If not stated differently, all variables in our analysis are scaled by the number of shares outstanding. We forecast earnings for up to five years ahead, i.e., for $\tau = 1 - 5$.[5]

[5] For our analysis, we primarily use one-, two- and three-year ahead forecasts. Four- and five-year ahead forecasts are needed for the ICC computation in Section 5.3.

In line with cross-sectional earnings forecast literature (e.g., Hou, Van Dijk and Zhang (2012) and Li and Mohanram (2014)), we use a rolling OLS regression approach with a ten-year window to generate the earnings forecasts.[6] First, at the end of June of each year of our sample period, data from year $t-9$ to year t is used to estimate the model parameters. Second, we multiply the computed parameters with the independent variables from year t to obtain firm-specific earnings estimates for year $t+\tau$. Out-of-sample earnings forecasts are available from 1979 onwards. To evaluate forecast accuracy, we use the PAFE (e.g., Hou, Van Dijk and Zhang (2012) and Li and Mohanram (2014)), which is defined as follows:

$$PAFE_{i,t+\tau} = \left| \frac{Earn_{t+\tau} - \widehat{Earn}_{t+\tau}}{prc_t} \right|, \tag{2}$$

where $\widehat{Earn}$ is the model-based earnings forecast and prc is the end-of-June stock price.

Earnings Management Measure

In line with previous literature (e.g., Frankel, Johnson and Nelson (2002), Klein (2002) and Bergstresser and Philippon (2006)), we use absolute discretionary accruals as a measure for the extent of a firm's EM. Discretionary accruals are defined as the residuals from the estimation of an accruals model. To compute non-discretionary accruals, we use the modified Jones (1991) model by Dechow, Sloan and Sweeney (1995):

$$TACC_{i,t} = \beta_0 + \beta_1(\Delta REV_{i,t} - \Delta REC_{i,t}) + \beta_2 PPE_{i,t} + \epsilon_{i,t}, \tag{3}$$

where ΔREV is the change in revenue, ΔREC is the change in receivables and PPE reflects property, plant and equipment. All variables are scaled by the number of shares outstanding.[7] Further, as specified by Jones (1991) and Dechow, Sloan and Sweeney (1995), the intercept is also scaled, i.e., the true constant term is suppressed (Peasnell, Pope and Young (2000)).

Following more recent studies (e.g., Chung and Kallapur (2003), Francis, LaFond, Olsson and Schipper (2005) and Bergstresser and Philippon (2006) among

[6] To lower data requirements, we start with a five-year window at the beginning of the sample period and expand the window to ten years successively.

[7] We scale our variables by the number of shares outstanding to be consistent with the variable definition of the earnings forecast model. Thereby, we deviate from the variable definition of Jones (1991) and Dechow, Sloan and Sweeney (1995). They scale all variables by lagged total assets to reduce heteroscedasticity. Following the approach of Jones (1991), untabulated tests show that the error term of the unscaled accruals model is also highly correlated with the number of shares outstanding. This indicates that scaling by the number of shares outstanding is also reasonable.

others), we implement a cross-sectional approach instead of time-series analysis initially employed by Jones (1991). Comparing cross-sectional to time-series accruals models, Bartov, Gul and Tsui (2000) find that only cross-sectional models are constantly able to detect EM. Further, accruals models are frequently estimated at industry level (Dechow, Ge and Schrand (2010)). We follow this approach and employ the Fama and French 48 industry classification.[8]

Similar to the model-based earnings forecasts, we use rolling OLS regressions with a ten-year window to estimate the model.[9] First, model parameters are computed using data from year $t-9$ to year t. Second, the computed parameters are multiplied with the independent variables from year t to obtain an estimate of non-discretionary accruals for year t, which in the following is represented by $\widehat{TACC}$. Lastly, subtracting this estimate from respective actual total accruals $TACC$ provides an estimate for discretionary accruals. The absolute value of discretionary accruals serves as our measure for the extent of a firm's EM, depicted in the following by EM. This measure is available from 1975 onwards and defined as follows:

$$EM_{i,t} = \left| TACC_{i,t} - \widehat{TACC}_{i,t} \right|. \tag{4}$$

The Relationship Between EM and Earnings Forecast Errors

First, we test the relation between the extent of a firm's EM and model-based earnings forecast accuracy using the following regression equation:

$$PAFE_{i,t+\tau} = \beta_0 + \beta_1 EM_{i,t} + \sum_{k=2}^{3} \beta_k Control_{i,t,k} + \epsilon_{i,t+\tau}. \tag{5}$$

We explicitly control for firm size by including the logarithm of total assets and for industry by adding industry dummies according to the Fama and French 48 industry classification.[10] We run annual cross-sectional OLS regressions for $\tau = 1-3$.[11]

Second, to examine whether the EM measure helps to improve the predictive ability of earnings forecast models, we use the following approach: We annually

[8] Untabulated tests show that the tenor of results is unchanged when we do not estimate the accruals model at industry level. However, we follow the approach which is dominantly used in the EM literature.

[9] Analogous to the earnings forecasts, we start with a five-year window at the beginning of the sample period and expand the window to ten years successively.

[10] Ecker, Francis, Olsson and Schipper (2013) identify firm size as a potentially important correlated omitted variable in tests for EM.

[11] Using the estimated EM measure as independent variable potentially induces an "error-in-variables" bias. That is, the regression coefficient of the EM measure might be biased towards zero (Griliches and Ringstad (1970)). Hence, our empirical results might understate the true effect of EM on forecast accuracy. However, an even higher true effect does not change the interpretation of our results.

rank firms into quintiles based on the extent of a firm's EM and create five dummy variables that indicate a firm's respective quintile. Next, we interact the earnings forecast model with the EM quintile dummy variables, i.e., we run a separate regression for each quintile subsample:

$$\begin{aligned} Earn_{i,t+\tau} = \sum_{k=1}^{5} Q_k(\beta_0 + \beta_1 Earn_{i,t} + \beta_2 NegE_{i,t} + \beta_3 NegExE_{i,t} \\ + \beta_4 BkEq_{i,t} + \beta_5 TACC_{i,t} + \epsilon_{i,t+\tau}) \end{aligned} . \tag{6}$$

The notation is analogous to equation 1 with the addition of the indicator variable Q representing the respective k^{th} EM quintile dummy variable. The variable is set to equal 1 if a firm belongs to the respective EM quintile and 0 otherwise. We rerun the analysis of the RI earnings forecast model and compare regression results and PAFEs of the initial RI model and our newly interacted model.

Third, we investigate if the earnings forecasts from the interacted model result in more reliable expected return proxies compared to the RI model. In line with earnings forecast literature (e.g., Hou, Van Dijk and Zhang (2012), Li and Mohanram (2014), and Azevedo, Bielstein and Gerhart (2021)), we use ICCs as a proxy for expected returns. The forecasted earnings are used as future cash flow proxies. Hence, more accurate forecasts should yield more reliable expected return proxies. Prior research has developed various ICC estimation methods. To guarantee that our results are not affected by any particular method, we follow the earnings forecast literature and use a composite ICC. Our ICC measure is the average of the following four commonly used ICC metrics (Azevedo, Bielstein and Gerhart (2021)). We use two ICCs based on a residual income model, i.e., metrics by Gebhardt, Lee and Swaminathan (2001) and Claus and Thomas (2001), and two ICCs based on an abnormal earnings growth model, i.e., metrics by Ohlson and Juettner-Nauroth (2005) and Easton (2004). We present a detailed description of the ICC metrics in section A.2 in the appendix. Following Hou, Van Dijk and Zhang (2012) and to increase coverage, we require only one ICC metric to be available to compute the composite ICC. We calculate the firm-specific composite ICC at the end of June of each year.

We analyze the relation of the composite ICC to future returns using two approaches commonly relied on in earnings forecast studies (e.g., Hou, Van Dijk and Zhang (2012) and Li and Mohanram (2014)). The first approach examines the relation at the firm-level using the following equation:

$$Ret_{i,t+1} = \beta_0 + \beta_1 ICC_{i,t} + \epsilon_{i,t+1}, \tag{7}$$

where Ret is the realized stock return at the end of June of the year $t+\tau$ and ICC is the one-year ahead composite ICC calculated at the end of June of the current year t for the end of June in year $t+\tau$. Using this equation, we run annual cross-sectional OLS regressions for $\tau = 1-3$. Values of β_1 closer to 1 imply a more reliable expected return proxy (Li and Mohanram (2014)).

The second approach evaluates the relation between the composite ICC and future returns on a portfolio level. In line with Hou, Van Dijk and Zhang (2012), we rank firms into decile portfolios based on the composite ICC at the end of June of each year. Next, we calculate the equally weighted buy-and-hold return for each decile portfolio for holding periods of up to three years. We mainly focus on the spread between the highest and lowest decile, i.e., implementing a long-short strategy (Azevedo, Bielstein and Gerhart (2021)). We test if this strategy results in significant returns and compare the realized returns based on the ICCs from the initial RI earnings forecast model to the ones retrieved from the interacted model. The idea behind this strategy is that more reliable ICCs result in a more accurate ranking of firms regarding their expected returns. Consequently, a more accurate ranking will yield higher returns from the long-short strategy.

4 Data

The sample we use for the empirical analysis consists of the intersection of the annual COMPUSTAT North American database and the monthly CRSP stock return file. It contains US American firms reporting in US dollar. The total sample period spans from 1971 to 2019. We implement a three-month reporting lag for firm fundamentals to become publicly available. Following previous literature (e.g., Dechow, Hutton, Kim and Sloan (2012)), we exclude financial firms (SIC codes 6,000 to 6,999) from our analysis as financial statements of these firms are subject to different regulatory frameworks.

The variables for the earnings forecast model are defined as follows. Earnings is income before extraordinary items (COMPUSTAT variable: IB) minus special items (SPI). Special items are set to zero if missing. Book equity is total common equity (CEQ). Total accruals are defined as income before extraordinary items (IB) minus cash flow from operations (OANCF). Since cash flow from operations is only available from 1988 onwards, we use the accruals definition of Richardson, Sloan, Soliman and Tuna (2005) in case of missing cash flow from operations (Li and

Mohanram (2014)).[12] To compute the PAFE, we take the price from the monthly CRSP stock return file (PRC). To estimate the accruals model, we use the following variables. Total accruals are defined analogously to the earnings forecast model. The change in revenue is current period's total revenue (REVT) minus total revenue from the previous period. Likewise, the change in receivables is current period's total receivables (RECT) minus total receivables from the previous period. Property, plant, and equipment is total gross property, plant, and equipment (PPEGT). For all models, variables are scaled by the number of common shares outstanding (CSHO). We require all relevant variables to be non-missing. Further, to mitigate the effect of outliers, we winsorize all variables annually at the 1st and 99th percentile.

To compute the ICC metrics, we further use the following variables. Earnings are defined analogous to the earnings forecast model. Book equity is total common equity (CEQ), dividends are common dividends (DVC) and total assets are set to be equal to the total assets measure (AT). These variables are scaled by the number of common shares outstanding (CSHO), too. The one-year buy-and-hold return is computed by compounding returns from the monthly CRSP stock return file (RET).

Table 1 presents descriptive statistics for the variables included in the earnings forecast model and for the EM measure.

Panel A shows summary statistics (cross-sectional mean, median, standard deviation and selected percentiles for firm-years with complete data) and Panel B displays Pearson and Spearman correlations. Our sample contains 164,337 firm-year observations. Similar to former studies, our sample includes around 30% of firms with negative earnings (e.g., Hou, Van Dijk and Zhang (2012) and Hess, Meuter and Kaul (2019)). Focusing on the EM measure, Panel A reveals that it is skewed to the right, i.e., the cross-sectional mean (1.09) is approximately twice as large as the median (0.54). Further, Panel B shows that Pearson (Spearman) correlations between the EM measure and the variables included in the earnings forecast model range between -0.24 and 0.07 (-0.16 and 0.41). We report positive correlations between the EM variable and earnings itself. Unsurprisingly, the EM variable is negatively correlated with the negative earnings dummy, which appears reasonable due to the negative correlation between earnings and the negative earnings dummy. Furthermore, total accruals are positively correlated with the EM variable.

[12]See table A7 for a more detailed description.

Table 1: Descriptive Statistics

Panel A: Summary Statistics

	N	Mean	Std	P1	P25	P50	P75	P99
Earn	164,337	1.00	12.84	-3.55	-0.07	0.57	1.65	7.63
NegE	164,337	0.29	0.45	0.00	0.00	0.00	1.00	1.00
NegExE	164,337	-0.24	0.80	-3.55	-0.07	0.00	0.00	0.00
BkEq	164,337	10.11	109.78	-2.62	2.25	6.11	12.82	50.32
TACC	164,337	0.55	2.89	-6.70	-0.05	0.12	1.14	8.48
EM	164,337	1.09	1.62	0.01	0.21	0.54	1.25	8.73

Panel B: Correlations

	Earn	NegE	NegExE	BkEq	TACC	EM
Earn	1.00	-0.78***	0.80***	0.71***	0.37***	0.27***
NegE	-0.09***	1.00	-0.98***	-0.50***	-0.35***	-0.16***
NegExE	0.09***	-0.48***	1.00	0.47***	0.35***	0.12***
BkEq	0.75***	-0.04***	0.01***	1.00	0.28***	0.41***
TACC	0.07***	-0.17***	0.19***	0.04***	1.00	0.02***
EM	0.05***	-0.06***	-0.24***	0.07***	-0.03***	1.00

Table 1 contains descriptive statistics for the pooled cross-section of firms from 1975 to 2019. Panel A displays summary statistics for the variables of the earnings forecast model and for the EM measure resulting from the accruals model by Dechow, Sloan and Sweeney (1995). Panel B presents the respective cross-correlations following Pearson (Spearman) below (above) the diagonal. ***, **, and * indicate significance at an alpha level of 1%, 5%, 10%, respectively.

5 Empirical Results

This section presents the empirical results. First, we provide evidence for a significant positive relation between the extent of a firm's EM and the respective model-based earnings forecast error. Second, we capitalize on this finding and use the EM measure to improve the predictive ability of earnings forecast models. Third, we show that the increased forecast accuracy results in more reliable expected return proxies. Lastly, we ensure that our findings are robust to different earnings forecast models.

5.1 The Relationship Between Earnings Management and Earnings Forecast Accuracy

First, we analyze the relation between the extent of a firm's EM and the accuracy of model-based earnings forecasts. We run annual cross-sectional regressions of PAFE on the EM measure while controlling for firm size and industry. Table 2 presents the results for forecast horizons of up to three years. It contains the time-series averages of the parameter estimates, Newey and West (1987) t-statistics and R^2 values.

Table 2: The Relationship Between EM and Earnings Forecast Accuracy

	PAFE_{t+1}	PAFE_{t+2}	PAFE_{t+3}
Coefficient	0.0204***	0.0189***	0.0182***
	(8.85)	(10.51)	(9.77)
R^2	0.1249	0.1334	0.1352
Controls	Yes	Yes	Yes

Table 2 depicts the relationship between EM and the RI model-based earnings forecast accuracy. It contains the time-series averages of the parameter estimates, Newey and West (1987) t-statistics and R^2 values from the annual regressions of PAFE on the EM measure. We control for firm size by including the logarithm of total assets and for industry by adding industry dummies according to the Fama and French 48 industry classification. ***, **, and * indicate significance at an alpha level of 1%, 5%, 10%, respectively.

In line with our expectations, the findings provide evidence for a significant positive relation between EM and forecast errors for all forecast horizons. That is, the higher the EM measure, the higher the PAFE, i.e., the lower the forecast accuracy. For one-, two-, and three-years ahead forecasts, the coefficient of the EM measure shows values of 0.0204, 0.0189, and 0.0182, respectively. Hence, the strength of the relation slightly decreases with an increasing forecast horizon. A possible explanation for such phenomenon includes two steps. First, the manipulation of earnings in the actual period negatively influences the earnings forecasts for the following periods, since those forecasts are made based on the modified earnings measure. Second, since forecasts tend to become less accurate with an increasing forecast horizon (e.g., Hou, Van Dijk and Zhang (2012) and Li and Mohanram (2014)), the proportion of the forecast error attributed to a management's modification of the earnings measure becomes less influential, which is reflected by the decreasing parameter estimates.

In general, the negative relation between EM and forecast accuracy we find indicates that managers' actions lower earnings' predictability. As pointed out in section 1, this could be related to an impaired quality of reported earnings due to opportunistic managerial discretion. Hence, our results are in line with previous studies finding that EM is performed with the intention of misleading stakeholders to obtain some personal gain (e.g., Perry and Williams (1994), Teoh and Wong (2002), and Bergstresser and Philippon (2006)), instead of aiming to increase the information content of reported earnings (Beneish (2001)).

Additionally, figure 1 plots the annual coefficients of the EM measure for one-, two-, and three-year ahead forecasts.

Figure 1: Relation Between Earnings Management and Earnings Forecast Accuracy Over Time

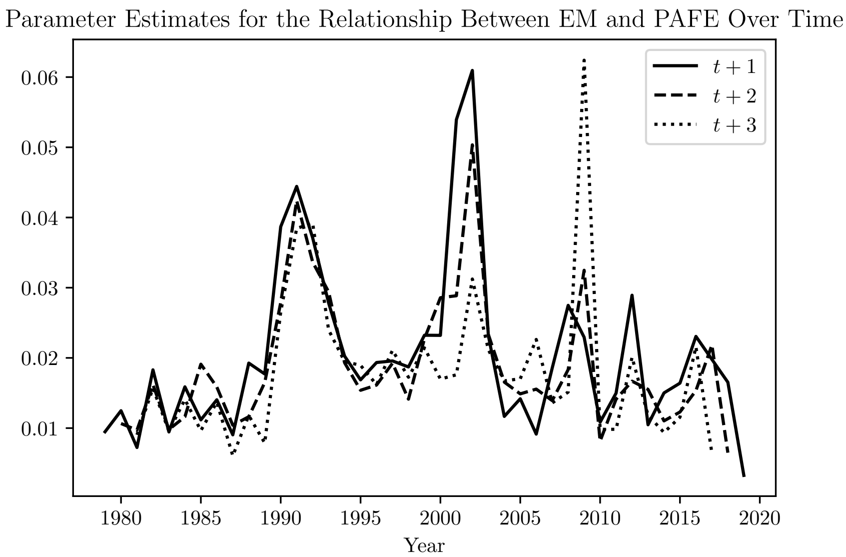

Figure 1 displays the influence of EM on model-based earnings forecast accuracy. It contains the annual parameter estimates from the regressions of PAFE on the EM measure for one-, two-, and three-year ahead forecasts. We further control for firm size and industry.

As can be seen, the coefficients approximately range between 0.01 and 0.06. Most importantly, figure 1 displays that the coefficients are entirely positive throughout the sample period, i.e., that the sign of the relation between EM and forecast accuracy is consistent. This further strengthens the significance of our findings. However, we add that it might bear interesting study opportunities to inspect the fluctuation of the coefficient's magnitude over time and leave that question for future research to be answered.

Although we tested and demonstrated a significant negative relationship between the firm's EM and the earnings forecast accuracy, future research might investigate the exact relationship in a more detailed way, i.e., the exact shape of such relationship. One could possibly as well find arguments for a U-shaped relationship. That is that firms in the extreme quantiles regarding their EM extent exhibit a lower earnings forecast accuracy compared to firms with a more moderate extent of EM. Thus, any extreme form of earnings management reduces the predictability of future earnings compared to a moderate earnings management. We leave testing this hypothesis or any other investigation of the exact form of the relationship between EM and forecast accuracy to future research.

5.2 Improving Earnings Forecasts with Information on Earnings Management

In this section, we make use of the insights gained from the previous section, i.e., that a higher level of EM is significantly related to larger earnings forecast errors. Based on this finding, we assume that firms' EM characteristics contain information that is important for predicting future earnings. More specifically, that the parameter estimates of earnings forecast models are influenced by the extent of a firm's EM. As outlined in section 3, we interact the RI model by Li and Mohanram (2014) with five EM quintile dummy variables to account for information about firms' EM. By interacting the model with the dummy variables, we allow for an additional variation of coefficients across EM quintiles. Thus, we expect to obtain more accurate parameter estimates for each subsample. We assume this approach translates to lower forecast errors on average compared to the initial RI earnings forecast model.

Table 3 on the next page presents results for the rolling earnings regressions for one, two, and three years ahead. It contains the time-series averages of the parameter estimates, Newey and West (1987) t-statistics and R^2 values for the initial RI model as well as for each of the five quintiles of the interacted model.

The first column covers the initial RI earnings forecast model, whereas columns two to six report results for each EM quintile subsample. Looking at the individual parameter estimates for each EM quintile, it becomes evident that they differ across each subsample as well as compared to the parameter estimates of the RI model. For example, for one-year ahead forecasts (Panel A), the RI model shows a lagged earnings parameter estimate of 0.79, whereas the EM quintiles exhibit larger coefficients ranging between 0.82 and 1. Similar patterns can be observed for two- and three-year ahead forecasts (Panel B and Panel C, respectively). For the negative earnings dummy, values for the EM quintiles are larger compared to the RI model,

Table 3: Parameter Estimates from the Earnings Forecast Regression

Panel A: $t + 1$

	RI	EM Q1	EM Q2	EM Q3	EM Q4	EM Q5
Intercept	0.26***	0.15**	0.01	0.01	0.01	0.13*
	(2.92)	(2.40)	(0.72)	(0.84)	(0.35)	(1.70)
Earn	0.79***	0.82***	1.00***	0.96***	0.94***	0.84***
	(21.43)	(14.01)	(171.57)	(80.19)	(168.63)	(29.74)
NegE	-0.44***	-0.21***	-0.08***	-0.07***	-0.08***	-0.28***
	(-5.36)	(-3.76)	(-6.72)	(-5.72)	(-5.79)	(-4.67)
NegExE	-0.33***	-0.01	-0.30***	-0.22***	-0.29***	-0.41***
	(-9.00)	(-0.10)	(-5.27)	(-3.64)	(-14.52)	(-11.45)
BkEq	0.00	0.01***	0.01***	0.01***	0.01***	0.01*
	(0.28)	(6.42)	(3.73)	(4.84)	(3.83)	(1.97)
TACC	0.06***	0.02***	-0.01***	0.01***	0.01***	0.04***
	(3.53)	(4.14)	(-4.68)	(3.05)	(3.73)	(3.37)
R^2	0.69	0.77	0.71	0.69	0.63	0.66

Panel B: $t + 2$

	RI	EM Q1	EM Q2	EM Q3	EM Q4	EM Q5
Intercept	0.20***	0.13***	0.00	0.04*	0.07*	0.16*
	(3.72)	(3.09)	(0.13)	(1.83)	(1.99)	(1.88)
Earn	0.74***	0.86***	0.94***	0.93***	0.88***	0.72***
	(22.76)	(24.15)	(56.88)	(127.42)	(82.63)	(20.09)
NegE	-0.44***	-0.20***	-0.12***	-0.13***	-0.14***	-0.40***
	(-8.29)	(-5.74)	(-4.76)	(-7.36)	(-5.58)	(-5.70)
NegExE	-0.49***	-0.17**	-0.35***	-0.40***	-0.49***	-0.50***
	(-13.19)	(-2.07)	(-5.59)	(-9.32)	(-15.13)	(-10.44)
BkEq	0.03***	0.02***	0.03***	0.02***	0.02***	0.03***
	(5.13)	(4.55)	(6.26)	(10.14)	(5.46)	(4.97)
TACC	0.01	-0.00	-0.04***	-0.01	0.01	-0.01
	(0.76)	(-0.12)	(-4.51)	(-1.57)	(1.10)	(-0.63)
R^2	0.62	0.65	0.54	0.51	0.46	0.55

Panel C: $t + 3$

	RI	EM Q1	EM Q2	EM Q3	EM Q4	EM Q5
Intercept	0.20***	0.13***	0.00	0.04*	0.07*	0.16*
	(3.72)	(3.09)	(0.13)	(1.83)	(1.99)	(1.88)
Earn	0.74***	0.86***	0.94***	0.93***	0.88***	0.72***
	(22.76)	(24.15)	(56.88)	(127.42)	(82.63)	(20.09)
NegE	-0.44***	-0.20***	-0.12***	-0.13***	-0.14***	-0.40***
	(-8.29)	(-5.74)	(-4.76)	(-7.36)	(-5.58)	(-5.70)
NegExE	-0.49***	-0.17**	-0.35***	-0.40***	-0.49***	-0.50***
	(-13.19)	(-2.07)	(-5.59)	(-9.32)	(-15.13)	(-10.44)
BkEq	0.03***	0.02***	0.03***	0.02***	0.02***	0.03***
	(5.13)	(4.55)	(6.26)	(10.14)	(5.46)	(4.97)
TACC	0.01	-0.00	-0.04***	-0.01	0.01	-0.01
	(0.76)	(-0.12)	(-4.51)	(-1.57)	(1.10)	(-0.63)
R^2	0.62	0.65	0.54	0.51	0.46	0.55

Table 3 contains the time-series averages of the parameter estimates, Newey and West (1987) t-statistics and R^2 values from the annual earnings regressions. Results are displayed for the RI model and the model interacted with the earnings managemant quintiles. Further, we show results for one-, two-, and three-year ahead forecasts. ***, **, and * indicate significance at an alpha level of 1%, 5%, 10%, respectively.

too. Values for the interaction term and for book equity vary, i.e., no clear pattern between the EM quintiles and the RI model is evident. Further, for all forecast horizons, the parameter estimate of total accruals is smaller for the EM quintiles in comparison to the RI model. This could be due to the fact that the EM measure is based on accruals, and thus, it already incorporates information about accruals into the model. To sum up, the findings support our assumption that the parameter estimates of earnings forecast models differ over the five subsamples characterized by the extent of a firm's EM. In other words, the relationship between the respective predictor variables and future earnings varies depending on the degree of EM a firm engages in.

Next, we assume that better fitting parameter estimates of the earnings forecast model for each EM quintile translates to lower forecast errors. Table 4 shows results of the forecasting performance of the RI model compared to the interacted model. We report mean and median PAFEs for earnings forecasts of up to three years ahead. Furthermore, we report the difference in PAFEs between both models and whether the difference is statistically significant.

Table 4: Earnings Forecast Error Comparison

	Median PAFE_{t+1}	Mean PAFE_{t+1}	Median PAFE_{t+2}	Mean PAFE_{t+2}	Median PAFE_{t+3}	Mean PAFE_{t+3}
RI	0.0372***	0.1330***	0.0488***	0.1437***	0.0641***	0.1690***
Model	(18.85)	(15.46)	(23.27)	(20.81)	(12.72)	(11.35)
Interacted	0.0318***	0.1176***	0.0458***	0.1335***	0.0564***	0.1470***
Model	(21.21)	(13.96)	(19.90)	(19.40)	(19.33)	(21.05)
Difference	-0.53***	-1.54***	-0.30***	-1.02***	-0.77*	-2.20*
	(−3.36)	(−4.34)	(−8.04)	(−6.15)	(−1.90)	(−1.97)

Table 4 compares time-series averages of median and mean PAFEs from the RI earnings forecast model and the model interacted with EM quintiles. One-, two-, and three-year ahead forecasts are analyzed. Further, we test if the difference in PAFE between both models (interacted model minus RI model) is statistically significant. ***, **, and * indicate significance at an alpha level of 1%, 5%, 10%, respectively.

Table 4 provides evidence that for both mean and median PAFEs, the interacted model significantly improves the predictive ability compared to the initial RI model. This finding holds for all forecast horizons. Forecasting one-year (two-year, three-year-) ahead leads to a median PAFE of 3.72% (4.88%, 6.41%) for the RI model compared to a significantly lower median PAFE of 3.18% (4.58%, 5.64%) for the interacted model. Results are similar when examining mean PAFE values, although mean PAFE values are generally higher than median PAFE values. Moreover, the differences in PAFEs between the RI and the interacted model are statistically significant at the 1% significance level for forecasts of up to two years ahead and at the 10% significance level for three-year ahead forecasts.

In conclusion, we provide evidence that incorporating information about the extent of a firm's EM into cross-sectional earnings forecast models leads to more accurate forecasts.

5.3 Evaluation of Implied Cost of Capital Estimates

The previous section provides evidence that interacting the earnings forecast model with quintile dummy variables based on the EM measure improves forecast accuracy. In this section, we follow recent research (e.g., Li and Mohanram (2014) and Hess, Meuter and Kaul (2019)) and analyze if the increased forecast accuracy results in more reliable ICC estimates. In line with the academic literature on the ICC (e.g., Gebhardt, Lee and Swaminathan (2001) and Hou, Van Dijk and Zhang (2012)), we evaluate ICCs by assessing their predictive ability for future realized returns. First, we perform firm-level tests to evaluate the relation between the computed composite ICC and realized future returns. Second, we test the predictive power of the composite ICC for future realized returns on a portfolio level.

Table 5 on the next page presents the results of the firm level-tests, showing the relation between the composite ICC and buy-and-hold returns for one-, two-, and three-years ahead. We annually regress realized future returns on the composite ICC, for both the initial earnings forecast model and the interacted model. The table shows time-series averages of parameter estimates, Newey and West (1987) t-statistics and R^2 values. We expect a positive and significant coefficient if ICCs are able to predict future returns. Further, a coefficient closer to 1 represents an ICC estimate that is on average closer to realized returns.

Table 5 reveals that the coefficients of the interacted model are closer to 1 in comparison to the RI model. For one-year ahead forecasts, the coefficient of the RI model is 0.1904 compared to 0.2176 for the interacted model. For two-year and three-year ahead forecasts, the values are 0.1659 compared to 0.1947 and 0.1472 compared to 0.1896, respectively. Further, the coefficients of the interacted model show higher t-statistics and thus higher significance. Moreover, for all forecast

Table 5: ICC Firm-Level Test

Panel A: Ret$_{t+1}$			
	Intercept$_{t+1}$	ICC$_{t+1}$	R^2
RI Model	0.1099***	0.1904**	0.0107
	(3.91)	(2.62)	
Interacted Model	0.1058***	0.2176***	0.0128
	(3.78)	(2.77)	

Panel B: Ret$_{t+2}$			
	Intercept$_{t+2}$	ICC$_{t+2}$	R^2
RI Model	0.0484**	0.1659**	0.0129
	(2.48)	(2.69)	
Interacted Model	0.0437**	0.1947***	0.0149
	(2.26)	(2.91)	

Panel C: Ret$_{t+3}$			
	Intercept$_{t+3}$	ICC$_{t+3}$	R^2
RI Model	0.0408**	0.1472**	0.0146
	(2.69)	(2.61)	
Interacted Model	0.0355**	0.1896***	0.0164
	(2.41)	(3.13)	

Table 5 depicts the relation between the composite ICC and buy-and-hold returns for one-, two-, and three-years ahead. We compare the RI earnings forecast model with the interacted model. The table show the time-series averages of the parameter estimates, Newey and West (1987) t-statistics and R^2 values from the annual regressions of realized future returns on the composite ICC. ***, **, and * indicate significance at an alpha level of 1%, 5%, 10%, respectively.

horizons, R^2 increases when interacting the RI earnings forecast model with the EM quintile dummy variables. In total, Table 5 provides evidence that ICCs based the interacted model are closer related to realized future returns than ICCs based on the RI model.

Table 6 below illustrates the results of the portfolio tests for the RI earnings forecast model and the interacted model. We annually rank firms into decile portfolios based on the respective composite ICC. For each decile portfolio, we calculate annualized equally weighted buy-and-hold returns for holding periods of up to three years. Further, we implement a long-short strategy by calculating the spread between the highest and lowest decile. A positive and significant return spread illustrates that the composite ICC has significant predictive power for future realized returns.

Table 6: ICC Portfolio Test

	Decile	ICC	Ret_{t+1}	Ret_{t+2}	Ret_{t+3}
RI Model	1	-0.0893	0.1022	0.0066	-0.0035
	2	-0.0200	0.1116	0.0437	0.0355
	3	0.0059	0.1130	0.0532	0.0471
	4	0.0259	0.1142	0.0540	0.0484
	5	0.0436	0.1217	0.0627	0.0575
	6	0.0615	0.1311	0.0748	0.0677
	7	0.0818	0.1363	0.0808	0.0716
	8	0.1095	0.1624	0.0917	0.0808
	9	0.1595	0.1747	0.0935	0.0807
	10	0.4834	0.2085	0.0786	0.0582
	H-L	**0.5727*** **	**0.1063*** **	**0.0720*** **	**0.0617*** **
		(13.74)	(3.24)	(3.19)	(3.20)
Interacted Model	1	-0.0948	0.0923	-0.0003	-0.0146
	2	-0.0246	0.1079	0.0365	0.0258
	3	0.0013	0.1128	0.0457	0.0435
	4	0.0209	0.1170	0.0524	0.0476
	5	0.0384	0.1251	0.0610	0.0551
	6	0.0554	0.1327	0.0803	0.0740
	7	0.0739	0.1441	0.0862	0.0757
	8	0.0975	0.1585	0.0970	0.0851
	9	0.1359	0.1698	0.0972	0.0863
	10	0.4461	0.2155	0.0835	0.0648
	H-L	**0.5410*** **	**0.1232*** **	**0.0838*** **	**0.0794*** **
		(11.10)	(3.48)	(3.50)	(3.94)

Table 6 reports time-series averages of annualized buy-and-hold returns of decile portfolios based on the composite ICC for one-, two-, and three-years ahead. We compare the RI earnings forecast model with the interacted model. For the high-minus-low (H-L) return spread, we further show Newey and West (1987) t-statistics. ***, **, and * indicate significance at an alpha level of 1%, 5%, 10%, respectively.

Table 6 reveals that for both models, annualized buy-and-hold returns for all holding periods increase almost monotonically from the first to last decile.[13] The corresponding high-minus-low return spreads are positive, statistically significant and economically meaningful for both models. However, the interacted model outperforms the RI model for all holding periods. For a one-year holding period, the buy-and-hold return spread for the RI model is 10.63%, while the interacted model yields a return spread of 12.32%. For a two-year (three-year) holding period, the return spread of the RI model is 7.20% (6.17%), whereas the interacted model shows a larger return spread of 8.38% (7.94%). Further, return spreads for the interacted model show larger t-statistics for all holding periods.

To summarize, Table 6 indicates that ICCs based on the interacted model have stronger predictive power for future realized returns on a portfolio level compared to ICCs based on the RI model. Combined with the results of Table 5, the findings provide evidence that the interacted model generates more reliable ICC estimates. Therefore, investors potentially benefit from using earnings forecasts that take information about the extent of a firm's EM into account. Furthermore, as the results from the previous section imply and in line with previous research, this gives additional arguments to establish cross-sectional earnings forecasts as an alternative to analysts' forecasts (e.g., Hou, Van Dijk and Zhang (2012) and Hess, Meuter and Kaul (2019)).

5.4 Robustness Check: Alternative Earnings Forecast Models

The previous sections provide evidence that the extent of a firm's EM is significantly negatively related to earnings forecast accuracy. We show that information about EM can be used to improve forecast accuracy and that this increased accuracy translates to more reliable ICCs. These findings are based on the RI earnings forecast model by Li and Mohanram (2014). To ensure that the findings are robust to alternative earnings forecast models, we further show results for the EP model by Li and Mohanram (2014) and the HVZ model by Hou, Van Dijk and Zhang (2012). Section A.3 and A.4 in the appendix display the results for the EP and HVZ model, respectively.

[13]With the exception of decile 10 for holding periods of two and three years.

First, tables A8 and A12 analyze the relation of the extent of a firm's EM to forecast accuracy, analogous to table 2. For both models, findings are similar to the RI model, i.e., we document a positive and significant relation between EM and forecast accuracy for all forecast horizons. Second, tables A9 (A12) compares forecast accuracy between the EP (HVZ) model and the EP (HVZ) model interacted with the EM quintile dummy variables. In line with our previous findings from Table 4, using the interacted models significantly improves forecast accuracy. Depending on the forecast horizon, the best performing model, i.e., RI, EP, or HVZ model, seems to vary. However, values for all models are rather close. Third, tables A10 and A14 show results for the firm-level ICC tests for the EP and the HVZ model, respectively, analogous to table 5. For both the EP and HVZ model, the interacted models show larger coefficients and t-statistics compared to the initial models. This confirms our previous findings. Further, while R^2 seems to be largest for the HVZ model, coefficients and t-statistics are largest for the RI model. Fourth, tables A11 and A15 display findings of the ICC portfolio tests. The results confirm our findings from Table 6, i.e., the interacted models yield larger return spreads for all holding periods. The only exception is the EP model for a one-year holding period. In general, return spreads for the RI and HVZ model seem rather similar, while the EP model performs worse.

In conclusion, sections A.3 and A.4 provide evidence that our results are robust to alternative cross-sectional earnings forecast models. This further strengthens our findings as it implies that not only the RI model by Li and Mohanram (2014) profits from incorporating information about firms' EM, but cross-sectional earnings forecast models in general.

6 Conclusion

Having accurate earnings forecasts is crucial as they are an important input for firm valuation, asset allocation or ICC calculation. Intuitively, the occurrence of EM, i.e., intentionally modifying earnings, should negatively affect forecast accuracy. Hence, the aim of this paper is to analyze the effect of firms' EM on model-based earnings forecast accuracy.

The analysis is structured as follows. First, we examine the general effect of EM on earnings forecast accuracy. We generate earnings forecasts for up to three years ahead with the RI model by Li and Mohanram (2014) and use the PAFE to evaluate forecast accuracy. Further, we compute the EM measure, i.e., absolute discretionary accruals, using the model of Dechow, Sloan and Sweeney (1995). We run annual cross-sectional regression of PAFE on the EM measure. In line with our expectations, we find a significantly positive relation between PAFE and EM for all forecast horizons. That is, with increasing EM, the PAFE increases, i.e., forecast accuracy decreases. Second, we capitalize on this finding and use the EM measure to improve forecast accuracy. We rank firms annually into quintiles based on the level of EM and create five dummy variables indicating a firm's respective quintile. Next, we interact the earnings forecast model with the EM quintile dummy variables. Again, we generate earnings forecasts for up to three years ahead and find that the forecasts of the interacted model show significantly lower PAFEs compared to the initial RI model. Third, we provide evidence that ICCs based on the interacted model are more reliable expected return proxies in comparison to the initial RI model. For the cross-section of firms, we annually regress realized future returns on the ICCs. We show that ICCs based on the interacted model exhibit higher correlations to realized future returns. Moreover, we annually rank firms into deciles based on the ICCs and implement a long-short-strategy, i.e., we compute the spread between the highest and lowest decile. We find that this portfolio approach yields higher returns for holding periods of up to three years when using ICCs based on the interacted model. Fourth, we ensure that the findings are robust to alternative earnings forecast models. We rerun the previous tests and provide evidence that the tenor of results is unchanged when using the EP model by Li and Mohanram (2014) or the HVZ model by Hou, Van Dijk and Zhang (2012).

We contribute to the literature by providing empirical evidence on the significantly negative relation between the extent of a firm's EM and the predictive ability of earnings forecast models. The negative relation indicates that managerial influence on earnings lowers earnings predictability. This is potentially related to an impaired quality of reported earnings due to opportunistic managerial discretion. Therefore, we support the findings of previous studies indicating that EM is performed for opportunistic reasons, i.e., with the intention of misleading stakeholders to obtain some personal gain (e.g., Perry and Williams (1994), Teoh and Wong (2002), and Bergstresser and Philippon (2006)), instead of aiming to increase the information content of reported earnings (Beneish (2001)). Further, we show that information about EM should be incorporated into earnings forecast models as it improves accuracy and results in more reliable ICCs that yield higher investment strategy returns. This supports previous research (e.g., Hou, Van Dijk and Zhang (2012) and Hess, Meuter and Kaul (2019)) and further establishes cross-sectional earnings forecasts as a viable alternative to analysts' earnings forecasts.

Future research on the relation between the extent of a firm's EM and forecast accuracy might focus on EM measures that are not based on accruals models. Some studies (e.g., Guay, Kothari and Watts (1996), McNichols (2000), and Thomas and Zhang (2000)) criticize the use of such EM measures as they argue that these models provide biased and noisy estimates of discretionary accruals. Alternatively, for instance, Stubben (2010) proposes to use revenue models instead of accruals models to estimate firms' EM or Dechow, Hutton, Kim and Sloan (2012) incorporate reversals of accruals accounting into their model. Further, we leave testing the hypothesis of a U-shaped or any other investigation of the exact form of the relationship between EM and forecast accuracy to future research as there are plausible arguments for a non-linear relationship as well. Finally, investigating the mechanism that leads to fluctuations in the magnitude of the relationship between the extent of EM and the respective forecast accuracy over time is a research question which due to focus limitations we did not touch on and thus as well leave for future researchers to answer.

To conclude, this study provides evidence that the extent of a firm's EM is significantly negatively related to the predictability of the respective firm's earnings. We use this finding and show that incorporating information about firms' EM into earnings forecast models increases forecast accuracy and improves ICC reliability. Therefore, future studies on model-based earnings forecasts should account for firms' EM.

References

Azevedo, V., P. Bielstein and M. Gerhart, 2021, Earnings Forecasts: The Case for Combining Analysts' Estimates with a Cross-Sectional Model, *Review of Quantitative Finance and Accounting* 56, 545–579.

Bartov, E., F. Gul and J. Tsui, 2000, Discretionary-Accruals Models and Audit Qualifications, *Journal of Accounting and Economics* 30, 421–452.

Beneish, M., 2001, Earnings Management: A Perspective, *Managerial Finance*.

Bergstresser, D. and T. Philippon, 2006, CEO Incentives and Earnings Management, *Journal of Financial Economics* 80, 511–529.

Botosan, C. and M. Plumlee, 2005, Assessing Alternative Proxies for the Expected Risk Premium, *The Accounting Review* 80, 21–53.

Chang, W., S. Monahan, A. Ouazad and F. Vasvari, 2021, The Higher Moments of Future Earnings, *The Accounting Review* 96, 91–116.

Chung, H. and S. Kallapur, 2003, Client Importance, Nonaudit Services, and Abnormal Accruals, *The Accounting Review* 78, 931–955.

Claus, J. and J. Thomas, 2001, Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stock Markets, *The Journal of Finance* 56, 1629–1666.

Dechow, P. and I. Dichev, 2002, The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors, *The Accounting Review* 77, 35–59.

Dechow, P., W. Ge and C. Schrand, 2010, Understanding Earnings Quality: A Review of the Proxies, Their Determinants and Their Consequences, *Journal of Accounting and Economics* 50, 344–401.

Dechow, P., A. Hutton, J. Kim and R. Sloan, 2012, Detecting Earnings Management: A New Approach, *Journal of Accounting Research* 50, 275–334.

Dechow, P. and C. Schrand, 2004, Earnings Quality, Working paper, Research Foundation of CFA Institute Charlottesville, VA.

Dechow, P. and D. Skinner, 2000, Earnings Management: Reconciling the Views of Accounting Academics, Practitioners, and Regulators, *Accounting Horizons* 14, 235–250.

Dechow, P., R. Sloan and A. Sweeney, 1995, Detecting Earnings Management, *Accounting Review* 70, 193–225.

Dechow, P., R. Sloan and A. Sweeney, 1996, Causes and Consequences of Earnings Manipulation: An Analysis of Firms Subject to Enforcement Actions by the SEC, *Contemporary Accounting Research* 13, 1–36.

Doyle, J., J. Jennings and M. Soliman, 2013, Do Managers Define Non-GAAP Earnings to Meet or Beat Analyst Forecasts?, *Journal of Accounting and Economics* 56, 40–56.

Easton, P., 2004, PE ratios, PEG ratios, and Estimating the Implied Expected Rate of Return on Equity Capital, *The Accounting Review* 79, 73–95.

Easton, P. and S. Monahan, 2005, An Evaluation of Accounting-Based Measures of Expected Returns, *The Accounting Review* 80, 501–538.

Easton, P. and S. Monahan, 2016, Review of Recent Research on Improving Earnings Forecasts and Evaluating Accounting-Based Estimates of the Expected Rate of Return on Equity Capital, *Abacus* 52, 35–58.

Ecker, F., J. Francis, P. Olsson and K. Schipper, 2013, Estimation Sample Selection for Discretionary Accruals Models, *Journal of Accounting and Economics* 56, 190–211.

Evans, M., K. Njoroge and K. Yong, 2017, An Examination of the Statistical Significance and Economic Relevance of Profitability and Earnings Forecasts from Models and Analysts, *Contemporary Accounting Research* 34, 1453–1488.

Fama, E. and K. French, 1997, Industry Costs of Equity, *Journal of Financial Economics* 43, 153–193.

Fama, E. and K. French, 2006, Profitability, Investment and Average Returns, *Journal of Financial Economics* 82, 491–518.

Francis, J., R. LaFond, P. Olsson and K. Schipper, 2005, The Market Pricing of Accruals Quality, *Journal of Accounting and Economics* 39, 295–327.

Frankel, R., M. Johnson and K. Nelson, 2002, The Relation Between Auditors' Fees for Nonaudit Services and Earnings Management, *The Accounting Review* 77, 71–105.

Gebhardt, W., C. Lee and B. Swaminathan, 2001, Toward an Implied Cost of Capital, *Journal of Accounting Research* 39, 135–176.

Gerakos, J. and R. Gramacy, 2013, Regression-Based Earnings Forecasts, *Chicago Booth Research Paper.*

Griliches, Z. and V. Ringstad, 1970, Error-in-the-variables bias in nonlinear contexts, *Econometrica: Journal of the Econometric Society* 38, 368–370.

Guay, W., S. Kothari and R. Watts, 1996, A Market-Based Evaluation of Discretionary Accrual Models, *Journal of Accounting Research* 34, 83–105.

Healy, P., 1985, The Effect of Bonus Schemes on Accounting Decisions, *Journal of Accounting and Economics* 7, 85–107.

Healy, P. and J. Wahlen, 1999, A Review of the Earnings Management Literature and its Implications for Standard Setting, *Accounting Horizons* 13, 365–383.

Hess, D., M. Meuter and A. Kaul, 2019, The Performance of Mechanical Earnings Forecasts, Working paper, University of Cologne and Saarland University.

Higashikawa, K., 2020, Earnings Quality and Earnings Forecasts Based on a Cross-Sectional Model, *Asia-Pacific Management Accounting Journal* 15, 141–163.

Hou, K. and M. Van Dijk, 2019, Resurrecting the Size Effect: Firm Size, Profitability Shocks, and Expected Stock Returns, *The Review of Financial Studies* 32, 2850–2889.

Hou, K., M. Van Dijk and Y. Zhang, 2012, The Implied Cost of Capital: A New Approach, *Journal of Accounting and Economics* 53, 504–526.

Jones, J., 1991, Earnings Management During Import Relief Investigations, *Journal of Accounting Research* 29, 193–228.

Klein, A., 2002, Audit Committee, Board of Director Characteristics, and Earnings Management, *Journal of Accounting and Economics* 33, 375–400.

Konstantinidi, T. and P. Pope, 2016, Forecasting Risk in Earnings, *Contemporary Accounting Research* 33, 487–525.

Lee, C., E. So and C. Wang, 2010, Evaluating Implied Cost of Capital Estimates, Working paper, Stanford University.

Lee, C., E. So and C. Wang, 2021, Evaluating Firm-Level Expected-Return Proxies: Implications for Estimating Treatment Effects, *The Review of Financial Studies* 34, 1907–1951.

Li, K. and P. Mohanram, 2014, Evaluating Cross-Sectional Forecasting Models for Implied Cost of Capital, *Review of Accounting Studies* 19, 1152–1185.

Lo, K., 2008, Earnings Management and Earnings Quality, *Journal of Accounting and Economics* 45, 350–357.

McNichols, M., 2000, Research Design Issues in Earnings Management Studies, *Journal of Accounting and Public Policy* 19, 313–345.

McNichols, M., 2002, Discussion of the Quality of Accruals and Earnings: The Role of Accrual Estimation Errors, *The Accounting Review* 77, 61–69.

Newey, W. and K. West, 1987, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703–708.

Ohlson, J. and B. Juettner-Nauroth, 2005, Expected EPS and EPS Growth as Determinants of Value, *Review of Accounting Studies* 10, 349–365.

Peasnell, K., P. Pope and S. Young, 2000, Detecting Earnings Management Using Cross-Sectional Abnormal Accruals Models, *Accounting and Business Research* 30, 313–326.

Perry, S. and T. Williams, 1994, Earnings Management Preceding Management Buy-Out Offers, *Journal of Accounting and Economics* 18, 157–179.

Richardson, S., R. Sloan, M. Soliman and I. Tuna, 2005, Accrual Reliability, Earnings Persistence and Stock Prices, *Journal of Accounting and Economics* 39, 437–485.

Stubben, S., 2010, Discretionary Revenues as a Measure of Earnings Management, *The Accounting Review* 85, 695–717.

Teoh, S., I. Welch and T. Wong, 1998, Earnings Management and the Long-Run Market Performance of Initial Public Offerings, *The Journal of Finance* 53, 1935–1974.

Teoh, S. and T. Wong, 2002, Why New Issues and High-Accrual Firms Underperform: The Role of Analysts' Credulity, *The Review of Financial Studies* 15, 869–900.

Thomas, J. and X. Zhang, 2000, Identifying Unexpected Accruals: A Comparison of Current Approaches, *Journal of Accounting and Public Policy* 19, 347–376.

Tian, H., A. Yim and D. Newton, 2021, Tail-Heaviness, Asymmetry, and Profitability Forecasting by Quantile Regression, *Management Science* 67, 5209–5233.

A Appendix

A.1 Variable Definition

Table A7: Variable Description

Panel A: Li and Mohanram (2014) and Hou, Van Dijk and Zhang (2012)		
Variable	Description	COMPUSTAT Variable
$Earn$	Earnings divided by number of shares outstanding.	IB-SPI
$NegE$	Indicator variable that equals 1 for firms with negative earnings and 0 otherwise.	
$NegExE$	Interaction term of E and NegE.	
$BkEq$	Book value of equity divided by number of shares outstanding.	CEQ
$TACC$	Sum of change in WC, change in NCO, and change in FIN, divided by number of shares outstanding.	WC= (ACT-CHE)-(LCT-DLC) NCO = (AT-ACT-IVAO)-(LT-LCT-DLTT) FIN = (IVST+IVAO)-(DLTT+DLC+PSTK)
Div	Common dividends divided by shares outstanding.	DVC
$DivD$	Indicator variable that equals 1 for dividend payers and 0 otherwise.	
AT	Total assets divided by number of shares outstanding.	AT

Panel B: Dechow, Sloan and Sweeney (1995)		
Variable	Description	COMPUSTAT Variable
$Rev - Rec$	Change in revenues minus change in receivables, divided by number of shares outstanding.	REV, REC
PPE	Gross total property, plants, and equipment divided by number of shares outstanding.	PPEGT

Table A7 contains the descriptions of the variables used throughout the paper. All variables refer to the current period and are scaled by number of shares outstanding (COMPUSTAT variable CSHO) if not defined otherwise. Panel A contains the variable descriptions of the earnings forecast models by Li and Mohanram (2014) and Hou, Van Dijk and Zhang (2012). Panel B contains the variable descriptions of the accruals model by Dechow, Sloan and Sweeney (1995).

A.2 Implied Cost of Capital Metrics

Appendix B presents the four ICC metrics used to compute the composite ICC. Our notation is akin to Hou, Van Dijk and Zhang (2012).

Gebhardt, Lee and Swaminathan (2001) derive the ICC using the following definition:

$$M_t = B_t + \sum_{\tau=1}^{11} \frac{E_t[(ROE_{t+\tau} - ICC)B_{t+\tau-1}]}{(1+ICC)^\tau} + \frac{E_t[(ROE_{t+12} - ICC)B_{t+11}]}{ICC(1+ICC)^{11}}, \quad (8)$$

where M is the market equity, ROE reflects the return on equity, $E_t[\,]$ are the market expectations based on information available in year t and $(ROE_{t+\tau} - ICC)B_{t+\tau-1}$ is the residual income in year $t+\tau$. To estimate the expected ROE for the years $t+1$ to $t+3$, we use the model-based earnings forecasts and book equity based on clean surplus accounting, i.e. $B_{t+\tau} = B_{t+\tau-1} + Earn_{t+\tau} - D_{t+\tau}$, where D reflects dividends. For firms with positive earnings, dividends are calculated using the current payout ratio. For firms with negative earnings, the payout ratio is estimated by dividing current dividends by $0.6 \times total\ assets$. We assume that following $t+3$, the expected ROE mean-reverts to the industry median value by year $t+11$.

Claus and Thomas (2001) estimate their ICC by solving the following equation:

$$\begin{aligned} M_t = B_t + \sum_{\tau=1}^{5} &\frac{E_t[(ROE_{t+\tau} - ICC)B_{t+\tau-1}]}{(1+ICC)^\tau} \\ &+ \frac{E_t[(ROE_{t+5} - ICC)B_{t+4}](1+g_a)}{(ICC - g_a)(1+ICC)^5}. \end{aligned} \quad (9)$$

To estimate expected ROE in years $t+1$ to $t+5$ we use the model-based earnings forecasts and book equity determined based on clean surplus accounting, analogous to the ICC metric by Gebhardt, Lee and Swaminathan (2001). In line with Azevedo, Bielstein and Gerhart (2021), we set the growth-rate g_a to the 10-year government bond yield minus an assumed real risk-free rate of 3%.

Ohlson and Juettner-Nauroth (2005) calculate the ICC as follows:

$$ICC = A + \sqrt{A^2 + \frac{E_t[Earn_{t+1}]}{M_t}(g_o - (\gamma - 1))} \quad (10)$$

where

$$A = 0.5\left((\gamma - 1) + \frac{E_t[D_{t+1}]}{M_t}\right),$$

$$g_o = 0.5\left(\frac{E_t[Earn_{t+3}] - E_t[Earn_{t+2}]}{E_t[Earn_{t+2}]} + \frac{E_t[Earn_{t+5}] - E_t[Earn_{t+4}]}{E_t[Earn_{t+4}]}\right).$$

Dividends are calculated analogous to Gebhardt, Lee and Swaminathan (2001). In this case, g_o is the short-term growth rate, estimated as the average of forecasted five-year and near-term growth calculated in line with Ohlson and Juettner-Nauroth (2005). Furthermore, they calculate the perpetual growth rate of abnormal earnings beyond the forecast horizon γ as the 10-year government bond yield minus an assumed real risk-free rate of 3%.

Easton (2004) shows that the ICC can be calculated using the following equation, with dividends calculated analogously to Gebhardt, Lee and Swaminathan (2001):

$$M_t = \frac{E_t[Earn_{t+2}] + ICC \times E_t[D_{t+1}] - E_t[Earn_{t+1}]}{ICC^2}. \tag{11}$$

A.3 Results for the EP Model

Table A8: The Relationship Between EM and Earnings Forecast Accuracy for the EP Model

	PAFE_{t+1}	PAFE_{t+2}	PAFE_{t+3}
Coefficient	0.0193***	0.0181***	0.0165***
	(8.28)	(10.54)	(11.19)
R^2	0.1198	0.1356	0.1390
Controls	Yes	Yes	Yes

Table A8 depicts the relationship between EM and the EP model-based earnings forecast accuracy. It contains the time-series averages of the parameter estimates, Newey and West (1987) t-statistics and R^2 values from the annual regressions of PAFE on the EM measure. We control for firm size by including the logarithm of total assets and for industry by adding industry dummies according to the Fama and French 48 industry classification. ***, **, and * indicate significance at an alpha level of 1%, 5%, 10%, respectively.

Table A9: Earnings Forecast Error Comparison for the EP Model

	Median PAFE_{t+1}	Mean PAFE_{t+1}	Median PAFE_{t+2}	Mean PAFE_{t+2}	Median PAFE_{t+3}	Mean PAFE_{t+3}
EP Model	0.0362***	0.1278***	0.0499***	0.1439***	0.0614***	0.1596***
	(23.07)	(15.18)	(21.03)	(21.33)	(19.62)	(20.76)
Interacted Model	0.0317***	0.1163***	0.0459***	0.1316***	0.0566***	0.1449***
	(21.72)	(14.21)	(19.04)	(19.74)	(17.90)	(20.96)
Difference	-0.45***	-1.15***	-0.40***	-1.23***	-0.48***	-1.47***
	(−5.59)	(−6.16)	(−11.32)	(−8.21)	(−10.01)	(−6.27)

Table A9 compares time-series averages of median and mean PAFEs from the EP earnings forecast model and the model interacted with EM quintiles. One-, two-, and three-year ahead forecasts are analyzed. Further, we test if the difference in PAFE between both models (interacted model minus EP model) is statistically significant. ***, **, and * indicate significance at an alpha level of 1%, 5%, 10%, respectively.

Table A10: ICC Firm-Level Test Based on the EP Earnings Forecast Model

Panel A: Ret$_{t+1}$			
	Intercept$_{t+1}$	ICC$_{t+1}$	R^2
EP Model	0.1120***	0.1484*	0.0095
	(3.90)	(2.02)	
Interacted Model	0.1099***	0.1571**	0.0106
	(3.87)	(2.07)	
Panel B: Ret$_{t+2}$			
	Intercept$_{t+2}$	ICC$_{t+2}$	R^2
EP Model	0.0541**	0.0915	0.0106
	(2.61)	(1.57)	
Interacted Model	0.0498**	0.1302**	0.0129
	(2.47)	(2.07)	
Panel C: Ret$_{t+3}$			
	Intercept$_{t+3}$	ICC$_{t+3}$	R^2
EP Model	0.0465***	0.0727	0.0112
	(2.89)	(1.45)	
Interacted Model	0.0413**	0.1251**	0.0137
	(2.68)	(2.22)	

Table A10 depicts the relation between the composite ICC and buy-and-hold returns for one-, two-, and three-years ahead. We compare the EP earnings forecast model with the interacted model. The table show the time-series averages of the parameter estimates, Newey and West (1987) t-statistics and R^2 values from the annual regressions of realized future returns on the composite ICC. ***, **, and * indicate significance at an alpha level of 1%, 5%, 10%, respectively.

Table A11: ICC Portfolio Test Based on the EP Earnings Forecast Model

	ICC	Ret$_{t+1}$	Ret$_{t+2}$	Ret$_{t+3}$
EP Model	**0.5234*** **	**0.1114*** **	**0.0554*** **	**0.0453** **
	(19.99)	(3.77)	(2.72)	(2.00)
Interacted Model	**0.5081*** **	**0.1069*** **	**0.0665*** **	**0.0641*** **
	(15.48)	(3.02)	(2.78)	(3.00)

Table A11 reports time-series averages of annualized buy-and-hold returns of decile portfolios based on the composite ICC for one-, two-, and three-years ahead. We compare the EP earnings forecast model with the interacted model. For the high-minus-low (H-L) return spread, we further show Newey and West (1987) t-statistics. ***, **, and * indicate significance at an alpha level of 1%, 5%, 10%, respectively.

A.4 Results for the HVZ Model

Table A12: The Relationship Between EM and Earnings Forecast Accuracy for the HVZ Model

	$PAFE_{t+1}$	$PAFE_{t+2}$	$PAFE_{t+3}$
Coefficient	0.0193***	0.0181***	0.0165***
	(8.28)	(10.54)	(11.19)
R^2	0.1198	0.1356	0.1390
Controls	Yes	Yes	Yes

Table A12 depicts the relationship between EM and the HVZ model-based earnings forecast accuracy. It contains the time-series averages of the parameter estimates, Newey and West (1987) t-statistics and R^2 values from the annual regressions of PAFE on the EM measure. We control for firm size by including the logarithm of total assets and for industry by adding industry dummies according to the Fama and French 48 industry classification. ***, **, and * indicate significance at an alpha level of 1%, 5%, 10%, respectively.

Table A13: Earnings Forecast Error Comparison for the HVZ Model

	Median $PAFE_{t+1}$	Mean $PAFE_{t+1}$	Median $PAFE_{t+2}$	Mean $PAFE_{t+2}$	Median $PAFE_{t+3}$	Mean $PAFE_{t+3}$
HVZ Model	0.0356***	0.1282***	0.0474***	0.1389***	0.0595***	0.1574***
	(19.14)	(13.58)	(21.03)	(18.22)	(20.63)	(16.88)
Interacted Model	0.0314***	0.1162***	0.0454***	0.1357***	0.0558***	0.1479***
	(21.47)	(13.99)	(19.49)	(18.32)	(18.95)	(19.86)
Difference	-0.42***	-1.20***	-0.20***	-0.32***	-0.38***	-0.95**
	(−4.07)	(−3.72)	(−6.14)	(−3.43)	(−3.45)	(−2.47)

Table A13 compares time-series averages of median and mean PAFEs from the HVZ earnings forecast model and the model interacted with EM quintiles. One-, two-, and three-year ahead forecasts are analyzed. Further, we test if the difference in PAFE between both models (interacted model minus HVZ model) is statistically significant. ***, **, and * indicate significance at an alpha level of 1%, 5%, 10%, respectively.

Table A14: ICC Firm-Level Test Based on the HVZ Earnings Forecast Model

Panel A: Ret$_{t+1}$			
	Intercept$_{t+1}$	ICC$_{t+1}$	R^2
HVZ Model	0.1124***	0.1588**	0.0114
	(3.92)	(2.07)	
Interacted Model	0.1093***	0.2007**	0.0120
	(3.79)	(2.50)	
Panel B: Ret$_{t+2}$			
	Intercept$_{t+2}$	ICC$_{t+2}$	R^2
HVZ Model	0.0503**	0.1465**	0.0139
	(2.50)	(2.23)	
Interacted Model	0.0461**	0.1930***	0.0161
	(2.33)	(2.75)	
Panel C: Ret$_{t+3}$			
	Intercept$_{t+3}$	ICC$_{t+3}$	R^2
HVZ Model	0.0422**	0.1347**	0.0149
	(2.70)	(2.35)	
Interacted Model	0.0381**	0.1805***	0.0170
	(2.50)	(2.96)	

Table A14 depicts the relation between the composite ICC and buy-and-hold returns for one-, two-, and three-years ahead. We compare the HVZ earnings forecast model with the interacted model. The table show the time-series averages of the parameter estimates, Newey and West (1987) t-statistics and R^2 values from the annual regressions of realized future returns on the composite ICC. ***, **, and * indicate significance at an alpha level of 1%, 5%, 10%, respectively.

Table A15: ICC Portfolio Test Based on the HVZ Earnings Forecast Model

	ICC	Ret$_{t+1}$	Ret$_{t+2}$	Ret$_{t+3}$
HVZ Model	**0.6100*** **	**0.1029*** **	**0.0559** **	**0.0503** **
	(13.80)	(3.12)	(2.35)	(3.00)
Interacted Model	**0.5688*** **	**0.1238*** **	**0.0884*** **	**0.0794*** **
	(7.73)	(3.45)	(3.61)	(4.00)

Table A15 reports time-series averages of annualized buy-and-hold returns of decile portfolios based on the composite ICC for one-, two-, and three-years ahead. We compare the HVZ earnings forecast model with the interacted model. For the high-minus-low (H-L) return spread, we further show Newey and West (1987) t-statistics. ***, **, and * indicate significance at an alpha level of 1%, 5%, 10%, respectively.