Three essays on the performance of earnings forecasts, bankruptcy predictions and textual analysis

Inauguraldissertation
zur
Erlangung des Doktorgrades
der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der
Universität zu Köln

2019

vorgelegt
von

Tobias Lorsbach, M.Sc.
aus
Prüm

Referent:	Univ.-Prof. Dr. Dieter Hess
Korreferent:	Univ.-Prof. Dr. Alexander Kempf
Vorsitz:	Univ.-Prof. Dr. Thomas Hartmann-Wendels
Tag der Promotion:	06. Juni 2019

This thesis consists of the following works:

Hess, Dieter and Tobias Lorsbach (2019):
Incorporating Quarterly Earnings Information into Cross-Sectional
Earnings Forecast Models,
Working Paper, University of Cologne.

Huettemann, Martin and Tobias Lorsbach (2019):
The Quality of Bankruptcy Data and its Impact on the Evaluation of Prediction Models:
Creating and Testing a German Database,
Working Paper, University of Cologne.

Lorsbach, Tobias (2019):
Word Power: Content Analysis in the Presence of Competing Information,
Working Paper, University of Cologne.

Acknowledgements

Die vorliegende Arbeit habe ich während meiner Zeit als Promotionsstudent der Cologne Graduate School in Management, Economics and Social Sciences und Wissenschaftlicher Mitarbeiter am Lehrstuhl für ABWL und Unternehmensfinanzen an der Universität zu Köln angefertigt.

In erster Linie möchte ich mich bei meinem Doktorvater Prof. Dr. Dieter Hess für die umfassende Betreuung meiner Arbeit in den vergangenen vier Jahren bedanken. Ich bedanke mich auch bei Prof. Dr. Alexander Kempf für die Erstellung des Zweitgutachtens, sowie Prof. Dr. Thomas Hartmann-Wendels für die Übernahme des Vorsitzes der Prüfungskommission.

Weiterhin bedanke ich mich bei der Cologne Graduate School in Management. Economics and Social Sciences für die umfangreiche Förderung, sowie insbesondere ihrer Direktorin Dr. Dagmar Weiler für die anhaltende administrative und auch persönliche Unterstützung im Rahmen meines Promotionsstipendiums.

Insbesondere bedanke ich mich bei Martin Hüttemann für die tolle Zusammenarbeit bei dem gemeinsamen Forschungsprojekt, sowie William Liu für die zahlreichen Diskussionen und Kommentare. Außerdem bedanke ich mich bei meinen weiteren Kollegen an der Cologne Graduate School und dem Lehrstuhl für ABWL und Unternehmensfinanzen an der Universität zu Köln: Niklas Blümke, Anke Gewand, Martin Meuter, Britta Plum, Tim Vater und Djarban Waning.

Zu guter Letzt, möchte ich mich herzlich bei meinem privaten Umfeld, Freunden und Familie, bedanken. Insbesondere bedanke ich mich bei meinen Eltern, sowie meinen beiden Geschwistern für die gesamte Unterstützung und Zuspruch während der Promotion, aber auch allen anderen bisherigen Lebensphasen und -situationen.

Tobias Lorsbach

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

In efficient capital markets, market prices reflect all available information. Market participants receive most of their financial information by annual and quarterly reports, that are the official statements of a company's profitability. By this, they include all information on company business operations and conditions, potential risks, litigation, profits, and current financial health. The information provided can be quantitative and qualitative in nature. Quantitative information is presented in graphs, tables, and most importantly, numbers and can be assessed easily because of its nature. The analysis of such information is a central theme in finance and capital markets research in accounting. For example, earnings figures and growth rates have been frequently used as fundamental inputs in several corporate finance settings, such as firm valuation or capital budgeting. Moreover, they take a crucial part to investment management practices of fund managers and investors. In contrast, qualitative information refers to unstructured data that is presented as written and spoken text and that is not readily quantified by numbers. In turn, an emerging discipline of science develops methods of content analyses to measure components, such as sentiment and readability, of textual data (see, e.g., Henry, 2008; Loughran and McDonald, 2011) that is typically interpreted by cognitive processing of each individual. However, human cognitive capabilities are limited, and, thus, financial intermediaries play a central role in collecting, processing and analyzing material data from corporate disclosures to provide capital markets with aggregated information.

Traditionally, auditors, financial analysts, and rating agencies have been the most important specialized information intermediaries in financial markets. Due to their reputation and regulations stakeholders used to have confidence in information provided by these intermediaries. However, recently, investor confidence evaporates because renowned auditors were involved in numerous accounting scandals and because rating agencies biased their appraisals in favor of issuer clients during the financial crisis. In addition, company analyses made by financial analysts appear to be substantially off-the-mark and forecasts are notoriously biased toward optimistic prospectuses (see, e.g., Abarbanell, 1991; O'Brien, 1988; Bradshaw et al., 2012). Therefore, in course of the digital revolution, company stakeholders, including investors, banks, regulators, and the media, have devoted significant resources to substitute major processes of financial intermediaries through technological alternatives. This includes, for example, fully automated investment advisory services using sophisticated computer algorithms, i.e., robo-advisers, sharing investments in social trading networks or raising funds through crowdfunding platforms. Besides these activities, from an investor's perspective, company analyses may also substantially benefit from fully automated analyses of quantitative data, such as the automation of the processes of forecasting firm-level earnings, financial insolvencies and liquidity ratings, or the creation of computerized analyses of qualitative content from company financial reports.

In Chapter 1 of my thesis, which is based on "Incorporating Quarterly Earnings information into Cross-Sectional Earnings Forecast Models" (2019) and is co-authored by Dieter Hess, I address this several aspects of this issue. In this article, we develop a novel framework to automatically update earnings forecast models in response to quarterly earnings results. This allows models' forecasts to be more informative and at the same time to provide high-frequency earnings expectations. Accordingly, we compare model forecasts and analyst forecasts and level the playing field between equity analysts and mechanical models. In addition, we evaluate the implied cost of capital (ICCs) estimates that are based on these forecasts. The ICC is the discount rate that equates a firm's future cash flows to its current stock price (see, e.g., Gordon and Gordon, 1997; Easton, 2004).

Most importantly, our analyses focus on four aspects. First, we introduce a new approach to incorporating intra-year information into individual cross-sectional forecast models. That is, we explicitly utilize quarterly earnings results to produce more accurate and timely earnings forecasts. Second, we assess whether model forecasts or analysts' forecasts perform better. Specifically, we analyze the changes in forecast accuracy throughout the financial year and particularly after structural breaks from quarterly earnings reports. Third, we evaluate ICC estimates that are based on these forecasts. Last, we also explore whether better earnings forecasts yield more reliable ICCs.

To innovate forecast models using fundamental interim earnings information, we develop a novel framework for updating model forecasts immediately after earnings release and incorporating that valuable information. Therefore, we adapt the way analysts anticipate quarterly earnings results. In particular, whenever a company discloses a quarterly earnings number, the forecast task for annual earnings reduces to predicting the results for the remaining fiscal quarters. We adjust forecast models to predict only the remaining unknown portion of the annual earnings results. Finally, we aggregate realized quarterly results and our forecasts for the yet unpublished quarterly results into earnings forecasts for the entire fiscal year. Using this concept, we formulate a parsimonious model that produces annual earnings forecasts throughout the year using a company's most recent financial statements, in a manner similar to analysts.

Our augmented framework produces substantially better earnings forecasts. That is, forecast errors of annual earnings forecasts made after the release of first quarter results are reduced by about 15.2%. Correspondingly, the forecasts made after the second and third quarters are more accurate by about 30.8% and 50.0%., respectively. Moreover, we find that once model forecasts and analysts' forecasts are compared systematically according to available public information, performance differences diminish. In detail, analysts' forecasts are more accurate only for the largest 10% of all firms. Hence, model forecasts efficiently improve forecast accuracy by using quarterly information. Moreover, we document that ICCs based on model forecasts better predict future returns than those constructed with analysts' forecasts. Despite the high accuracy of analysts' short-term forecasts, we show that analysts' forecasts, in general, are very optimistic and grossly inaccurate in long-term horizons. Accordingly, analyst-based ICCs are not significantly related to future returns. However, we still find that better forecasts from the earnings model yield better predictors for future stock returns. In fact, model forecasts that incorporate interim earnings information produces better expected return proxies than the initial model specifications. Hence, better forecast quality yields ceteris paribus (c.p.) better proxies for future returns.

This study most closely relates to the work done by Hou et al. (2012) and Li and Mohanram (2014). We apply our framework of adjusting forecast models according to available quarterly earnings information to the model of Hou et al. (2012), hereafter HVZ, and to the residual income (RI) and earnings persistence (EP) models specified by Li and Mohanram (2014). First, we benchmark our extended models against the initial model to test whether model forecasts improve from quarterly earnings information. In previous studies, analysts' forecasts anticipated fundamental quarterly information, while initial models incorporate only annual earnings information. Accordingly, there is no study that comprehensively compares the earnings forecast performance of equity analysts with that of cross-sectional forecast models (e.g., Feng, 2014). Thereby, we clarify the confounding effects from previous studies that use analyst forecasts and consider quarterly earnings statements by also incorporating this information into model forecasts.

Our results contribute to the literature on earnings forecasts and future return proxies in several ways. First, we find that earnings forecasts can be potentially automated using cross-sectional models and high frequency data. In addition, we show that more accurate earnings expectations translate into more reliable proxies of future returns. Finally, our results may motivate researchers to further innovate forecast models to incorporate additional high-frequency data. Traditionally, mechanical models have exclusively incorporated financial accounting data. However, our findings suggest that the performance of model earnings forecasts may further improve from additional information, such as macroeconomic indicators or stock market data, that also carries valuable information and help to beat equity analysts even in short-term horizons.

While a company's earnings results are a key input in several asset pricing models, for example to estimate the implied cost of capital (see Chapter 2), they are also a primary indicator of a company's current and future financial strength. The second essay (Chapter 3) is based on my working paper "The Quality of Bankruptcy Data and its Impact on the Evaluation of Prediction Models: Creating and Testing a German Database" (2019) and is co-authored by Martin Huettemann. In this paper, we apply several bankruptcy prediction models in the context of German companies. While the methodology of bankruptcy predictions and choice of predictor variables that may indicate financial weakness have been intensively addressed in previous studies (see, e.g., Altman, 1968; Ohlson, 1980; Shumway, 2001; Bharath and Shumway, 2008), this article focuses on another central aspect: the primary quality of bankruptcy information. In detail, we explore whether imprecise bankruptcy indicators produce confounding results by comparing different bankruptcy models. For this purpose, we create an alternative database of German bankruptcies by systematically collecting information from public sources, testing differences among existing commercial databases and, finally, examining how bankruptcy prediction models are driven by these divergences.

Existing studies commonly use commercial databases to obtain bankruptcy information. In the U.S., LoPucky, the SDC Platinum Database, Moody's Default and Recovery Database, Capital IQ, and the Fixed Investment Securities Database (FISD) are examples of such databases. Even though the SDC Platinum Database and Moody's Default and Recovery Database contain some data on European bankruptcies, there is sparse data available for non-U.S. firms. Similarly, there is little literature that captures the predictability of bankruptcies outside the U.S. (see, e.g., Dahiya and Klapper, 2007; Altman, Iwanicz-Drozdowska, Laitinen, Suvas, 2017; Tian and Yu, 2017). In spite of this, there is no study that provides guidance on how to obtain bankruptcy data from public services (particularly outside the U.S.). None of the existing studies discuss the underlying data of bankruptcy events and, consequently, how this may drive bankruptcy models to perform differently or whether firm characteristics are connected to higher probabilities of default.

The first contribution of this study is that we generate a dataset of bankruptcy information using automated web scraping and textual analysis. We develop a systematic methodology to obtain bankruptcy information from corporate news releases and public sources. In detail, we implemented an automated web-scraper to download all publicly available corporate disclosures from online news archives and, furthermore, we crawled relevant public databases to obtain bankruptcy information for public firms in Germany. Applying this methodology, we compile a bankruptcy database that is more complete and accurate regarding bankruptcy events and dates than the most frequently used commercial databases, i.e., Bureau van Dijk and Compustat Global. Most importantly, our bankruptcy database, hereafter HL, includes bankruptcies that are not reflected in those databases and, hence, it cannot be reproduced by any combination of these commercial databases.

Furthermore, this article contributes to the voluminous literature on bankruptcy prediction research in several ways. That is, we comprehensively compare of popular bankruptcy prediction models for the German market using our solid bankruptcy information database. We use our bankruptcy data to conduct two empirical analyses. First, we compare the performance of several bankruptcy prediction models in the context of German firms. Second, we compare our database with Bureau van Dijk data and find that the quality of bankruptcy data significantly affects parameter estimates and the out-of-sample evaluation of bankruptcy prediction models. It is likely that previous studies that used incomplete bankruptcy information provided by Bureau van Dijk or Compustat Global present biased parameters for commonly used bankruptcy prediction models. As a result, the out-of-sample assessment based on these biased parameters may not be informative. In detail, we find that when Bureau van Dijk information is used, the out-of-sample evaluation yields completely different inferences than evaluations using our HL data, i.e., the in-sample tests set to other model specifications and the out-of-sample evaluation recommends another prediction model. This suggests that results of previous studies should be reviewed in light of the relevance of underlying bankruptcy data quality.

This article adds important findings to the existing literature on bankruptcy predictability. In a case study for the German capital market, we show that the underlying quality of bankruptcy data fundamentally matters to the integrity of bankruptcy prediction models. This serves investors and researchers equally. On the one hand, researchers may extend our web-scraping approach to compile even broader data sets of bankruptcy events, for example, for non-public firms or roll-outs to other capital markets. In addition, accurate bankruptcy information is crucial for several other applications, such as, analyzing systemic risk or credit spreads. Hence, investors, banks, rating agencies, and creditors can better validate their bankruptcy prediction models using our German data and, thus, better quantify default risks in credit contracts or improve liquidity classifications and market participants' ratings using basic financial statement information from annual earnings reports, such as quantitative earnings, asset and liability numbers.

The information content of earnings reports is generally a focal point for many measurement controversies in accounting and finance (e.g., Beaver, 1968). As an alternative to analyzing quantitative data from corporate disclosures (as discussed in Chapters 2 and 3), a growing body of literature derives a new class of data by analyzing textual content within financial news or corporate disclosures. Therefore, turning from direct processing of quantitative information to analyzing textual content, the third essay (Lorsbach, 2019, "Word Power: Content Analysis in the Presence of Competing Information") introduces a novel framework for quantifying the qualitative content of financial disclosures according to immediate stock market responses. While most researchers use readily available quantitative information, such as earnings numbers or balance sheet items, today's vast availability of descriptive information about firms found in corporate disclosures, the financial press, and social media have led to an extensive number of opportunities for qualitative content analysis. In general, qualitative information refers to unstructured data from written and spoken texts that are not readily quantified and are generally analyzed through cognitive processing by individuals. However, complex algorithms and technological progress have made it feasible to process this voluminous unstructured text data using automated procedures.

Traditionally, financial economics has predominantly used "bag-of-words" methods, such as dictionary-based classifications, to categorize single words as either positive or negative (see, e.g., Henry, 2008; Li, 2008; Loughran and McDonald, 2011). In contrast, Jegadeesh and Wu (2013) developed an approach that determines the relative content of words using investors' cognitive language processing by gauging market reactions to 10-K filings. However, several studies find that investor responses to 10-K filings are not significant per se and are essentially meaningless when information is preempted by previous earnings announcements (see, e.g., Easton and Zmijewski, 1993;

Li and Ramesh, 2009). Therefore, there remain several open questions in the literature. For example, there is no study that encounters how market responses can be affected by competing information events in the context of textual content analyses. Although most studies claim that earnings announcements preempt other corporate disclosures, such as 10-K filings, no study exists that evaluates the reliability of tone measures based on earnings announcements. Furthermore, none of the existing studies consider investor responses to earnings announcements to measure the language processing of financial statements, such as 10-K filings, in capital markets.

Pivotal to my framework is the focus on earnings announcements dates (EAD) as structural breaks in the information environment of capital market participants compared to subsequent SEC-filings of annual results, that is, Form 10-K. Earnings announcements are the first notice of a firm's financial performance and, as a result, preempt the informational value of subsequent 10-K filings.

This study evaluates tone measures that apply earnings announcement returns or 10-K filing returns to interpret the relative strength of positive and negative words used in annual reports. I find that the market reaction to earnings announcements is more informative than 10-K filing returns, as these filing returns misclassify the relative strength of words and, consequentially, tone measures. Market reaction to earnings announcements allows to better quantify textual tone and elaborates the interpretation of financial language use. In fact, my approach yields a stronger correlation of qualitative information to first and second moments of stock returns, that is, future returns and volatility, and future profitability.

The study results have important implications for both investors and academics. First, this study shows that investors predominantly respond to earnings announcements rather than 10-K filings. Hence, interpreting the textual information of financial disclosures, based on observable market reactions, may help advise financial institutions, companies, and customers to evaluate qualitative content in financial reports from a broader perspective. Most importantly, it may aid in obtaining better corporate valuation or investment decisions. Second, future research can consider examining whether stock markets react in a similar way to other financial disclosures and perform additional content analysis based on observable price movements. Thus, researchers may apply my approach of content analysis to other public disclosures, namely macroeconomic disclosures (central banks, consumer indices), analysts' reports, or other news releases. Thereby, this essay fundamentally contributes to the flourishing literature on computerized financial text analysis.

In summary, Chapter 2 of my thesis finds that quantitative information from interim financial earnings disclosures fundamentally improves the earnings forecast accuracy of mechanical models and levels the playing field when comparing model forecasts to analyst forecasts. Similarly, quantitative financial disclosure information is a key input in most bankruptcy prediction models. Using web crawling techniques to aggregate bankruptcy information of German companies, Chapter 3 shows that the reliability of widely used bankruptcy prediction models is determined by the quality of the underlying bankruptcy data. Accordingly, bankruptcy prediction models can be improved by collecting accurate bankruptcy data and discarding incorrect information. Complementary, Chapter 4 applies a new perspective to information provided in corporate disclosures by examining qualitative or textual content. This section introduces a novel framework focusing on analysis of the textual content of annual reports. Using the immediacy of market reactions and investor responses to textual information enables quantifying the qualitative content of financial disclosures.

Chapter 2

Incorporating Quarterly Earnings Information into Cross-Sectional Earnings Forecast Models[1]

2.1 Introduction

Corporate earnings are a key indicator of a company's financial strength and therefore used as key input in many models, for example, to assess a firm's value or to infer its bankruptcy probability. Managers, investors, banks, and analysts allocate substantial resources to acquire timely and accurate forecasts of future earnings, typically produced by equity analysts. Recently, Hou, van Dijk and Zhang (2012) and others propagate mechanical earnings forecasts as a substitute for equity analysts' earnings predictions. Most importantly, such forecasts are obtained from cross-sectional

[1] Chapter 2 is based on the research article "Incorporating Quarterly Earnings Information into Cross-Sectional Earnings Forecast Models" written by Dieter Hess and Tobias Lorsbach, as of February 2019. Thanks are due to Martin Huettemann, Ashok Kaul, William Liu, Martin Meuter, Nandu Nayar and seminar participants at the 2017 Annual Meeting of the European Accounting Association (EAA), the 2018 European Conference of the Financial Management Association (FMA) and members of the University of Cologne, Saarland University and Cologne Graduate School for insightful comments and helpful suggestions.

regressions on the basis of large samples and therefore are not prone to behavioral biases. Nevertheless, equity analyst's forecasts are typically found to be substantially more precise than regression-based earnings forecasts. We show that this precision advantage is solely due to the fact that previous cross-sectional model specifications neglect important information in quarterly earnings reports while analysts strongly benefit from this information. Our paper levels the playing field by developing a framework to incorporate quarterly earnings releases into cross-sectional models. This allows us to update earnings forecasts more frequently, and at the same time, increase their precision substantially. As a result, our augmented cross-sectional model largely closes the performance gap to analysts and, thus, allows to provide reasonably precise earnings forecasts for the huge number of firms which are not covered by analysts.

Quarterly earnings results provide partial realizations of annual earnings that are observable by investors, financial analysts, and other market participants during the financial year. Unfortunately, the state-of-the-art models of Hou, van Dijk and Zhang (2012) (thereafter HVZ) and Li and Mohanram (thereafter LM) can consider only annual accounting information. In contrast, our model incorporates valuable quarterly earnings results in order to update annual earnings forecasts on a higher frequency. Our approach works as follows: Whenever a company discloses a quarterly earnings number, the forecast task reduces to predicting results for the remaining fiscal quarters. Therefore, in a first step, we disaggregate next year's annual earnings into its already published and its yet unpublished quarterly components. In a second step, we then generate a prediction for the yet unpublished quarterly results. Finally, we aggregate the already published quarterly results and our forecasts for the yet unpublished quarterly results into a forecast for the entire fiscal year. An important aspect is that, our extended approach allows to update mechanical earnings forecasts on a higher frequency, i.e., directly after the release of new quarterly earnings reports. We document that incorporating the valuable new information from quarterly releases into the mechanical models' annual earnings forecasts strongly improves their accuracy. For example, the accuracy of annual earnings forecasts made after the release of first quarter results improve by about 15.2%. Correspondingly, the accuracy of forecasts made after the second quarter improves by about 30.8%, and after the third quarter by about 50.0%. In addition, our model extension also yields superior 2- and 3-year-ahead earnings predictions, i.e., being more accurate by about 16.6% and 11.0%, respectively. Comparing forecast accuracy, we find that analysts cannot outperform models anymore, once we allow the models to draw on quarterly earnings information. Comparing forecast performance across coverage and size portfolios indicates that analyst forecasts can only beat our model forecasts for the very large firms, typically followed by a larger number of analysts. However, model forecasts for smaller firms and less covered firms are superior to analyst forecasts.

Intuitively, more accurate earnings forecasts should also lead to better investment decisions. In fact, we document that our superior earnings forecasts yield better estimates for implied cost of capital (ICC), i.e., the rate of return that equates future earnings forecasts and current stock prices. For example, we find that an investment strategy buying stocks from the upper decile of estimated ICCs and selling stocks from the bottom ICC decile yields a return of 5.64% in the year after portfolio formation, outperforming corresponding strategies based on conventional model forecasts, e.g., from Li and Mohanram (2014), by 0.96% and those based on analyst forecast by 2.69%. And most importantly, we show that this additional return is not due to a higher return volatility. Applying our model forecasts in settings where firms are not covered by analysts (e.g., private firms, small firms, developing countries), we find that portfolio returns are substantially larger. In fact, our model forecasts provide ICCs that yields an excess return of 11.75% outperforming corresponding strategies based on conventional model forecasts, e.g., from Li and Mohanram (2014), by 2.84%.

Our approach is related to two major strands in the finance and accounting literature, i.e., the literature on model-based earnings forecasts and analysts' earnings forecasts. Most closely related to our study, Hou et al. (2012) introduce a cross-sectional model that uses annual financial statements to forecast earnings. They show that their earnings forecasts are less biased than analyst forecasts and provide a stronger link between ICC estimates and future realized returns. Building on this approach, Li and Mohanram (2014) propose two different model specifications and employ a different earnings definition, i.e., earnings-per-share excl. special items. We follow Li and Mohanram (2014) for two reasons: First, their earnings definition comes closest to the "Street" earnings definition that is generally used by financial analysts and, thus, yield a level playing field for benchmarking model to analyst forecasts. Second, their models apparently produce somewhat more accurate predictions. However, an important issue in both Hou et al. (2012) and Li and Mohanram (2014) is the timing of model and analyst forecasts. Both studies compute model forecasts at the end-of-June and compare them to most recently available analyst forecasts at that point in time. However, for most firms this procedure grants analysts a substantial information advantage over models. For instance, firms with December fiscal-year-end have already published their first quarterly result in June. While this information helps analysts to improve their forecasts (see, e.g., Bluemke, Hess and Stolz, 2017), it cannot be picked up by the models of Hou et al. (2012) or Li and Mohanram (2014). In contrast, we show how to exploit this information to improve cross-sectional forecast models.

Early approaches of model-based earnings forecasts use time-series models to predict quarterly earnings results, as such provide longer time-series (see, e.g., Bradshaw, Drake, Myers and Myers, 2012). Ball and Ghysels (2017) are the most recent researchers attempting time-series models to predict future earnings. However, those models have commonly very high data requirements, i.e., implementation generally requires at least ten years of quarterly or even 40 years of annual data. And for the lack of firms with sufficiently long historical data, those models cover only very small subsamples and reflect survivorship and success biases, which makes these approaches impractical in the context of asset pricing and market efficiency tests.

Traditionally, equity analysts, in their function as information catalysts, provide capital markets with future earnings projections. However, the extant literature concludes that analyst forecasts are strongly over-optimistic and driven by incentives, such as career concerns (e.g., Abarbanell, 1991; O'Brien, 1988; Bradshaw, Drake, Myers and Myers, 2012; Richardson, Teoh and Wysocki, 2004). In addition, long-term forecasts from analysts tend to be optimistically biased, grossly inaccurate and, from a valuation perspective, essentially meaningless (e.g., La Porta, 1996; Chan, Karceski and Lakonishok, 2003). Hence, it is not surprising that little to no evidence of a relation between analysts-based return proxies and future stock returns is found (e.g., Easton and Monahan, 2005; Easton and Sommers, 2007; Hou et al., 2012; Lee, So and Wang, 2015; Penman, Reggiani, Richardson and Tuna, 2015). More recent studies (e.g., Gode and Mohanram, 2013; Larocque, 2013) try to adjust analysts' earnings forecasts for their firm-specific bias in forecast errors. However, these adjustments moderately reduce the absolute analysts' forecast errors, and at the same time, their high data demands dramatically reduce sample sizes. In contrast, we confirm previous findings that only ICCs based on model's earnings forecasts are correlated to stock returns (e.g., Hou et al., 2012; Li and Mohanram, 2014; Hess, Kaul and Meuter, 2018). Furthermore, research services are costly and, hence, only subject to very large firms with potential trading sales for brokers and banks.

A growing body of finance and accounting research uncover the strong predictive power of accounting information, such as profitability and firm-level earnings, on the cross-section of average returns (see for example, Fama and French, 2006; Hou et al., 2012; Novy-Marx, 2013; Lewellen 2014). Earlier studies use current/trailing accounting information to predict future returns, mostly for the lack of reliable future earnings predictions (see, e.g., Lee, Ng and Swaminathan, 2009; Botosan, Plumlee and Wen, 2011; Lyle and Wang, 2015). However, Hou et al. (2012) develop a cross-sectional model to forecast annual earnings for a very broad set of firms using only current financial statements data. More recent studies already use the explicit forward-looking information from earnings forecast models in asset pricing and market efficiency tests, commonly firm-level earnings forecasts (see, e.g., Larocque and Lyle, 2013; Rusticus, 2014; Wang, 2015; Lee, So and Wang, 2016) or even higher moments of future earnings (e.g., Chang, Monahan, Quazad and Vasvari, 2014). Several other model specifications follow that employ additional accounting variables (e.g., Li and Mohanram, 2014; Ashton and Wang, 2013; Chang et al., 2014). But, most importantly those models have in common that they can solely utilize annual statements and neglect recent and more frequent intra-year information, such as quarterly earnings results. In this paper, we show how this information can be infused to future earnings expectations and whether this yields more reliable proxies for future stock returns.

With our approach we overcome the limitations of previous studies: One the one hand, our model allows to incorporate intra-year information, i.e., quarterly earnings results, and produces substantially better earnings predictions when quarterly earnings information becomes available. This allows us to make models' forecasts more informative and at the same time to provide high-frequency earnings expectations. Thus, annual earnings forecasts contain the most recent firm-level information for asset pricing and market efficiency tests. On the other hand, we overcome the limitations of previous studies in comparing model and analyst forecasts. We level the playing field against analysts as we facilitate the use of information in forecast models that is broader in scope and more frequently observed. We reduce analysts' information advantages by incorporating quarterly earnings data into model forecasts and show that the accuracy of model forecasts is competitive to analysts. This is striking as analysts should still benefit from further public non-accounting information that is so far neglected in forecast models, such as stock returns or macroeconomic indicators. Hence, we offer evidence that model and analyst forecasts contain complementary information.

Furthermore, our extended framework produces annual earnings forecasts on a high frequency. While this study computes earnings forecasts for each month, we could, in fact, compute forecasts for every day or hour. Our approach enables portfolio managers to adjust their portfolio strategies directly after earnings releases, even before equity analysts can issue new estimates. Such a timing advantage is particularly appealing to practitioners and shows that the process of forecasting firm-level earnings and anticipation of new stock market information can be automated. This is, in fact, most relevant in settings where firms are not covered by analysts (e.g., private firms, small firms, developing countries). In addition, our framework is fundamental, as it allows future research to add further valuable firm- and market-level data to earnings forecast models, for example, daily stock returns, monthly GDP, or ad-hoc disclosures. In addition, model forecasts can be adjusted to predict financial statement numbers that are not covered by professional analysts, such as sales, EBIT, cash flows or accounting accruals.

Moreover, this study has further implications on the ongoing discussion about the reliability of accounting-based return proxies. Our results offer evidence that more accurate earnings forecasts provide better estimates of future stock returns. Hence, the performance of accounting-based return proxies is conditional on the reliability of the input factors used in various ICC models. We show that only model forecasts provide reasonable measures of future earnings expectations and yields significant correlations to future stock returns. Therefore, future research should focus on the use of model forecasts and whether forecast quality can be further improved.

In addition, several implications of our study are more applied. Foremost, our results shed light on the potential of forecast models to provide high frequency earnings expectations. These earnings expectations are relevant for a broad set of event studies and applicable to settings with reliable public data, but no analyst coverage (e.g., small public firms, private firms or developing countries). Furthermore, we highlight the potential of model-based forecasts to obtain earnings expectations that apparently provide complementary information to those of analysts and to automate the forecasting procedure. Especially the latter is an appealing idea to incorporate increasing information and data availability, in situation where human cognitive capabilities do not suffice for information processing. Overall, our results strengthen the claim that earnings forecasts models provide an appealing approach by enlarging the research coverage of companies and, thus, broadening the investment scope for investors. However, our forecasting model may also serve research firms to potentially digitalize the process of forecasting and restructure costs to meet the increasing competition from regulatory challenges such as MiFID II.

Section 2.2 explains our model extension for mechanical earnings forecasts. Section 2.3 describes our dataset and the estimation methodology. In Section 2.4, we evaluate the resulting improvements in forecast performance and benchmark forecast models to financial analysts. We outline our conclusions and the direction of future research in Section 2.5.

2.2 Model development

Hou, van Dijk and Zhang (2012) introduce cross-sectional regressions to predict corporate earnings. Li and Mohanram (2014) and Konstantinidi and Pope (2016) propose to use modified specifications. Like equity analysts these models focus on predicting earnings for a given fiscal year. In all models lagged annual earnings is the core variable, driving the models' forecast performance for the most part.[2] For ease of exposition, we start with a strongly reduced model version, which represents the least common ground of the previous models, i.e., we regress a company's current annual earnings, A_t, on its corresponding annual earnings during the previous fiscal year:

$$A_t = \alpha_0 + \alpha_1 A_{t-1} + \varepsilon_t \tag{2.1}$$

[2] While Hou et al. (2012) and Li and Mohanram (2014) directly regress current earnings on lagged earnings and accruals, Konstantinidi and Pope (2016) decompose lagged earnings into cash flows and accruals.

In such a model, one-year-ahead forecasts are simply obtained as

$$A_{t+1} = \hat{\alpha}_0 + \hat{\alpha}_1 A_t \tag{2.2}$$

The main innovation of our approach is that, besides lagged annual earnings, we include recent quarterly earnings information from the companies' interim reports. This allows for meaningful forecast updates throughout the year as new information about the current fiscal year becomes available. We show that this interim information improves forecasts via two different channels.

First, it allows us to cancel out forecast errors by substituting quarterly forecast with their corresponding realizations, one after another. For example, if the results for the first quarter are already known, we just have to forecast the remaining three quarters and, thus, avoid a forecast error for the first quarter. Hence, a forecast for a given fiscal year becomes more and more precise as quarterly results are published one after another. While equity analysts naturally use quarterly results to update their forecasts, so far existing cross-sectional models cannot incorporate this information and, therefore, provide inferior forecasts.

Second, using recent quarterly information allows us to pick up new trends early. For example, if first quarter results turn out to be surprisingly strong, it may signal more strength for the quarters ahead. To exploit persistence in quarterly results, we track innovations in recently published quarterly earnings by including a variable that captures year-over-year changes in quarterly results. This second mechanism yields substantial forecast error improvements, as well.

Before we start delineating our model, it is important to note that a firm's fiscal year may differ from the calendar year. Therefore, at any given month we find some firms that have just published their annual statements, while others have already reported earnings for their first, second or third fiscal quarter. In general, we can distinguish four different types of firms, corresponding to number of quarterly reports, q, they have already disclosed for the current fiscal year, i.e., $q = 0, 1, 2, 3$.

Let Q_t^i denote a firm's earnings per share in quarter i of fiscal year t and $Q_t^{\{i,\dots,j\}}$ denote the sum of quarterly earnings from quarter i to quarter j. Then we can rewrite earnings for an entire fiscal year as the sum of the corresponding four quarterly earnings:

$$A_t = Q_t^1 + Q_t^2 + Q_t^3 + Q_t^4 = Q_t^{\{1,\ldots,4\}} \tag{2.3}$$

Substituting $A_t = Q_t^{\{1,\ldots,4\}}$ into (2.1) yields:

$$Q_t^{\{1,\ldots,4\}} = \alpha_0 + \alpha_1 Q_{t-1}^{\{1,\ldots,4\}} + \varepsilon_t \tag{2.4}$$

For firms that have just published their previous annual report ($q = 0$), it makes no difference whether we use equation (2.4) or (2.1) to obtain regression coefficients and, in a second step, predictions. However, think of firms that have already published first quarter results, i.e., $q = 1$. Since we already know their first quarter earnings, we simply need to forecast earnings for the remaining three fiscal quarters. In order to obtain such forecasts, we would like to run a different regression for these firms, i.e., regressing the sum of their earnings for the last three fiscal quarters on the earnings for the corresponding three quarters of the previous fiscal year:

$$Q_t^{\{2,\ldots,4\}} = \alpha_0 + \alpha_1 Q_{t-1}^{\{2,\ldots,4\}} + \varepsilon_t \tag{2.5}$$

To calculate now a forecast for the entire fiscal year t+1, we take the three-quarter forecast, i.e., $\hat{\alpha}_0 + \hat{\alpha}_1 Q_t^{\{2,\ldots,4\}}$ and add this to the already published first quarter earnings Q_{t+1}^1:

$$\hat{A}_{t+1} = Q_{t+1}^1 + \left(\hat{\alpha}_0 + \hat{\alpha}_1 Q_t^{\{2,\ldots,4\}}\right) \tag{2.6}$$

We can use the same approach for firms having already published results for their second quarter ($q = 2$) or their third quarter ($q = 3$). Consequently, we get four different regression equations, corresponding to the number of quarterly results a firm has already published. Let $Q_{t+1}^{\{1,\ldots,q\}} \equiv \sum_{i=1}^{q} Q_{t+1}^i$ describe the sum of already published quarterly results for fiscal year t+1, and let $Q_{t+1}^{\{q+1,\ldots,4\}} \equiv \sum_{i=q+1}^{4} Q_{t+1}^i$ denote the sum of the remaining quarterly results. Then the four regression equations for $q = 0, 1, 2, 3$ can be written more generally as

$$Q_t^{\{q+1,\ldots,4\}} = \alpha_{0,q} + \alpha_{1,q} Q_{t-1}^{\{q+1,\ldots,4\}} + \varepsilon_t \tag{2.7}$$

and correspondingly. the prediction equations for $q = 0, 1, 2, 3$ as

$$\hat{A}_{t+1} = Q_{t+1}^{\{1,\ldots,q\}} + \left[\hat{\alpha}_0(1-q) + \hat{\alpha}_1 Q_t^{\{q+1,\ldots,4\}}\right] \tag{2.8}$$

Note that if we would estimate equation (2.7) separately for $q = 0, 1, 2, 3$, we would quadruple the number of estimated parameters. That is, we would get four different intercepts ($\alpha_{0,q}$ one for each q) as well as four different slope parameters ($\alpha_{1,q}$). However, it makes little sense to assume different slope parameters for each fiscal quarter, as this would imply that the persistence of earnings is different across the four quarters. Moreover, an increasing number of parameters typically comes with a strong disadvantage. While more parameters would allow for a better in-sample fit, they tend to impair out-of-sample forecasts at the same time.[3] In order to avoid an unnecessary inflation of parameters, we impose adequate restrictions on the parameters. First, we require that the earnings persistence parameters are identical across all four quarters, i.e., we assume $\alpha_{1,q} = \alpha_1$ for $q = 0, 1, 2, 3$. Second, we also impose adequate restrictions on the intercept. Note however that the intercept is proportional to the time interval to which the earnings are related. For example, for $q = 0$ we model profits for the entire fiscal year, while for $q = 1$ we only model three quarters of the year. Hence, without restricting the intercepts we should observe that $\alpha_{0,1}$ is approximately $\frac{3}{4}\alpha_{0,0}$. Similarly, we should find that $\alpha_{0,2} \approx \frac{2}{4}\alpha_{0,0}$ for $q = 2$ and $\alpha_{0,3} \approx \frac{1}{4}\alpha_{0,0}$ for $q = 3$. Hence, we can easily impose a restriction on the proportions of intercepts for different q by defining a factor $f_q = \frac{(4-q)}{4}$ which can guarantee $\alpha_{0,q} = f_q \alpha_{0,0}$. Based on these restrictions for the intercept and the slope parameters, we can now combine the four individual equations for all into a single (restricted) regression:

$$Q_t^{\{q+1,\dots,4\}} = \alpha_0 \cdot f_q + \alpha_1 Q_{t-1}^{\{q+1,\dots,4\}} + \varepsilon_t \tag{2.9}$$

This allows us to run a single regression for the all firms while at the same time incorporating all of their most recently disclosed quarterly earnings information. Accordingly, we can write the prediction equation as

$$\hat{A}_{t+1} = Q_{t+1}^{\{1,\dots,q\}} + \left[\hat{\alpha}_0 \cdot f_q + \hat{\alpha}_1 Q_t^{\{q+1,\dots,4\}}\right] \tag{2.10}$$

By differentiating between the number of already published quarterly reports, equations (2.9) and (2.10) allow to incorporate successively incoming interim earnings

[3] In fact, we find that restricting the parameter space increase the precision of our out-of-sample forecasts significantly.

information and, thus, to update fiscal year earnings forecasts over time. Note that such an updating behavior is common among equity analysts. For example, Bluemke et al. (2018), report that the majority of equity analysts update their annual forecasts within 20 days after an interim release. Moreover, they observe a significantly higher precision for those forecast that were updated. Similarly, we find that allowing cross-sectional models to pick up interim information significantly improves their out-of-sample forecasts.

In essence equation (2.9) yields substantially improved forecasts, simply by avoiding predictions of already announced quarterly results. Unsurprisingly, substituting quarterly forecasts with corresponding quarterly realizations reduces the overall error in annual forecasts. In addition to that, we think that picking up new quarterly earnings releases should provide a second benefit. Specifically, we assume that information about (surprisingly) high or low growth in the first fiscal quarters helps to improve predictions for the remaining fiscal quarters of the same year. To model possible persistence in quarterly earnings growth rates, we construct a variable that accounts for year-over-year changes of quarterly earnings during the first q quarters. Precisely, we define $\Delta Q_{t,t-1}^{\{1,\dots,q\}} = Q_t^{\{1,\dots,q\}} - Q_{t-1}^{\{1,\dots,q\}}$ for $q = 1, 2, 3$ and $\Delta Q_{t,t-1}^{\{1,\dots,q\}} = 0$ for $q = 0$. To preserve the parsimony of the model, again we need a factor that accounts for the fact that our new right-hand side variable, $\Delta Q_{t,t-1}^{\{1,\dots,q\}}$, relates to a different time frame than our left-hand side variable $Q_t^{\{q+1,\dots,4\}}$. For example, for $q = 1$ the difference term relates to just a single quarter, while $Q_t^{\{q+1,\dots,4\}}$ covers three quarters. That is, we would try to project the growth of the first quarter onto the remaining three quarters. In contrast, for q we project the growth of the first two quarters on the remaining two quarters. To account for these changing proportions, we therefore define the factor g_q as

$$g_q = \begin{cases} 0 & q = 0 \\ \frac{4-q}{q} & q = 1, 2, 3 \end{cases} \tag{2.11}$$

Adding $g_q \cdot \Delta Q_{t,t-1}^{\{1,\dots,q\}}$ to equation (2.9) yields the regression equation

$$Q_t^{\{q+1,\dots,4\}} = \alpha_0 \cdot f_q + \alpha_1 \cdot Q_{t-1}^{\{q+1,\dots,4\}} + \alpha_2 \cdot g_q \cdot \Delta Q_{t,t-1}^{\{1,\dots,q\}} + \varepsilon_t \tag{2.12}$$

Correspondingly, we obtain the prediction equation

$$\hat{A}_{t+1} = Q_{t+1}^{\{1,\dots,q\}} + \left[\hat{\alpha}_0 \cdot f_q + \hat{\alpha}_1 \cdot Q_t^{\{q+1,\dots,4\}} + \hat{\alpha}_2 \cdot g_q \cdot \Delta Q_{t+1,t}^{\{1,\dots,q\}}\right] \quad (2.13)$$

In a final step, we include some additional variables to facilitate comparisons with previous models. For example, besides lagged earnings the Hou et al. (2012) model contains lagged total assets (TA_{t-1}), annual dividend payments (D_{t-1}), annual accruals (ACC_{t-1}), a negative earnings dummy (NED_{t-1}) and a dividend payment dummy (DD_{t-1}). In order to add these variables into our extended model, again we need to account for possible mismatches in their timeframes with the timeframe of our left-hand side variable. Fortunately, there is one pattern for all these variables, i.e., they are fixed for a given fiscal year while our left-hand variable, $Q_t^{\{q+1,\dots,4\}}$, decreases monotonically with q. Therefore, we can adjust all the above-mentioned variables using the same proportionality factor as for our intercept, i.e., f_q. Then an augmented version of the Hou et al. (2012), hereafter HVZ, model can be written as

$$\begin{aligned} Q_t^{\{q+1,\dots,4\}} = {} & \alpha_0 \cdot f_q + \alpha_1 \cdot Q_{t-1}^{\{q+1,\dots,4\}} + \alpha_2 \cdot g_q \cdot \Delta Q_{t,t-1}^{\{1,\dots,q\}} \\ & + \alpha_3 \cdot f_q \cdot NED_{t-1} + \alpha_4 \cdot f_q \cdot TA_{t-1} \\ & + \alpha_5 \cdot f_q \cdot DD_{t-1} + \alpha_6 \cdot f_q \cdot D_{t-1} \\ & + \alpha_7 \cdot f_q \cdot ACC_{t-1} + \varepsilon_t \end{aligned} \quad (2.14)$$

Correspondingly, an augmented version of the earnings persistence (EP) model of Li and Mohanram (2014) is obtained as

$$\begin{aligned} Q_t^{\{q+1,\dots,4\}} = {} & \alpha_0 \cdot f_q + \alpha_1 \cdot Q_{t-1}^{\{q+1,\dots,4\}} + \alpha_2 \cdot g_q \cdot \Delta Q_{t,t-1}^{\{1,\dots,q\}} \\ & + \alpha_3 \cdot f_q \cdot NED_{t-1} \cdot Q_{t-1}^{\{q+1,\dots,4\}} \\ & + \alpha_4 \cdot f_q \cdot NED_{t-1} + \varepsilon_t \end{aligned} \quad (2.15)$$

In addition, we obtain an augmented version of the residual income model (RI) of Li and Mohanram (2014), with BV_{t-1} denoting a firm's lagged book value of equity:

$$Q_t^{\{q+1,\dots,4\}} = \alpha_0 \cdot f_q + \alpha_1 \cdot Q_{t-1}^{\{q+1,\dots,4\}} + \alpha_2 \cdot g_q \cdot \Delta Q_{t,t-1}^{\{1,\dots,q\}}$$
$$+\alpha_3 \cdot f_q \cdot NED_{t-1} \cdot Q_{t-1}^{\{q+1,\dots,4\}}$$
$$+\alpha_4 \cdot f_q \cdot NED_{t-1} + \alpha_5 \cdot f_q \cdot BV_{t-1}$$
$$+\alpha_6 \cdot f_q \cdot ACC_{t-1} + \varepsilon_t \tag{2.16}$$

Appendix A provides a detailed description of the regression and prediction equations for all models.

To summarize, the above described extensions of the standard cross-sectional earnings regression models changes the way how historical earnings information is utilized. Instead of restricting the models to draw solely on annual earnings releases, our approach allows to consistently include quarterly earnings information as well. At the same time, we show how to remain the parsimony of the models. Instead of estimating the model separately for each quarter, we introduce appropriate scaling factors, f_q and g_q, that account for differences in the timeframes of the left- and right-hand side variables. These scaling factors facilitate a (restricted) estimation for all four quarters at the same time. Untabulated results show that this parsimonious model yields more stable parameter estimates and, in addition, improved out-of-sample forecasts.

2.3 Methods and data

2.3.1 Mechanical earnings forecasts

Hou et al. (2012) and Li and Mohanram (2014) has developed three main models, i.e., HVZ, EP and RI model, that utilize a firms’ financial statements to forecast their future earnings results. However, in their former version, these models incorporate only the latest available annual financial statement. In contrast, we show how to incorporate very informative and more frequent quarterly earnings releases into these models. This extension strongly improves the forecast quality of model predictions and is applicable to all former versions of the HVZ, EP and RI model. For reasons of plausibility, we implement the HVZ, EP and RI model in their standard version used in

Hou et al. (2012) and Li and Mohanram (2014) and in an extended version that adapts our innovative approach. Thus, we consistently trace the forecast quality of our extended model forecasts to those from prior versions. The comparisons between the standard and extended versions are very similar across all three models. Therefore, we abbreviate our section on the empirical results and provide detailed results primarily for the RI model. However, Appendix B contains a concise overview of our results regarding the forecast quality of the EP and HVZ models in their extended versions.

We follow Hou et al. (2012) and run rolling regressions on windows of ten-years of data. However, in contrast to Hou et al. (2012), Li and Mohanram (2014) and others who run the regressions once a year in June, we estimate the models each month in between January 1982 and December 2014. This procedure is more in line with practitioners demands and how financial analysts issue forecasts as it yields annual earnings forecasts for each month. Furthermore, it allows us to incorporate new information more frequently in each month. Models that utilize new information on a higher frequency produce better expectations of firm's financials. However, only if the models can pick up new information arriving during the year, such as quarterly earnings reports. By incorporating quarterly earnings information into cross-sectional regressions, we also level the playing field for comparisons of professional analysts and mechanical models.

Therefore, we compute annual earnings forecasts for horizons of one- to five-years ahead and examine the accuracy and the bias of model-based earnings forecasts and those of analysts. In detail, forecast bias is defined as the signed price-scaled forecast error (PFE), i.e. the difference between actual earnings per share and forecasted earnings per share, whereas forecast accuracy is defined as the absolute price-scaled forecast error (PAFE). Both signed and absolute forecast errors are scaled by stock prices at the previous fiscal year-end. To benchmark our results with analysts' earnings forecasts, we partition our sample into firms covered by analysts and those not covered.

2.3.2 Implied cost of capital (ICC) models

Implied cost of capital (ICC) is the rate of return that equates the present value of expected future earnings to the current stock price. In line with Hou et al. (2012), we use the following five valuation models to generate ICC estimates: the Ohlson and Juettner-Nauroth (2005) abnormal earnings growth model (OJ) , the Gebhardt, Lee and Swaminathan (2001) and Claus and Thomas (2001) residual income valuation models (GLS and CT, respectively), the Easton (2004) price-earnings to growth model (PEG) and the Gordon and Gordon (1997) dividend discount model without growth (GG). Appendix C provides a detailed description of these ICC models. Like Hou et al. (2012), we also employ the average of the five individual ICC estimates to obtain a robust composite ICC. Moreover, we follow previous studies such as Lewellen (2014) and Lee et al. (2015) in excluding ICCs ranging outside 0% and 100%.[4] For explicit formulas and detailed explanations of each individual ICC model see Appendix C.

Some ICC models require longer-term earnings forecasts to solve for cost of capital, for example, GLS requires forecasts with horizons up to three years, CT requires even 5-year-ahead forecasts. While this posits no problem to mechanical earnings forecasts, as in general they can be generated for longer-term horizons as well, explicit analysts' long-term forecasts are rarely available for horizons beyond two years. To circumvent this problem, we apply the I/B/E/S consensus long-term growth projection to construct earnings forecasts for three-, four- and five-year ahead forecast horizons. For one- and two-year forecast horizons we use the consensus earnings forecast from I/B/E/S. Like Hou et al. (2012) we use actual (street) earnings from I/B/E/S to calculate forecast errors of analysts. To evaluate the models, we follow Li and Mohanram (2014) and employ a comparable "Street" earnings definition, i.e., earnings before extraordinary items less special items from Compustat.

2.3.3 Data and sample characteristics

Our sample consists of all U.S. firms in the intersection of the Compustat's North-America fundamentals annual file and the CRSP's monthly stock file. We exclude observations for which we cannot obtain the required variables to estimate model-based earnings forecasts, in particular, annual and quarterly earnings, book equity and total

[4] Alternatively, Gode and Mohanram (2013) and Larocque (2013) trim their ICC estimates at each estimation date across the cross-section of firms for the 1% percentiles. But this step seems to exclude observations more inconsistent, i.e., it may also eliminate economically valid ICC estimates within the range of 0-100%.

assets.[5] In addition, we screen the data for obvious errors.[6] We assume that quarterly reports become available within three months, as the SEC requires even small firms to file quarterly results (10-Q) at least within 90 days. To benchmark forecast performance of the cross-sectional models to that of equity analysts, we collect most recent annual EPS median estimates from I/B/E/S summary files. As a reasonable coverage of I/B/E/S analysts' earnings forecasts starts in 1982, we select this point in time as our sample start.

Table 2.1 Descriptive sample statistics of cross-sectional regression variables,1982-2014 (N = 167,921)

Variable	Mean	1%	25%	Median	75%	99%	Std
AE_t	0.90	-4.00	-0.07	0.59	1.64	7.45	1.73
$QE_{t,1}$	0.20	-1.00	-0.02	0.12	0.36	1.89	0.43
$QE_{t,2}$	0.24	-1.04	-0.01	0.15	0.41	2.12	0.47
$QE_{t,3}$	0.24	-1.20	-0.01	0.15	0.43	2.18	0.49
$QE_{t,4}$	0.22	-1.95	-0.04	0.15	0.46	2.46	0.60
BV_t	9.89	-2.47	2.43	6.80	14.02	52.28	10.30
TA_t	42.64	0.18	5.19	15.32	41.12	508.10	80.61
D_t	0.37	0.00	0.00	0.02	0.48	3.33	0.66
ACC_t	-1.01	-12.83	-1.46	-0.39	-0.01	5.37	2.38
NED_t	0.28	0.00	0.00	0.00	0.78	1.00	0.45
DD_t	0.48	0.00	0.01	0.20	1.00	1.00	0.50
p_t	18.66	0.23	4.56	12.38	25.72	91.40	24.74
RoE_t	0.13	0.09	0.12	0.13	0.14	0.17	0.02
pr_t	0.20	0.00	0.00	0.01	0.32	1.00	0.31

Note: This table reports the descriptive statistics of the variables in the cross-sectional earnings models and variables used in the ICC estimation. The reported values are time-series averages of monthly cross-sectional means, standard deviations and respective percentiles. All values are on per-share basis except of the dummy variables ($NED_t$and DD_t), price p_t, return-on-equity roe_t and the payout ratio pr_t. Financial statement information is obtained from Compustat, whereas stock market information stems from CRSP. A detailed description of the variables and data sources is provided in Appendix D. The number of observations is the number of firm-years within the sample.

Table 2.1 provides summary statistics of the required accounting and stock market variables used in the regressions and the ICC models. For ease of exposition, we compress the interaction term of the earnings variable and negative earnings dummy as IAT_t. Moreover, to minimize the effect of outliers, we winsorize all variables at the 1st and 99th percentiles. We report time-series averages with Newey-West corrected t-statistics. Even though our data requirements are slightly more demanding as we require quarterly earnings information for the current year t, we collect a large sample of 167,921 firm-years. In line with previous research, our sample contains accruals which are

[5] On the other hand, we set missing accruals to zero. The reason is that accruals contribute virtually nothing to improving forecasts but dropping observations would result in a substantial reduction of the sample.

[6] In particular, we exclude observations for which Compustat reports zero total assets, zero shares outstanding or zero stock prices.

negative on average and we find several firms with negative earnings. According to total assets, our sample also covers a broad range of small and large firms. Overall, Table 2.1 shows that our sample data is comparable to those used by Hou et al. (2012) or Li and Mohanram (2014). A detailed description of all variables is provided in Appendix D.

2.4 Empirical results

2.4.1 Regression coefficients and earnings predictions

Table 2.2 Coefficients from residual income (RI) regression model

		EPS_{t+1}				EPS_{t+2}				EPS_{t+3}			
	q	**Extended RI**		**Standard RI**		**Extended RI**		**Standard RI**		**Extended RI**		**Standard RI**	
Int		-0.005		-0.005		0.028		0.028	***	0.077		0.077	***
$Q_{t-\tau}^{\{q+1\dots4\}}$	0	0.831	***	0.885	***	0.724	***	0.816	***	0.633	***	0.777	***
	1	0.767	***			0.684	***			0.626	***		
	2	0.706	***			0.651	***			0.633	***		
	3	0.587	***			0.565	***			0.627	***		
$\Delta Q_{t+1-\tau,t}^{\{1\dots q\}}$	1	0.392	***			0.249	***			0.169	***		
	2	0.486	***			0.191	***			0.098	***		
	3	0.580	***			0.030	***			-0.195	***		
$NED_{t-\tau}$		-0.214	***	-0.146	***	-0.265	***	-0.214	***	-0.329	***	-0.283	***
$IAT_{t-\tau}$		-0.350	***	-0.304	***	-0.483	***	-0.454	***	-0.579	***	-0.570	***
$BV_{t-\tau}$		0.025	***	0.017	***	0.045	***	0.031	***	0.066	***	0.046	***
$ACC_{t-\tau}$		-0.018	***	-0.021	***	-0.049	***	-0.042	***	-0.061	***	-0.051	***
Adj. R^2		68.25%		64.48%		47.54%		48.78%		41.01%		40.35%	

Note: This table reports the estimations from pooled cross-sectional regressions for each estimation date from 1982 to 2014 using the previous ten years of data. It tabulates the time-series averages of the regression coefficients for each model version and different forecast horizons τ=1,2,3 The time-series t-statistics are estimated trough Newey-West error corrections. The significance of the Newey-West time-series t-statistics are given as *** for the 1%-level, ** for the 5%-level and * for the 10%-level. We estimate q-specific coefficients for earnings variables in the extended RI model. $\Delta Q_{t+1-\tau,t-\tau}^{\{1\dots q\}}$ incorporates recent quarterly earnings information to the forecast model. Facilitating the interpretation, we standardized this coefficient to the value of one to account for the different level of earnings to the LHS variable. In detail, we multiplied the $\Delta Q_{t+1-\tau,t-\tau}^{\{1\dots q\}}$variable by three for the estimation after the first quarterly earnings announcement as alignment to the three remaining quarterly earnings on the LHS in the regression model, whilst we divide by three after the Q3 announcement, respectively. A detailed description of the variables is provided in Appendix D. The regression and prediction equations for the extended and standard version of cross-sectional models is formulated in Appendix A. The coefficient estimates of the EP and HVZ model are consistently similar and available upon requests.

Table 2.2 reports the estimated coefficients from our rolling window regressions. We contrast the coefficients for our extended version of the RI model with those for the standard RI model for forecast horizons from one to three years. Like previous studies, signs and sizes of the coefficients are consistent with economic intuition (see, e.g., Hou et al., 2012; Li and Mohanram, 2014; Hess et al., 2018). In the extended model, the coefficient for lagged earnings ($Q_{t-\tau}^{\{q+1,\dots,4\}}$) is largest for $q = 0$, i.e., 0.831. Intuitively, it is almost identical to the lagged earnings coefficient in the standard model that uses only annual earnings (0.885). This coefficient decreases for longer forecast horizons, i.e., to 0.729 for two-year and 0.633 for three-year horizons. For $q = 1, 2, 3,$ we observe monotonically decreasing lagged earnings coefficient $Q_{t-\tau}^{\{q+1,\dots,4\}}$, for example, for the one-year forecast horizon, we observe coefficient of 0.767 for $q = 1$, 0.706 for $q = 2$ and 0.587 for $q = 3$. Thus, the coefficient changes when companies release further quarterly earnings information. The picture is very similar for the other forecast horizons. While this may be astonishing at first glance, note that at the same time we obtain an increasing coefficient for quarterly earnings growth, e.g., for the one-year forecast horizon, we observe 0.392 for $q = 1$, 0.486 for $q = 2$ and 0.580 for $q = 3$. This supports the notion that information about growth in the first quarters of a year helps to improve the earnings predictions for the remaining quarters of the same year. In fact, the results suggest that recent growth is more informative than lagged earnings values. Hence, the extended model allows better coefficient estimates for earnings variables as it differentiates between changing intra-year forecast horizons.

Negative coefficients for the loss dummy $NED_{t-\tau}$ and the interaction term $IAT_{t-\tau}$ supports the notion that losses are less persistent than profits (see, e.g., Li, 2011). The sizes and signs of the coefficients are in line with the standard model. Further accounting variables, such as book equity $BV_{t-\tau}$ and accruals $ACC_{t-\tau}$ are also related to future earnings. The coefficient of book equity is persistently larger for the extended model, we observe 0.025 for one-year forecast horizons, 0.045 for two-year forecast horizons and 0.066 for three-year forecast horizons, respectively. Thus, a company's current profitability is increasingly important for predicting earnings at longer horizons. Accruals are a proxy for capital expenditures and positively related with future profitability (e.g., Fama and French, 2000, 2006). Since there are many firms with negative accruals $ACC_{t-\tau}$, the coefficient is also negative and equivalent across both models.[7] Our model extension improves the explanatory power of the regression model and explains an additional portion of variation in earnings. In detail, the adjusted R^2 of

[7] See Table 2.1, which summarizes the descriptive statistics of the regression variables.

our extended RI model is 68.3%, whereas the standard RI model explains 64.5% of the variation in earnings.

Table 2.3 Descriptive statistics of earnings predictions, 1982-2014

Variable	Mean	1%	25%	Median	75%	99%	Std	Avg. N
Extended Residual Income Model Predictions								
EPS_{t+1}	0.96	-3.32	-0.09	0.63	1.70	7.65	1.72	167,921
EPS_{t+2}	1.08	-2.56	-0.02	0.72	1.81	7.67	1.68	167,921
EPS_{t+3}	1.23	-2.26	0.05	0.85	1.98	8.04	1.73	167,921
Standard Residual Income Model Predictions								
EPS_{t+1}	1.00	-2.31	-0.10	0.66	1.69	7.35	1.61	167,921
EPS_{t+2}	1.15	-1.37	-0.03	0.77	1.83	7.60	1.60	167,921
EPS_{t+3}	1.33	-0.78	0.08	0.91	2.03	8.09	1.67	167,921
Analysts Earnings Predictions								
EPS_{t+1}	1.41	-2.29	0.39	1.13	2.17	7.54	1.63	108,484
EPS_{t+2}	1.81	-1.42	0.68	1.46	2.57	8.41	1.70	102,647
EPS_{t+3}	2.24	-1.13	0.97	1.86	3.07	9.62	1.89	88,699

Note: This table reports the descriptive statistics of the earnings-per-share predictions from cross-sectional earnings models and financial analysts for annual results of fiscal year t+1, t+2 and t+3. The reported values are time-series averages of cross-sectional means, standard deviations and respective percentiles. A detailed description of the regression models is provided in Appendix A. The number of observations is the number of firm-years within the sample.

Table 2.3 reports earnings predictions from cross-sectional models and financial analysts for future fiscal year t+1 (FY1), t+2 (FY2) and t+3 (FY3). We find that forecasts from our extended model are spread more widely with a larger standard deviation. In addition, the mean and median of forecasts is slightly smaller than those of the standard model. That is, one-year ahead forecasts from the extended model are \$0.96 per share in mean and \$0.63 in median. The predictions from the standard model are \$1.00 per share in mean and \$0.66 in median, respectively. This may stem from anticipation of earnings surprises in already realized quarterly earnings results. For instance, negative surprises within the first quarters of the current year yields smaller forecasts through our extended model, while standard model forecasts are kept unchanged. Note that analyst forecasts are only available for a subset of our sample, i.e., long-term forecasts (FY3) of analysts exist only for 53% of firms. The median EPS forecast of analysts is \$1.13 for one-year forecast horizons (FY1), \$1.46 for two-year forecast horizons (FY2) and \$1.86 for three-year forecast horizons (FY3), respectively. This is not surprising, as the existing literature extensively documents that analysts generally cover large firms and issue forecasts that tend to be too optimistic.

2.4.2 Forecast errors – Extended versus Standard model

In this part of the analysis, we assess whether our extension improves accuracy and bias of model forecasts by comparing the predictions of our extended model with those from the standard model version.

Panel A of Table 2.4 compares models in terms of the absolute forecast error scaled by stock price (PAFE), i.e., forecast accuracy. The PAFE of extended model forecasts substantially decreases with each quarterly earnings result, i.e., for current-year earnings forecasts EPS_{t+1} (FY1). For example, at the beginning of the year both median PAFEs are virtually identical. That is, the median PAFE of extended model forecasts is 3.71%. Using quarterly earnings information, the median PAFE reduces by appr. 17% with each quarterly release that becomes available. In case that a firm released already three quarterly earnings results median PAFE reduces to 1.72%. In fact, this result shows that model forecasts can be improved by up to 50% using our augmented framework. Similarly, the mean PAFE decreases from 9.65% to 5.26% using extended model forecasts. The picture is very similar for longer forecast horizons. We find that PAFE decreases for longer forecast horizons, i.e., for earnings results in fiscal year t+2 (FY2) and t+3 (FY3). Model forecasts for earnings results in fiscal year t+2 (FY2) are more accurate by up to 16.6% and model forecasts for three years ahead earnings (FY3) by up to 11.0%. This is also very important, because several applications, such as valuation and cost of capital estimation, rely on such reasonable long-term growth predictions. Thus, using model predictions may yield ceteris paribus (c.p.) superior estimates of the terminal value of a company, respectively.

Table 2.4 Forecast quality with the extended framework for all firms, 1982-2014

Panel A: Forecast accuracy of mechanical earnings forecasts

		Extended RI				**Standard RI**				**Difference**				**N**
	After	**Mean**		**Median**		**Mean**		**Median**		**Mean**		**Median**		
$\boldsymbol{EPS}_{t+1,q}$	$Q3_{t+1}$	5.26%	***	1.72%	***	9.75%	***	3.44%	***	4.48%	***	1.72%	***	151,959
	$Q2_{t+1}$	7.03%	***	2.39%	***	9.68%	***	3.46%	***	2.64%	***	1.07%	***	150,789
	$Q1_{t+1}$	8.45%	***	2.92%	***	9.63%	***	3.44%	***	1.18%	***	0.52%	***	150,549
	A_t	9.65%	***	3.71%	***	9.59%	***	3.60%	***	-0.06%	*	-0.10%	***	149,554
$\boldsymbol{EPS}_{t+2,q}$	$Q3_{t+1}$	9.19%	***	3.83%	***	10.80%	***	4.60%	***	1.61%	***	0.76%	***	133,385
	$Q2_{t+1}$	9.96%	***	4.24%	***	10.96%	***	4.66%	***	1.00%	***	0.42%	***	132,304
	$Q1_{t+1}$	10.46%	***	4.43%	***	10.88%	***	4.63%	***	0.42%	***	0.20%	***	132,084
	A_t	10.81%	***	4.86%	***	10.83%	***	4.84%	***	0.02%		-0.02%		131,224
$\boldsymbol{EPS}_{t+3,q}$	$Q3_{t+1}$	10.62%	***	4.81%	***	11.48%	***	5.41%	***	0.86%	***	0.59%	***	117,202
	$Q2_{t+1}$	11.13%	***	5.03%	***	11.67%	***	5.42%	***	0.54%	***	0.39%	***	116,214
	$Q1_{t+1}$	11.42%	***	5.30%	***	11.60%	***	5.45%	***	0.18%	***	0.15%	***	116,012
	A_t	11.86%	***	5.84%	***	11.41%	***	5.43%	***	-0.45%	***	-0.41%	***	115,279

Panel B: Forecast bias of mechanical earnings forecasts

		Extended RI				**Standard RI**				**Difference**				**N**
	After	**Mean**		**Median**		**Mean**		**Median**		**Mean**		**Median**		
$\boldsymbol{EPS}_{t+1,q}$	$Q3_{t+1}$	0.26%	*	0.25%	***	0.88%	***	0.09%		0.62%	**	-0.16%	*	151,959
	$Q2_{t+1}$	0.92%	***	0.27%	***	0.88%	***	0.09%		-0.04%		-0.18%	**	150,789
	$Q1_{t+1}$	1.07%	***	0.20%	***	0.86%	***	0.10%		-0.21%		-0.10%	*	150,549
	A_t	0.75%	**	0.08%		0.79%	**	0.19%		0.04%		0.11%	***	149,554
$\boldsymbol{EPS}_{t+2,q}$	$Q3_{t+1}$	0.66%	**	0.07%		-0.25%		-0.64%	***	-0.92%	***	-0.72%	***	133,385
	$Q2_{t+1}$	0.40%		-0.24%		-0.27%		-0.75%	***	-0.67%	***	-0.51%	***	132,304
	$Q1_{t+1}$	-0.32%		-0.52%	***	-0.27%		-0.62%	***	0.05%		-0.11%	**	132,084
	A_t	-1.41%	***	-1.04%	***	-0.22%		-0.50%	*	1.19%	***	0.54%	***	131,224
$\boldsymbol{EPS}_{t+3,q}$	$Q3_{t+1}$	0.54%		-0.39%	**	-1.69%	***	-1.66%	***	-2.24%	***	-1.28%	***	117,202
	$Q2_{t+1}$	-0.44%		-1.03%	***	-1.64%	***	-1.72%	***	-1.21%	***	-0.70%	***	116,214
	$Q1_{t+1}$	-1.96%	***	-1.65%	***	-1.78%	***	-1.68%	***	0.18%	**	-0.04%		116,012
	A_t	-3.84%	***	-2.64%	***	-1.90%	***	-1.75%	***	1.94%	***	0.89%	***	115,279

Note: This table reports the forecasts performance of our extended and the existing standard RI model in terms of the forecast bias and accuracy for one-year (FY1), two-year (FY2) and three-year (FY3) ahead annual earnings. Bias is the price-scaled forecast error and reported as time-series averages of the median and mean forecast bias. Accuracy is the absolute price-scaled forecast error and reported as time-series averages of the median and mean forecast accuracy. The evaluation is performed with respect to the number of quarterly earnings announcements (q) that are public at the estimation date. The time-series t-statistics are estimated trough Newey-West error corrections. The significance of the Newey-West time-series t-statistics are given as *** for the 1%-level, ** for the 5%-level and * for the 10%-level. In Panel A, we provide the results of the absolute price-scaled forecast error (ACCURACY) for all firms in our sample. Whereas in Panel B, we provide the results of the price-scaled forecast error (BIAS) for all firms in our sample.

In the next step, we compare our extended model with the existing standard model in terms of forecast bias, i.e., signed forecast errors. The sign of forecast errors is particularly interesting to explore whether earnings forecasts are optimistic or pessimistic and its reasons. Several studies (e.g., Abarbanell, 1991; Bradshaw et al., 2012; O'Brien, 1988; Richardson et al., 2004) find that analysts actively bias earnings forecasts to preserve management contacts. In contrast, we expect that model-based forecasts are free of incentive-driven bias. In Panel B of Table 2.4, we show that model's forecasts for one-year horizons are slightly downward biased and, thus, of pessimistic nature. That is, forecasts of the standard model are pessimistic by appr. 0.1% in median and 0.88% in mean. However, since forecast bias in both model's predictions is economically insignificant, our results are in line with the expectation. While two-years ahead forecasts tend to appr. unbiased, long-term forecasts are increasingly optimistic. In detail, the median forecast bias (PFE) is appr. -1.70% for the standard model. The median PFE of the extended model ranges from -0.39% to -2.64%. This bias in longer-term forecasts stems from smaller samples to evaluate long-term forecasts, as more historical data is required. The difference between in-sample and out-of-sample increases for longer forecast horizons as new firms enter the sample while others quit. This result is in line with previous studies, e.g., Hou et al. (2012) and Li and Mohanram (2014).

In further analyses, we evaluate whether the forecast quality of our model is consistent across diverse firm characteristics, i.e., size, age, analysts following, dispersion in analyst forecasts and earnings smoothness (see, e.g., Bradshaw et al., 2011; Guay et al., 2011; Hou et al., 2012). This analysis allows to identify characteristics on a firm-level that may drive the accuracy of model's forecasts. Additionally, we test how model's forecast quality relates to subperiods and macroeconomic conditions when forecasts are issued. To conserve space, we do not report the results for forecast bias and focus on forecast accuracy (PAFE) for one-year ahead forecasts (FY1).

Table 2.5 shows the median PAFE of our extended model and the standard model for one-year ahead forecasts (FY1) according to different subsamples. In this test, we select all forecasts across different forecast horizons, i.e., $q = 0, 1, 2, 3$ and sort firms into quintiles for each year based on one of the following five firm-characteristics: (1) Size, (2) age, (3) analysts following, (4) analysts' dispersion and (5) earnings volatility. Size, age and analyst coverage are proxies for a firms' information environment and visibility. In general, larger, older and more covered firms tend to be more visible, issue more public information and are covered more by the media. Dispersion across analyst forecasts accounts for the business complexity. That is, earnings are more difficult to predict. In addition, we use earnings volatility to infers from the quality of accounting information of a company.[8]

Table 2.5 Relation between model-based forecast accuracy (EPS_{t+1}) and firm characteristics, subperiods and subsamples

Panel A: Median forecast accuracy across portfolios of firm-characteristics

Quintiles			Size		Age		Analysts following		Analyst dispersion		Earnings volatility	
		Ext. RI	7.25%	***	2.40%	***	2.66%	***	1.09%	***	1.61%	***
Low	1	**St. RI**	10.61%	***	3.44%	***	3.95%	***	1.51%	***	2.12%	***
		rel. Diff	-31.7%	***	-30.3%	***	-32.5%	***	-27.5%	***	-24.1%	***
		Ext. RI	3.01%	***	2.55%	***	1.97%	***	1.15%	***	1.44%	***
	2	**St. RI**	4.43%	***	3.69%	***	2.86%	***	1.52%	***	1.87%	***
		rel. Diff	-32.1%	***	-30.8%	***	-31.0%	***	-24.0%	***	-23.1%	***
		Ext. RI	1.94%	***	2.47%	***	1.60%	***	1.35%	***	1.79%	***
Medium	3	**St. RI**	2.84%	***	3.48%	***	2.19%	***	1.78%	***	2.41%	***
		rel. Diff	-31.7%	***	-28.9%	***	-27.0%	***	-24.0%	***	-25.4%	***
		Ext. RI	1.41%	***	2.01%	***	1.37%	***	1.75%	***	2.54%	***
	4	**St. RI**	1.92%	***	2.76%	***	1.85%	***	2.42%	***	3.75%	***
		rel. Diff	-26.4%	***	-27.1%	***	-26.2%	***	-27.8%	***	-32.3%	***
		Ext. RI	1.21%	***	1.68%	***	1.22%	***	2.94%	***	3.88%	***
High	5	**St. RI**	1.52%	***	2.26%	***	1.54%	***	4.17%	***	5.80%	***
		rel. Diff	-20.6%	***	-26.0%	***	-21.0%	***	-29.5%	***	-33.2%	***

Panel B: Median forecast accuracy across subperiods

	1982 - 1989		1990 - 1997		1998 - 2006		2007 - 2014	
Ext. RI	2.77%	***	2.21%	***	1.81%	***	1.92%	***
St. RI	3.75%	***	3.07%	***	2.56%	***	2.82%	***
rel. Diff	-26.3%	***	-27.9%	***	-29.1%	***	-32.1%	***

Panel C: Median forecast accuracy for economic cycle peaks and troughs

	NBER				CFNAI					
	troughs		others		peaks		troughs		others	
Ext. RI	2.68%	***	2.17%	***	2.15%	***	2.29%	***	2.12%	***
St. RI	3.85%	***	3.08%	***	2.94%	***	3.28%	***	2.98%	***
rel. Diff	-30.3%	***	-29.5%	***	-27.1%	***	-30.2%	***	-29.0%	***

[8] In untabulated tests, we also sort into three groups (“Low”, “Medium” and “Large”) and decile portfolios of firm-level characteristics. The results are consistent across these different approaches. Results for mean forecast accuracy are essentially similar.

Table 2.5 Relation between model-based forecast accuracy (EPS_{t+1}) and firm characteristics, subperiods and subsamples

Note: This table reports the median forecast accuracy of our extended and the existing standard RI model for one-year (FY1), two-year (FY2) and zhree3-year (FY3) ahead annual earnings. Accuracy is the absolute price-scaled forecast error and reported as time-series averages of the median forecast accuracy. Differences in forecast accuracy are illustrated as percentage relative to the accuracy of the standard model. In this analysis forecasts across different information sets, e.g. $q = 0, 1, 2, 3$, are pooled for all firms and evaluated against several firm-characteristics and time components. The time-series t-statistics are estimated trough Newey-West error. The significance of the Newey-West time-series t-statistics are given as *** for the 1%-level, ** for the 5%-level and * for the 10%-level. In Panel A, we provide the results of the absolute price-scaled forecast error (ACCURACY) for all firms in our sample regarding individual firm-characteristics. Size is the market equity at the estimation date; age is the number of months a firm occurs in CRSP; analyst coverage is the number of financial analysts within the I/B/E/S consensus estimate; analyst dispersion is the standard deviation across analyst forecasts; earnings volatility is the standard deviation of earnings divided by the standard deviation of cash flow from operations over the last five years. We form portfolios based on these variables for each year. The reported values are time-series averages of annual median forecast accuracy. In Panel B, we evaluate the forecast performance across for subperiods of our sample. In Panel C, we examine whether economic cycle can impact the quality of mechanical earnings forecasts. The NBER index equals one in years around turning points and zero in normal periods. The CFNAI index is divided into three categories. A value above 0.7 indicates and economic peak and a value below -0.7 indicates an economic downturn. The data of U.S. business cycles is obtained from The National Bureau of Economic Research (NBER) at http://www.nber.org/cycles/cyclesmain.html. The data of the Chicago Fed National Activity Index (CFNAI) is collected from the Federal Reserve Bank of Chicago at https://www.chicagofed.org/research/data/cfnai/current-data.

We find that for smaller, younger and less covered firms the prediction of future annual earnings is less precise. That is, the PAFE of our extended model is 1.21% for large firms, while 7.25% for very small firms. In general, smaller firms have higher earnings volatilities that strongly reduces the predictability of future earnings. Likewise, firms that are less followed by analysts or dispersion across analysts is larger, generally, have more complex business models and, thus, producing reliable earnings expectations is more intricate.

In addition, this analysis shows that extended model forecasts persistently outperform forecasts from the standard model. Specifically, the PAFE is appr. 20% smaller relatively to the standard model. The differences in PAFEs are relatively larger for small and younger firms with less professional accounting departments, i.e., where earnings management is less established and, thus, the volatility of earnings is essentially higher. In detail, the PAFE reduces by appr. 30% using quarterly earnings data for firms with high earnings volatilities, i.e., the median PAFE drops from 5.80% to 3.88%. Hence, our extension improves forecasts when it matters the most. For instance, firms that have either complex business models or less smoothed earnings results and companies covered with controversial earnings expectation from research services, i.e., higher dispersion across equity analysts.

In Panel B and C of Table 2.5, we divide our sample according to different subperiods and economic conditions under which forecasts are performed and compute the pooled forecast accuracy for those groups. In general, model forecasts are more accurate for more recent years and under normal economic conditions. Both analyses show that our extension yields better forecasts across different years and, particularly, through different economic conditions, using the NBER or CFNAI classification.[9] In detail, the advantage in forecast quality of our extended model is somewhat more pronounced for more recent years and periods of economic recessions. That is, the PAFE reduces from 2.82% to 1.92% for the recent subperiod from 2007 to 2014, which is particularly surprising as this period covers the effects from the financial crisis. Hence, our augmented framework appears to be most valuable in periods, where earnings expectations are more heterogeneous.

2.4.3 Forecast errors – Extended model versus equity analysts

As the next part of our forecast performance analysis, we compare our newly-developed extended forecast model against professional analysts. Thus, we evaluate whether model-based forecasts can beat or compete with those from analysts. To compare model's earnings forecasts with those of analysts, we take following steps to compare both in-depth. First, it is important to compare forecasts along the same forecast horizon. Hou et al. (2012) and Li and Mohanram (2014) compute model forecasts that solely utilize annual statements and compare those with recent analyst forecasts available in June. This grants analysts a substantial information advantage over models. In fact, 95% of our firms already released quarterly earnings results in June.[10] Previously reported differences in forecast quality are therefore strongly biased due to the richer information incorporated in analysts' earnings forecasts. For example, Hou et al. (2012) reports that absolute forecast errors (PAFE) of their level-earnings models are three-times larger than those of equity analysts by the end of June. Incorporating quarterly earnings results into our model strongly reduce the information lead of analysts. Thus, our augmented framework is

[9] The data of U.S. business cycles is obtained from The National Bureau of Economic Research (NBER) at http://www.nber.org/cycles/cyclesmain.html. The data of the Chicago Fed National Activity Index (CFNAI) is collected from the Federal Reserve Bank of Chicago at https://www.chicagofed.org/research/data/cfnai/current-data.

[10] For example, firms with FYE in December released first quarter results and firms with FYE in June released already third quarter results.

crucial for a consistent comparison of model and analyst forecasts.[11] Second, we utilize a analogous earnings definition model forecasts that analysts generally employ, i.e., "Street"-EPS that exclude special items. Third, it is important to reduce the sample to firms where both model and analyst's forecasts are available, e.g., Easton and Monahan (2016). We note that, while we generate model forecasts for all firms, requiring analyst forecasts reduces the sample by appr. 40%. Thus, this analysis captures only a fraction of publicly traded companies that is covered by model forecasts.

Panel A of Table 2.6 compares absolute price-scaled forecast errors (PAFE) of our extended model and analysts. Most importantly, we find that model-based forecasts compete with analyst forecast at most forecast horizons. Specifically, long-term forecasts from our extended model are significantly more accurate having median PAFEs of 3.18% to 3.47% for the three-years ahead forecasts compared to analysts' PAFEs of 3.73% to 4.65%. Similarly, PAFEs of the two-years ahead forecasts are of 2.62% to 2.93% compared to analysts' PAFEs of 2.56% to 3.43%. Surprisingly, analysts' forecasts can beat model-based forecasts only at short forecast horizons, i.e., six months or less (after the release of the second quarter results). The PAFE of analyst forecasts is significantly smaller by 0.30% when firms already disclosed their second quarter results. In general, we find that previously reported differences in forecast accuracy diminish once we compare models and financial analysts at the same forecast horizons. We also note that analysts improve their forecasts stronger during the year. This is presumably because financial analysts may benefit from additional information, such as direct communication with the management (or example, conference calls), macroeconomic news, technological transition or even earnings guidance from the company itself (e.g., Brown, Call, Clement and Sharp, 2015). Our results are surprising in terms that model forecasts are strongly competitive to the performance of professional analysts. It appears that analysts do not take into account the financial performance of all firms in the cross-section. Hence, our results indicate that our forecast model provides complementary

[11] Alternatively, researchers could indeed trim the comparison of both earnings forecasts, i.e., from mechanical forecast models and financial analysts, to forecast horizons at the very beginning of their financial year. In this case both forecasts would be formed only on previous year financial results. However, a method of solely annual forecasts adjustments and comparisons appears to be very unsuitable for practical applications.

information to analysts. We believe that analysts are too focused on a specific stock and, thus, do not fully anticipate the overall economic conditions.

Table 2.6 Comparison of model's forecast accuracy against professional equity analysts

Panel A: Forecast accuracy of extended model-based vs analysts' earnings forecasts

		Extended RI				**Analysts**				**Difference**				**N**
	After	**Mean**		**Median**		**Mean**		**Median**		**Mean**		**Median**		
$EPS_{t+1,q}$	$Q3_{t+1}$	2.63%	***	1.17%	***	2.47%	***	0.75%	***	-0.16%	*	-0.42%	***	94,540
	$Q2_{t+1}$	3.53%	***	1.58%	***	3.27%	***	1.28%	***	-0.26%	**	-0.30%	***	93,204
	$Q1_{t+1}$	4.28%	***	1.86%	***	4.16%	***	1.82%	***	-0.12%		-0.03%		92,240
	A_t	5.18%	***	2.24%	***	4.84%	***	2.20%	***	-0.34%	**	-0.04%		90,636
$EPS_{t+2,q}$	$Q3_{t+1}$	5.30%	***	2.62%	***	4.80%	***	2.56%	***	-0.50%	***	-0.06%		78,599
	$Q2_{t+1}$	5.47%	***	2.82%	***	5.21%	***	2.94%	***	-0.26%	*	0.12%	*	76,501
	$Q1_{t+1}$	5.58%	***	2.79%	***	5.62%	***	3.24%	***	0.04%		0.45%	***	73,952
	A_t	5.71%	***	2.93%	***	5.89%	***	3.43%	***	0.18%		0.49%	***	69,686
$EPS_{t+3,q}$	$Q3_{t+1}$	5.85%	***	3.18%	***	5.90%	***	3.73%	***	0.05%		0.55%	***	59,889
	$Q2_{t+1}$	5.96%	***	3.27%	***	6.17%	***	4.12%	***	0.21%		0.85%	***	58,076
	$Q1_{t+1}$	5.89%	***	3.34%	***	6.58%	***	4.44%	***	0.69%	***	1.11%	***	57,396
	A_t	6.04%	***	3.47%	***	6.87%	***	4.65%	***	0.82%	***	1.18%	***	56,004

Panel B: Forecast bias of extended model-based vs analysts' earnings forecasts

		Extended RI				**Analysts**				**Difference**				**N**
	After	**Mean**		**Median**		**Mean**		**Median**		**Mean**		**Median**		
$EPS_{t+1,q}$	$Q3_{t+1}$	-0.08%		0.16%	***	-1.36%	***	-0.20%	***	-1.29%	***	-0.36%	***	94,540
	$Q2_{t+1}$	-0.02%		0.13%	**	-1.99%	***	-0.55%	***	-1.96%	***	-0.68%	***	93,204
	$Q1_{t+1}$	-0.23%		0.08%		-2.63%	***	-0.98%	***	-2.40%	***	-1.06%	***	92,240
	A_t	-0.65%	***	-0.09%		-3.24%	***	-1.34%	***	-2.58%	***	-1.25%	***	90,636
$EPS_{t+2,q}$	$Q3_{t+1}$	0.24%		-0.08%		-3.10%	***	-1.70%	***	-3.34%	***	-1.62%	***	78,599
	$Q2_{t+1}$	-0.12%		-0.32%	**	-3.42%	***	-2.10%	***	-3.29%	***	-1.78%	***	76,501
	$Q1_{t+1}$	-0.49%	**	-0.43%	***	-3.62%	***	-2.38%	***	-3.13%	***	-1.94%	***	73,952
	A_t	-1.33%	***	-0.70%	***	-3.95%	***	-2.59%	***	-2.62%	***	-1.89%	***	69,686
$EPS_{t+3,q}$	$Q3_{t+1}$	-0.13%		-0.53%	***	-4.12%	***	-2.93%	***	-3.99%	***	-2.40%	***	59,889
	$Q2_{t+1}$	-0.59%	**	-0.87%	***	-4.33%	***	-3.27%	***	-3.74%	***	-2.41%	***	58,076
	$Q1_{t+1}$	-1.24%	***	-1.10%	***	-4.57%	***	-3.60%	***	-3.32%	***	-2.50%	***	57,396
	A_t	-2.26%	***	-1.54%	***	-4.82%	***	-3.84%	***	-2.55%	***	-2.30%	***	56,004

Note: This table reports the forecasts performance of our extended and RI model and financial analysts' consensus forecasts available in I/B/E/S in terms of the forecast bias and accuracy for one-year (FY1), two-year (FY2) and three-year (FY3) ahead annual earnings. Bias is the price-scaled forecast error and reported as time-series averages of the median and mean forecast bias. Accuracy is the absolute price-scaled forecast error and reported as time-series averages of the median and mean forecast accuracy. The evaluation is performed with respect to the number of quarterly earnings announcements (q) that are public at the estimation date. The time-series t-statistics are estimated trough Newey-West error corrections. The significance of the Newey-West time-series t-statistics are given as *** for the 1%-level, ** for the 5%-level and * for the 10%-level. In Panel A, we provide the results of the absolute price-scaled forecast error (ACCURACY) for firms with coverage by equity analysts in our sample. Whereas in Panel B, we provide the results of the price-scaled forecast error (BIAS) for firms with coverage by equity analysts in our sample.

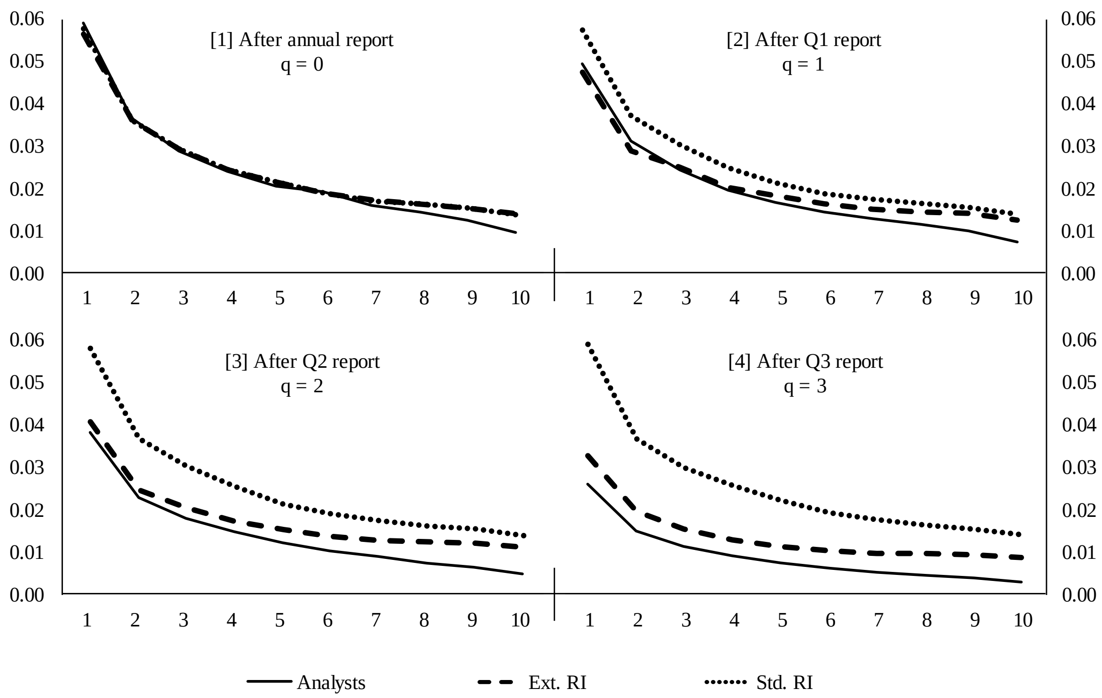
In further untabulated tests, we similarly evaluate the forecasts that are available by the end of June to apply the concept of the previous studies, e.g., Hou et al. (2012), Li and Mohanram (2014) and Hess et al. (2018). Our results for analysts' earnings forecasts and the standard models are virtually identical to the extant literature. In contrast, our extended model forecasts are much more precise in June, as we incorporate available quarterly earnings results. Hence, a comparison of analysts' and extended model forecasts yields the same findings.

Panel B of Table 2.6 reports the price-scaled forecast errors (PFE), i.e., forecast bias. First, we find a substantially reduced forecast bias of our extended model in contrast to analyst forecasts. Model's one-year ahead forecasts (FY1) are not significantly biased. In contrast, bias in analysts' earnings forecasts is large and highly significant. For instance, the PFE of analysts' forecasts ranges from -0.20% to -1.34% in median and -1.36% to -3.24% in mean for the one-year ahead forecasts, while model-based PFEs are much smaller and even insignificant, respectively. Model's forecast bias is also much smaller for longer forecast horizons. The negative bias in analyst forecasts implies an overly optimistic expectation of future corporate earnings. In line with previous studies, we find that optimism strongly increases with longer forecast horizons (e.g. Abarbanell, 1991; Bradshaw et al., 2012; O'Brien, 1988; Richardson et al., 2004). The bias in long-term forecasts from financial analysts is, in general, three times larger than the bias of our forecast model, e.g., PFE of analysts ranges from -2.93% to -3.84% for the three-years ahead forecasts (-0.53% to -1.54% PFE of model forecasts).

In general, our model extension allows models to compete with forecasts from professional equity analysts in terms of forecast accuracy (PAFE) and provides less biased forecasts (PFE). This is very important as previous studies, e.g., Hou et al. (2012) and Li and Mohanram (2014), that use only annual financials and neglect information lead in analyst forecasts, report that model's forecasts are substantially less precise. Overall, our results strengthen the claim that earnings forecasts models provide an appealing approach to provide complementary information to those of analysts and to automate the forecasting procedure.

In an additional step of our performance analysis, we analyze the surprisingly small differences in forecast accuracy across model-based and analysts' earnings forecasts. In detail, we assess whether these differences are robust across firm characteristics, for example, market capitalization.

Figure 2.1 Differences in forecast accuracy across firm size portfolios



Note: Figure 2.1 illustrates earnings forecast accuracy of our extended RI model and financial analysts for companies' annual earnings in fiscal year t relative to its market capitalization. In this analysis, we sort all earnings forecasts into decile portfolios of companies' market equity for each year and according to the number of quarterly earnings results ($q = 0,1,2,3$) that are available at a time. Then, we compute time-series averages of forecast accuracy and bias for each portfolio group. The reported figures show the median forecast accuracy for covered firms, where model-based and analysts' forecasts are available. The forecast accuracy is a strong intersection within firm-size and, hence, predictability of annual earnings results. Using other proxies for firm-size, i.e., trading voluminal or total assets, provide similar results. The differences between model-based and analysts' earnings forecasts are only significant for very large firms.

Figure 2.1 illustrates the forecast accuracy (PAFE) across firm-size, i.e., market capitalization at the fiscal year-end.[12] Firm size is generally a good proxy for earnings volatility and most importantly for the information environment and visibility of a company. First, we observe that PAFE is a strongly related to firm-size for both analysts' and model-based forecasts. This supports the results from previous analyses that earnings of smaller firms are generally less predictable and persistent. Finding this pattern for both model and analyst forecasts is very important, because it suggests that large firms are not covered by more skilled analysts (or small firms are assigned to less skilled analysts). In

[12] To conserve space, we report only results for market equity as a proxy for firm-size, however, measures as trading volumes and total assets show similar results.

fact, earnings of large firms are easier to predict by nature and, thus, that analysts have less issues to estimate their future results. Related to this point, we see that the performance gap depends on firm size: analyst forecasts are only more precise for larger firms, where they have an information advantage. The differences in forecast accuracy are only significant for firms falling into the largest size decile particularly when most of the quarterly results are already announced. One potential reason is that larger firms seem to issue substantially more forward-looking information within their quarterly disclosures or other financial releases (see, e.g., Anilowski, Feng and Skinner, 2007). More precise and frequent earnings guidance and larger media coverage helps analysts to estimate more uniform earnings expectations. In additional tests considering the full sample (untabulated), we find that analysts cover mostly the very large firms. Hence, analysts cover mainly firms which are easier to forecast. For less visible firms, analysts' forecast quality is worse than those of extended models, because small firms may provide less detailed information within their quarterly disclosures, i.e., such as information about cash flows, accruals and note disclosures. This would explain that financial analysts barely cover firms with less comprehensible business models, where coverage is not efficient and cross-sectional forecast methods apparently provide a better forecast quality, i.e., accuracy and bias. This is very important, because cross-sectional models are a cost-efficient tool to forecast earnings for firms that are not covered by analysts.

2.4.4 Implications for implied cost of capital models

This section addresses whether earnings forecasts from our extended model produce implied cost of capital (ICC) estimates that are better predictors for future returns. After we calculate model forecasts and observe analyst forecasts, we use these forecasts to solve for a company's ICC. We follow the previous literature, e.g., Hou et al. (2012), Li and Mohanram (2014), and calculate a robust composite ICC estimate by taking the average of ICC estimates from five individual ICC models.[13]

[13] The individual pricing equations are described in detail within Appendix C. The tenor is unaltered when using each individual ICC model. The composite ICC estimate is the mean of the five individual ICC estimates (GLS, CT, OJ, MPEG, GG). In detail, we focus on the abnormal earnings growth model from Ohlson and Juettner-Nauroth (2005) (OJ), the modified version from Easton (2004) (MPEG), the residual income valuation models from Gebhardt et al. (2001) (GLS) and Claus and Thomas (2001) (CT) and the simple expected return model from Gordon and Gordon (1997) (GG).

Table 2.7 Descriptive statistics of composite ICC estimates, 1982-2014

Panel A: COVERED FIRMS - Summary Statistics

	Avg. N	Mean	1%	25%	Median	75%	99%	Std
$ICC^{Ext.\ RI}$	103,000	0.0852	0.0134	0.0604	0.0796	0.1027	0.2213	0.0383
$ICC^{St.\ RI}$	104,643	0.0882	0.0193	0.0612	0.0817	0.1060	0.2419	0.0412
$ICC^{Analysts}$	103,338	0.0991	0.0225	0.0776	0.0944	0.1150	0.2230	0.0361

Panel B: COVERED FIRMS - Correlation between model-based and analyst-based ICCs

	$ICC^{Ext.\ RI}$	$ICC^{St.\ RI}$	$ICC^{Analysts}$	*Ret*
$ICC^{Ext.\ RI}$	-	0.83	0.56	0.07
$ICC^{St.\ RI}$	0.81	-	0.55	0.06
$ICC^{Analysts}$	0.52	0.51	-	0.02
Ret	0.04	0.03	0.01	-

Panel C: NON-COVERED FIRMS - Summary Statistics

	Avg. N	Mean	1%	25%	Median	75%	99%	Std
$ICC^{Ext.\ RI}$	49,080	0.1059	0.0045	0.0614	0.0888	0.1244	0.4218	0.0766
$ICC^{St.\ RI}$	50,935	0.1117	0.0122	0.0654	0.0934	0.1319	0.4268	0.0775
$ICC^{Analysts}$	-	-	-	-	-	-	-	-

Panel D: NON-COVERED FIRMS - Correlation between model-based and analyst-based ICCs

	$ICC^{Ext.\ RI}$	$ICC^{St.\ RI}$	$ICC^{Analysts}$	*Ret*
$ICC^{Ext.\ RI}$	-	0.73	-	0.08
$ICC^{St.\ RI}$	0.72	-	-	0.06
$ICC^{Analysts}$	-	-	-	-
Ret	0.04	0.03	-	-

Note: This table reports the summary statistics of composite ICC estimates from standard and extended model-based and analysts' earnings forecasts. The composite ICC estimate is the mean of the five individual ICC estimates (GLS, CT, OJ, MPEG, GG). In detail, we focus on the abnormal earnings growth model from Ohlson and Juettner-Nauroth (2005) (*OJ*), the modified version from Easton (2004) (*MPEG*), the residual income valuation models from Gebhardt et al. (2001) (*GLS*) and Claus and Thomas (2001) (*CT*) and the simple expected return model from Gordon and Gordon (1997) (*GG*). A detailed description of the individual ICC model equation and terminal value assumptions is provided in Appendix C. Panel A and Panel B presents the summary statistics and correlations of the subsample of covered firms. Panel C and Panel D presents the summary statistics and correlations of the subsample of non-covered firms. Realized returns is the future twelve-months holding return subsequently to the estimation date on a firm-level. The reported values are time-series averages. The correlation matrices in Panel B and D reports the Pearson-correlations within the lower corner and Spearman-correlations in the upper corner.

Table 2.7 presents the characteristics of our model-based and analyst-based ICC estimates. Panel A shows the univariate statistics of ICCs and Panel B displays the correlations between ICCs for the subsample of analyst covered firms. In Panel A, we observe that the two model-based ICCs differ only slightly in mean, median and selected percentiles. Nevertheless, our model extension ($ICC^{Ext.RI}$) reduces the standard deviation of ICC estimates to 0.0383. Analyst-based ICCs ($ICC^{Analysts}$) are larger in mean, median and most percentiles as compared to both model-based ICCs. Hence, optimism in analyst forecasts leads to generally larger ICC estimates (e.g., Easton and Sommers, 2007; Guay et al., 2011; Hou et al., 2012). Panel B reports the Pearson and Spearman correlation of

ICC estimates and subsequent twelve-months returns. However, ICCs based on the extended model are interestingly higher correlated with future realized returns than analyst-based ICCs. For example, the Spearman correlation between $ICC^{Ext.RI}$ and Ret is 0.07, but only 0.02 between $ICC^{Analysts}$ and Ret. Similarly, the correlations of standard model-based ICCs ($ICC^{Std.RI}$) and future returns are slightly smaller, i.e., 0.06 for Spearman's and 0.03 for Pearson's correlation.

Panel C and Panel D of Table 2.7 reports the summary statistics for firms not covered by analysts. If market participants want to include firms from this large segment into their investment process, they require model forecasts to compute earnings expectations. Comparing the sample of covered and non-covered firms, we find a larger variation of ICCs for the latter. In fact, this pattern may result from a decline in forecast accuracy for this sample and, thus, a higher uncertainty for these firms. Nevertheless, ICCs for non-covered firms are larger in mean and median than ICCs from the covered market segment. Moreover, we find positive correlations of ICCs and future realized returns. Overall, the ICCs are larger for firms, where the risk premium is higher, i.e., smaller firms with higher earnings volatility, less public information and not covered by financial analysts. The larger correlations of model-based ICCs to future returns indicate that our extended model produces forecasts that better infers future stock returns.

2.4.5 Evaluating ICCs as accounting-based expected return proxies

In the next step, we analyze whether our ICC estimates allow to distinguish between out- and underperforming stocks in subsequent periods. Therefore, we rank firms into deciles according to their ICC estimates at the end of each month. Afterwards, we construct zero investment strategies buying stocks in the upper decile and selling stocks in the bottom decile and calculate the return of equally weighted portfolios for the subsequent twelve months holding period.[14] Significant return spreads indicate whether investors can use our augmented framework to create excess returns with this investment strategy. This analysis is predominantly used in the extant literature to determine the reliability of ICC estimates (see, e.g., Hou et al., 2012; Li and Mohanram, 2014).

[14] In line with previous studies, we neglect transaction costs for simplicity in this test. However, since model-based ICCs are mainly adjusted after quarterly earnings release and, thus, mainly four times a year, the portfolio turnover for ICCs based on mechanical forecast models is low.

Table 2.8 Portfolio strategies based on ICC estimates

Panel A: COVERED FIRMS - Twelve months returns of composite ICC estimates portfolios

	$ICC^{Ext.\ RI}$					$ICC^{St.\ RI}$					$ICC^{Analysts}$				
PF	**ICC**	**Ret**	**Std**	**Shp**	**t-stat**	**ICC**	**Ret**	**Std**	**Shp**	**t-stat**	**ICC**	**Ret**	**Std**	**Shp**	**t-stat**
1	3.50	12.61	30.84	0.40	4.34	3.69	12.93	30.13	0.42	4.58	4.77	11.81	26.85	0.43	4.67
2	5.24	12.03	22.08	0.53	5.86	5.30	12.22	22.19	0.54	5.95	6.94	12.33	20.55	0.58	6.42
3	6.16	12.80	20.27	0.61	6.82	6.25	13.27	20.15	0.64	7.08	7.83	13.70	19.35	0.69	7.63
4	6.94	13.67	19.80	0.67	7.45	7.07	13.86	19.69	0.69	7.57	8.52	14.35	18.53	0.76	8.37
5	7.68	14.34	19.65	0.71	7.82	7.85	14.57	19.64	0.72	7.98	9.15	15.09	18.40	0.80	8.82
6	8.43	15.03	19.82	0.74	8.16	8.65	15.13	19.79	0.75	8.19	9.80	16.00	19.89	0.79	8.66
7	9.27	15.73	20.11	0.76	8.36	9.54	15.69	20.07	0.76	8.32	10.53	16.88	21.24	0.78	8.56
8	10.29	16.34	21.01	0.76	8.31	10.61	16.27	20.91	0.76	8.31	11.45	17.04	22.56	0.74	8.09
9	11.82	17.76	23.35	0.75	8.08	12.22	17.01	22.78	0.73	7.94	12.84	16.60	25.31	0.64	7.01
10	16.37	18.25	31.09	0.58	6.28	17.23	17.62	32.29	0.53	5.83	16.88	14.76	33.22	0.43	4.74
10-1	**12.87**	**5.64**	**28.43**	**0.19**	**2.10**	**13.54**	**4.68**	**27.48**	**0.16**	**1.80**	**12.11**	**2.95**	**26.45**	**0.10**	**1.18**

Panel B: NON-COVERED FIRMS - Twelve months returns of composite ICC estimates portfolios

	$ICC^{Ext.\ RI}$					$ICC^{St.\ RI}$				
PF	**ICC**	**Ret**	**Std**	**Shp**	**t-stat**	**ICC**	**Ret**	**Std**	**Shp**	**t-stat**
1	2.48	10.39	30.63	0.33	3.62	2.98	12.56	31.99	0.38	4.20
2	4.94	10.78	24.99	0.42	4.62	5.33	12.13	26.68	0.44	4.84
3	6.34	13.22	23.51	0.55	6.01	6.71	12.68	22.82	0.54	5.89
4	7.46	13.83	21.23	0.64	6.92	7.83	13.76	21.24	0.63	6.90
5	8.47	14.87	21.40	0.68	7.39	8.87	15.53	22.37	0.68	7.38
6	9.50	17.14	22.37	0.75	8.12	9.95	16.83	22.60	0.73	7.86
7	10.72	18.02	23.20	0.76	8.21	11.26	16.83	22.67	0.73	7.85
8	12.47	19.21	24.92	0.76	8.18	13.13	18.27	23.82	0.75	8.14
9	15.85	19.90	28.29	0.69	7.44	16.64	19.16	28.01	0.67	7.20
10	27.76	22.14	34.26	0.64	6.85	28.40	21.47	34.17	0.62	6.67
10-1	**25.28**	**11.75**	**19.20**	**0.59**	**7.09**	**25.42**	**8.91**	**19.40**	**0.44**	**5.24**

Note: This table reports the results from portfolio tests of our composite ICC estimates. Panel A presents the results for the market segment of firms with analyst coverage, while Panel B shows the results for other firms that are not covered by professional equity analysts. The composite ICC estimate is the mean of the five individual ICC estimates (GLS, CT, OJ, MPEG, GG). In detail, we focus on the abnormal earnings growth model from Ohlson and Juettner-Nauroth (2005) (*OJ*), the modified version from Easton (2004) (*MPEG*), the residual income valuation models from Gebhardt et al. (2001) (*GLS*) and Claus and Thomas (2001) (*CT*) and the simple expected return model from Gordon and Gordon (1997) (*GG*). A detailed description of the individual ICC model equation and terminal value assumptions is provided in Appendix C. In this analysis, we sort our firms into decile portfolios based on the ICC estimates from model-based and analysts' earnings forecasts at each month in our sample range. We then compute twelve-months holding returns for each portfolio and return spreads of an investment strategy that buys the upper portfolio and sells the bottom portfolio (10-1). We tabulate the average ICC estimate *ICC*, average realized return *Ret*, standard deviation of realized returns *Std*, the Sharpe-Ratio *Shp* and Newey-West corrected time-series t-statistics *t-stat* for the time-series of investments within our sample range. The Sharpe-Ratio quantifies

Table 2.8 Portfolio strategies based on ICC estimates

the trade-off between realized return and volatility for each portfolio strategy, i.e., the portfolio excess return per unit of the portfolios' standard deviation. The t-statistic indicates whether investment returns are significant from zero across the overall time-series.

In Table 2.8, we tabulate the results from these portfolio tests. We find that our extended model yields larger return spreads for annual holding periods. For covered firms, we obtain an average realized spread of 5.64% p.a. with our model extension, whereas the standard model is barely able to yield significant 4.68% p.a. Hence, our approach provides an additional return of 0.96% p.a. due to more accurate earnings forecasts and, thus, more precise ICC estimates. This finding is supported by higher t-statistics and larger Sharpe ratios (see, e.g., Sharpe, 1994). In addition, we find strong advantages for the subsample of non-covered firms. We generate an average return spread of 11.75% p.a., when we use our extended model-based ICC estimates ($ICC^{Ext.RI}$). Hence, our approach significantly outperforms existing alternatives that yields an average return of 8.91% p.a. Hence, the better quality and timeliness of earnings expectation is particularly valuable for firms that are not covered by financial analysts. In general, the higher returns for this market segment implies that smaller firms are riskier and investors expectation are less uniform, since analysts' recommendations and reports as important information catalysts are not available.

Likewise Hou et al. (2012), we show that an investment strategy based on ICCs from analyst forecasts does not yield significant portfolio spreads. As we find in the previous sections, analyst forecasts are strongly biased and grossly inaccurate in longer forecast horizons. Hence, the reasons that analyst forecasts do not produce reliable ICC estimates is twofold. On the one hand, poor long-term forecasts and growth assumptions may impair the construction of terminal values within the ICC models. On the other hand, optimistic bias in analyst forecasts produce larger ICC estimates. Indeed, the latter should not be problematic in portfolio tests, if biases are consistent across all forecasts and firms. For example, if all forecasts are 1% to optimistic, this would not affect the ranking of firms into portfolios. However, forecast biases of analysts are not homogeneous across firms. In fact, previous research shows that forecast biases are predominantly driven by firm characteristics, such as firm-size, analysts following and earnings growths (see, e.g., Gu and Wu, 2003; Bradshaw et al., 2012). For example, analysts actively bias earnings forecasts to preserve management contacts, which is most important for larger companies.

Additionally, we conduct a second analysis, in which we test whether ICCs match future returns on a company-level, thus, whether the ICCs can predict individual stock returns. Therefore, we run monthly regressions of twelve-months returns of a stock on its corresponding ICCs:

$$r_{t,1,12} = \alpha + \beta \cdot ICC_t + \varepsilon_t \tag{2.17}$$

The interpretation of coefficient estimates is straight forward. If ICCs can perfectly predict future stock returns, coefficient β equals one. Likewise, the intercept α would be zero. In contrast, a coefficient β of zero indicates that ICCs cannot predict future returns at all.

Table 2.9 provides the results from monthly regressions of realized future return on firm-specific ICC estimates. For the subsample of covered firms, we notice that our extended model-based ICCs ($ICC^{Ext.RI}$) provide strongly superior correlations to future realized returns. That is, the coefficient is significantly large, i.e., β is 0.4420, and the intercept is the smallest across all alternatives, i.e., α is 0.100. Moreover, for analysts-based ICCs and standard model-based ICCs, both T- and F-Tests show that the relation between ICCs and future realized returns is weak. For the subsample of non-covered firms, we find that both model-based ICCs are correlated with future realized returns. However, ICCs based on our extended forecast approach persistently outperform ICCs that are based on the pre-existing standard models. In regressions using $ICC^{Ext.RI}$, the coefficient β is significantly closer to one according to both T- and F-Tests. The intercept a is closer to zero and the overall regression yields a larger adjusted R^2. Hence, studies on ICCs as accounting-based return proxies should consider our model extension to find more reliable relations between ICCs and future returns.

In general, we find that our model extension yields significant correlation of ICCs and future returns. We demonstrate that standard model forecasts deliver persistently inferior ICCs, i.e. smaller correlations to future returns and smaller portfolio returns from ICC-based sorting. Additionally, we find no evidence that analyst-based ICCs are connected to future returns. Optimism in annual earnings forecasts and rigorously imprecise long-term growth forecasts of financial analysts deteriorate the relation between ICCs and future realized returns (e.g., Botosan and Plumlee, 2005; Easton and Monahan, 2005; Easton and Sommers, 2007; Guay et al., 2011).

Table 2.9 Firm-level regression of returns and ICCs

Panel A: COVERED FIRMS - Regressions of annual returns on implied cost of capital (ICC)

	$ICC^{Ext.\ RI}$			$ICC^{St.\ RI}$			$ICC^{Analysts}$		
	a	**b**	**Adj. R²**	**a**	**b**	**Adj. R²**	**a**	**b**	**Adj. R²**
	0.100	**0.4420**	1.63%	*0.113*	**0.2697**	1.48%	*0.124*	**0.1246**	1.43%
	[4.47]	*[2.73]*		*[5.35]*	*[1.88]*		*[6.31]*	*[0.78]*	
		F-stat	*p-val*		*F-stat*	*p-val*		*F-stat*	*p-val*
WALD-Test	b = 0	*7.43*	*0.006*	b = 0	*3.53*	*0.060*	b = 0	*0.60*	*0.438*
	b = 1	*11.85*	*0.001*	b = 1	*25.85*	*0.000*	b = 1	*29.68*	*0.000*

Panel B: NON-COVERED FIRMS - Regressions of annual returns on implied cost of capital (ICC)

	$ICC^{Ext.\ RI}$			$ICC^{St.\ RI}$		
	a	**b**	**Adj. R²**	**a**	**b**	**Adj. R²**
	0.098	**0.5155**	0.63%	*0.108*	**0.3913**	0.58%
	[4.65]	*[7.68]*		*[5.06]*	*[6.12]*	
		F-stat	*p-val*		*F-stat*	*p-val*
WALD-Test	b = 0	*58.96*	*0.000*	b = 0	*37.48*	*0.000*
	b = 1	*52.10*	*0.000*	b = 1	*90.69*	*0.000*

Note: This table reports the results from firm-level tests of our composite ICC estimates. Panel A presents the results for the market segment of firms with analyst coverage, while Panel B shows the results for other firms that are not covered by professional equity analysts. The composite ICC estimate is the mean of the five individual ICC estimates (GLS, CT, OJ, MPEG, GG). In detail, we focus on the abnormal earnings growth model from Ohlson and Juettner-Nauroth (2005) (*OJ*), the modified version from Easton (2004) (*MPEG*), the residual income valuation models from Gebhardt et al. (2001) (*GLS*) and Claus and Thomas (2001) (*CT*) and the simple expected return model from Gordon and Gordon (1997) (*GG*). A detailed description of the individual ICC model equation and terminal value assumptions is provided in Appendix C. In this analysis, we run monthly regressions of a firms estimated implied cost of capital (ICC) and its compute subsequent twelve-months stock return:

$$r_{t,1,12} = \alpha + \beta \cdot ICC_t + \varepsilon_t$$

If ICCs can perfectly predict future stock returns, coefficient β equals one. Likewise, the intercept α is zero. In contrast, a coefficient β of zero indicates that ICCs cannot predict future returns at all. The displayed values are time-series average with Newey-West corrected t-statistics. In addition, we performed Wald-Tests to test whether the coefficient β is zero (no relation between future return and ICC estimates) or whether the coefficient β is one (perfect relation between future return and ICC estimates).

2.5 Conclusion

Earnings forecasts are a key input of asset pricing models and a primary indicator of a companies' future profitability. Therefore, financial economists have exerted increasing effort to quantify determinants of future earnings. The most promising approach stems from Hou et al. (2012) that develops a cross-sectional model to forecast annual earnings for a very broad set of firms using only current financial statements data. Recent related studies on model forecasts find that model forecasts are less precise compared to analyst forecasts (see, e.g., Li and Mohanram, 2014; Ashton and Wang, 2013; Chang et al., 2014). However, we assume that these findings are due to the lack of additional intra-year information within existing forecast models whereas analysts benefit from a large information lead from recent quarterly earnings results.

We level the playing field by extending the approach of Hou et al. (2012) to incorporate essential quarterly earnings results into model forecasts of annual earnings. Using our augmented approach, we strongly improve forecast accuracy and likewise reduces the forecast bias in both short- and long-term horizons. In addition, we close the performance gap between model and analyst forecasts. That is, we find that once quarterly earnings results are incorporated into forecast models, differences in forecast accuracy against financial analysts diminish. Financial analysts should benefit from a broad spectrum of other public information and, therefore, should provide more accurate forecasts for a wide range of firms. But, we find that equity analysts can only provide more accurate earnings forecasts for very large firms that appear to be more visible.

In addition, our findings have implications on the ongoing discussion whether ICCs serve as a valid proxy for expected stock returns (see, e.g., Botosan et al., 2011; Penman, 2015). That is, we show that ICCs estimated from model forecasts are a reliable proxy, whereas ICCs estimated from analyst forecasts are not. In addition, we show that correlations between ICCs and future realized returns are particularly strong for firms without analyst coverage. This is very important, because asset managers and empirical studies rely on our forecast model for firms that are not covered by analysts. Furthermore, we examine the notion that more accurate earnings forecasts provide better ICC estimates (e.g., Easton and Monahan, 2005; Easton and Sommers, 2007; Gode and Mohanram, 2013; Larocque, 2013). In fact, we show that ceteris paribus (c.p.) higher forecast performance from our model extension translates into more reliable ICCs.

Our approach is very important in various settings in practice. Our extension to interim financial results allows to update financial conditions, i.e., earnings, sales or leverage, frequently and regularly during the year. The results in this study show that the process of forecasting a firm's financials on a high frequency can be fully automated. This is striking as model-based forecasting can potentially initiate further digitalization of a broad spectrum of financial services. This may serve investors and institutions with better expectations of future earnings and financials of a company.

Most notably, we illustrate the potential of our model forecasts for (1) sell-side analysts or brokers and (2) portfolio managers. First, analysts may benefit from model forecasts as a benchmark or cost-efficient alternative to predict earnings for small firms or even to increase the coverage of their research services. This is particularly important in presence of further regulatory challenges, i.e., the European MiFID II regulation, that may disrupt cost structures in the research services industry. Second, portfolio managers may consider model forecasts to estimate more reliable expected stock returns for their security selection as well as to substantially enlarge the investment universe to firms that are not covered by analysts' research (e.g., small or private firms, developing economies).

Similarly, banks and reinsurers can utilize the projections from forecast models to justify or adjust credit ratings based on updates from quarterly earnings reports. In addition, corporate managers can use our models for financial planning. Furthermore, researchers may apply our approach to other financial statement performance measures, such as predicting sales, gross profits, cash flows (see, e.g., Heinrichs, Hess, Homburg and Sievers, 2013) or accounting accruals to detect earnings management.

Chapter 3

The Quality of Bankruptcy Data and its Impact on the Evaluation of Prediction Models: Creating and Testing a German Database[15]

3.1 Introduction

For decades, academics and practitioners have been tasked with the prediction of corporate bankruptcies. While considerable efforts have been made to improve the methodologies used in bankruptcy prediction models (e.g., Altman, 1968; Ohlson, 1980; Shumway, 2001; Vassalou and Xing, 2004), previous studies have paid little attention to

[15] This Chapter 3 is based on the academic article "The Quality of Bankruptcy Data and its Impact on the Evaluation of Prediction Models: Creating and Testing a German Database" written by Martin Huettemann and Tobias Lorsbach, as of February 2019. We are grateful to Dieter Hess, Martin Meuter, and William Liu for their insightful discussions and suggestions. This paper has also greatly benefitted from comments made by seminar participants at the University of Cologne and an anonymous reviewer. Moreover, we acknowledge the help from the customer service teams at Bureau van Dijk, Creditreform, EQS Group (DGAP) and APA OTS to clarify questions regarding the data availability of bankruptcy information.

the quality of the underlying bankruptcy data. Quality bankruptcy data can be defined as complete and correct information regarding bankruptcy events and explicit bankruptcy dates. Accurate bankruptcy data is crucial for two main reasons. First, it allows to obtain unbiased parameter estimates for bankruptcy models, because incorrect data can affect the significance and size of the coefficients and, thus, the variable setup of models. Second, validation of bankruptcy prediction models strongly depends on the integrity of the bankruptcy data. In fact, inaccurate information could affect the evaluation of out-of-sample performance. Consequently, we investigate the impact of data quality on the evaluation of bankruptcy prediction models.

Studies commonly use commercial databases to collect bankruptcy information. In the U.S., the SDC Platinum Database, Moody's Default and Recovery Database, Capital IQ, and Fixed Investment Securities Database (FISD) are examples of such databases. Even though SDC Platinum Database and Moody's Default and Recovery Database contain some data on European bankruptcies, the data availability is relatively sparse for non-U.S. firms. SDC Platinum reports 250 recent bankruptcies that are outside of the U.S. and Moody's Default and Recovery Database lists 108 bankruptcy events in Germany since 1980.[16] Therefore, we focus on the most frequently used European databases: Compustat Global (e.g., Dahiya and Klapper, 2007; Tian and Yu, 2017) and Bureau van Dijk (BvD) (e.g., Altman, Iwanicz-Drozdowska, Laitinen, and Suva, 2017; Filipe, Grammatikos, and Michala, 2016; Lohmann and Ohliger, 2017). Few studies have examined the quality of these popular bankruptcy databases. For instance, BvD deletes bankruptcy information after five years of inactivity. Moreover, requesting data directly from Creditreform, BvD's provider of bankruptcy information for German firms, would not rectify this central limitation since Creditreform also deletes a firm's bankruptcy information after bankruptcy proceedings areterminated. Therefore, one aim of this study is to quantify the amount of erroneous bankruptcy information in the databases generally used in earlier bankruptcy studies: Bureau van Dijk and Compustat.

[16] This information is retrieved from SDC Platinum and Moody's Analytics customer service. See also: Moody's default & recovery database, retrieved from https://www.moodys.com/sites/products/productattachments/drd_brochure.pdf and https://financial.thomsonreuters.com/en/products/data-analytics/market-data/sdc-platinum-financial-securities.html.

Unlike commercial databases such as BvD and Compustat Global, the UCLA-LoPucki Bankruptcy Research Database (BRD) consists of data on U.S. bankruptcies retrieved directly from sources such as court files or Securities and Exchange Commission (SEC) filings.[17] However, there are no guidelines for producing a bankruptcy database that derives information from public sources. This study constructs these guidelines by describing a methodology to systematically collect accurate bankruptcy data from public sources and applying it to a specific stock market. We focus on one country because parameter estimates may differ across countries for two reasons. First, the definition of bankruptcy may vary by regulatory requirement and, second, administrative firm-level data and, thus, the definition of variables used for bankruptcy prediction, differ by country. For example, Altman et al. (2017) argue that differences in financial statements can be attributed to variances in fiscal systems across countries. Nevertheless, previous studies (see, e.g., Altman et al., 2017) must use the entire European market to obtain sufficient bankruptcy data. We choose Germany as a case country for two reasons. First, Germany lacks an academic bankruptcy database that contains data from public sources, similar to the U.S. data in the UCLA-LoPucki BRD. Second, Germany is one of the largest stock markets in Europe. It is noteworthy that while our methodology can be applied to other countries, it is critical that disclosure obligations and their public availability be checked when doing so; in the case of the United States and United Kingdom, this would mean referring to SEC filings and the Regulatory News Service (RNS) as the national news provider, respectively.

Next, we compare our bankruptcy database (hereafter, HL) with the most commonly used databases for German bankruptcies, Compustat Global and BvD. In particular, we analyze the completeness and accuracy of the bankruptcy event and date information. We then conduct a two-part empirical analysis of public German firms. In the first part, we compare the bankruptcy prediction models of Altman (1968), Ohlson (1980), Shumway (2001), Bharath and Shumway (2008), and Hess and Huettemann (2018). There is voluminous bankruptcy prediction research, but a majority of existing studies focus on U.S. corporations, while research that presents international evidence remains relatively sparse. For example, Altman et al. (2017) assess the performance of

[17] See UCLA-LoPucki Bankruptcy Research Database: http://lopucki.law.ucla.edu/ (accessed July 20, 2018).

Altman's (1983) Z-score for 31 European and three non-European countries. Tian and Yu (2017) investigate the significance of ratios for bankruptcy prediction in Japan and selected European countries, while Dahiya and Klapper (2007) compare key industrial nations. All these studies use either BvD or Compustat Global as commercial databases. Note that because we have a sufficient number of bankruptcy events, we can focus on a single European country, giving our study the advantages noted above. In the second part, we investigate how using the HL database, instead of BvD data, affects the results of bankruptcy prediction models. More specifically, we analyze the parameter estimates and out-of-sample performances when we use our bankruptcy data compared to BvD data. We also compare the ability of the respective bankruptcy dataset to produce unbiased parameter estimates by applying them to a validation sample with the same bankruptcy dummies. Note that the fact that BvD deletes firm information after five years of inactivity does not alter the results of this comparison as we restrict our sample to the period with full BvD data coverage.

The empirical results of this study are as follows. First, more than 80% of all public German firms' bankruptcies can be extracted from easily accessible corporate disclosures. Second, HL bankruptcy events are more complete and accurate than those listed by BvD and Compustat Global. While our HL database includes 277 bankruptcies, BvD and Compustat cover only 63 and 27 events, respectively. BvD and Compustat Global's incomplete data applies not only to small- and medium-sized enterprises but also to large firms. For example, BvD does not include the 2009 bankruptcy case of Arcandor AG, a warehouse business valued at 500 million euros. Surprisingly, BvD declares bankruptcy for firms that never filed for insolvency and continue to exist, such as Suedzucker AG. We further find that only a few bankruptcies are captured solely by Compustat Global or Bureau van Dijk and, not by HL. Third, the bankruptcy dates for HL-listed events are more accurate than those contained in BvD and Compustat Global. For 25% of firms, HL reports bankruptcies two months earlier than BvD and for 50% of firms, HL reports bankruptcies 24 months earlier than Compustat Global. Fourth, the choice of bankruptcy database affects parameter estimates. If we use the inaccurate bankruptcy events reported in BvD, the parameters change in terms of significance and size. We show that HL information produces more realistic parameter estimates than BvD data. Fifth, we demonstrate that using HL data, instead of BvD information, has a major impact on out-of-sample results. When researchers use models estimated based on BvD data, they cannot effectively predict true bankruptcy outcomes, that is, out-of-sample results for bankruptcies in the HL database. For example, using BvD bankruptcy information would yield similar out-of-sample performances for the Altman (1968) and Ohlson (1980) models. However, using HL's precise information reveals that the Ohlson (1980) model significantly outperforms that of Altman (1968). Finally, we show that, opposed to models that use only accounting-based variables, market-based bankruptcy prediction models (Bharath and Shumway, 2008; Hess and Huettemann, 2018; Shumway, 2001) are a better fit for the German market.

For our scope of application, we find that the quality of bankruptcy data has a significant impact on the interpretation of bankruptcy prediction model results. Specifically, we speak to the consequences of training bankruptcy models with noisy bankruptcy data. For example, using BvD information instead of HL data suggests that other bankruptcy prediction models may be more appropriate. This study is the first to show that frequently used commercial bankruptcy databases of Compustat Global (e.g., Dahiya and Klapper, 2007; Tian and Yu, 2017) and BvD (e.g., Altman et al., 2017; Filipe et al., 2016; Lohmann and Ohliger, 2017) are inaccurate. We describe a systematic methodology to gather more precise bankruptcy information free of charge and create the first academic bankruptcy database for Germany. Furthermore, using this database we are the first to compare bankruptcy prediction models for Germany based on valid data.

In Section 3.2 describes how we compile our bankruptcy data. Afterwards, Section 3.3 and 3.4 explains the methods for bankruptcy prediction models and our empirical results. Section 3.5. outlines our contribution to future research.

3.2 Our bankruptcy database

3.2.1 German insolvency proceedings

According to Germany's 2009 insolvency statute ("Insolvenzverordnung"), a company or creditor has the right to file a request at the local court ("Amtsgericht"), which is the court of first instance, if there are reasons for insolvency.[18] Such reasons

[18] In fact, insolvency is the state of companies being able to their liabilities, bankruptcy commonly refers to the legal status as being declared to be unable to serve liabilities. However, since both words have

could be a company's illiquidity (inability to meet the obligations that are due), imminent illiquidity, or over-indebtedness (obligations exceed assets). The company is obliged to file for insolvency within the first weeks of experiencing illiquidity or over-indebtedness.

After a company files for insolvency, the responsible court may take protective measures, which include appointment of an interim insolvency administrator. If the administrator verifies that the company's funds are sufficient to cover the costs of a proceeding, he or she initiates insolvency proceedings; otherwise, the company is liquidated. The insolvency administrator takes over the company's administration and is responsible for restructuring measures, liquidating business units, and collecting outstanding receivables to partially service creditors' claims. Exchange-listed firms are required to immediately report material events to their stakeholders and the public at large. They must submit ad-hoc statements when insolvency is imminent, or an application is submitted to the corresponding court. The corresponding court issues additional statements when taking proactive actions, openings, or further information on bankruptcy proceedings. Both sources are used in this study to compile a dataset of bankruptcy information.

3.2.2 Extracting German bankruptcy data from public sources

The most commonly used database for German companies in the finance and accounting literature stems from Compustat Global or BvD as commercial databases. However, using both databases for bankruptcy prediction raises several issues. For instance, Compustat Global only contains delisting dates and the reasons for the delisting. However, a delisting is often requested at a later stage during the bankruptcy proceedings. Thus, delisting dates are often determined several years after a firm has applied for bankruptcy. More recent studies predominantly are based on BvD that deletes a firm's financial data when it has not published annual reports for five consecutive years. This may apply to firms in bankruptcy proceedings and it is likely that the database includes firms that filed for bankruptcy more than five years ago. For instance, Filipe et al. (2016) uses a sample period from 2000 to 2009 but find no bankruptcies for 2000. In fact, studies

similar connotations and financial research can only measure the probability of bankruptcies, as form of insolvency, both words are used synonymously.

that apply BvD's bankruptcy information can use only the training and validation samples from the past five years. To obtain a sufficient number of observations for a coherent bankruptcy prediction analysis, previous studies have utilized data from various countries, although the definition for bankruptcy events and prediction variables tend to differ by country (see, e.g., Altman et al., 2017).

This study is the first to describe a methodology to systematically collect accurate bankruptcy data from public sources and apply it in the context of Germany. We aggregate our bankruptcy data from multiple online and free-access sources: (i) financial disclosures from "Deutsche Gesellschaft für Ad-hoc-Publizität mbH (DGAP)", and "APA Originaltext-Service GmbH (APA OTS)", (ii) the German business register, and (iii) InsolNet, which is a specific bankruptcy online database. Our approach is straightforward and follows three steps. First, we parse corporate news releases for bankruptcy-related news. Second, we crawl online releases by the German bankruptcy courts, which are compiled in the German business register. Finally, we validate our results by obtaining data from an explicit bankruptcy database. To apply this methodology to other countries, it is important to check for public availability of the disclosure statements.

3.2.2.1 Corporate news releases

In regulated stock markets, companies must immediately inform investors about material events, particularly when they apply for bankruptcy. Financial disclosures and company news releases are mainly distributed by professional ad-hoc service providers. The German market is highly concentrated in the DGAP, which currently distributes approximately 98% of all news releases in Germany. Further, we consider APA OTS as an additional source for corporate news releases because it covered several German firms from 2007 to 2011 before it stopped reporting on German firms in 2013. Another advantage of news releases by ad-hoc providers is that they are free of charge.

In the first step, we use Python, a script-based programming language, to direct web queries to the DGAP and APA OTS web servers. We request all news releases and download the full-text information of each document. We collect the complete archives for both DGAP and APA OTS, containing 363,282 news releases for listed companies from 1997 to 2016. For each article, we process the full document into tokens of single words to evaluate if words are related to bankruptcy news. We use dynamic regular expressions to test if the root of each word contains insolvency wordings. These regular expressions create a word list of 150 German words connected to news releases about bankruptcy (see Appendix B for the full list). This procedure reduces the overall set of documents to a concise sample of 462 disclosures. We then manually check the news releases to aggregate the bankruptcy information, most importantly, the dates of bankruptcy filings and openings.

3.2.2.2 German business register (Unternehmensregister)

The German business register is a government entity that provides free public access to key corporate information such as annual reports, court statements, or register keys. It is the central platform for storing company data. The register also serves as a distributor of key statements from bankruptcy courts containing information about bankruptcy dates, decisions, status, meetings, and further proceedings. Notably, information is available for both public and private firms. However, researchers should be aware of the official deletion of proceedings of online bankruptcy statements.[19] To process this information, we create web queries for information about each firm to check for any bankruptcy court statements. In addition, we manually review the results obtained from our automated web queries.

3.2.2.3 InsolNet

As a robustness check, we submit similar web requests for each firm to *InsolNet.de*'s web servers. InsolNet is a commercial data provider that compiles statements from bankruptcy courts and presents them in a structured manner. Therefore, we examine the correctness and completeness of our bankruptcy data using InsolNet. However, since it provides only the opening date for bankruptcy proceedings, we prefer data from other sources to obtain the initial dates when bankruptcy information was made public. Even though this study focuses on companies listed on stock exchanges to include bankruptcy prediction models requiring capital market information, such as Shumway (2001), our data collection approach can be used to extract bankruptcy information for all private and public companies in Germany.

[19] Bankruptcy statements are deleted six months after the bankruptcy proceedings are completed.

3.2.3 Summary statistics

We apply a straightforward definition for bankruptcy: A firm is bankrupt if it has filed a request to initiate an insolvency proceeding. We generally use the date of bankruptcy filing as an indicator because it is the earliest mention of a company's financial distress. Further, we exclude firms in liquidation, since they may have ceased activities for reasons other than failure, for example, shareholder decisions, mergers, or discontinuation of operations by an allied company or foreign branch. Thus, we include only firms that are seriously financially distressed.

Using this straight definition of bankruptcy, we apply our predefined framework and collect voluminous data on bankruptcies of exchange-listed firms in Germany during the last decades. Table 3.1 illustrates the bankruptcy data we collected from multiple sources mentioned above and shows numbers of bankruptcy in total and, in addition, those used in our sample. In general, our data includes all 1,711 securities listed in the Compustat Global company and security files, either incorporated or headquartered in Germany. Most importantly, Compustat Global provides us with the international securities identification number (ISIN) as primary, and database-independent, identifier for German stocks. To perform web queries with the German business register and InsolNet, we reference to the ISIN to merge corporate news releases with company's stocks and names. As stated above we focus on public firms and, thus, we find that most firms release public disclosure statements immediately after filing for bankruptcy. That is, over 80% of our bankruptcy data originates from corporate disclosures by the DGAP and APA OTS. Our sample data on bankruptcy filings even consists by 97% of information retrieved from corporate disclosures. This is not surprising, as German disclosure obligations require firms to disclose material events, such as bankruptcy filings, immediately to its stockholders. However, the extend of information that can be obtained from such disclosures suggests that researchers should closely examine corporate news in other countries to likewise obtain more reliable data, for instance on bankruptcies. In contrast, data from other sources tends to be limited. Nevertheless, it is noteworthy that the business register is also an interesting source of data. Most notably, it covers all disclosures made by bankruptcy courts even for private firms. However, we note that firm-level bankruptcy information data compiled by us is available upon request.

Table 3.1 Data sources of HL bankruptcy data

		All firms		Sample firms	
Id	**Source**	**N**	**%**	**N**	**%**
1	DGAP & APA OTS news releases, Filing date	230	83.0%	135	97.1%
2	Unternehmensregister, Filing date	9	3.2%	1	0.7%
3	Unternehmensregister, Earliest date	16	5.8%	3	2.2%
4	DGAP & APA OTS news releases, Opening date	6	2.2%		0.0%
5	Unternehmensregister, Opening date	4	1.4%		0.0%
6	Insolnet, Opening date	8	2.9%		0.0%
7	Insolnet, Opening date (with historical names)	3	1.1%		0.0%
8	Web search	1	0.4%		0.0%
	Total bankruptcies	**277**	**19.3%**	**139**	**20.6%**
	Total non-bankruptcies	**1,434**		**674**	

Notes: This table reports the data sources used to create the HL bankruptcy data along with the proportion of firms. "All firms" includes all entities in Compustat Global that are either incorporated or headquartered in Germany (i.e., 1,711). The "Sample firms" are companies with sufficient accounting and stock market data to predict the probabilities of several bankruptcy models. In general, the bankruptcy data stems from either ad-hoc disclosure (e.g., DGAP or APA OTS) or bankruptcy notifications in German business register or Insolnet. We consistently use the earliest available data of bankruptcy notifications (i.e., filing for bankruptcy proceedings). Since most of our observations are directly obtained from ad-hoc disclosures, our bankruptcy data commonly refer to the date of filing for insolvency proceedings.

3.2.4 Comparison with other bankruptcy databases

We evaluate our collected bankruptcy data (HL data) against the predominantly used data sources, including the delistings on Compustat Global and BvD status codes. Broadly, Compustat Global provides delisting dates and the reasons for delisting. We follow Dahiya and Klapper (2007) and Tian and Yu (2017), who classify bankrupt firms based on reason 2 ("bankruptcy") and reason 3 ("liquidation"). Altman et al. (2017) and Filipe et al. (2016) use BvD's status code to indicate if firms are in liquidation or bankruptcy proceedings. We call a firm bankrupt if it has been assigned the BvD status code "Active (insolvency proceedings)." There are two reasons we do not include the status levels "Active (default of payment)", "Active (dormant)", "Dissolved", "Dissolved (liquidation)" or "In liquidation", which also apply to German public firms. The notional reason is that we aim for a consistent definition of bankruptcy across all databases, most importantly in our HL data. However, the practical reason is that BvD does not provide any date for status levels other than "Active (insolvency proceeding)". This is critical as there is no information on whether a firm's financial statements can be used in an empirical analysis of out-of-sample data.

Table 3.2 Bankruptcy frequencies across diverse databases

	All			Sample		
	Compustat Global *Delisting (Liquidation, Bankruptcy)*	**Bureau van Dijk** *Status Code (Bankruptcy proceedings)*	**HL**	**Compustat Global** *Delisting (Liquidation, Bankruptcy)*	**Bureau van Dijk** *Status Code (Bankruptcy proceedings)*	**HL**
1996	1	-	-	-	-	-
1997	1	-	1	-	-	-
1998	1	-	-	-	-	-
1999	-	-	3	-	-	-
2000	1	-	3	-	-	-
2001	1	-	23	-	-	10
2002	2	-	44	-	-	23
2003	1	-	20	1	-	11
2004	2	-	16	2	-	11
2005	-	-	6	-	-	5
2006	3	-	7	2	-	3
2007	1	-	8	1	-	5
2008	2	-	16	-	-	10
2009	-	5	27	-	2	15
2010	1	6	20	-	5	12
2011	2	6	11	-	4	5
2012	2	12	15	1	5	6
2013	2	12	19	-	4	8
2014	1	10	12	1	3	5
2015	2	8	16	1	6	6
2016	1	4	10	-	2	4
	27	**63**	**277**	**9**	**31**	**139**

Note: This table reports the numbers of corporate bankruptcies in each year on the basis of different databases (i.e., Compustat delisting codes, Bureau van Dijk status codes, and our HL databases). Compustat delisting codes for liquidation or bankruptcy are obtained from Compustat Global. Bureau van Dijk data is taken from the Amadeus subscription. The HL database is created using the approach described in this study. While the information by BvD is available only for the most recent years, data from Compustat and HL date back to 1996 and cover large bankruptcies that occurred during the early 2000s recession.

Table 3.2 summarizes the distribution of bankruptcies from 1996 to 2016. Our approach identifies 277 bankrupt firms, whereas BvD includes 63 bankrupt firms and Compustat provides only 27 delistings. Note that, for this study, we gather bankruptcy information only for firms covered by Compustat Global. In this case, the difference in number of bankruptcies cannot be attributed to a different firm coverage. Our final sample, as constructed in Table 3.1, consists of 139, 31, and 9 bankruptcies that arise from data collected from HL, BvD, and Compustat, respectively. Table 3.2 clearly shows that Compustat Global delisting codes cannot be used to conduct a valid bankruptcy prediction analysis. Delisting codes are generally a bad proxy for bankruptcies. Foremost, Compustat Global categorizes many firms as delisted for “Other reasons” without providing further details. In addition, some firms experienced turnaround under bankruptcy administration and restructuring and, thus, were not delisted. Vice versa, Compustat does not delist several bankrupt firms undergoing bankruptcy proceedings because they still trade at penny levels. BvD provides somewhat better bankruptcy data, although the coverage is less than 50% of our bankruptcy data. BvD deletes firm history five years after bankruptcy and, thus, BvD data that is requested in 2017 contains only bankruptcies between 2013 and 2017. Note that this limitation does not affect the results of our comparison analysis in Section 3.2 since it only uses the time period covered by BvD bankruptcies. Because we also have access to BvD's vintage data that we extracted in the years 2013, 2014, 2015, and 2016, we can artificially extend BvD's horizon and identify bankruptcies in earlier years. However, these vintage data can no longer be requested through WRDS or directly from BvD.[20] Despite these measures, our methodology yields significantly more bankruptcies than Compustat and BvD.

Table 3.3 shows the number of bankruptcies captured by Compustat Global and BvD respectively, but not by the HL database. First, we find that Compustat Global reports 12 firms that filed for bankruptcy, which are not covered in HL. Vice versa, this means that HL exclusively captures 262 firms.[21] Most delistings exclusively included in Compustat Global stems from firms in liquidation and not bankruptcies and, thus, firms that not fit the straight definition of bankruptcy in the HL data. Moreover, we find that the other commercial database, BvD, captures data for two firms that are not present in the HL database. Most noteably, one of the two companies is Suedzucker AG, a renowned MDAX company. According to an extensive analysis of Suedzucker AG, this firm never issued a bankruptcy application, filing, or statement at all. Accordingly, the information provided information in BvD is partially wrong. However, given these two firms that are not included in HL and the 63 bankruptcies covered in BvD and HL, HL exclusively captures 216 firms.[22] Furthermore, it is noteworthy that in our final sample, neither Compustat Global nor BvD exclusively capture any bankruptcy that are not already covered in HL.

[20] In fact, several information on German firms within BvD is obtained from Creditreform as the original data provider. We also directly contacted Creditreform to request bankruptcy information, however, Creditreform also deletes bankruptcy information three years after a firm's bankruptcy.

[21] In detail, we combine the results from table 3.2 and 3.3 and find that 277 firms are covered in HL, 27 firms are covered in Compustat Global. In addition, only 12 firms of these 27 firms are exclusively covered by Compustat Global. Thus, only 15 firms that are covered by HL are also covered by Compustat Global and, therefore, 262 firms of 277 are exclusively covered in HL.

[22] Once again, we combine the results from table 3.2 and 3.3 and find that 277 firms are covered in HL, 63 firms are covered in Bureau van Dijk. Given table 3.3, only two firms of these 63 firms are exclusively included in Bureau van Dijk data. Thus, 61 firms that are covered by Bureau van Dijk are also covered in HL and, therefore, 216 firms of 277 are exclusively covered in HL, respectively.

Table 3.3 Bankruptcy frequencies across databases (not covered in HL)

	All		Sample	
	Compustat Global	**Bureau van Dijk**	**Compustat Global**	**Bureau van Dijk**
	Delisting (Liquidation, Bankruptcy)	***Status Code (Bankruptcy proceedings)***	***Delisting (Liquidation, Bankruptcy)***	***Status Code (Bankruptcy proceedings)***
1996	1	-	-	-
1997	1	-	-	-
1998	1	-	-	-
1999	-	-	-	-
2000	-	-	-	-
2001	1	-	-	-
2002	-	-	-	-
2003	-	-	-	-
2004	-	-	-	-
2005	-	-	-	-
2006	1	-	-	-
2007	-	-	-	-
2008	2	-	-	-
2009	-	-	-	-
2010	1	-	-	-
2011	2	-	-	-
2012	-	1	-	-
2013	1	-	-	-
2014	-	1	-	-
2015	-	-	-	-
2016	1	-	-	-
	12	**2**	**-**	**-**

Note: This table reports the number of corporate bankruptcies for each year that were listed in Compustat (delisting codes) and Bureau van Dijk (status codes) respectively, but not captured by HL databases. Compustat delisting codes for liquidation and bankruptcy are obtained from Compustat Global. Bureau van Dijk data stems from the Amadeus subscription. The HL database is created using the approach described in this study. While the information from BvD is available only for the most recent years, the data from Compustat and HL date back to 1996 and cover large bankruptcies of the early 2000s recession.

Table 3.4 provides further details on the firms that go bankrupt each year. We report the largest bankrupt firms in terms of market equity at the preceding fiscal year-end. While Compustat Global delistings do not provide valid bankruptcy data (not even for large-scale firms), BvD is somewhat consistent with data for certain years. For example, in 2013, the largest bankruptcy reported in the BvD database is that of Praktiker AG and this is consistent with our HL database. However, unlike HL, BvD does not account for Arcandor AG, which filed for bankruptcy in 2009. We also note that BvD data has some serious errors. For Solar Millennium AG, BvD reports 2012 as the bankruptcy year, whereas the company went bankrupt in 2011. Similarly, it claims that Suedzucker AG, a renowned MDAX company, had been going through bankruptcy proceedings since 2012, even though this firm never filed for bankruptcy. The results show that our approach not only provides further bankruptcy data for small-scale firms but also proves that commercial data sources are inaccurate for even the largest firms.

Table 3.4 Largest bankruptcies covered in diverse databases

Panel A: Largest bankruptcies in Germany: HL versus. Bureau van Dijk

	HL		Bureau van Dijk	
	Company name	*MkEq*	*Company name*	*MkEq*
2001	KINOWELT MEDIEN AG	1542.2	-	-
2002	ISION INTERNET AG	775.3	-	-
2003	MEDIA AG	77.0	-	-
2004	AGIV REAL ESTATE AG	108.6	-	-
2005	PGAM ADVANCED TECHNOLOGIE AG	46.3	-	-
2006	HUCKE AG	22.9	-	-
2007	KOEHLER & KRENZER FASHION AG	35.8	-	-
2008	THIELERT AKTIENGESELLSCHAFT	354.1	-	-
2009	ARCANDOR AG	5188.4	EDOB ABWICKLUNGS AG	478.3
2010	PRIMACOM AG	189.6	PRIMACOM AG	189.6
2011	SOLAR MILLENNIUM AG	272.3	AGIV REAL ESTATE AG	108.6
2012	CENTROTHERM INTERNATIONAL AG	570.1	SOLAR MILLENNIUM AG	272.3
2013	PRAKTIKER AG	79.3	PRAKTIKER AG	79.3
2014	HANSA GROUP AG	144.2	MIFA MITTELDEUTSCHE FAHRRADWERKE	73.4
2015	JOYOU AG	307.3	JOYOU AG	307.3
2016	KTG ENERGIE AG	72.5	HELIOCENTRIS FUEL ENERGY SOL	53.5

Panel B: Largest bankruptcies in Germany: HL versus Compustat Global

	HL		Compustat Global	
	Company name	*MkEq*	*Company name*	*MkEq*
2001	KINOWELT MEDIEN AG	1542.2	-	-
2002	ISION INTERNET AG	775.3	-	-
2003	MEDIA AG	77.0	TELESENS KSCL AG	616.0
2004	AGIV REAL ESTATE AG	108.6	DAS WERK AG	245.0
2005	PGAM ADVANCED TECHNOLOGIE AG	46.3	-	-
2006	HUCKE AG	22.9	UMWELTKONTOR RENEWABLE ENERGY	66.4
2007	KOEHLER & KRENZER FASHION AG	35.8	ADORI AG	17.3
2008	THIELERT AKTIENGESELLSCHAFT	354.1	-	-
2009	ARCANDOR AG	5188.4	-	-
2010	PRIMACOM AG	189.6	-	-
2011	SOLAR MILLENNIUM AG	272.3	-	-
2012	CENTROTHERM INTERNATIONAL AG	570.1	PHENOMEDIA AG	96.6
2013	PRAKTIKER AG	79.3	-	-
2014	HANSA GROUP AG	144.2	CONERGY AG	50.3
2015	JOYOU AG	307.3	TRIA IT-SOLUTIONS AG	3.2
2016	KTG ENERGIE AG	72.5	-	-

Note: This table contrasts the largest bankruptcy for each calendar year in our HL database against existing commercial databases (i.e., BvD and Compustat Global). This analysis exclusively covers our sample firms to provide scalable data on the economic relevance and size of bankrupt firms. Most importantly, the test provides information on whether the coverage of databases is restricted by size and years and proves that extensive coverage of our HL database is not created by solely considering the bankruptcies of smaller firms with less economic relevance.

Table 3.5 Differences in bankruptcy dates across HL, Compustat and Bureau van Dijk data

Panel A: Time differences for all firms								
Differences in days	Mean	P1	P25	Median	P75	P99	Std	N
D (HL vs. Compustat Global)	1,064	145	354	723	1,374	3,870	1,111	15
D (HL vs. Bureau van Dijk)	259	−95	0	2	56	3,599	817	61
Differences in months	Mean	P1	P25	Median	P75	P99	Std	N
D (HL vs. Compustat Global)	35	5	12	24	46	129	37	15
D (HL vs. Bureau van Dijk)	9	-3	0	0	2	120	27	61
Panel B: Time differences for sample firms								
Differences in days	Mean	P1	P25	Median	P75	P99	Std	N
D (HL vs. Compustat Global)	1,219	354	723	816	1,374	3,870	1,089	9
D (HL vs. Bureau van Dijk)	261	0	0	1	69	3,378	787	31
Differences in months	Mean	P1	P25	Median	P75	P99	Std	N
D (HL vs. Compustat Global)	35	12	12	35	59	59	33	9
D (HL vs. Bureau van Dijk)	9	0	0	0	2	113	9	31

Note: This table reports the differences between the dates of initial bankruptcy filings. To elaborate, we compare the lags in days and months between the explicit dates of bankruptcy announcements by different databases. Compustat delisting codes for liquidation or bankruptcy are obtained from Compustat Global. Bureau van Dijk data is taken from the Amadeus subscription. The HL database is created using the approach described in this study.

Table 3.5 highlights the differences in bankruptcy dates by database. Panel A reports the results for the full sample, while Panel B contrasts the results for the sample used in the bankruptcy models. The median difference in dates between Compustat and HL is 24 months for all firms and 35 months for sample firms. That is, for half the firms, Compustat reports a bankruptcy date that is more than 24 months (35 months) after the bankruptcy date registered in our database. The 75th percentile is 46 months for all firms and 59 months for the sample firms. That is, for 25% of the sample firms, Compustat reports a delisting date that is 59 months or more after our date. This is because stock delistings generally happen several years after firms file for bankruptcy. For bankruptcy events in BvD, half the dates are somewhat congruent with those in our HL database. The median distance between the BvD and HL bankruptcy dates is two days for the full sample and one day for our sample. This small lag is because BvD relies on court announcements that slightly lag the direct corporate announcements that are used to determine bankruptcy filing dates in the HL database. However, the 75th percentile is 56 days for all firms and 69 days for the sample firms. Thus, BvD dates substantially lag behind the data we collected. Note that we can compare only a few HL events with those in BvD, given the poor coverage of the latter. For comparison, we can use 31 of the 139 bankruptcies (22%) in the HL dataset for our sample. Overall, the results in Table 3.5 support the notion that commercial bankruptcy data is inaccurate in terms of bankruptcy dates.

Overall, the HL bankruptcy dataset outperforms both BvD and Compustat in terms of coverage, correctness of bankruptcy events, and accuracy of bankruptcy dates.

This indicates that our dataset cannot be reproduced by simply gathering information from the two commercial bankruptcy databases. In the following sections, we investigate if this higher quality of bankruptcy information influences the interpretation of bankruptcy prediction model results.

3.3 Data and method

3.3.1 Sample description and summary statistics

Our initial sample includes all firms listed in Compustat Global's company and security files that are either incorporated or headquartered in Germany between 1995 and 2015. While coverage of bankruptcies in BvD starts in 2009, our HL data allows to cover a longer history of data. This is very important as several bankruptcy prediction models require a reliable validation dataset. For instance, the model of Hess and Huettemann (2018) requires several years to predict future profitability measures and derive bankruptcy probabilities. We delete observations with data errors and missing values. In detail, we exclude all observations that do not contain all variables that are required for any bankruptcy prediction model. This allows us to exclude erroneous observations where fundamental information, such as earnings or total assets, are not available. Therefore, we require the variable sets in Altman (1968), Ohlson (1980), Shumway (2001), Bharath and Shumway (2008), and Hess and Huettemann (2018) for each observation.

Appendix A describes the variable construction for these bankruptcy prediction models in detail. To reduce the effect of outliers, we winsorize all variables (except indicator variables and probabilities) annually at the 1st and 99th percentile. Since we require five years of training data to perform cross-sectional earnings regressions for Hess and Huettemann (2018), which is based on Hou, van Dijk, and Zhang (2012) and Li and Mohanram (2014), our sample for bankruptcy prediction begins in 2000.

Table 3.6	Summary statistics of variables within selected bankruptcy models (N = 95,431)							
Variable	Model	Mean	STD	1%	25%	Median	75%	99%
$WCTA_t$	*A / O*	0.222	0.258	−0.520	0.064	0.218	0.385	0.830
$RETA_t$	*A*	−0.143	1.075	−5.955	−0.016	0.025	0.168	0.596
$EBITTA_t$	*A / HH*	0.009	0.186	−0.783	−0.009	0.047	0.091	0.317
$METL_t$	*A*	2.856	7.141	0.038	0.446	1.023	2.458	30.273
STA_t	*A*	1.179	0.719	0.016	0.718	1.081	1.479	3.917
$SIZE_t$	*O / HH*	5.208	2.100	1.184	3.766	4.929	6.332	11.554
$TLTA_t$	*O / S*	0.569	0.238	0.074	0.401	0.581	0.723	1.241
$CLCA_t$	*O*	0.712	0.613	0.060	0.383	0.585	0.846	3.333
$OENEG_t$	*O*	0.024	0.154	0.000	0.000	0.000	0.000	1.000
$NITA_t$	*O / S / BS*	−0.025	0.203	−1.018	−0.028	0.023	0.057	0.260
$FUTL_t$	*O*	−0.011	0.626	−2.885	−0.038	0.061	0.169	1.343
$INTWO_t$	*O*	0.205	0.404	0.000	0.000	0.000	0.000	1.000
$CHIN_t$	*O*	0.005	0.597	−1.000	−0.357	0.039	0.350	1.000
$RSIZE_t$	*S*	−25.377	2.643	−31.160	−27.049	−25.535	−23.882	−18.625
ER_t	*S / BS / HH*	−0.379	0.726	−1.715	−0.812	−0.402	−0.042	2.179
$STDER_t$	*S / HH*	0.117	0.109	0.016	0.061	0.092	0.142	0.462
$PNBE_t$	*HH*	0.210	0.204	0.000	0.001	0.172	0.391	0.705
$Neg\ EarnFrc_t$	*HH*	0.346	0.476	0.000	0.000	0.000	1.000	1.000
$CAPXTA_t$	*HH*	0.050	0.052	0.000	0.017	0.036	0.065	0.267
TXTt	*HH*	40.822	169.453	−13.306	0.037	1.503	10.067	979.000
MLR_t	*HH*	0.490	0.255	0.034	0.281	0.489	0.694	0.973
$PD\text{-}Merton_t$	*BS*	0.317	0.318	0.000	0.005	0.223	0.576	0.999
$LNME_t$	*BS*	4.589	2.111	0.691	3.111	4.238	5.784	10.479
$LNBD_t$	*BS*	4.530	2.301	0.023	2.907	4.314	5.832	11.256
$VOLME_t$	*BS*	0.729	0.557	0.123	0.377	0.587	0.898	3.226

Note: This table reports the summary statistics for the following forecast variables (all values except dummy variables and probability values are in million dollars). Each observation represents one firm in a given year. Specifically, it shows variables used to forecast bankruptcy. For more details, see data construction in Appendix A. WCTA is working capital over total assets, RETA is retained earnings over total assets, EBITTA is earnings before interest and taxes over total assets, METL is the market value of equity over the book value of total debt, STA is sales over total assets, SIZE is the logarithm of total assets, TLTA is total liabilities over total assets, CLCA is current liabilities over current assets, OENEG is a dummy that takes the value of one if total liabilities exceed total assets and zero otherwise, NITA is net income over total assets, FUTL is funds provided by operations over total liabilities, INTWO is a dummy that takes the value of one if net income has been negative for the past two years and zero otherwise, CHIN is change in net income, RSIZE is the logarithm of market equity divided by the value-weighted market equity of the index, ER is excess return, STDER is the standard deviation of return, PNBE is the probability that losses deplete current book equity, NegEarnFrc is a dummy for negative earnings forecast, CAPXTA is capital expenditure over total assets, TXT is taxes, MLR is the market leverage ratio, PD-Merton is the KMV probability, LNME is the logarithm of market equity, LNBD is the logarithm of the book value of debt, and VOLME is the inverse of market equity volatility. The reported values are the time series averages of yearly cross-sectional means, medians, standard deviations, and respective percentiles. To treat extreme outliers and data errors, all variables (except indicator variables and probability values) are winsorized annually at the 1st and 99th percentile. The column labeled “Model” indicates in which model the variable has been used, where “A” is Altman (1968), “O” is Ohlson (1980), “S” is Shumway (2001), “BS” is Bharath and Shumway (2008), and “HH” is Hess and Huettemann (2018). The sample period is 2000-2016. The summary statistics are reported for observations in which all the models’ variables are available.

For each firm-year observation, we construct twelve monthly observations to enable market participants to perform bankruptcy predictions for each month. All bankruptcy measures are lagged by three months to ensure that they are observable when used for estimation. For example, the first observation for a firm-year with a fiscal year-end of December 31, 2009 has an estimation date of March 31, 2010, and the last observation for the respective firm-year has an estimation date of February 28, 2011. A firm observation is defined as bankrupt if the firm files for bankruptcy exactly twelve months after the date of estimation. Thus, in such cases, the dependent variable equals one; otherwise, it equals zero. Since we account for bankruptcies until the end of 2016, our sample includes firm months with an estimation date before or at the end of December 2015.[23]

Table 3.6 provides the summary statistics for all variables used to forecast bankruptcy. We report the mean, median, standard deviation, and certain percentiles of 95,431 firm months with complete data availability for 2000-2015. The results indicate a significant cross-sectional variation among these variables. For example, RETA has a standard deviation of 1.975. In addition, its 1st and 99th percentiles are -5.955 and 0.596, respectively. Interestingly, the probability of book equity becoming negative, PNBE has a mean of 0.210, which is twice that found in Hess and Huettemann's (2018) study on US firms (0.107). This suggests that German firms operate with negative book equity more frequently than US firms.

3.3.2 Method

Shumway (2001) demonstrates that the likelihood function of hazard models is equivalent to that of logistic regressions with multiple observations per firm. We follow Shumway (2001), Chava and Jarrow (2004), and Campbell, Hilscher, and Szilagyi (2008) and estimate the hazard model as a multi-period logistic regression. Thus, the probability of a firm becoming bankrupt follows a logistic distribution with parameters (α, β) and is equal to

$$P_t(y_{it} = 1) = \frac{1}{1+exp(-\alpha-\beta x_{i,t})} \tag{3.1}$$

where y_{it} is a bankruptcy dummy that equals one if the firm fails in twelve months and zero otherwise, and $x_{i,t}$ is the vector of explanatory variables that are known at time t. The higher the term $\alpha + \beta x_{i,t}$, the greater the estimated probability of bankruptcy. The estimates and their significance levels are calculated using a maximum likelihood technique. Shumway (2001) points out that the test statistics produced by a logistic regression are incorrect for the hazard model. Correct test statistics are calculated

[23] We also perform empirical tests that predict bankruptcy for a forecast horizon of one month rather than twelve months and find that the results remain robust regardless of horizon change.

by dividing them by the average number of observations per firm. The statistics reported in this study have been adjusted accordingly.

We conduct two empirical analyses. First, we compare the bankruptcy prediction models. To produce strictly out-of-sample forecasts, we estimate the parameters using data between 2000 and 2007 and apply the resulting coefficients to predict bankruptcies from 2008 to 2015. Second, we compare the HL and BvD databases. Given the data restrictions in the BvD database, we are limited to a shorter period. We estimate the parameters with data from 2009 to 2012 and then use the coefficients to predict bankruptcies from 2013 to 2015.

Static models are based on a single observation per firm and, thus, result in sample selection bias. In contrast, our approach uses all available firm observations to estimate the logistic regression. In fact, our estimation technique exploits more information and eliminates any sample selection bias. Note that applying such a technique to Altman's (1968) and Ohlson's (1980) static models already improves their performance as compared to adopting the estimation techniques originally suggested.

3.4 Empirical results

3.4.1 Comparison across models

3.4.1.1 Estimation results

As the first part of our performance analysis, we assess which of the commonly used bankruptcy prediction model better fits the German stock market. Therefore, we estimate all previously mentioned bankruptcy models using the accounting information in Compustat Global and bankruptcy data in our HL database. Table 3.7 reports the estimation results for the hazard models of Altman (1968), Ohlson (1980), Shumway (2001), Bharath and Shumway (2008), and Hess and Huettemann (2018) including parameter estimates as well as their significance. In addition, it presents the likelihood ratio test for each model.

Table 3.7 Parameter estimates of common bankruptcy models

Variable	Altman (1968)	Ohlson (1980)	Shumway (2001)	Bharath and Shumway (2008)	Hess and Huetttemann (2018)
Constant	-6.159 ***	-6.400 ***	-12.665 ***	-6.783 ***	-8.968 ***
	[622.75]	[71.00]	[77.20]	[145.32]	[211.75]
$WCTA_t$	-0.674	-1.054			
	[2.70]	[1.28]			
$RETA_t$	0.000				
	[0.02]				
$EBITTA_t$	-1.394 ***				-0.614
	[20.47]				[2.35]
$METL_t$	0.015				
	[1.63]				
STA_t	-0.216				
	[1.59]				
$SIZE_t$		-0.115			-0.019
		[2.30]			[0.04]
$TLTA_t$		0.869	0.766 *		
		[1.64]	[3.05]		
$CLCA_t$		-0.407			
		[0.91]			
$OENEG_t$		-0.163			
		[0.07]			
$NITA_t$		0.397	-0.811 **	-0.450	
		[0.40]	[6.16]	[1.60]	
$FUTL_t$		-0.230			
		[0.99]			
$INTWO_t$		1.181 ***			
		[18.50]			
$CHIN_t$		-0.716 ***			
		[12.10]			
$RSIZE_t$			-0.176 ***		
			[10.94]		
ER_t			-1.312 ***	-1.037 ***	-1.227 ***
			[49.86]	[38.08]	[35.84]
$STDER_t$			-0.243		0.145
			[0.34]		[0.10]
$PNBE_t$					2.734 ***
					[10.95]
$Neg\ EarnFrc_t$					0.072
					[0.08]
$CAPXTA_t$					2.128
					[2.69]
TXTt					-0.012 *
					[3.13]
MLR_t					0.868 *
					[3.63])
$PD\text{-}Merton_t$				1.058 **	
				[4.06]	
$LNME_t$				-0.229	
				[6.34]	
$LNBD_t$				0.004	
				[0.00]	
$VOLME_t$				-0.439	
				[0.76]	
N	47,738	47,738	47,738	47,738	47,738
LRT	25.97 ***	56.07 ***	93.25 ***	103.30 ***	112.81 ***

Note: This table reports the results of the hazard models for the bankruptcy indicators for Altman (1968), Ohlson (1980), Shumway (2001), Bharath and Shumway (2008), and Hess and Huettemann's (2018) market model. Parameter estimates for all the variables in each model are reported along with their chi-square statistics in parentheses. The hazard model is estimated for 2000-2007 with 47,738 observations and 78 bankruptcies. The chi-square of the likelihood ratio test for each model is reported in the row labeled LRT. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

The likelihood ratio test is significant for each model. Thus, for all models, the independent variables have explanatory power. Not all variables are statistically significant, which is in contrast to studies on U.S. firms. However, if the parameters are statistically significant, the signs of these coefficients are consistent with economic intuition and those found in previous studies. For Altman (1968), lower profitability (EBITTA) yields higher estimated probability of bankruptcy. In the case of Ohlson (1980), the probability of bankruptcy rises if net income is negative for the past two years (INTWO) and if the change in net income (CHIN) is negative. For Shumway (2001), firms that are more leveraged (TLTA), less profitable (NITA), and smaller (RSIZE) are more likely to become bankrupt. In Bharath and Shumway (2008), lower excess return (ER), lower market equity (LNME), and higher PD-Merton yield higher estimated default probability. Finally, for Hess and Huettemann (2018), firms with higher PNBE are more likely to fail. The higher the market leverage ratio (MLR) and the lower the tax (TXT) and excess return (ER), the greater the estimated probability of bankruptcy.

3.4.1.2 Out-of-sample results

Table 3.8 presents the out-of-sample accuracies. Panel A reports the goodness-of-fit deciles. To create this table, we rank firms into deciles based on their fitted bankruptcy probability values for each year in our validation sample (i.e., 2008 to 2015). That is, firms most likely to default in the subsequent year are sorted into the first decile and those with the lowest estimated default probabilities are assigned to the tenth decile. We report the percentage of bankrupt firms that fall under each of the ten probability deciles. A model is accurate if it estimates a high default probability for bankrupt firm-years and assigns many bankrupt firms into low deciles.

Hess and Huettemann’s (2018) model classifies 59.02% of all bankrupt firms into the highest default probability decile (decile one). That is, a bank can exclude 59.02% of all bankruptcies if it does not lend money to the 10% of firms with the highest expected default measures. Shumway (2001) and Bharath and Shumway (2008) classify 57.38% and 54.1% of all bankrupt firms into the first decile, respectively. As a result, models using a combination of accounting and market information strongly outperform Altman’s (39.34%) and Ohlson’s (32.79%) accounting-based models.

Table 3.8 Out-of-sample results: Comparison across models

Panel A: Goodness-of-fit deciles

Decile	Altman (1968)	Ohlson (1980)	Shumway (2001)	Bharath and Shumway (2008)	Huettemann and Hess (2018) – Market
1	39.34	32.79	57.38	54.10	59.02
2	11.48	24.59	24.59	24.59	14.75
3	9.84	16.39	8.20	9.84	11.48
4	6.56	6.56	4.92	6.56	6.56
5	6.56	1.64	0.00	3.28	1.64
6	3.28	1.64	3.28	0.00	1.64
7	1.64	3.28	0.00	0.00	1.64
8	9.84	1.64	0.00	0.00	0.00
9	6.56	4.92	0.00	0.00	0.00
10	4.92	6.56	1.64	1.64	3.28

Panel B: Area under the ROC curve

Model	Mean	95% Confidence interval		STD
Altman (1968)	0.675 ***	0.597	0.752	0.039
Ohlson (1980)	0.731 ***	0.656	0.804	0.037
Shumway (2001)	0.851 ***	0.804	0.898	0.024
Bharath and Shumway (2008)	0.854 ***	0.81	0.899	0.023
Huettemann and Hess (2018)	0.842 ***	0.789	0.895	0.027

Panel C: Comparison of Area under the ROC curve

	Altman (1968)	Ohlson (1980)	Shumway (2001)	Bharath and Shumway (2008)	Huettemann and Hess (2018) – Market
Altman (1968)	-	0.057	0.176 ***	0.180 ***	0.167 ***
Ohlson (1980)	-0.057	-	0.120 ***	0.123 ***	0.111 ***
Shumway (2001)	-0.176 ***	-0.120 ***	-	0.003	-0.009
Bharath and Shumway (2008)	-0.180 ***	-0.123 ***	-0.003	-	-0.012
Huettemann and Hess (2018)	-0.167 ***	-0.111 ***	0.009	0.012	-

Note: This table compares the out-of-sample accuracy of various bankruptcy prediction models. Parameter estimates from the training sample (2000–2007) are used to predict bankruptcies for the validation period 2008–2015. This validation sample includes 47,693 firm-years and 61 bankruptcies. All the models are estimated with a multi-period logistic regression. For Panel A, we rank firms into deciles based on their fitted bankruptcy probability values for every year, where the firms with the highest values fall into the first decile. We report the percentage of bankrupt firms that are classified into each probability deciles. Panel B reports the mean of the area under the ROC curve (AUC) along with its significance to be greater than 0.5 its standard deviation and the 95% Wald confidence interval. Panel C compares the means of the AUC across the models reporting their mean differences and their significances. *** denotes significance at the 1% level.

For the top two deciles (in aggregate), the correct predictions are 81.97% for Shumway (2001), 78.69% for Bharath and Shumway (2008), 73.77% for Hess and Huettemann (2018), 57.38% for Ohlson (1980), and 50.82% for Altman (1968). Panel B reports the distribution of the area under the receiver operating characteristic (ROC) curve, also referred to as area under the curve (AUC), for the validation sample. The ROC curve plots the true positive rate against the false positive rate for all cut-off points. The AUC is measured relative to the area of the unit square. A value of 0.5 indicates a random model with no predictive ability and a value of 1.0 denotes perfect discrimination. To compute the AUC, we estimate the parameters for each model using the training sample (2000-2007) and adopt these parameters to predict bankruptcies in our validation sample (2008-2015). Chi-squared tests for the differences in the means of the AUC across all models are shown in Panel C.

For each model we test the hypothesis that the AUC is equal to 0.5, that is that the model is a purely random classifier. This hypothesis is rejected for all models. Bharath and Shumway's (2008) model has an average AUC of 0.854, which is insignificantly higher than 0.851 in Shumway (2001) and 0.842 in Hess and Huettemann (2018). Given this, the market-based models significantly outperform the accounting-based ones in Ohlson (1980) and Altman (1968) with an average AUC of 0.731 and 0.675, respectively. These results are consistent with those reported using goodness-of-fit deciles: Shumway (2001), Bharath and Shumway (2008), and Hess and Huettemann (2018) have similar out-of-sample performance for German firms, followed by Ohlson (1980) and Altman (1968). This study's results are consistent with those in studies such as Shumway (2001), Hillegeist, Keating, Cram, and Lundstedt (2004), and Campbell et al. (2008), who demonstrate that market variables can improve the accuracy of bankruptcy predictions. In contrast, for example, Reisz and Perlich (2007), and Agarwal and Taffler (2008) show that accounting-based models have similar performance.

3.4.2 Comparison across bankruptcy databases

In the second part of our performance analysis, we analyze the effect of different bankruptcy databases on parameter estimation and validation of bankruptcy prediction models. Since the BvD database deletes firm histories, we reduce our sample period to the range from 2009 to 2015. While this eliminates the effect of BvD's deletion procedure on the analysis results, it ensures a fair test across databases. However, this test obviously includes substantially fewer bankruptcy events for both estimation and validation. In addition, we exclude Compustat Global because it has only two delisting events for this period, which is insufficient for a reasonable analysis.

Table 3.9 Parameter estimates of bankruptcy models using sample ranges covered in Bureau van Dijk (N = 24,457)

Variable	Altman (1968)		Ohlson (1980)		Shumway (2001)		Bharath and Shumway (2008)		Hess and Huettemann (2018)	
	HHL	BvD	HHL	BvD	HHL	BvD	HHL	BvD	HHL	BvD
Constant	-5.80 ***	-6.85 ***	-10.50 ***	-13.40 ***	-14.46 ***	-12.68 ***	-5.11 ***	-5.98 ***	-10.10 ***	-10.88 ***
	[236.56]	[183.98]	[73.95]	[63.49]	[53.40]	[25.17]	[30.92]	[28.61]	[153.57]	[111.31]
$WCTA_t$	-0.96	-0.23	1.04	3.40 **						
	[1.66]	[0.05]	[0.77]	[5.26]						
$RETA_t$	-0.07	-0.05								
	[0.30]	[0.07]								
$EBITTA_t$	-1.57	-1.24							-0.08	0.20
	[2.26]	[0.62]							[0.01]	[0.01]
$METL_t$	-0.78 **	-0.57 *								
	[6.27]	[2.98]								
STA_t	0.01	0.25								
	[0.00]	[0.82]								
$SIZE_t$			-0.05	0.09					0.01	0.20
			[0.24]	[0.56]					[0.13]	[1.85]
$TLTA_t$			5.09 ***	6.41 ***	1.62 ***	1.74 **				
			[18.09]	[14.76]	[8.06]	[5.12]				
$CLCA_t$			0.33	0.69 **						
			[1.51]	[5.50]						
$OENEG_t$			-2.81 ***	-3.86 **						
			[8.31]	[6.39]						
$NITA_t$			-0.07	-0.16	-0.60	-0.13	-1.13 *	-0.89		
			[0.01]	[0.01]	[1.02]	[0.02]	[3.64]	[1.00]		
$FUTL_t$			-0.46	-0.10						
			[1.25]	[0.01]						
$INTWO_t$			0.62	1.08 *						
			[1.99]	[3.65]						
$CHIN_t$			-0.35	-0.48						
			[1.23]	[1.33]						
$RSIZE_t$					-0.22 ***	-0.14				
					[8.24]	[1.98]				
ER_t					-1.32 ***	-0.95 **	-1.42 ***	-1.25 **	-0.86 ***	-0.60
					[14.68]	[4.87]	[12.79]	[6.09]	[6.80]	[2.03]
$STDER_t$					3.46 ***	3.24 **			2.87 **	2.22
					[9.01]	[3.94]			[5.22]	[1.51]

Table 3.9 Parameter estimates of bankruptcy models using sample ranges covered in Bureau van Dijk (N = 24,457)

$PNBE_t$									1.60 *	2.37 **
									[3.76]	[5.21]
$NegEarnFrc_t$									0.51	-0.10
									[0.97]	[0.02]
$CAPXTA_t$									2.43	-3.07
									[0.52]	[0.25]
TXTt									-0.01	-0.01
									[0.80]	[0.98]
MLR_t									3.26 ***	3.11 **
									[8.94]	[4.93]
$PD\text{-}Merton_t$							-0.61	-1.35		
							[0.41]	[1.15]		
$LNME_t$							-0.72 ***	-0.73 ***		
							[25.13]	[15.30]		
$LNBD_t$							0.56 ***	0.72 ***		
							[16.66]	[16.00]		
$VOLME_t$							-3.09 **	-2.64 *		
							[6.24]	[3.61]		
LRT	32.88 ***	11.18 ***	47.84 ***	30.49 ***	70.74 ***	24.84 ***	83.03 ***	34.80 ***	81.10 ***	36.77 ***
Wilks' Lambda F-value	2.59 **		2.94 **		7.46 ***		5.22 ***		4.48 ***	

Note: This table reports the results of the hazard models of the bankruptcy indicators for the market models proposed by Altman (1968), Ohlson (1980), Shumway (2001), Bharath and Shumway (2008), and Hess and Huettemann's (2018) market model. Parameter estimates for all the variables in each model are reported along with their chi-square statistics in parentheses. The hazard model is estimated for 2009–2012 with 24,578 observations. In the training sample, there are 31 bankruptcies for the HHL database and 18 bankruptcies for the BvD database. The chi-square of the likelihood ratio test for the hypothesis that each parameter is equal to zero is reported. Furthermore, the F-value of Wilks' Lambda for the hypothesis that the parameter estimates created by HHL data and BvD data are equal is reported. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

3.4.2.1 Estimation results

Likewise in the first part of our performance analysis, we estimate the commonly used bankruptcy models. However, in this analysis we separately use the HL data and BvD information to estimate parameters and compare those with one another. Table 3.9 reports the hazard model results for all models when using the HL or BvD bankruptcy databases. In addition to the parameter estimates, it presents their significance, and the likelihood ratio test.

For each model, the chi-squared statistic of the likelihood ratio test is higher if we utilize the HL bankruptcies as opposed to the BvD bankruptcies. For example, in Shumway (2001), HL data yields a chi-squared statistic of 70.74 and BvD data produces a value of 24.84. For Ohlson (1980), the variables WCTA, CLCA, and INTWO are only significant if we use the BvD database for the estimation. As for Shumway (2001), the variable RSIZE is significant if we base our estimation on HL bankruptcy events, but not if we use BvD bankruptcies. In Bharath and Shumway (2008), the variable NITA is significant for parameter estimation on HL data, and in Hess and Huettemann (2008), the variables ER and STDER are significant if we use HL for parameter estimation.

In addition to this analysis, we conduct formal tests on the differences in coefficients across the two training samples. In detail, we test whether parameter estimates emerging from the use of the HL data equal those of the BvD data. Given the estimates from our Wilks' Lambda F-statistics, this hypothesis is rejected for each model. Hence, the quality of underlying bankruptcy data does affect the parameter estimates for each model. For instance, if we apply the inaccurate BvD bankruptcy events, the parameter estimates are different in terms of significance and size. While this may only be relevant for in-sample estimation and the set of variables that can explain bankruptcies ex-post, we also test whether the parameter estimates using more accurate data from HL does also provide better results in out-of-sample results and, thus, outperform those estimated by BvD. That is, using the HL data we better predict a better probability of bankruptcy for each company ex-ante.

Table 3.10 Out-of-sample results: Comparison across bankruptcy databases

Panel A: Goodness-of-fit deciles

Decile	Altman (1968)				Ohlson (1980)				Shumway (2001)				Bharath and Shumway (2008)				Huettemann and Hess (2018) - Market			
Training	*HHL*	*BvD*	*HHL*	*BvD*	*HHL*	*BvD*	*HHL*	*BvD*	*HHL*	*BvD*	*HHL*	*BvD*	*HHL*	*BvD*	*HHL*	*BvD*	*HHL*	*BvD*	*HHL*	*BvD*
Validation	*HHL*		*BvD*		*HHL*		*BvD*		*HHL*		*BvD*		*HHL*		*BvD*		*HHL*		*BvD*	
1	40.0	33.3	33.3	33.3	33.3	20.0	25.0	16.7	60.0	53.3	75.0	66.7	46.7	33.3	58.3	41.7	46.7	33.3	50.0	41.7
2	6.7	13.3	8.3	8.3	20.0	20.0	25.0	16.7	13.3	20.0	8.3	16.7	20.0	26.7	16.7	25.0	20.0	6.7	25.0	8.3
3	6.7	6.7	0.0	0.0	26.7	20.0	25.0	16.7	13.3	6.7	8.3	8.3	6.7	13.3	0.0	8.3	13.3	26.7	0.0	16.7
4	6.7	6.7	8.3	8.3	0.0	0.0	0.0	0.0	0.0	13.3	0.0	0.0	6.7	13.3	8.3	8.3	13.3	20.0	16.7	25.0
5	6.7	6.7	8.3	8.3	6.7	13.3	0.0	16.7	6.7	0.0	0.0	0.0	13.3	6.7	8.3	8.3	0.0	6.7	0.0	0.0
6	6.7	6.7	8.3	16.7	0.0	13.3	8.3	16.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.7	6.7	8.3	8.3
7	13.3	20.0	16.7	8.3	0.0	0.0	0.0	0.0	6.7	0.0	8.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	6.7	0.0	8.3	8.3	0.0	6.7	0.0	8.3	0.0	6.7	0.0	8.3	6.7	0.0	8.3	0.0	0.0	0.0	0.0	0.0
9	6.7	6.7	8.3	8.3	6.7	6.7	8.3	8.3	0.0	0.0	0.0	0.0	0.0	6.7	0.0	8.3	0.0	0.0	0.0	0.0
10	0.0	0.0	0.0	0.0	6.7	0.0	8.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Panel B: Area under ROC curve

Model	Mean							
Training	*HHL*	*BvD*			*HHL*	*BvD*		
Validation	*HHL*		*Diff*		*BvD*		*Diff*	
Altman (1968)	0.670	0.682	*-0.011*		0.622	0.630	*-0.008*	
Ohlson (1980)	0.733	0.673	*0.060*	*	0.690	0.630	*0.060*	**
Shumway (2001)	0.841	0.827	*0.014*		0.870	0.861	*0.010*	
Bharath and Shumway (2008)	0.800	0.782	*0.017*		0.821	0.795	*0.026*	**
Huettemann and Hess (2018)	0.831	0.775	*0.056*	**	0.842	0.803	*0.039*	*

Note: This table compares the out-of-sample accuracy for various bankruptcy prediction models with yearly observations. We estimate and validate the sample with HHL and BvD bankruptcy events. The parameter estimates from the training sample (2009–2012) are used to predict bankruptcies for the validation period 2013–2015. This validation sample includes 16,629 firm-years and 15 bankruptcies in the HHL database and 12 bankruptcies in the BvD database. All the models are estimated using a hazard model. For Panel A, we rank firms into deciles based on their fitted bankruptcy probability values for every year, where firms with the highest values are categorized as the first decile. We report the percentage of bankrupt firms that are classified into each probability decile. Panel B reports the mean of the area under the ROC curve (AUC), the differences, and the significance of the differences. ** and * denote significance at the 5% and 10% levels, respectively.

3.4.2.2 Out-of-sample results

Table 3.10 presents the out-of-sample accuracies for all models. The parameters are estimated using the training sample and accuracy is evaluated with the validation sample. Both samples warrant bankruptcy dummies as the dependent variable. Dummies that emerge from the two bankruptcy databases, HL and BvD, are used for both parameter estimation and validation of the models. Thus, we have four out-of-sample results for each model, with which we conduct two empirical tests.

First, we evaluate the ability of each database to produce unbiased parameter estimates. If the parameter estimates from the bankruptcy dummies of one dataset yield better out-of-sample results, we can conclude that this bankruptcy dataset produces better parameter estimates. To derive this information, we compare the two results obtained using different bankruptcy dummies in the training sample but the same HL bankruptcy dummies in the validation sample.

Panel A of Table 3.10 shows the goodness-of-fit deciles. If we use the HL dummies for validation, the rate of bankrupt firms in the highest default probability decile estimated in Altman's (1968) model is 40.00% when using the HL parameter estimates and 33.33% with the BvD estimates. In general, estimating the parameters with HL rather than BvD bankruptcies produces greater accuracy in bankruptcy predictions. Likewise, in Ohlson (1980), the proportion of bankrupt firm-years in decile one is 33.33% if we estimate the parameters with HL dummies, which is higher than the 20% obtained if we estimate parameters with BvD information. We observe the same pattern for Shumway (2001) (HL: 60%, BvD: 53.33%), Bharath and Shumway (2008) (HL: 46.67%, BvD: 33.33%), and Hess and Huettemann (2018) (HL: 46.67%, BvD: 33.33%).

Panel B reports the mean of the area under the ROC curve (AUC) for all models and HL information is again used for the validation sample. Ohlson (1980) has an average AUC of 0.733 if we estimate the parameters with HL data, which significantly exceeds the average AUC of 0.673 when parameters are estimated with BvD data. For Hess and Huettemann (2018), the average AUC is 0.831 with HL estimates, which is significantly higher than the AUC of 0.775 obtained using BvD estimates. The chi-squared tests for the differences in AUC in correlated samples show that these two differences are statistically significant. We find the same pattern for Shumway (2001) (HL: 0.841, BvD:

0.827) and Bharath and Shumway (2008) (HL: 0.800, BvD: 0.782). An exception is Altman (1968), where the average AUC is 0.670 with HL estimates and slightly higher with BvD estimates (0.682).

We observe higher accuracy when using HL data instead of BvD data for parameter estimation. Note that we obtain consistent results when using BvD dummies instead of HL dummies in the validation sample. We conclude that the more accurate bankruptcy information in the HL database compared to the inaccurate BvD database produces more realistic parameter estimates. This analysis speaks to the consequences of training bankruptcy models with noisy bankruptcy data. When models are estimated using BvD data, researchers cannot effectively predict true bankruptcy outcomes, that is, out-of-sample results for bankruptcies in the HL database.

Second, we compare the out-of-sample results when using HL and BvD information for both parameter estimation and validation. When performing parameter estimation with the more accurate HL database, Shumway (2001) has the highest average AUC of 0.841, followed by Hess and Huettemann (2018), Bharath and Shumway (2008), Ohlson (1980), and Altman (1968) with 0.831, 0.800, 0.733 and 0.670, respectively. If we estimate the parameters using the inaccurate BvD database, Shumway (2001) has the highest average AUC of 0.861, followed by Hess and Huettemann (2018) with 0.803, Bharath and Shumway (2008) with 0.795, and Ohlson (1980) and Altman (1968) equal at 0.630.

Thus, previous studies that use BvD information would conclude that Altman and Ohlson’s models have the same out-of-sample performance and are equally effective in predicting bankruptcies. However, we reach a different conclusion when using more accurate HL information: Ohlson has significantly higher performance and, thus, is the better bankruptcy prediction model. Likewise, studies using BvD data would conclude that Shumway (2001) significantly outperforms Hess and Huettemann (2018). In reality, however, if we use HL data, both models perform almost equally well. Specifically, data quality significantly affects the reliability of results for bankruptcy prediction models and the inferences from comparing alternative model specifications.

3.5 Conclusion

In this study, we show that the quality of bankruptcy data has a significant impact on the estimation and evaluation of bankruptcy prediction models. We introduce an alternative database of German bankruptcies by systematically collecting information from public sources. In doing so, we show that our bankruptcy database has more complete and accurate data on bankruptcy events and dates than the most frequently used databases, BvD and Compustat Global. In other words, our bankruptcy database cannot be reproduced using these two commercial databases. To the best of our knowledge, we are the first to make a comprehensive comparison of several bankruptcy prediction models for the German market using an appropriate database. Most importantly, in our analysis of German public firms we demonstrate that the higher quality of our bankruptcy database produces significantly better parameter estimates and out-of-sample results for bankruptcy prediction models compared to the use of BvD information.

The implication for studies that utilize bankruptcy information is huge. It is likely that previous studies that used incorrect bankruptcy information provided by BvD or Compustat present biased parameters for factors that are supposed to drive a company’s financial condition. As a result, the out-of-sample assessment based on these biased parameters is not informative. For example, BvD information would recommend a model whose performance deteriorates when more accurate HL data is used. Therefore, the conclusions drawn in previous studies may need to be reviewed in light of accurate bankruptcy data. In addition, accurate bankruptcy information is crucial to several other applications, such as, analyzing systemic risks or credit spreads.

Further research should compile the bankruptcy events of German non-public firms and extend our methodology to extract complete bankruptcy information for other countries if regulatory requirements are fulfilled. Finally, investigating whether data quality affects results in other countries may be of special interest.

Chapter 4

Word Power: Content Analysis in the Presence of Competing Information[24]

4.1 Introduction

In efficient capital markets, investors exploit available information, including quantitative and qualitative data. Most researchers utilize readily quantifiable data, such as accounting numbers, to examine the flow of information in stock markets and the timeliness and scope of market reaction to such information. However, in addition to technical improvement, the bulk of descriptive information that is available today about firms, which is included in corporate disclosures, financial press, research services, or even micro blogs such as Twitter, has created extensive opportunities for analyzing qualitative information. Presently, few studies have explored in detail how investors efficiently incorporate descriptive information into prices and convert qualitative information into quantitative measures.

[24] This final chapter is based on my single-authored manuscript, namely "Word Power: Content Analysis in the Presence of Competing Information" as of January 2019. I gratefully acknowledge the helpful comments by Dieter Hess, Martin Huettemann, William Liu as well as seminar participants at the University of Cologne. This paper has also greatly benefitted from insightful comments and suggestions of two anonymous reviewers.

The qualitative content that researchers have analyzed in finance is predominantly found in corporate disclosures such as 10-K filings or earnings releases. In addition to audited financial statements and future prospectuses, Form 10-K provides a comprehensive overview of a company's business. Since regulatory disclosure requirements established by the U.S. Securities and Exchange Commission (SEC) drive Form 10-K filing, the structure is clear and the rules are robust. The SEC has strict guidelines about information that must be included and how it must be organized. Financial statements and presentation of numbers are predefined: for example, all non-GAAP (generally accepted accounting principles) numbers are reconciled to GAAP equivalents. While numbers are clearly defined, written information leaves more scope for interpretation, that is, the provision of descriptive information is provided and how readers' cognitive processes it into individual interpretation of content.

Sentiment analysis attempts to transform qualitative information, mostly textual, into quantitative measures and examines how individuals interpret written information. The research's financial domain has predominantly applied "bag-of-words" methods, such as dictionary-based word classification, to categorize information content as either positive or negative. In contrast to financial economists who used either relative word weights (e.g., Henry, 2008; Li, 2008) or term frequencies (e.g., Loughran and McDonald, 2011) from predefined word lists, Jegadeesh and Wu (2013) use an innovative framework for sentiment analysis, referred to as the "word power" approach. They determine relative word strength based on the market's response to 10-K filings. More importantly, this approach eliminates researchers' subjectivity in classifying words as either positive or negative and allows different strengths for each word. In addition, this approach considers cognitive language processing by interpreting word usage based on the observable reaction of individuals to those words. However, previous studies (e.g., Easton and Zmijewski, 1993) find that investor response to 10-K filings is not significant per se and seems essentially meaningless when information is preempted by previous earnings announcements. This is very important, because Li and Ramesh (2009) document that nearly 95% of 10-K filing events are preceded by preliminary earnings announcements. Hence, this study evaluates if market reactions to previous earnings announcements reveal relevant information to quantify the qualitative information in financial disclosures effectively, namely 10-K filings.

This study contributes to the existing literature in several ways. Clarifying the performance of the "word power" approach of Jegadeesh and Wu (2013), a new method for content analysis, redefines the interpretation of common dictionary-based approaches. In addition, I contrast the information in earnings announcement returns and stock price reactions during subsequent 10-K filings. Based on a Bayesian learning model, I show that market reactions to 10-K filings are not informative enough to quantify the document tone. More importantly, I find that filing returns are determined by the timing of previous earnings announcements. Furthermore, other factors such as the quarterly rebalancing of a portfolio of mutual funds produce noise in observable market reactions to 10-Ks. Besides, filing returns are (1) on average only half of the reactions observed to previous earnings announcements, and (2) not correlated to their direction and, hence, yield erroneous measures of tone. On the other hand, investor responses to earnings announcements are substantially stronger. Notably, using primary investor responses from earnings announcements yields reliable measures of tone and a better interpretation of the relative strengths of words used in financial contexts. In fact, these measures of tone provide a better indication of changes in futures return volatility and predict future stock returns.

My approach is related to two major strands in the literature on finance and accounting, namely the literature on investor responses to information events and analysis of concepts to quantify sentiment from a textual content. First, the extant literature investigates if information events trigger commutated investor responses. The theoretical framework for empirical tests on investor responses is developed from early Bayesian learning models, which constitute that the quality or informativeness of events is conditional pre-disclosure information (e.g., Holthausen and Verrecchia, 1988; Kim and Verrecchia, 1991). Besides the large body of literature that examines only earnings announcements or 10-K filings (e.g., Landsman and Maydew, 2002; Beaver, 1968), Francis et al. (2002) use the implications from Bayesian learning to investigate if competing information from analysts' reports reduces the usefulness of earnings announcements. This study tests a similar concept of retrieving competing information but uses the timing of earnings announcements against 10-K filings that contain virtually identical quantitative information. In the same context of earnings announcements and 10-K filing dates, Easton and Zmijewski (1993) document that reactions to 10-Ks are insignificant results conditional on preliminary announcements. Li and Ramesh (2009) also consider the framework where earnings announcements reduce investor responses to 10-K and 10-Q filings. They support the notion that observed reactions to 10-K filings are noisy and relate to other patterns such as concurrent adjustments to a portfolio of mutual funds at the end of calendar quarters. In addition, they show that most firms release earnings information earlier during announcements and conference calls.

The second strand of literature relates to content analysis and text mining as emerging disciplines of science with applications in various areas. Tetlock, Saar-Tsechansky, and Macskassy (2008) examine market reactions to gauge the sentiment behind news article. In addition, Price, Doran, Peterson, and Bliss (2012) and Demers and Vega (2009) analyze the tone of earnings press releases and earnings conference calls. Feldman, Govindaraj, Livnat, and Segal (2010) and Loughran and McDonald (2011) examine the tone of 10-K reports filed with the SEC and their relationship with stock returns. Several studies consider the evaluation of word lists to trigger improved classification of financial content. Several wordlists emerged from this research; most notable are the General Inquirer (GI), Harvard IV-4 Psychological Dictionary, DICTION dictionary as well as Henry's (2008) and LM word lists developed from Loughran McDonald (2011). Since the latter was developed to analyze content in the context of particular information in 10-Ks, I apply the LM word list aggregating different word reflections into groups similar to Jegadeesh and Wu (2013). However, my approach can be easily implemented using other dictionaries. Further studies on text mining models and methods in finance can be found in Kearney and Liu (2014) and Loughran and McDonald (2016).

The remainder of the paper is organized as follows: Section 4.2 describes the content analysis methods utilized in the study and contrasts them with those used by Jegadeesh and Wu (2013). Section 4.3 explains the theoretical framework to analyze the composition of earning announcements and 10-K filings using a Bayesian learning model and derives appropriate hypotheses. Section 4.4 describes data retrieval and data processing as well as implementation. Section 4.5 reports the empirical results. Finally, Section 4.6 concludes and offers guidance for future research.

4.2 Approach of content analysis

This section introduces the approach of content analysis using readers' reaction, namely shareholders and investors, to corresponding textual information available in 10-K filings. Jegadeesh and Wu (2013) determine the relative strength of words based on observable investor responses to the corresponding documents, or 10-K filings. Notably, this technique allows quantifying textual tone based on cognitive processing of information, that is, word usage, into stock prices and individual investors' expectations. In addition, this approach eliminates researchers' subjectivity classification for each word (i.e., positive or negative) and bias from different word lists or dictionaries. In addition, it allows gauging the different sentiment connotations of words. For example, "defaults" or "bankruptcies" appear to be more negative in financial terms than "disagreements," even though they appear equally often in 10-K documents. This method is different from prior approaches, which use either simple word proportions with equal weight for each word in the lexicon or term weighting schemes based on the relative frequency of a word in a document (e.g., Loughran and McDonald, 2011).

Under the bag-of-words framework, Jegadeesh and Wu (2013) construct their "word power" scores (WP) for each document from the following functional form:

$$WP_i = \sum_{j=1}^{J} (w_j F_{i,j}) \frac{1}{a_i} \tag{4.1}$$

where w_j is the weight for word j, $F_{i,j}$ is the number of occurrences of word j in document i, and a_i is the total number of words in document i. The latter can be computed by counting words in a document and determining the frequency of appearance for each word (and its inflections). Thus, a_i is the length (in terms of words) of document i, while $F_{i,j}$ is the number of appearances of a single word in the text. Both terms express the relative frequency of a word in a document, that is, words appear more often if documents are long. In contrast, the strength of each word w_j and, therefore, the relevance is strongly subjective and can vary with the context and even individuals or cultures. Hence, this term needs to be computed using the explicit perception of individuals to certain words and, thus, an observable quantifiable measure that reflects the immediate reaction to new information.

Broadly, this quantitative sentiment score satisfies the following properties: (1) it is inversely proportion to document length, (2) it is positively related to the strength of the word, and (3) it is positively related to word frequency. Since word frequencies and document length can be computed directly by screening a document, only the relative strength of words w_j needs to be derived using Bayesian learning methods. For this, Jegadeesh and Wu (2013) use a regression-based approach under the assumption that document scores should be correlated with immediate market reactions, such as stock returns to 10-K filings:

$$r_i = a + \left(\sum_{j=1}^{J} (b_j F_{i,j}) \frac{1}{a_i}\right) + \varepsilon_i \tag{4.2}$$

This regression yields an estimate of $\hat{b}_j$ and its standard deviation for each word in a predefined dictionary of words that ought to be positively or negatively afflicted. Subsequently, the strength for each word $\hat{w}_j$ is obtained as follows:

$$\hat{w}_j = \frac{\hat{b}_j - \bar{b}}{Std(\hat{b}_j)} \tag{4.3}$$

where $\hat{w}_j$ is the estimate of w_j, $\hat{b}_j$ is the slope coefficient estimate from equation (4.2), and $\bar{b}$ is the mean of $\hat{b}_j$ across all words. This step standardizes individual word strengths across all words. Since $\hat{w}_j$ is only a proxy for "true" word strength w_j, its reliability strongly depends on the consistency of correlation between stock returns around 10-K filing dates, and "true" document tone.

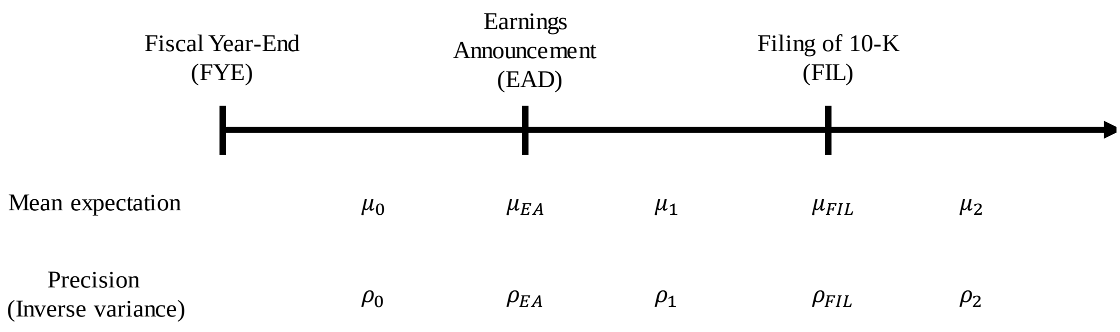
Based on current literature, I analyze the content of 10-Ks as they provide a comprehensive overview of a company's business, audited financial statements, and information. Jegadeesh and Wu (2013) use 10-K filings to analyze the content of annual reports. However, the literature is widely dispersed if 10-K filings trigger any investor response. Given that most firms disclose preliminary earnings results, observable reactions to 10-K filings are not a good proxy for interpreting word characteristics such as sentiment, direction, and strength. This study examines if earnings announcement returns, as the first information issued, may better relate to qualitative information in 10-Ks and if tone measures become more reliable.

4.3 Theoretical framework and hypothesis development

Building on the recent advancements in the emerging field of content analysis and text mining, this section introduces the theoretical framework for this study. This section focuses in detail on the extent to which different information events cause different investor reactions. This is very significant, as market reactions are important to anticipate the cognitive processing of qualitative information in financial disclosures and, thus, quantify a textual tone. In this study, I am interested in investors’ reaction to Form 10-K, specifically if earnings announcements preempt new information in 10-Ks. A wide range of Bayesian learning models are used in the literature on financial markets to model investor responses to events and analyze the information content (e.g., Holhausen and Verrecchia, 1988; Kim and Verrecchia, 1991; Blume, Easley and O’Hara, 1994; Hautsch and Hess, 2007).

Hess and Niessen (2010) developed a specific Bayesian learning model that fits the requirement of sequential arrival and similar information content of two events. However, their context is the emergence of similar macroeconomic indicator releases and the relevance of the first issuance of information to observe related market reaction. This study uses their model to determine the informativeness and investor responses across earnings announcements and 10-K filings. The timeline and setting of the information arrival to investors from both events are illustrated in Figure 4.1.

Figure 4.1 Bayesian learning from earnings announcements (EAD) and 10-K filings (FIL)



Note: This figure illustrates the expectation of market participants in the framework before earnings announcements (*EAD*), between earnings announcements and 10-K filings, and after filing 10-K (*FIL*) returns. Both announcements contain information regarding some economic variable X, that is, the "true" value of a company. According to Hess and Niessen’s (2010) model, assume that market participants form homogenous and normally distributed expectations with respect to X, for example, the mean expectation and its precision (inverse variance).

Earnings announcements are the first piece of definitive information released on the previous fiscal year's operating performance. The filing of Form 10-K, as mandated by the SEC to meet regulatory requirements, follows the earnings announcement. Notably, 10-Ks have very strong regulatory requirements on the content and amount of information that firms must provide. Both events, earnings announcements and 10-K filings, provide an interesting setting for analyzing the informativeness of similar content through different disclosures.

Hess and Niessen (2010) assume a linear relationship between asset prices P and traders' expectations μ with respect to X, or the "true" value of the company. The linear relationship between asset prices P and mean μ is constructed with sensitivity ν:

$$P = \begin{cases} \nu \cdot \mu_0 & \textit{before earnings announcement} \\ \nu \cdot \mu_1 & \textit{after earnings announcement} \\ \nu \cdot \mu_2 & \textit{after 10-K filing} \end{cases} \tag{4.4}$$

Given the two informational events, expectations of interest are as follows: (1) μ_0 is the expectation before earnings announcement, (2) μ_1 is the expectation after earnings announcement but before 10-K filings, and (3) μ_2 is expectation after market participants process both informational events. The expectation at each time $t = 1, 2$ is the weighted average of the prior expectation and new information from the announcement scaled by their information precision ρ, the inverse variance of expectation. Hence, the mean expectation after the first earnings announcement (*EA*) is as follows:

$$\mu_1 = \mu_0 \frac{\rho_0}{\rho_0+\rho_{EA}} + \mu_{EA} \frac{\rho_{EA}}{\rho_0+\rho_{EA}} \tag{4.5}$$

In this equation, μ_{EA} is the mean expectation and ρ_{EA} is the information precision that stems from the earnings announcement. Likewise, μ_0 is the mean expectation and ρ_0 is the information precision that preexists before market participants receive the information from earnings announcements. In the next step, the mean expectation after the 10-K filings (*FIL*), or the second informational event, is as follows:

$$\mu_2 = \mu_1 \frac{\rho_1}{\rho_1+\rho_{FIL}} + \mu_{EA} \frac{\rho_{FIL}}{\rho_1+\rho_{FIL}} \tag{4.6}$$

The notation μ_{FIL} is the mean expectation and ρ_{FIL} is the information precision that stems from the 10-K filing. Likewise, μ_1 is the mean expectation and ρ_1 is the information precision that is formed by market participants in response to the previous earnings announcement and existing information. Given these formulas, changes in stock prices or investor responses are defined as:[25]

$$\Delta P_{EAD} = \nu \cdot (\mu_{EA} - \mu_0) \cdot \frac{\rho_{EA}}{\rho_0} \tag{4.7}$$

$$\Delta P_{FIL} = \nu \cdot (\mu_{FIL} - \mu_1) \cdot \frac{\rho_{FIL}}{\rho_1} \tag{4.8}$$

Broadly, market reactions can be observed across two instances. First, changes in the mean expectations across all investors can move stock prices. Second, convergence of the variance in expectations across investors, referred to hereafter as the precision of expectations, can cause market reactions. Under the assumption that mean expectations from earnings, announcements and 10-Ks are similar ($\mu_{FIL} = \mu_{EA}$), a price reaction at 10-K filing can only be explained by higher information precision ρ_{FIL}.

This assumption is valid, since both events refer exactly to the same content, that is, financial results for the previous fiscal year and, therefore, identical in operating results. Even the use of non-GAAP statements such as those used in earnings announcements explicitly requires reconciliation of statements to the corresponding GAAP values. However, reporting inconsistencies between earnings announcements and 10-K filings verges on defrauding and misleading investors and, hence, leads to major enforcement actions by the SEC. In addition, there are numerous examples where irregularities in financial disclosures aids hedge funds to increase pressure on firms' shares or leads to major accounting scandals. Hence, it is very unlikely that firms actively state differences across both statements.

Given this Bayesian learning model, market reactions to 10-K filings are persistently small, compared to those after the first earnings announcement date (*EAD*), and depend on the precision in information. As 10-Ks provide detailed information on the corresponding balance sheet and income statements, investor responses to 10-Ks imply that previous earnings announcements have been less precise. Thus, the quality of

[25] A systematic derivation of those formulas is extensively described in Hess and Niessen (2010).

informativeness in earnings announcements is inversely related to investors' reliance on subsequent 10-K filings.

In this context, it seems likely that earnings announcement returns should provide incremental information beyond that provided by smaller returns around 10-K filing dates. Building on these notions, I test the following hypotheses for market reactions to both events:

Hypothesis A1: *Investor responses to earnings announcements are substantially larger than observable market reactions around 10-K filing dates.*

Hypothesis A2: *Larger market reactions to 10-Ks are noisy because they are filed directly after earnings announcements.*

Hypothesis A3: *Even though earnings announcements predominantly receive investors' attention, market reactions to 10-K filings are significantly different from zero due to higher information precision.*

To test Hypothesis A1, I analyze the signed and unsigned returns around both events. As news can be either positive or negative, I expect signed excess returns are not significantly different from zero across all firms. In contrast, unsigned (absolute) returns around both events are significantly different from zero. However, earnings announcement returns (*EADRet*) are substantially larger than 10-K filing returns (*FILRet*). Overall, market reactions can be noisy and attributable to several other reasons than 10-K filings. For example, Li and Ramesh (2009) note that volatility and trading volumes are large at the end of the calendar quarter because fund managers adjust their portfolios before quarterly closing. However, market reactions to 10-Ks are observable, if firms file their 10-Ks directly after the earnings announcements. Consequently, I test Hypothesis A2 by analyzing market reactions across a subsample, where 10-Ks are filed several days after the earnings announcements. In addition, to support hypothesis A3, I test if market reactions to 10-Ks are persistently observable after correcting for other factors such as the timing of filings and other control variables.

Whereas analysis of market reactions identifies the relevance of both events to investors, the following hypotheses explicitly address the informativeness of returns around 10-K filings and earnings announcements in the context of the emerging discipline in content analysis:

Hypothesis B1: *Given that large returns are associated with larger appearance of positive and negative words in financial documents, reliable tone measures are unbiased and should correlate with recurrent use of positive and negative words.*

Hypothesis B2: *Since earnings announcement returns (EADRet) are a better proxy for true reaction to annual financial information, EADRet yields better estimates of the relative strengths of words in financial contexts.*

To test hypotheses B1 and B2, I parse each document into the relevant word frequencies and compute a tone measure based on earnings announcement returns and 10-K filing returns. As the first step, I analyze the tone measures on their properties, that is, biases; in the second step, I control to see if these tone measures coincide with the larger appearance of positive and negative word counts. In particular, I expect negative investor reactions to result from a large number of negatively afflicted words in a document and, thus, the tone measure to be also negative. On the other hand, positive surprises within financial disclosures are reflected by positive market reactions and, likewise, a positive tone measure. A good estimate of a document's tone should reflect these properties. Given that earnings announcement returns are large, it seems likely that their tone measure better quantifies the qualitative information in the financial disclosure, that is, 10-Ks.

Furthermore, using the information in earnings announcement returns (*EADRet*), beyond that in 10-K filing returns (*FILRet*), not only alters the explicit tone measures but also yields quantitative measures that substantially explain future performances of a company. Therefore, I test the following hypothesis:

Hypothesis B3: *Tone measures based on earnings announcement returns (EADRet) explain ex-ante the future stock performance and operating business of a company.*

If a tone measure accounts for the true qualitative information within financial texts, it should describe the firm's future performance more effectively. I test hypothesis

B3 using multivariate regressions to explain the first and second moments of future excess returns and a firm's future profitability.

4.4 Data and implementation

4.4.1 Processing 10-K documents from EDGAR

After downloading all the 10-K documents from the SEC EDGAR database, the raw text files are processed using Hering's (2016) parsing procedure. First, I remove all the graphics, MS Excel, and PDF files, and XBRL instances from the document. Next, SEC header information and document type information is removed. All the tables and exhibits in the 10-K documents are excluded. The removal of HTML tags and attributes yields cleaned 10-Ks that include only the main section of the document. In the next step, I removed encoded characters from the document to process only proper words for the subsequent analysis. To compute vectors of word frequencies and document length, each document is stripped into individual tokens words. These tokens are matched computationally with the positive or negative words in the dictionary. I do not count positive or negative words that are preceded by a negator within a distance of three words.[26] In this study, I use the updated negative and positive word list constructed by Loughran and McDonald (2011) containing 355 positive and 2,355 negative words.[27] This list considers different inflections as separate words. Under the assumption that different inflections of root words (such as *"loss," "losses,"* …) are attributed to the same word strength, I follow Jegadeesh and Wu (2013) and merge individual words and inflections in this dictionary into groups of the same root words. This yields a dictionary and, hereafter, the JW dictionary reduces the list to 123 positive and 718 negative words.[28]

[26] I use *no*, *never*, and *not* as negators.

[27] The LM dictionary data is obtained from the LoughranMcDonald_MasterDictionary_2014.csv downloaded at https://www3.nd.edu/~mcdonald/Word_Lists.html.

[28] The JW dictionary is downloadable from Andrew Di Wu's website at https://sites.google.com/site/diandrewwu/data. This list contains identical number of words as the LM dictionary, that is, 355 positive and 2,355 negative words, but classified into 123 positive and 718 negative groups of the identical root word.

Table 4.1 Sample construction

Step	Description	N	Δ
1	Complete sample of EDGAR 10-Ks filed 1994-2016 (excluding duplicates)	159,337	
2	With match in WRDS merging tables	97,780	- 61,557
3	With available CRSP stock market data	86,758	- 11,022
4	Non-financial firms	67,119	- 19,639
5	Price on FYE at least $3	52,125	- 14,994
6	Share code 10 or 11	52,121	- 4
7	NYSE, AMEX or NASDAQ exchange listings	52,084	- 37
8	Compustat book equity > 0	50,742	- 1,342
9	Number of words in Form 10-K larger than 1,000	50,563	- 179
10	With dependent variables (event returns)	50,102	- 461
11	With control variables (size, volatility, turnover, BM, accruals, …)	49,744	- 358
12	With computable word-power scores	48,869	- 875
	Firm-year sample	**48,869**	
	Number of unique firms	**7,321**	
	Average number of years per firm	**7**	

Note: This table reports the process of sample construction starting with all 10-K documents in SEC's EDGAR database from 1994 to 2016. The SEC EDGAR includes 168,716 10-K filings after deletion of duplicates in accession numbers. Similar to Loughran and McDonald (2011), I keep only the first filings in a given year, following the previous filings by at least 180 days because most information in the 10-K may be revealed in the first filing. Using the WRDS linking table, I match 10-K filings from ordinary firms to corresponding accounting information from Compustat through a companies' GVKEY and stock market data from CRSP through its PERMNO. Several 10-Ks with missing GVKEY or PERMNO are based on asset-backed partnerships, real estate, and nonoperating firms that are required to report to the SEC. Moreover, I exclude financial firms (Standard Industrial Classification code from 6000 through 6999) because connotations of words could be different across financial and non-financial firms. The stock price at the fiscal year-end must be at least $1.00 to reduce the effect of bid-ask-bounces in the market responses. In addition, all securities must be ordinary shares (Share code 10 or 11) and not ADRs (American depositary receipts) or other instruments and must be listed on regulated exchanges, such as NYSE, AMEX, or NASDAQ. Finally, I require all regression and dependent variables to be available for firm-year observations to analyze tone scores, determinants, and market responses.

4.4.2 Sample construction

The sample construction starts with downloading all the raw 10-K documents from the SEC EDGAR database that were filed between 1994 and 2016. Table 4.1 documents the sample selection process systematically.

After the parsing process, I exclude some 10-Ks with multiple filings in the same calendar year: for example, secondary 10-K filings due to changes in fiscal years and multiple filings that appear within six months to account for fiscal year changes across calendar years. In the next step, I use the WRDS linking table to match CIK from EDGAR to GVKEY from Compustat and PERMNO from CRSP. With this step, I exclude many firms that are required to file with the SEC, but they are essentially real estate, nonoperating, or asset-backed partnerships/trusts that are not ordinary securities within CRSP. This yields an appropriate database of operative companies regularly filing 10-Ks to disclose their annual financials. In further steps, the sample construction again follows the criteria of Jegadeesh and Wu (2013). All financial firms (Standard Industrial Classification codes from 6000 through 6999) are excluded because words have different connotations, for example, risk, or casualty between financial and non-financial firms. Furthermore, I also require firms to be listed on the NYSE, Amex, or NASDAQ, with an ordinary share class (share code of 10 or 11) and eliminate the role of bid-ask bounces by requiring a stock price of at least $1. As noted by Loughran and McDonald (2011), some 10-K documents include only references to other filings within EDGAR. Thus, I only consider 10-K documents that include more than 1,000 words.

Lastly, I require data to be available for regression variables, such as excess returns in the announcement windows and filing periods, control variables used in the regressions (e.g., size, book-to-market (BM), leverage, turnover, volatility and accruals), and estimated lagged word weights to compute word scores using Jegadeesh and Wu's (2013) approach. As Table 4.1 shows, these sample selection criteria yield a sample of 50,229 firm-year observations consisting of 7,321 unique firms.

4.4.3 Announcement and filing returns

My primary tests examine stock returns relative to earnings announcement and 10-K filing dates. I use the Bayesian learning model from Section 4.3 to contrast both event-returns and draw conclusions about the informativeness of both disclosures using content analysis. Griffin (2003) notes that the event period of stock returns relating to 10-K documents is four days, that is, +0 to +3 days after the 10-K filing date. This is due to a short lag between when the filing is received and when it is publicly posted. To facilitate comparability, I compute stock returns as in the previous studies:

$$r_i = \prod_{t=0}^{3} ret_{i,t} - \prod_{t=0}^{3} ret_{vwi,t} \qquad (4.9)$$

The excess return r_i is the product of the return on stock i, or $ret_{i,t}$, minus the product of the return on the CRSP value-weighted index vwi or $ret_{vwi,t}$, from a certain event date t to $t+3$. The filing return, hereafter *FILRet*, is the excess return within four days (+0 to +3) after the 10-K filing date (*FIL*). The earnings announcement return, hereafter *EADRet*, is the excess return within four days (+0 to +3) after the earnings announcement date (*EAD*). A majority of the literature finds that return volatility during these event windows is larger than in normal periods (see, e.g., Griffin (2003) for the informativeness of 10-Ks and Francis et al. (2002) on the usefulness of earnings announcements).

4.4.4 Control variables

In the empirical analysis, the following control variables are inserted in the regressions. To account for different patterns across firm characteristics, size, book-to-market (BM), leverage, accruals, turnover, and volatility are used. These variables are similar to those used in related studies (e.g., Tetlock et al., 2008; Jegadeesh and Wu, 2013; Loughran and McDonald, 2011). The detailed definition for each variable is as follows:

Size is the natural logarithm of market capitalization at the end of the fiscal year, that is, share price (PRCC_F) times the number of shares (CSHPRI). Further, *BM* is the book-to-market ratio computed as shareholders' equity (SEQ) divided by market capitalization at the end of the fiscal year. Leverage is the sum of short-term (DLC) and long-term liabilities (DLTI) divided by total assets (AT). *Accruals* are calculated using the cash flow method, in other words, income before extraordinary items (IBC) minus cash flow from operating activities (OANCF) divided by total assets (AT) at the end of the fiscal year. *Turnover* is the natural logarithm of the number of shares traded in the period of 250 trading days prior to the earnings announcement date divided by the number of total outstanding shares (CSHPRI). *Volatility* is the standard deviation of the firm-specific component of returns estimated, such as excess return, using the previous 60 months of data as of the end of the month and before the filing date and earnings announcements.

4.5 Empirical results

4.5.1 Summary statistics

Table 4.2 provides summary statistics for the accounting and stock market variables required for sentiment calculation and empirical tests. Overall, I collected a large sample of 50,229 firm-years from 1994 to 2016, including 7,321 unique firms. In Panel A, the statistics of earnings announcement returns (*EADRet*) and returns within the filing window of 10-Ks (*FILRet*) are reported. Both measures of investor responses are used to quantify the sentiment of 10-K documents. *EADRet* and *FILRet* are distributed around zero, as news can be either positive or negative because both median and mean

Table 4.2 Sample summary statistics

Panel A: Dependent Variables

Dependent Variables	Mean	St.Dev.	1%	10%	25%	50%	75%	90%	99%
FILRet	0.001	0.072	-0.195	-0.067	-0.030	-0.001	0.028	0.069	0.220
EADRet	0.004	0.104	-0.257	-0.107	-0.049	-0.000	0.051	0.114	0.328

Panel B: Control Variables

Control Variables	Mean	St.Dev.	1%	10%	25%	50%	75%	90%	99%
Size	6.122	1.798	2.492	3.911	4.856	5.997	7.247	8.521	10.898
BM	0.593	0.482	0.054	0.172	0.288	0.485	0.764	1.115	2.365
Leverage	0.181	0.176	0.000	0.000	0.014	0.145	0.298	0.429	0.670
Turnover	0.200	0.941	-2.275	-1.039	-0.351	0.270	0.827	1.327	2.198
Volatility	0.139	0.072	0.050	0.070	0.090	0.124	0.168	0.222	0.386
Accruals	-0.057	0.139	-0.448	-0.148	-0.089	-0.047	-0.010	0.035	0.184

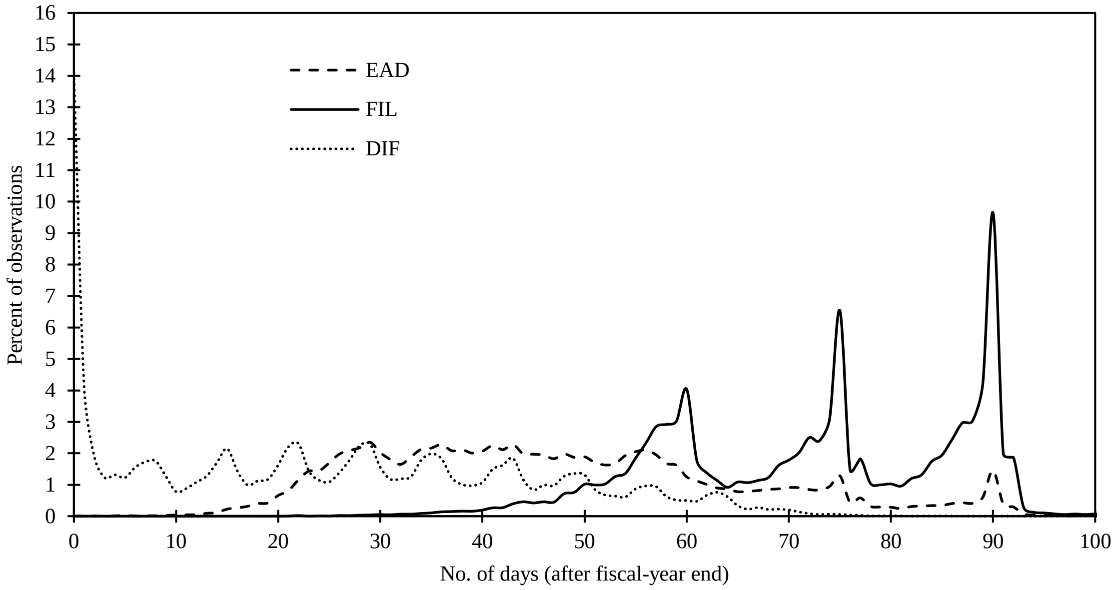
Note: This table reports descriptive statistics of dependent and control variables for the sample of firm-years from 1994 to 2016. Variables are defined in Section 4.4 of the main text. Panel A shows summary statistics for variables that are used as dependent variables in word power regressions and for interpreting document tone. *FILRet* is the market response to 10-Ks in a four-day window [+0, +3] measured as excess return over the CRSP market return. *EADRet* is the market response to earnings announcement dates (*EAD*) and measured consistently to *FILRet*. Panel B shows summary statistics for variables that are used to determine document tone and to control for different patterns across firm characteristics. *Size* is the natural logarithm of market capitalization at the fiscal year-end. *BM* is book-to-market value. *Leverage* is the aggregation of short- and long-term liabilities scaled by total assets. *Turnover* is the natural logarithm of number of shares traded in the previous period of 250 trading days divided by the number of total outstanding shares. *Volatility* is the root mean squared error (RSME) from Fama-French three-factor model using the previous 250 trading days. *Accruals* are computed using the cash flow method.

are close to zero. However, both distributions are skewed. Indeed, negative reactions are generally smaller; for example, 1% percentile of *EADRet* is -25.7% whereas 99% percentile is 32.8%. Panel B shows the descriptive information for control variables that may determine the tone of the document. In line with previous research, my sample contains firms with negative accruals and different leverage, turnover, and volatility. According to firm size, the sample also covers a broad range of small and large firms. Overall, Table 4.2 shows that the sample data are comparable to those used by Loughran and McDonald (2011) or Jegadeesh and Wu (2013).

4.5.2 Timing of Form 10-K filing dates (FIL) and earnings announcement dates (EAD)

In this analysis, I examine the timing of earnings announcement dates (*EAD*) and 10-K filing dates (*FIL*). It is important to validate the assumption of the Bayesian learning model that earnings announcements are the first announcements to disclose information of the prior fiscal year's operating results. The results are illustrated in Figure 4.2.

Figure 4.2 Timing of 10-K filings (FIL) and earnings announcement dates (EAD)



Note: This figure illustrates when companies disclose annual results in earnings announcements and when they file their statutory 10-K document with the SEC after the fiscal year-end by showing the percent of observations in the sample for each day. The solid line is the time difference between fiscal year-end (FYE) and the date on which companies file their 10-K Forms in SEC's EDGAR system (*FIL*). The dashed line is the difference in days between fiscal year-end (FYE) and the date on which a firm discloses its earnings announcement (*EAD*). The dotted line is the difference between 10-K filing dates (*FIL*) and earnings announcement dates (*EAD*).

The SEC requires companies to file Form 10-K within 60 days, 75 days, or 90 days depending on the public float.[29] Figure 4.2 illustrates that the majority of companies do no more than meet the minimum requirements of the SEC regulations. Generally, only very few firms file 10-Ks before the deadline. This casts a doubt on the timeliness of information in the 10-Ks and if the concurrent stock price reactions are informative.

[29] Large accelerated filers are companies with public float of more than $700m with Form 10-K deadline of 60 days. Accelerated filers are companies with public float between $700m and $75m with Form 10-K deadline of 75 days. Non-accelerated filers are companies with public float below $75m with Form 10-K deadline of 90 days.

Earnings announcements appear substantially earlier. Specifically, 50% of the firms had already disclosed earnings announcement within the first 45 days after the end of the fiscal year. The median time lag between both events is 22 days. Events with the same date are firms that filed 10-Ks synonymously with the earnings announcement or forewent a separate earnings announcement. The assumptions of the Bayesian learning model are generally fulfilled. Hence, investor responses are triggered particularly at the time of earnings announcements due to a shift in mean expectations. Given that both events provide similar information, reactions to 10-K filing is rational if the content is significantly more detailed and precise (see hypothesis A3). Even in this case, extracting word strengths from filing reactions are not based on the changes in operating results and conditions but only the details of a company’s earnings announcements.

4.5.3 Determinants of market responses to 10-Ks and earnings announcements

In this analysis, I elaborate on the hypothesis that earnings announcements trigger investor reactions, while price movements from 10-Ks are generally small or explainable by recent earnings announcements. Table 4.3 presents the univariate statistics of announcement and filing returns across the full sample and a subsample where earnings announcements are released at least four days before Form 10-K is filed.

Panel A of Table 4.3 presents the time-series statistics of signed and unsigned event returns and compares them. The results in Table 4.3 provide evidence in favor of hypothesis A1, that is, the differences between signed stock returns in both event windows are small across the two samples. In fact, positive and negative reactions offset each other which yields expected returns of zero on average. In contrast, unsigned reactions are significantly large for earnings announcements, as formulated in hypothesis A1. For example, the median four-day price shift is 5.01% for earnings announcements and only 2.93% for 10-K filings in the full sample. For the subsample, the median unsigned reaction to 10-K filings reduces to 2.60% for the four-day window. Hence, a portion of market responses that researchers have observed from 10-K filings comes from the earlier reaction to preceding earnings announcements. This is supported by Panel B of Table 4.3, which shows that the correlation between filing and announcement returns is generally large, but nonexistent if the lag between both events is sufficiently large.

Table 4.3 Univariate analysis – Market reactions to competing information

Panel A: Univariate analysis

	Full sample				Subsample excl. concurrent EA and Filings (N=37,455)			
	MEAN		MEDIAN		MEAN		MEDIAN	
FILRet	0.0007		-0.0010	*	0.0009		-0.0007	
EADRet	0.0038	***	-0.0002		0.0045	***	0.0006	
Difference	-0.0031	***	-0.0008		-0.0037	***	-0.0013	
|FILRet|	0.0457	***	0.0293	***	0.0376	***	0.0260	***
|EADRet|	0.0714	***	0.0501	***	0.0704	***	0.0502	***
Difference	-0.0258	***	-0.0208	***	-0.0328	***	-0.0242	***

Panel B: Correlations

	Full sample		Subsample excl. concurrent EA and Filings (N=37,455)	
	FILRet	EADRet	FILRet	EADRet
	Spearman	*Pearson*	*Spearman*	*Pearson*
FILRet	-	0.3244***	-	0.0180**
EADRet	0.2452***	-	0.0188**	-

Note: This table reports the time series averages and medians of returns in the filing period and earnings announcement window for the full sample (N=50,229) and the subsample where both windows are not overlapping from 1994 to 2016. The subsample that excludes noncurrent earnings announcements and 10-K filings consists only of firm-years where 10-Ks are filed at least four days after the preceding earnings announcement. *FILRet* is the market response to 10-Ks in a four-day window [+0, +3] measured as excess return over the CRSP market return. *EADRet* is the market response to earnings announcement dates (*EAD*) and measured consistently to *FILRet*. The reported values are time-series statistics for each quarter with corrected Newey-West errors and t-statistics. Panel A shows the differences between market responses to 10-K filings and earnings announcements in mean and median for excess returns and absolute excess returns, respectively. Panel B shows the correlations between the excess returns in both windows for the full sample and the subsample where both windows do not consist similar return data.

Table 4.4 shows the multivariate analysis of 10-K filings; *FILRet* is relative to earnings announcement returns and the timing of both disclosures. Specifically, I use three variables, earnings announcement returns (*EADRet*), *Time* between both disclosures, and *Quarter End*, to explain filing returns to 10-Ks. Li and Ramesh (2009) note that volatility and trading volumes are larger at the end of the calendar quarter because fund managers adjust their portfolios before quarterly closing. Therefore, I include the dummy variable *Quarter End* to account for these price movements.

Table 4.4 Multivariate analysis – Market reactions to competing information

	FILRet				FILRet			
	Full sample				Subsample excl. concurrent EA and Filings (N=37,455)			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
EADRet	0.219***	0.220***			0.007	0.008		
	[10.88]	[10.97]			[1.36]	[1.59]		
Time	-0.002		0.001		-0.001		-0.000	
	[-1.54]		[0.44]		[-0.59]		[-0.16]	
Quarter End	0.007**			0.005*	0.007*			0.006*
	[2.25]			[1.70]	[1.94]			[1.83]
Control variables:								
Size	0.000	0.000	-0.000	-0.000	0.000	-0.000	0.000	0.000
	[0.99]	[0.73]	[-0.24]	[-0.13]	[0.33]	[-0.11]	[-0.05]	[0.07]
BM	-0.001	-0.001	0.001	0.001	-0.000	-0.000	-0.000	0.000
	[-0.54]	[-0.51]	[0.37]	[0.53]	[-0.14]	[-0.10]	[-0.03]	[0.11]
Turnover	-0.002***	-0.002***	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	[-2.64]	[-2.68]	[-1.25]	[-1.22]	[-1.26]	[-1.31]	[-1.30]	[-1.28]
Leverage	-0.001	-0.000	0.001	0.000	0.000	0.001	0.001	0.001
	[-0.34]	[-0.17]	[0.33]	[0.14]	[0.15]	[0.23]	[0.23]	[0.31]
Volatility	-0.004	0.000	-0.008	-0.013	0.002	0.007	0.007	0.001
	[-0.39]	[0.02]	[-0.67]	[-1.04]	[0.21]	[0.56]	[0.59]	[0.10]
Accruals	-0.008	-0.006	-0.009	-0.010	-0.007	-0.005	-0.004	-0.006
	[-1.57]	[-1.12]	[-1.38]	[-1.60]	[-1.20]	[-0.78]	[-0.71]	[-1.06]
R^2	18.7%	17.6%	4.7%	5.2%	7.9%	5.9%	5.9%	6.6%
Adj. R^2	16.2%	15.6%	2.4%	2.9%	3.7%	2.6%	2.6%	3.3%
Obs.	48,869				37,366			

Note: This table reports the results from multivariate regression tests to explain filing returns from earnings announcement returns and control variables for the full sample (N=50,229) and the subsample (N=37.455) where both windows are not overlapping from 1994 to 2016. The subsample that excludes non-current earnings announcements and 10-K filings consists only of firm-years where 10-Ks are filed at least four days after the preceding earnings announcement. *FILRet* is the market response to 10-Ks in a four-day window [+0, +3] measured as excess return over the CRSP value-weighted market return. *EADRet* is the market response to earnings announcement dates (*EAD*), which is measured consistently to *FILRet*. *Time* represents the months between the earnings announcements and the filing date, that is, number of days divided by 30 days per month. This variable accounts for longitudinal earnings announcements drifts in stock returns. *Quarter-End* is a dummy and equals one if a firm filed a 10-K form at the end of the calendar quarter otherwise it is zero. This variable accounts for abnormal returns at the calendar quarter caused by portfolio adjustment and extraordinary trading activities by mutual funds and, hence, controls for excess returns that cannot be attributed to an investor response to 10-Ks. The reported values are time-series statistics for each quarter with corrected Newey-West errors and t-statistics from 87 quarterly Fama-MacBeth regressions. The definition of control variables is described in Section 4.4.

The results in Table 4.4 provide additional empirical evidence in support of hypothesis A2. That is, I find that a large portion of the variation in filing returns can be explained by announcement returns; in other words, the adjusted R^2 increases to 18.7% when *EADRet* is included in the regression. If I exclude observations where *EADRet* and *FILRet* windows overlap, the correlation shrinks. However, the correlation between both event returns is still positive for this subsample, excluding concurrent earnings announcements and 10-K filings. This supports the notion that earnings announcement drifts may affect the measurement of filing returns; in other words, capital markets are slowly processing earnings news into stock prices. Interestingly, *Time* between both events shows insignificant correlation to *FILRet* after correcting for *EADRet*. On average, the *Quarter End* dummy is significant across all regression equations and indicates that *FILRet* is larger by 0.5-0.7% due to the concurrent portfolio adjustments from mutual

fund managers. This table demonstrates that 10-K filing returns are a noisy measure for interpretation of "true" news of operating results. This is very important in content analysis that draws sentiment and relative word strengths from the observable market reaction to their disclosures.

4.5.4 Textual analysis of Form 10-K

In this section, I contrast the results of textual content and sentiment analysis of 10-K documents based on price movements to earnings announcements against the market reactions to the filing of 10-Ks. I denote WP^{FILRet} as the tone score based on 10-K filing returns and WP^{EADRet} as the tone score based on earnings announcement returns. Table 4.5 shows the descriptive statistics of the different measures of tone.

Panel A of Table 4.5 reports the summary statistics of both tone measures, namely WP^{FILRet} and WP^{EADRet}. I find that tone measures based on filing returns are strongly optimistically biased. That is, time-series averages of WP^{FILRet} are generally positive. This is surprising, since I show in the previous section that both *FILRet* and *EADRet* do not show significantly positive or negative bias. Most notably, WP^{FILRet} is positive even for financially disruptive periods such as the financial crisis in 2008. This is surprising and inconsistent given the fact that the number of negative word counts is particularly large during these periods (untabulated). The results support the notion of Hypothesis B1. While WP^{EADRet} fulfills the properties of a good proxy for document tone, WP^{FILRet} is optimistically biased and less correlated to the qualitative information from positive or negative language use. This phenomenon can be explained by the finding that large negative filing returns are not related to larger word counts in 10-Ks, which are the same across WP^{FILRet} and WP^{EADRet}. Hence, extremely negative reactions in the filing windows are noisy and cannot be associated with word counts in 10-Ks, whereas the reactions to earnings announcements create reliable tone measures. Consistent with this, in Panel B, I find little correlation between WP^{FILRet} and WP^{EADRet}. Again, this is striking since both tone measures rest on identical 10-K documents and word counts. Differences arise only from the relationship between word counts and earnings announcement returns, respectively.

Table 4.5 Tone measure based on earnings announcement (EADRet) and filing returns (FILRet)

Panel A: Summary statistics

	Mean	St. Dev.	1%	10%	25%	50%	75%	90%	99%
WP^{FILRet}	0.013	0.010	-0.012	0.001	0.007	0.013	0.018	0.024	0.038
WP^{EADRet}	0.004	0.011	-0.023	-0.009	-0.003	0.004	0.010	0.016	0.031

Panel B: Correlation between WP-Scores

	WP^{FILRet}	WP^{EADRet}
	Spearman	*Pearson*
WP^{FILRet}	-	0.3183
WP^{EADRet}	0.3039	-

Panel C: Determinants of tone measures

	Intercept	Size	BM	Turnover	Leverage	Volatility	Accruals	Adj. R^2
WP^{FILRet}	0.007**	0.000***	0.001***	0.000	0.001	0.007	-0.001	7.00%
	[2.17]	[4.09]	[3.60]	[1.48]	[1.31]	[1.50]	[-0.69]	
WP^{EADRet}	0.002	0.000***	0.001***	0.000	0.001	-0.003	0.002*	6.40%
	[0.39]	[2.79]	[4.38]	[0.30]	[1.50]	[-0.59]	[1.96]	

Note: This table reports descriptive statistics of tone scores using the "word power" approach of Jegadeesh and Wu (2013) with different dependent regression variables to interpret the sentiment from annual reports for the sample of firm-years from 1994 to 2016. WP^{FILRet} is the score from word weights that are estimated based on the filing of excess return (*FILRet*). WP^{EADRet} is the score from word weights that are estimated based on the excess return around the announcement date (*EADRet*). Panel A shows the time-series statistics of the different tone scores. Panel B reports the time-series averages of the Pearson and Spearman correlations across different scores. Panel C tabulates the results from regression of different tone measures on control variables in this study. The reported values are time-series statistics for each quarter with corrected Newey-West errors and t-statistics from 87 quarterly Fama-MacBeth regressions. All the regressions include the control variables used in this study, that is, *Size*, *BM*, *Turnover*, *Leverage*, *Volatility*, and *Accruals*. The definition of control variables is described in Section 4.4.

In Panel C of Table 4.5, I examine the factors that may affect the tone of 10-Ks using quarterly regressions of tone measures and control variables. I consider the following set of factors *Size*, *BM*, *leverage*, *turnover*, *volatility*, and *accruals*. *Size*, *leverage*, and *volatility* proxy for the risk firms face. Risky firms may state potentially more negative consequences of their risk in the statutory 10-K filings than firms that are relatively safe. Growth firms (small *BM*) are valued more for their growth opportunities and, hence, are likely to be more cautious in their 10-Ks. Firms with high stock turnovers attract more investor attention; hence, managements would likely be more cautious in setting investor expectations. Similarly, firms that had recent poor performance such as lower accruals are more likely to use negative wording to justify such performances in 10-Ks. Notably, WP^{FILRet} and WP^{EADRet} are significantly related to *BM* and *Size*. This suggests that large firms disclose positive information, as they are more experienced in communicating information to investors. Interestingly, the intercept in the regression on WP^{FILRet} is significantly positive, which indicates that the tone measure has an optimistic bias across all the firms and could misclassify word counts.

Table 4.6 Comparison of word power weights across different models

	Most impactful positive words		**Most impactful negative words**	
Rank	WP^{FILRet}	WP^{EADRet}	WP^{FILRet}	WP^{EADRet}
1	**revolutionize**	**ingenuity**	imperil	immoral
2	**ingenuity**	**revolutionize**	**dismal**	disgorge
3	regain	**acclaimed**	**disorderly**	**disorderly**
4	informative	unmatched	deface	**vitiate**
5	**excited**	dream	insubordination	dispossess
6	**acclaimed**	beautiful	irreconcilable	**abdication**
7	unparalleled	destined	denigrate	**dismal**
8	valuable	**excited**	**abdication**	unexcused
9	insightful	superior	**vitiate**	derelict
10	conducive	plentiful	condone	inimical

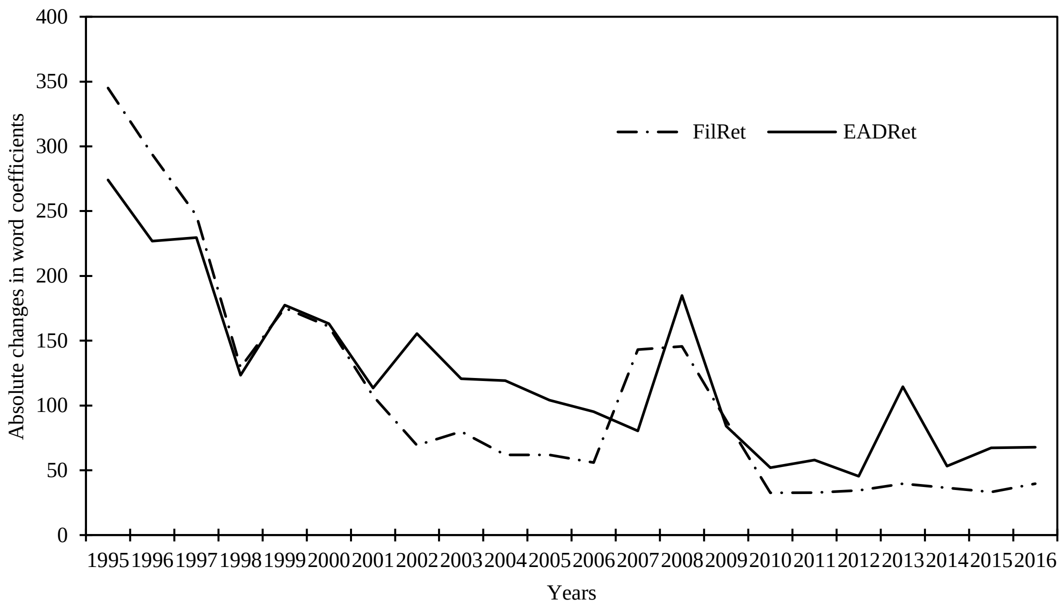
Note: This table reports the ten most impactful positive and negative words used in 10-K filings using the "word power" approach of Jegadeesh and Wu (2013) with different dependent regression variables, in other words, earnings announcements (*EADRet*) or filing returns (*FILRet*), for the sample of firm-years from 1994 to 2016. WP^{FILRet} is the score from word weights that are estimated based on filing excess returns (*FILRet*). WP^{EADRet} is the score from word weights that are estimated based on the excess return around the announcement date (*EADRet*). Words that occur as one of the most impactful words across both scores, WP^{FILRet} and WP^{EADRet}, are indicated in bold.

In the next step, I compare relative word strengths across the different approaches using announcement returns (*EADRet*) and filing returns (*FILRet*). Table 4.6 reports the most impactful positive and negative words in the JW dictionary, i.e., words with the largest coefficients, using the sample from 1994 to 2016.

While very impactful words are somewhat consistent across both approaches, only three positive and four negative words appear among the top ten of both sets of the strongest words. In particular, similar positive words that relate strongly to the positive document tone are *"ingenuity"*, *"revolutionize"*, *"acclaimed"*, and *"excited"* (in addition to their reflections). Likewise, the negative words are *"dismal"*, *"disorderly"*, *"abdication"*, and *"vitiate"* (in addition to their reflections). The results back hypothesis B2 showing that the interpretation of relative word strengths strongly depends on the market reactions that researchers associate with the textual documents. Therefore, using less significant 10-K filing returns may bias the results of language analysis.

I test the robustness of estimated word weights across time. Figure 4.3 illustrates the absolute changes in word weights aggregated for all the words in the JW dictionary for each year.

Figure 4.3 Sensitivity of words strengths from year to year



Note: This figure illustrates the sensitivity of word weights from the "word power" approach of Jegadeesh and Wu (2013) estimated based on earnings announcement returns (*EADRet*) and filing period returns (*FILRet*). I compute absolute changes in the weight of each word in the dictionary from year-to-year. The values shown in Figure 4.3 are the sum of absolute changes across all the words in the JW dictionary, in other words, 123 positive and 718 negative words, for each year from 1995 to 2016. The solid line indicates total absolute changes in word weights estimated based on the values of earnings announcement returns (*EADRet*). The dashed line is total absolute changes in word weights estimated based on the values of filing period returns (*FILRet*).

Similar to Jegadeesh and Wu (2013), I use a recursive window to compute word strengths from regressions based on event returns. Figure 4.3 visualizes that the estimated coefficients have become more robust during recent years, when sufficient data are observed to compute valid word weights. In particular, the approach proposed by Jegadeesh and Wu (2013) requires at least a cross-section of almost ten years to yield valid results on the relative strength of a word's sentiment. After 2005, the changes in word weights are virtually constant. Accordingly, word weights have been somewhat consistent between years. However, the remaining changes in word weights can be explained by changing language use and, hence, changing perception, interpretation, and sensitivity of a word’s context. Obviously, changes in word usage are more pronounced for estimations based on earnings announcement returns because they are generally much larger than for the filing window. Broadly, performance analysis of document tone may consider that computation is very noisy in the earlier years, where data availability, namely the number of 10-K documents and observable market reactions, are weak. I exclude observations before 2005 in the evaluation samples and proceed with the rest in the remainder of the study.

In the next analysis, I explore if the current document tone contains valuable information about a firm’s future performance, namely future excess returns, return volatility, and future profitability. To test this proposition, I compute the market-adjusted excess returns and volatility of stock prices after the 10-K filing and return-on-assets of the next fiscal year. Given that 10-K documents provide meaningful information to investors, the current tone classification should predict future performance. Hence, I expect that a reliable tone measure correlates significantly to future firm performance. Foremost, I elaborate on whether the choice of relevant market reactions, either filing or earnings announcement returns, to compute tone may change the inferences between document tone and future firm performance.

In Table 4.7, I estimate multivariate regressions using indicators for future firm performance as the dependent variables on different tone measures and additional control variables as the explanatory variables. Stock return is the monthly excess return following earnings announcement and filing date adjusted for market returns from CRSP. The relationship between future returns and document tone is expected to be positive, as positive (negative) news and outlook trigger stocks to outperform (underperform) the market portfolio. A positive and significant coefficient indicates that tone measures have significant predictive power for future realized returns. In Panel A of Table 4.7, I find that tone that is based on earnings announcements returns (WP^{EADRet}) correlates significantly to future excess returns. In particular, the coefficient is 0.464 in a regression including WP^{FILRet} as a tone measure and 0.385 including only WP^{EADRet} with significant t-statistics. In contrast, coefficients of WP^{FILRet} are economically small and insignificant in both regressions. In support of hypothesis B3, a tone measure based on filing returns conveys no meaningful information to investors about future firm performance. Overall, this finding is consistent with Demers and Vega (2010) and Feldman et al. (2010) that tone contains valuable information for predicting future excess returns. However, I show that correlations between tone and future returns are significant only when "true" market reactions at earnings announcements are used with the tone of 10-Ks.

Table 4.7 Textual tone and future performance

Panel A: Multivariate regressions on future excess returns

Dependent Variable:	Excess return		
	(1)	(2)	(3)
WP^{FILRet}	0.025	0.251	
	[0.13]	[1.43]	
WP^{EADRet}	0.464**		0.385**
	[2.29]		[2.41]

Panel B: Multivariate regressions on future stock volatility

Dependent Variable:	Volatility		
	(1)	(2)	(3)
WP^{FILRet}	-0.034*	-0.044**	
	[-1.95]	[-2.08]	
WP^{EADRet}	-0.049**		-0.054**
	[-2.06]		[-2.32]

Panel C: Multivariate regressions on future profitability

Dependent Variable:	Operating results		
	(1)	(2)	(3)
WP^{FILRet}	0.241	0.830	
	[0.46]	[1.55]	
WP^{EADRet}	1.910***		1.641***
	[6.76]		[4.58]

Note: This table reports the results from multivariate regression tests to explain future firm performance from current document tone measures. Firm performance is measured using three instances: future excess return, future return volatility, and future operating results. Stock return is the monthly excess return following earnings announcement and filing date adjusted for market returns from CRSP. Future return volatility is defined as the root mean square error (RSME) from market adjusted returns in the window [+10, +250] relative to the 10-K filing date. Finally, the future operating result is return-on-assets (RoA) of the next fiscal year computed as operating income after depreciation (OIADP) divided by total assets (AT). WP^{FILRet} is the score from word weights that are estimated based on filing excess return (*FILRet*). WP^{EADRet} is the score from word weights that are estimated based on the excess return around the announcement date (*EADRet*). The reported values are time-series statistics for each quarter with corrected Newey-West errors and t-statistics from 48 quarterly Fama-MacBeth regressions from 2005 to 2016. All the regressions include the control variables used in this study: *Size, BM, Turnover, Leverage, Volatility* and *Accruals*. The definition of control variables is described in Section 4.4.

In the next step, I perform an additional test to examine if tone also explains the higher moments of future excess returns, for instance, future return volatility. I define future return volatility as the root mean square error (RSME) from market adjusted returns in the window [+10, +250] relative to the 10-K filing date. Post event volatilities proxy the changes in risk after 10-K disclosures. Consistent with Jegadeesh and Wu (2013), I expect that the 10-K filings of risky firms are more likely to have a negative tone, whereas a positive tone reduces investors' uncertainty. That is, the coefficient of tone should be negative in the regression. In Panel B of Table 4.7, I find that the coefficients of WP^{FILRet} and WP^{EADRet} are negative and, thus, consistent with rational expectations. Overall, future return volatility increases in response to negative 10-Ks computed based on filing returns (*FILRet*) and earnings announcement returns (*EADRet*). In addition, I find that the coefficients of WP^{FILRet} are generally small and less significant. Again, this shows that market reactions around 10-K filings are not a reliable source of information to interpret negative and positive word counts.

Lastly, I also test if tone measures are related to the future operating results of a firm and future accounting information. Future operating results are return-on-assets of the next fiscal year computed as operating income after depreciation (OIADP) divided by total assets (AT). Overall, positive (negative) information increases (decreases) future earnings expectations of a firm and prompts analysts to adjust their earnings forecasts. Given that tone is correctly measured in either WP^{FILRet} or WP^{EADRet}, I expect that both measures relate positively to future profitability. In Panel C of Table 4.7, I find that only WP^{EADRet} yields statistically and economically large coefficients in the regressions. I find those of WP^{FILRet} and WP^{EADRet} are positive but insignificantly related to future profitability. This is reasonable as even 10-Ks with large negative word counts may be assigned to positive returns (as illustrated in Table 4.5) and, thus, a positive document tone. Hence, these results corroborate the findings from the previous analysis that use of filing returns yields tone measures that misclassify textual information.

4.5.5 Robustness tests

Textual analysis is an emerging discipline in financial research. Therefore, various studies use different word lists, documents, and tone measures. To account for other approaches in the extant literature, I performed additional analysis to test the robustness of the results. First, I computed tone based *FILRet* and *EADRet* using alternative word lists. In particular, I used a word list developed by Loughran and McDonald (2011) that consists of either negative or positive words. Other word lists, such as Harvard General Inquirer, DICTION, or Henry (2008), are not specified for textual analysis in a financial context and, thus, are not applicable to the analysis of financial documents such as 10-Ks. In addition, tone measures were created that are strictly computed from word counts in the MD & A section of 10-Ks, Item 7, “Management discussion and analysis." The results are consistent with those from full 10-K filings. However, to be consistent with the major strand of the literature, that is Loughran and McDonald (2011), I use the complete textual information from 10-Ks. This has several advantages. First, the MD & A section is only mandatory for large firms and, thus, reduces the sample size systematically. Second, textual information is also conveyed in other sections of the 10-K. Finally, Loughran and McDonald (2011) have already verified that the usage of either the main text of 10-Ks or the MD & A subsection yields consistent results.

4.6 Conclusion and future research

This study contributes to the emerging field of content analysis and text mining in financial literature. My empirical analysis provides evidence suggesting that prior studies miss valuable implications of the competing effects of market reactions in response to events such as earnings announcements and 10-K filings and, thus, the quality of subsequent tone measures.

First, I find that market reactions to earnings announcements are substantially larger than for 10-K filings. Most importantly, observable market reactions around the 10-K filing dates are noisy and can be explained by the fact that filings directly follow the previous earnings announcement or other factors such that mutual fund investors adjust their portfolios before quarterly closing. This is because earnings announcements preempt most of the information content of 10-Ks and, given the efficient capital markets reaction, the information is already incorporated into stock prices when firms file their 10-Ks. In addition, investor responses to 10-Ks relate to the increasing precision of information and, thus, depend fundamentally on the accuracy of the previous earnings announcements. Second, building on the evidence from tone measures based on earnings announcements and filing returns, the results reveal that the earnings announcements window yields reliable stock market reactions that are consistent with the information retrieved from corporate 10-K filings. Specifically, large positive and negative returns as response to earnings announcement coincide with more frequent appearances of positive and negative words in financial documents. Hence, tone measures are more reliable and unbiased. Furthermore, since earnings announcement returns (*EADRet*) are a better proxy for "true" reactions to annual financial information, *EADRet* yields improved estimates of the relative strength of words in financial contexts. Finally, using *EADRet* not only shifts the interpretation of using financial language in content analysis, their tone measure also provides a better indication of the changes in future return volatility and predict future stock returns more effectively.

The study results have important implications for both investors and academics. First, this study shows that investors predominantly respond to earnings announcements rather than 10-K filings. Hence, interpreting the textual information of financial disclosures, based on observable market reactions, may help to advise financial institutions, companies, and costumers to evaluate qualitative content in financial reports from a broader perspective. Moreover, it may help to obtain better corporate valuation or investment decisions. Second, future research can consider examining if stock markets react in a similar way to other financial disclosures and perform additional content analysis based on observable price movements. Consequently, researchers can apply such content analysis to other public disclosures, namely macroeconomic disclosures (central banks, consumer indices), analysts' reports, or other news releases.

Bibliography

Abarbanell, J., 1991. Do analysts' earnings forecasts incorporate information in prior stock price changes? *Journal of Accounting and Economics* 14, 147-165.

Agarwal, V., Taffler, R., 2008. Comparing the performance of market-based and accounting-based bankruptcy prediction models. *Journal of Banking and Finance* 32, 1541-1551.

Altman, E. I., 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance* 23, 589-609.

Altman, E. I., 1983. Corporate financial distress. New York, Wiley.

Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., Suvas, A., 2017. Financial distress prediction in an international context: a review and empirical analysis of Altman's z-score model. *Journal of International Financial Management and Accounting* 28, 131-171.

Anilowski, C., Feng, M., D., Skinner, D., 2007. Does earnings guidance affect market returns? The nature and information content of aggregate earnings guidance. *Journal of Accounting and Economics* 44, 36-63.

Ashton, D., Wang, P., 2012. Terminal valuations, growth rates and the implied cost of capital. *Review of Accounting Studies* 18, 261-290.

Ball, R., Ghysels, E., 2017. Automated earnings forecasts: Beat analysts or combine and conquer? Working paper, University of Michigan.

Beaver, W., 1968. The information content of annual earnings announcements. *Journal of Accounting Research* 6, 67-92.

Bharath, S. T., Shumway, T., 2008. Forecasting default with the Merton distance to default model. *The Review of Financial Studies* 21, 1339-1369.

Bluemke, N., Hess, D., Stolz, A., 2017. Predicting sell-side analysts' relative earnings forecast accuracy when it matters most. Working paper, University of Cologne.

Blume, L., Easley, D., O'Hara, M., 1994. Market statistics and technical analysis: The role of volume. *The Journal of Finance* 49, 153-181.

Botosan, C., Plumlee, M., 2005. Assessing alternative proxies for the expected risk premium. *The Accounting Review* 80, 21-53.

Botosan, C., Plumlee, M., Wen, H., 2011. The relation between expected returns, realized returns, and firm risk characteristics. *Contemporary Accounting Research* 28, 1085-1122.

Bradshaw, M., Drake, M., Myers, J., Myers, L., 2012. A re-examination of analysts' superiority over time-series forecasts of annual earnings. *Review of Accounting Studies* 17, 944-968.

Brown, L., Hagerman, R., Griffin, P., Zmijewski, M., 1987. Security analyst superiority relative to univariate time-series models in forecasting quarterly earnings. *Journal of Accounting and Economics* 9, 61-87.

Brown, L., Call, A., Clement, M., Sharp, N., 2015. Inside the "Black Box" of sell-side financial analysts. *Journal of Accounting Research* 53, 1-47.

Campbell, J. Y., Hilscher, J., Szilagyi, J., 2008. In search of distress risk. *The Journal of Finance* 63, 2899-2939.

Chan, L., Karceski, J., Lakonishok, J., 2003. The level and persistence of growth rates. *The Journal of Finance* 58, 643-684.

Chang, W., Monahan, S., Ouazad, A., Vasvari, F., 2014. The higher moments of future earnings. Working paper, INSEAD.

Chava, S., Jarrow, R. A., 2004. Bankruptcy prediction with industry effects. *Review of Finance* 8, 537-569.

Claus, J., Thomas, J., 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.

Dahiya, S., Klapper, L., 2007. Who survives? A cross-country comparison. *Journal of Financial Stability* 3, 261-278.

Demers, E., Vega, C., 2009. Soft information in earnings announcements: News or Noise? Working paper, INSEAD.

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., Monahan, S., 2005. An evaluation of accounting-based measures of expected returns. *The Accounting Review* 80, 501-538.

Easton, P., Sommers, G., 2007. Effect of analysts' optimism on estimates of the expected rate of return implied by earnings forecasts. *Journal of Accounting Research* 45, 983-1015.

Easton, P., Monahan, S., 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.

Easton, P., Zmijewski, M. E., 1993. SEC Form 10K/10Q reports and annual reports to shareholders: Reporting lags and squared market model prediction errors. *Journal of Accounting Research* 43, 113-129.

Fama, E. F., French, K., 1993. Common risk factors in the returns on bonds and stocks. *Journal of Financial Economics* 47, 427-465.

Fama, E., French, K., 2000. Forecasting profitability and earnings. *The Journal of Business* 73, 161-175.

Fama, E., French, K., 2006. Profitability, investment and average returns. *Journal of Financial Economics* 82, 491-518.

Feldman, R., Govindaraj, S., Livnat, J., Segal, B., 2010. Management's tone change, post earnings announcement drift and accruals. *Review of Accounting Studies* 15, 915-953.

Filipe, S. F., Grammatikos, T., Michala, D., 2016. Forecasting distress in European SME portfolios. *Journal of Banking and Finance* 64, 112-135.

Francis, J., Schipper, K., Vincent, L., 2002. Earnings announcements and competing information. *Journal of Accounting and Economics* 33, 313-342.

Gebhardt, W., Lee, C., Swaminathan, B., 2001. Toward an implied cost of capital. *Journal of Accounting Research* 39, 135-176.

Griffin, P. A., 2003. Got information? Investor response to Form 10-K and Form 10-Q EDGAR filings. *Review of Accounting Studies* 8, 433-460.

Gordon, J., Gordon, M., 1997. The finite horizon expected return model. *Financial Analysts Journal* 53, 52-61.

Guay, W., Kothari, S., Shu, S., 2011. Properties of implied cost of capital using analysts' forecasts. *Australian Journal of Management* 36, 125-149.

Gu, Z., Wu, J. S., 2003. Earnings skewness and analyst forecast bias. *Journal of Accounting and Economics* 35, 5-29.

Hautsch, N., Hess, D., 2007. Bayesian learning in financial markets: Testing for the relevance of information precision in price discovery. *Journal of Financial and Quantitative Analysis* 42, 189-208.

Heinrichs, N., Hess, D., Homburg, C., Lorenz, M., Sievers, S., 2013. Extended dividend, cash flow, and residual income valuation models: Accounting for deviations from ideal conditions. *Contemporary Accounting Research* 30, 42-79.

Henry, E., 2008. Are investors influenced by how earnings press releases are written? *The Journal of Business Communication* 45, 363-407.

Hering, J., 2016. The annual report algorithm: Retrieval of financial statements and extraction of textual information. Working paper, University of Erlangen-Nuremberg.

Hess, D., Huettemann, M., 2018. Predicting bankruptcy via cross-sectional earnings forecasts. Working Paper, University of Cologne.

Hess, D., Meuter, M., Kaul, A., 2018. The performance of mechanical earnings forecasts. Working paper, University of Cologne.

Hess, D., Niessen, A., 2010. The early news catches the attention: On the relative price impact of similar economic indicators. *Journal of Financial and Quantitative Analysis* 30, 909-937.

Hillegeist, S. A., Keating, E. K., Cram, D. P., Lundstedt, K. G., 2004. Assessing the probability of bankruptcy. *Review of Accounting Studies* 9, 5-34.

Holthausen, R. W., Verrecchia, R. E., 1988. The effect of sequential information releases on the variance of price changes in an intertemporal multi-asset market. *Journal of Accounting Research* 26, 82-106.

Hou, K., van Dijk, M., 2011. Profitability shocks and the size effect in the cross-section of expected stock returns. Working paper, Ohio State University and Rotterdam School of Management.

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

Jegadeesh, N., Wu, D., 2013. Word power: A new approach for content analysis. *Journal of Financial Economics* 110, 712-729.

Kearney, C., Liu, S., 2014. Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis* 33, 171-185.

Kim, O., Verrecchia, R. E., 1991. Market reaction to anticipated announcements. *Journal of Financial Economics* 30, 273-309.

Konstantinidi, T., Pope, P. F., 2016. Forecasting risk in earnings. *Contemporary Accounting Research* 33, 487-525.

La Porta, R., 1996. Expectations and the cross-section of stock returns. *The Journal of Finance* 51, 1715-1742.

Landsman, W. R., Maydew, E. L., 2002. Has the information content of quarterly earnings announcements declined in the past three decades? *Journal of Accounting Research* 40, 797-808.

Larocque, S., 2013. Analysts' earnings forecast errors and cost of equity capital estimates. *Review of Accounting Studies* 18, 135-166.

Lee, C., Ng, D., Swaminathan, B., 2009. Testing international asset pricing models using implied cost of capital. *Journal of Financial and Quantitative Analysis* 44, 307-335.

Lee, C., So, E., Wang, C., 2015. Evaluating firm-level expected return proxies. Working paper, Stanford University.

Lewellen, J., 2014. The cross-section of expected stock returns. Working paper, Dartmouth College.

Li, F., 2008. Annual report readability, current earnings, and earnings persistence. *Journal of Financial Economics* 45, 221-247.

Li, E. X., Ramesh, K., 2009. Market reaction surrounding the filing of periodic SEC reports. *The Accounting Review* 84, 1171-1208.

Li, K., 2011. How well do investors understand loss persistence? *Review of Accounting Studies* 16, 630-667.

Li, K., Mohanram, P., 2014. Evaluating cross-sectional forecasting models for implied cost of capital. *Review of Accounting Studies* 19, 1152-1185.

Lohmann, C., Ohliger, T., 2017. Nonlinear relationships and their effect on the bankruptcy prediction. *Schmalenbach Business Review* 18, 261-287.

Loughran, T., McDonald, B., 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance* 66, 35-65.

Loughran, T., McDonald, B., 2016. Textual analysis in accounting and finance: A survey. *Journal of Accounting Research* 54, 1187-1230.

Lyle, M. R., Wang, C., 2015. The cross-section of expected holding period returns and their dynamics: A present value approach. *Journal of Financial Economics* 116, 505-525.

Mohanram, P., Gode, D., 2013. Removing predictable analysts forecast errors to improve implied cost of equity estimates. *Review of Accounting Studies* 18, 443-478.

Moody's Analytics (2018, July 20). Retrieved from https://www.moodys.com/sites/products/productattachments/drd_brochure.pdf.

Novy-Marx, R., 2013. The other side of value: The gross profitability premium. *Journal of Financial Economics* 108, 1-28.

O'Brien, P., 1988. Analysts' forecasts as earnings expectations. *Journal of Accounting and Economics* 10, 53-83.

Ohlson, J., Juettner-Nauroth, B., 2005. Expected EPS and EPS growth as determinants of value. *Review of Accounting Studies* 10, 349-365.

Ohlson, J., 1980. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research* 18, 109-131.

Penman, S., Reggiani, F., Richardson, S., Tuna, I., 2015. An accounting-based characteristic model for asset pricing. Working paper, Columbia Business School.

Price, M. S., Doran, J. S., Peterson, D. R., Bliss, B. A., 2012. Earnings conference calls and stock returns: The incremental informativeness of textual tone. *Journal of Banking and Finance* 36, 992-1011.

Reisz, A. S., Perlich, C., 2007. A market-based framework for bankruptcy prediction. *Journal of Financial Stability* 3, 85-131.

Richardson, S., Teoh, S., Wysocki, P., 2004. The walk-down to beatable analyst forecasts: The role of equity issuance and insider trading incentives. *Contemporary Accounting Research* 21, 885-924.

Rusticus, T., 2011. Market inefficiency and implied cost of capital models. Working paper, Northwestern University.

SDC Platinum (2018, July 20). Retrieved from https://financial.thomsonreuters.com/en/products/data-analytics/market-data/sdc-platinum-financial-securities.html.

Sharpe, W., 1994. The Sharpe Ratio. *Journal of Portfolio Management* 21, 49-58.

Shumway, T., 2001. Forecasting bankruptcy more accurately: A simple hazard model. *The Journal of Business* 74, 101-124.

Tetlock, P. C., Saar-Tsechansky, M., Mackassy, S., 2008. More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance* 63, 1437-1467.

Tian, S., Yu, Y., 2017. Financial ratios and bankruptcy predictions: An international evidence. *International Review of Economics and Finance* 51, 510-526.

UCLA-LoPucki Bankruptcy Research Database (2018, July 20). Retrieved from http://lopucki.law.ucla.edu/.

U.S. Securities and Exchange Commission (2017, August 28). Form 10-K. Retrieved from https://www.sec.gov/answers/form10k.htm.

Vassalou, M., Xing, Y., 2004. Default risk in equity returns. *The Journal of Finance* 59, 831-868.

Wang, C., 2015. Measurement errors of expected-return proxies and the implied cost of capital. Working paper, Harvard Business School.

APPENDICES

Table of Contents

Appendix to Chapter 2

Appendix A: Regression models and prediction equations

Residual income (RI) model (see also Li and Mohanram (2014)):

Standard RI model

Regression equation:

$$A_t = \alpha_0 + \alpha_1 \cdot A_{t-\tau} + \alpha_2 \cdot NED_{t-\tau} + \alpha_3 \cdot IAT_{t-\tau} + \alpha_4 \cdot BV_{t-\tau} + \alpha_5 \cdot ACC_{t-\tau} + \varepsilon_t$$

Prediction equation:

$$\hat{A}_{t+\tau} = \hat{\alpha}_0 + \hat{\alpha}_1 \cdot A_t + \hat{\alpha}_2 \cdot NED_t + \hat{\alpha}_3 \cdot IAT_t + \hat{\alpha}_4 \cdot BV_t + \hat{\alpha}_5 \cdot ACC_t$$

In those formulas, A_t is the annual earnings, NED_t is a negative earnings dummy, IAT_t is the interaction term of the earnings variable and negative earnings dummy, BV_t is book equity and ACC_t is accruals. All variables are computed on per-share basis. A detailed description of the variable definition is provided in Appendix D. We run this regression model for each month within the sample range 1982 to 2014 for the cross-section of all firms within the rolling ten-year windows (120 months) prior to the estimation date. From the model's coefficients, we compute predictions for future annual earnings per share in fiscal year t+τ, with $\tau = 1, \dots, 5$.

Extended RI model

Regression equation:

$$Q_t^{\{q+1,\dots,4\}} = \alpha_0 \cdot f_q + \sum_{j=0}^{3} QD_j^{(q)} \left[\alpha_{1,n} Q_{t-\tau}^{\{q+1,\dots,4\}} + \alpha_{2,n} \cdot g_q \cdot \Delta Q_{t-\tau+1,t-\tau}^{\{1,\dots,q\}}\right] + \alpha_3 \cdot f_q \cdot NED_{t-\tau} + \alpha_4 \cdot f_q \cdot IAT_{t-\tau} + \alpha_5 \cdot f_q \cdot BV_{t-\tau}$$
$$+ \alpha_6 \cdot f_q \cdot ACC_{t-\tau} + \varepsilon_t$$

Prediction equation:

$$\hat{A}_{t+\tau} = Q_{t+1}^{\{1,\dots,q\}} + \hat{\alpha}_0 \cdot f_q + \sum_{j=0}^{3} QD_j^{(q)} \left[\hat{\alpha}_{1,j} Q_t^{\{q+1,\dots,4\}} + \hat{\alpha}_{2,j} \cdot g_q \cdot \Delta Q_{t+1,t}^{\{1,\dots,q\}}\right] + \hat{\alpha}_3 \cdot f_q \cdot NED_t + \hat{\alpha}_4 \cdot f_q \cdot IAT_t + \hat{\alpha}_5 \cdot f_q \cdot BV_t + \hat{\alpha}_6 \cdot f_q \cdot ACC_t$$

The dummy variable $QD_j^{(q)}$ equals one, if $j = q$, and otherwise zero.

$Q_t^{\{q+1,\dots,4\}}$ are the remaining quarterly earnings results of the previous fiscal year t, while $Q_{t+1}^{\{1,\dots,q\}}$ are quarterly earnings results that are already announced for the fiscal year t+1. NED_t is a negative earnings dummy, IAT_t is the interaction term of the earnings variable and negative earnings dummy, BV_t is book equity and ACC_t is accruals (details on the variable definitions are provided in Appendix D). All variables are on per-share basis. Since we divide our sample into four groups per available quarterly earnings information q, we estimate group-specific regression coefficients. We utilize dummy variables $QD_j^{(q)}$ to estimate the four regression models within a single regression equation. We run this regression model for each month within the sample range 1982 to 2014 for the cross-section of all firms within the rolling ten-year windows (120 months) prior to the estimation date. From the model's coefficients, we compute predictions of remaining quarterly earnings per share in fiscal year t+τ, with $\tau = 1, \dots, 5$. We then aggregate forecasts of annual earnings $\hat{A}_{t+\tau}$ from already realized quarterly earnings results $Q_{t+1}^{\{1,\dots,q\}}$ and predicted remaining quarterly results $\hat{Q}_{t+\tau}^{\{q+1,\dots,4\}}$. To reduce the number of coefficients in our model we interact only the earnings variables with the dummy variable $QD_j^{(q)}$. Other accounting variables, such as $BVPS_t$, are then aligned to the number of remaining quarterly earnings $\hat{Q}_{t+\tau}^{\{q+1,\dots,4\}}$ to estimate a single coefficient. For example, if we predict 100% of the future annual earnings results, we also use 100% of the accounting variables, such as book equity or accruals. However, if quarterly earnings information is released and our regression is used to predict the remaining three quarterly earnings, thus, only 75% of the annual result, we also multiply book equity and accruals by 75%. Hence, the regression model is balanced in proportion of dependent and independent variables. After this adjustment, the coefficients of the additional variables are virtually identical. Therefore, we reduce the number of coefficients and estimate single coefficients for each variable, e.g., book equity, accruals. This yields a more parsimonious model and avoids in-sample overfitting by a large number of coefficients. However, interacting all variables with the dummy variable $QD_j^{(q)}$ does not change the results, but creates a large set of coefficients.

Earnings persistence (EP) model (see also Li and Mohanram (2014)):

Standard EP model

Regression equation:

$$A_t = \alpha_0 + \alpha_1 \cdot A_{t-\tau} + \alpha_2 \cdot NED_{t-\tau} + \alpha_3 \cdot IAT_{t-\tau} + \varepsilon_t$$

Prediction equation:

$$\hat{A}_{t+\tau} = \hat{\alpha}_0 + \hat{\alpha}_1 \cdot A_t + \hat{\alpha}_2 \cdot NED_t + \hat{\alpha}_3 \cdot IAT_t$$

In those formulas, A_t is the annual earnings, NED_t is a negative earnings dummy and IAT_t is the interaction term of the earnings variable and negative earnings dummy. All variables are computed on per-share basis. A detailed description of the variable definition is provided in Appendix D. We run this regression model for each month within the sample range 1982 to 2014 for the cross-section of all firms within the rolling ten-year windows (120 months) prior to the estimation date. From the model's coefficients, we compute predictions for future annual earnings per share in fiscal year t+τ, with $\tau = 1, \ldots, 5$.

Extended EP model

Regression equation:

$$Q_t^{\{q+1 \ldots 4\}} = \alpha_0 \cdot f_q \cdot + \sum_{j=0}^{3} QD_j^{(q)} \left[\alpha_{1,j} Q_{t-\tau}^{\{q+1,\ldots,4\}} + \alpha_{2,j} \cdot g_q \cdot \Delta Q_{t-\tau+1,t-\tau}^{\{1,\ldots,q\}}\right] + \alpha_3 \cdot f_q \cdot NED_{t-\tau} + \alpha_4 \cdot f_q \cdot IAT_{t-\tau} + \varepsilon_t$$

Prediction equation:

$$\hat{A}_{t+\tau} = Q_{t+1}^{\{1,\ldots,q\}} + \hat{\alpha}_0 \cdot f_q \cdot + \sum_{j=0}^{3} QD_j^{(q)} \left[\hat{\alpha}_{1,j} QE_t^{\{q+1,\ldots,4\}} + \hat{\alpha}_{2,j} \cdot g_q \cdot \Delta Q_{t+1,t}^{\{1,\ldots,q\}}\right] + \hat{\alpha}_3 \cdot f_q \cdot NED_t + \hat{\alpha}_4 \cdot f_q \cdot IAT_t$$

The dummy variable $QD_j^{(q)}$ equals one, if $j = q$, and otherwise zero.

$Q_t^{\{q+1,\ldots,4\}}$ are the remaining quarterly earnings results of the previous fiscal year t, while $Q_{t+1}^{\{1,\ldots,q\}}$ are quarterly earnings results that are already announced for the fiscal year t+1. NED_t is a negative earnings dummy and IAT_t is the interaction term of the earnings variable and negative earnings dummy (details on the variable definitions are provided in Appendix D). All variables are on per-share basis. Since we divide our sample into four groups per available quarterly earnings information q, we estimate group-specific regression coefficients. We utilize dummy variables $QD_j^{(q)}$ to estimate the four regression models within a single regression equation. We run this regression model for each month within the sample range 1982 to 2014 for the cross-section of all firms within the rolling ten-year windows (120 months) prior to the estimation date. From the model's coefficients, we compute predictions of remaining quarterly earnings per share in fiscal year t+τ, with $\tau = 1, \ldots, 5$. We then aggregate forecasts of annual earnings $\hat{A}_{t+\tau}$ from already realized quarterly earnings results $Q_{t+1}^{\{1,\dots,q\}}$ and predicted remaining quarterly results $\hat{Q}_{t+\tau}^{\{q+1,\dots,4\}}$. To reduce the number of coefficients in our model we interact only the explicit earnings variables with the dummy variable $QD_j^{(q)}$. For example, if we predict 100% of the future annual earnings results, we also use 100% of the additional variables, such as negative earnings dummy or intercept. However, if quarterly earnings information is released and our regression is used to predict the remaining three quarterly earnings, thus, only 75% of the annual result, we also multiply the additional variables by 75%. Hence, the regression model is balanced in proportion of dependent and independent variables. After this adjustment, the coefficients of the additional variables are virtually identical. Therefore, we reduce the number of coefficients and estimate single coefficients for each variable, e.g., loss dummies, interaction terms. This yields a more parsimonious model and avoids in-sample overfitting by a large number of coefficients. However, interacting all variables with the dummy variable $QD_j^{(q)}$ does not change the results, but creates a large set of coefficients. However, interacting all variables with the dummy variable $QD_j^{(q)}$ does not change the inferences, but creates a large set of coefficients.

HVZ model (see also Hou, van Dijk and Zhang (2012)):

Standard HVZ model

Regression equation:

$$A_t = \alpha_0 + \alpha_1 \cdot A_{t-\tau} + \alpha_2 \cdot NED_{t-\tau} + \alpha_3 \cdot TA_{t-\tau} + \alpha_4 \cdot D_{t-\tau} + \alpha_5 \cdot DD_{t-\tau} + \alpha_6 \cdot ACC_{t-\tau} + \varepsilon_t$$

Prediction equation:

$$\hat{A}_{t+\tau} = \hat{\alpha}_0 + \hat{\alpha}_1 \cdot A_t + \hat{\alpha}_2 \cdot NED_t + \hat{\alpha}_3 \cdot TA_t + \hat{\alpha}_4 \cdot D_t + \hat{\alpha}_5 \cdot DD_t + \hat{\alpha}_6 \cdot ACC_t$$

In those formulas, A_t is the annual earnings, NED_t is a negative earnings dummy, TA_t is total assets, D_t is cash dividends, DD_t is a dummy for dividend-paying firms and ACC_t is accruals. All variables are computed on per-share basis. A detailed description of the variable definition is provided in Appendix D. We run this regression model for each month within the sample range 1982 to 2014 for the cross-section of all firms within the rolling ten-year windows (120 months) prior to the estimation date. From the model's coefficients, we compute predictions for future annual earnings per share in fiscal year t+τ, with $\tau = 1, \dots, 5$.

Extended HVZ model

Regression equation:

$$Q_t^{\{q+1,\dots,4\}} = \alpha_0 \cdot f_q + \sum_{j=0}^{3} QD_j^{(q)} \left[\alpha_{1,j} Q_{t-\tau}^{\{q+1,\dots,4\}} + \alpha_{2,j} \cdot g_q \cdot \Delta Q_{t-\tau+1,t-\tau}^{\{1,\dots,q\}}\right] + \alpha_3 \cdot f_q \cdot NED_{t-\tau} + \alpha_4 \cdot f_q \cdot TA_{t-\tau} + \alpha_5 \cdot f_q \cdot D_{t-\tau} + \alpha_6 \cdot f_q \cdot DD_{t-\tau} + \alpha_7 \cdot f_q \cdot ACC_{t-\tau} + \varepsilon_t$$

Prediction equation:

$$\hat{A}_{t+\tau} = Q_{t+1}^{\{1,\dots,q\}} + \hat{\alpha}_0 \cdot f_q + \sum_{j=0}^{3} QD_j^{(q)} \left[\hat{\alpha}_{1,j} Q_t^{\{q+1,\dots,4\}} + \hat{\alpha}_{2,j} \Delta Q_{t+1,t}^{\{1,\dots,q\}}\right] + \hat{\alpha}_3 \cdot f_q \cdot NED_t + \hat{\alpha}_4 \cdot f_q \cdot TA_t + \hat{\alpha}_5 \cdot f_q \cdot D_t + \hat{\alpha}_6 \cdot f_q \cdot DD_t + \hat{\alpha}_7 \cdot f_q \cdot ACC_t$$

The dummy variable $QD_j^{(q)}$ equals one, if $j = q$, and otherwise zero.

$Q_t^{\{q+1,\dots,4\}}$ are the remaining quarterly earnings results of the previous fiscal year t, while $Q_{t+1}^{\{1,\dots,q\}}$ are quarterly earnings results that are already announced for the fiscal year t+1. NED_t is a negative earnings dummy, TA_t is total assets, D_t is cash dividends, DD_t is a dummy for dividend-paying firms and ACC_t is accruals (details on the variable definitions are provided in Appendix D). All variables are on per-share basis. Since we divide our sample into four groups per available quarterly earnings information q, we estimate group-specific regression coefficients. We utilize dummy variables $QD_j^{(q)}$ to estimate the four regression models within a single regression equation. We run this regression model for each month within the sample range 1982 to 2014 for the cross-section of all firms within the rolling ten-year windows (120 months) prior to the estimation date. From the model's coefficients, we compute predictions of remaining quarterly earnings per share in fiscal year t+τ, with $\tau = 1, \dots, 5$. We then aggregate forecasts of annual earnings $\hat{A}_{t+\tau}$ from already realized quarterly earnings results $Q_{t+1}^{\{1,\dots,q\}}$ and predicted remaining quarterly results $\hat{Q}_{t+\tau}^{\{q+1,\dots,4\}}$. To reduce the number of coefficients in our model we interact only the earnings variables with the dummy variable $QD_j^{(q)}$. Other accounting variables, such as TA_t, are then aligned to the number of remaining quarterly earnings $\hat{Q}_{t+\tau}^{\{q+1,\dots,4\}}$to estimate a single coefficient. For example, if we predict 100% of the future annual earnings results, we also use 100% of the accounting variables, such as book equity or accruals. However, if quarterly earnings information is released and our regression is used to predict the remaining three quarterly earnings, thus, only 75% of the annual result, we also multiply book equity and accruals by 75%. Hence, the regression model is balanced in proportion of dependent and independent variables. After this adjustment, the coefficients of the additional variables are virtually identical. Therefore, we reduce the number of coefficients and estimate single coefficients for each variable, e.g., book equity, accruals. This yields a more parsimonious model and avoids in-sample overfitting by a large number of coefficients. However, interacting all variables with the dummy variable $QD_j^{(q)}$ does not change the results, but creates a large set of coefficients.

Appendix B: Summarized results – EP & HVZ model

Table B-1 Median forecast accuracy and bias of alternative forecast models (1982-2014)

Panel A: ALL FIRMS - Median forecast accuracy

	q	*After*	Extended HVZ Model		Standard HVZ Model		Extended EP Model		Standard EP Model		N
$EPS_{t+1,q}$	3	$Q3_{t+1}$	1.67%	***	3.50%	***	1.75%	***	3.47%	***	151,959
	2	$Q2_{t+1}$	2.38%	***	3.51%	***	2.44%	***	3.48%	***	150,789
	1	$Q1_{t+1}$	2.94%	***	3.51%	***	2.98%	***	3.48%	***	150,549
	0	A_t	3.77%	***	3.68%	***	3.80%	***	3.64%	***	149,554
$EPS_{t+2,q}$	3	$Q3_{t+1}$	3.85%	***	4.90%	***	3.88%	***	4.80%	***	133,385
	2	$Q2_{t+1}$	4.34%	***	5.00%	***	4.34%	***	4.88%	***	132,304
	1	$Q1_{t+1}$	4.68%	***	4.97%	***	4.63%	***	4.83%	***	132,084
	0	A_t	5.33%	***	5.06%	***	5.45%	***	4.91%	***	131,224
$EPS_{t+3,q}$	3	$Q3_{t+1}$	4.83%	***	6.04%	***	4.90%	***	5.89%	***	117,202
	2	$Q2_{t+1}$	5.28%	***	6.05%	***	5.29%	***	5.95%	***	116,214
	1	$Q1_{t+1}$	5.85%	***	6.13%	***	5.85%	***	5.96%	***	116,012
	0	A_t	6.55%	***	6.08%	***	6.78%	***	5.91%	***	115,279

Panel B: ALL FIRMS - Median forecast bias

	q	*After*	Extended HVZ Model		Standard HVZ Model		Extended EP Model		Standard EP Model		N
$EPS_{t+1,q}$	3	$Q3_{t+1}$	-0.08%		-0.59%	***	0.19%	***	-0.24%	**	151,959
	2	$Q2_{t+1}$	-0.24%	***	-0.59%	***	0.05%		-0.25%	**	150,789
	1	$Q1_{t+1}$	-0.41%	***	-0.58%	***	-0.22%	**	-0.25%	**	150,549
	0	A_t	-0.59%	***	-0.49%	***	-0.55%	***	-0.15%		149,554
$EPS_{t+2,q}$	3	$Q3_{t+1}$	-0.46%	***	-1.71%	***	-0.09%		-1.33%	***	133,385
	2	$Q2_{t+1}$	-1.03%	***	-1.79%	***	-0.69%	***	-1.45%	***	132,304
	1	$Q1_{t+1}$	-1.48%	***	-1.70%	***	-1.38%	***	-1.34%	***	132,084
	0	A_t	-1.99%	***	-1.57%	***	-2.57%	***	-1.35%	***	131,224
$EPS_{t+3,q}$	3	$Q3_{t+1}$	-0.96%	***	-2.82%	***	-0.62%	***	-2.80%	***	117,202
	2	$Q2_{t+1}$	-1.93%	***	-2.92%	***	-1.70%	***	-2.85%	***	116,214
	1	$Q1_{t+1}$	-2.70%	***	-2.86%	***	-2.97%	***	-2.86%	***	116,012
	0	A_t	-3.61%	***	-2.89%	***	-4.54%	***	-2.84%	***	115,279

Note: This table reports the accuracy and bias of earnings forecasts from the HVZ and EP model in standard version and with our model extensions. Bias is defined as price-scaled forecast error. Accuracy is defined as price-scaled absolute forecast errors. The reported values are time-series averages with Newey-West corrected t-statistics. We evaluate the forecast performance with respect to forecast horizon, $\tau = 1,2,3.$ and the available quarterly earnings results, $q = 0, 1, 2, 3$. Panel A displays the forecast accuracy for all forecasts within the sample range from 1982 to 2014. Likewise, Panel B displays the forecast bias for all forecasts within the sample range from 1982 to 2014.

Table B-2 Performance of alternative model-based expected return proxies (1982-2014)

Panel A: Annual portfolio returns of composite ICC estimates

		COVERED FIRMS					NON-COVERED FIRMS				
	PF	**ICC**	**Ret**	**Std**	**Shp**	**t-stat**	**ICC**	**Ret**	**Std**	**Shp**	**t-stat**
	1	3.57	12.57	31.59	0.39	4.23	2.95	9.38	31.51	0.29	3.16
ICC$^{\text{Ext. HVZ}}$	***10***	17.45	18.20	30.03	0.59	6.47	35.19	24.14	37.97	0.63	6.73
	10-1	13.88	5.62	29.31	0.18	2.02	32.23	14.76	21.65	0.67	7.53
	1	3.68	13.10	30.65	0.42	4.57	3.66	11.25	31.61	0.34	3.77
ICC$^{\text{St. HVZ}}$	***10***	18.63	17.34	31.55	0.54	5.87	38.66	23.91	39.52	0.60	6.39
	10-1	14.95	4.24	28.41	0.14	1.58	35.00	12.66	24.72	0.50	5.59
	1	3.60	12.88	30.65	0.41	4.45	2.89	10.17	29.25	0.34	3.68
ICC$^{\text{Ext. EP}}$	***10***	16.91	17.71	31.79	0.55	6.00	35.06	23.15	38.93	0.59	6.32
	10-1	13.31	4.83	26.96	0.17	1.90	32.17	12.98	20.88	0.61	7.08
	1	3.88	12.98	29.38	0.43	4.70	3.48	11.84	32.32	0.36	3.89
ICC$^{\text{St. EP}}$	***10***	17.76	16.78	32.76	0.50	5.49	35.29	24.07	38.81	0.61	6.61
	10-1	13.87	3.80	26.42	0.13	1.53	31.80	12.23	22.94	0.52	5.80

Panel B: Regressions of annual returns on implied cost of capital estimates

	COVERED FIRMS			NON-COVERED FIRMS		
	a	**b**	**Adj. R²**	**a**	**b**	**Adj. R²**
ICC$^{\text{Ext. HVZ}}$	*0.105*	**0.3844**	1.71%	*0.100*	**0.3836**	0.64%
	[4.56]	*[2.42]*		*[4.65]*	*[6.79]*	
		F-stat	*p-val*		*F-stat*	*p-val*
WALD-Test	b = 0	*5.85*	*0.016*	b = 0	*46.07*	*0.000*
	b = 1	*15.00*	*0.000*	b = 1	*118.93*	*0.000*
	a	**b**	**Adj. R²**	**a**	**B**	**Adj. R²**
ICC$^{\text{St. HVZ}}$	*0.118*	**0.2110**	1.50%	*0.109*	**0.2607**	0.66%
	[5.51]	*[1.57]*		*[5.13]*	*[4.80]*	
		F-stat	*p-val*		*F-stat*	*p-val*
WALD-Test	b = 0	*2.47*	*0.116*	b = 0	*23.05*	*0.000*
	b = 1	*34.49*	*0.000*	b = 1	*185.42*	*0.000*
	a	**b**	**Adj. R²**	**a**	**b**	**Adj. R²**
ICC$^{\text{Ext. EP}}$	*0.110*	**0.3388**	1.28%	*0.107*	**0.3528**	0.54%
	[5.18]	*[2.38]*		*[5.32]*	*[5.99]*	
		F-stat	*p-val*		*F-stat*	*p-val*
WALD-Test	b = 0	*5.65*	*0.017*	b = 0	*35.89*	*0.000*
	b = 1	*21.50*	*0.000*	b = 1	*120.75*	*0.000*
	a	**b**	**Adj. R²**	**a**	**B**	**Adj. R²**
ICC$^{\text{St. EP}}$	*0.124*	**0.1601**	1.17%	*0.110*	**0.3178**	0.58%
	[6.13]	*[1.24]*		*[5.26]*	*[5.68]*	
		F-stat	*p-val*		*F-stat*	*p-val*
WALD-Test	b = 0	*1.53*	*0.216*	b = 0	*32.31*	*0.000*
	b = 1	*42.07*	*0.000*	b = 1	*148.94*	*0.000*

Note: This table reports the results from portfolio and firm-level tests of our composite ICC estimates when using HVZ or EP model specifications. Panel A presents the results from monthly decile portfolio sorts on ICC estimates and subsequently realized twelve-month portfolio returns. The composite ICC estimate is the mean of the five individual ICC estimates (GLS, CT, OJ, MPEG, GG). In detail, we focus on the abnormal earnings growth model from Ohlson and Juettner-Nauroth (2005) (OJ), the modified version from Easton (2004) (MPEG), the residual income valuation models from Gebhardt et al. (2001) (GLS) and Claus and Thomas (2001) (CT) and the simple expected return model from Gordon and Gordon (1997) (GG). A detailed description of the individual ICC model equation and terminal value assumptions is provided in Appendix C. In this analysis, we sort our firms into decile portfolios based on the ICC estimates from model-based at each month in our sample range. We then compute twelve-months holding returns for each portfolio and return spreads of an investment strategy that buys the upper portfolio and sells the bottom portfolio (10-1). We tabulate the average ICC estimate *ICC*, average realized return *Ret*, standard deviation of realized returns *Std*, the Sharpe-Ratio *Shp* and Newey-West corrected time-series t-statistics *t-stat* for the time-series of investments within our sample range. The Sharpe-Ratio quantifies the trade-off between realized return and volatility for each portfolio strategy, i.e., the portfolio excess return per unit of the portfolios' standard deviation. The t-statistic indicates whether investment returns are significant from zero across the overall time-series. In Panel B, we run monthly regressions of a firms estimated implied cost of capital (ICC) and its compute subsequent twelve-months stock return:

$$r_{t,1,12} = \alpha + \beta \cdot ICC_t + \varepsilon_t$$

The displayed values are time-series average with Newey-West corrected t-statistics for the subsamples of covered firms and non-covered firms. If ICCs can perfectly predict future stock returns, coefficient β equals one. Likewise, the intercept α is zero. In contrast, a coefficient β of zero indicates that ICCs cannot predict future returns at all. In addition, we performed Wald-Tests to test whether the coefficient β is zero (no relation between future return and ICC estimates) or whether the coefficient β is one (perfect relation between future return and ICC estimates).

Appendix C: Implied cost of equity capital models

ICC	Formula and assumptions	Source
GLS	$p_t = bvps_t + \sum_{\tau=1}^{11} \frac{E_t[(roe_{t+\tau} - r) \cdot bvps_{t+\tau-1}]}{(1+r)^{\tau}} + \frac{E_t[(roe_{t+12} - r) \cdot bvps_{t+11}]}{(1+r)^{12}}$ where p_t is stock price at the estimation date in year t, r is implied cost of capital (ICC), $bvps_t$ is book value of equity per share and $roe_{t+\tau}$ is the expected return on equity based on explicit EPS forecasts for year t+1, t+2, t+3 and the clean surplus relation for year t+4, …, 12. The expected return on equity is assumed to be constant beyond the horizon t+12. We estimate future book value per share using the clean surplus relation: $bvps_{t+\tau} = bvps_{t+\tau-1} + (1 - pr_t) \cdot E_t[eps_{t+\tau}]$, where payout ratio pr_t is estimated as dividends divided by earnings or divided by 0.06*total assets if earnings are negative.	Gebhardt et al. (2001)
CT	$p_t = bvps_t + \sum_{\tau=1}^{4} \frac{E_t[(roe_{t+\tau} - r) \cdot bvps_{t+\tau-1}]}{(1+r)^{\tau}} + \frac{E_t[(roe_{t+5} - r) \cdot bvps_{t+4}] \cdot (1+g)}{(1+r)^{5} \cdot (r-g)}$ where p_t is stock price at the estimation date in year t, r is implied cost of capital (ICC), $bvps_t$ is book value of equity per share and $roe_{t+\tau}$ is the expected return on equity based on explicit EPS forecasts for year t+1, t+2, t+3, t+4 and t+5. The expected return on equity grows by g in perpetuity. It is set to the current risk-free rate minus 3%. We estimate future book value per share using the clean surplus relation: $bvps_{t+\tau} = bvps_{t+\tau-1} + (1 - pr_t) \cdot E_t[eps_{t+\tau}]$, where payout ratio pr_t is estimated as dividends divided by earnings or divided by 0.06*total assets if earnings are negative.	Claus and Thomas (2001)
OJ	$p_t = \frac{E_t[eps_{t+1}] \cdot \left(stg - (\gamma - 1)\right)}{(r - A^2) - A^2}$ with A $= \frac{1}{2}\left((\gamma - 1) \frac{E_t[eps_{t+1}] \cdot pr_t}{p_t}\right)$ and $stg = \frac{1}{2}\left(\frac{E_t[eps_{t+3}] - E_t[eps_{t+2}]}{E_t[eps_{t+2}]} - \frac{E_t[eps_{t+5}] - E_t[eps_{t+4}]}{E_t[eps_{t+4}]}\right)$,	Ohlson and Juettner-Nauroth (2005)

ICC	Formula and assumptions	Source
	where p_t is stock price at the estimation date in year t, r is implied cost of capital (ICC), pr_t is current payout ratio and $eps_{t+\tau}$ are is explicit EPS forecasts for year t+τ. Payout ratio is estimated as dividends divided by earnings or divided by 0.06*total assets if earnings are negative. The short-term growth *stg* is approximated by the arithmetic mean of predicted earnings growth in year t+3 and t+5. γ represents the perpetual growth rate in abnormal earnings beyond the forecast horizon. It is set to the risk-free rate minus 3%.	
MPEG	$p_t = \frac{E_t[eps_{t+2}] + (r \cdot pr_t - 1) \cdot E_t[eps_{t+1}]}{r^2}$ where p_t is stock price at the estimation date in year t, r is implied cost of capital (ICC), pr_t is current payout ratio and eps_{+1} and eps_{+2} are explicit EPS forecasts for year t+1 and t+2. Payout ratio is estimated as dividends divided by earnings or divided by 0.06*total assets if earnings are negative.	Easton (2004)
GG	$p_t = \frac{E_t[eps_{t+1}]}{r}$ where p_t is stock price at the estimation date in year t, r is implied cost of capital (ICC), and eps_{t+1} is the explicit EPS forecast for year t+1. This formula can be solved for r by the incerse of the forward P/E ratio.	Gordon and Gordon (1997)

Appendix D: Variable descriptions and formulas

Notation	Variable description	Formula
A	Earnings are defined as income before extraordinary items (Compustat item #18) minus special items (#17). If firms do not report any special items in their annual report, special items equal zero. We standardize on per share basis (#25).	$= \frac{ib - spi}{csho}$
Q	Quarterly earnings are defined as income before extraordinary items (#8) minus special items (#32). If firms do not report any special items in their annual report, special items equal zero. We standardize on per share basis (#25).	$= \frac{ibq - spiq}{csho}$
NED	The negative earnings dummy equals one for firms where the corresponding earnings variable is negative, whereas this variable equal to zero for firms with non-negative annual or remaining quarterly earnings.	$= \{0,1\}$
IAT	This is the interaction term of the negative earnings dummy NED and the earnings variable of the regression, i.e., annual earnings or the remaining quarterly earnings.	$= NED_t \cdot AE_t$
BV	The book value of equity is defined as stockholder's equity (#216). If not reported, we reconcile book equity from total assets (#6), total liabilities (#181) and minority interest (#38). We standardize on per share basis (#25).	$= \frac{seq}{csho}$
TA	Total asset is obtained from Compustat item #6.	$= \frac{at}{csho}$
D	Dividends are defined as total cash dividends (#127). If not available, we sum up dividends on common and preferred stocks (#21 & #19). We standardize on per share basis (#25).	$= \frac{dvt}{csho}$
DD	This dividend dummy equals one for dividend-paying firms, whereas it equals zero for firms without dividend payments.	$= \{0,1\}$

Notation	Variable description	Formula
ACC	Starting in 1988, accruals are calculated using the cash flow statement method as earnings (#123) minus cash flow from operations (#308). Prior to 1988, accruals are calculated using the balance sheet method as the changes in non-cash current assets (#4 & #1) less changes in current liabilities (#5) excluding the change in short-term debt (#34) and changes in taxes payable (#71) minus depreciation and amortization expenses (#14). Like other variables, we standardize accruals on per share basis (#25).	*Since 1988:* $= \frac{ib - oancf}{csho}$ *Until 1988:* $= \frac{\Delta act - \Delta che}{csho} - \frac{\Delta dlc - \Delta dlc - \Delta txp}{csho} - \frac{dp}{csho}$
RoE	The industry return on equity is the rolling ten-year median return on equity for Fama-French 49-industry classification. For firms with positive earnings and book equity, we calculate return on equity as income before extraordinary items (#18) divided by lagged book equity.	$= \frac{AE}{BVPS}$
pr	Payout ratio is defined as cash dividends divided by income before extraordinary items (#18). For loss firms, we divide by 0.06*total assets (#6).	$= \frac{D}{AE}$
p	Stock price is defined as absolute price (prc) from CRSP monthly files at the estimation date. Negative values are just an indicator that if closing prices are not available, CRSP inserts the bid/ask average.	$= \lvert prc \rvert$

Appendix to Chapter 3

Appendix A: Construction of variables for earlier bankruptcy prediction models

In this appendix, we describe the construction of variables used in Altman (1968), Ohlson (1980), Shumway (2001), Bharath and Shumway's (2008) most effective model, and Hess and Huettemann's (2018) market model. Most notations follow the Compustat item description, respectively.

Altman (1968) bankruptcy prediction model

Altman (1968) obtains a Z-score using a linear weighted sum of five ratios:

$$Z = \beta_0 + \beta_1 WCTA + \beta_2 RETA + \beta_3 EBITTA + \beta_4 METL + \beta_5 STA,$$

where $WCTA$ is working capital (WCAP) divided by total assets (AT), $RETA$ is retained earnings (RE) divided by total assets (AT), $EBITTA$ is earnings before interest and taxes (EBIT) divided by total assets (AT), $METL$ is the market value of equity (PRCC_F multiplied by CSHO) divided by the book value of total debt (LT), STA is sales (SALE) divided by total assets (AT) and Z is the Z-score (overall index). $WCTA$ is a proxy for firm liquidity. $RETA$ and $EBITTA$ measure different aspects of profitability. $METL$ is a widely used measure of leverage and STA describes the firm's efficiency in using assets to generate sales. The Z-score characterizes the financial strength of a firm by aggregating these five accounting ratios into one figure using the estimated coefficients $\beta_1, \dots, \beta_5$.

Ohlson (1980) bankruptcy prediction model

Ohlson (1980) finds nine variables to be significant and defines his O-score model as:

$$O = \beta_0 + \beta_1 SIZE + \beta_2 TLTA + \beta_3 WCTA + \beta_4 CLCA + \beta_5 OENEG$$
$$+\beta_6 NITA + \beta_7 FUTL + \beta_8 INTWO + \beta_9 CHIN,$$

where $SIZE$ is the logarithm of total assets (AT), $TLTA$ is total liabilities (LT) over total assets (AT), $WCTA$ is working capital (WCAP) over total assets (AT), $CLCA$ is current liabilities (LCT) over current assets (ACT), $OENEG$ is a dummy that takes the value of one if total liabilities (LT) exceed total assets (AT) and zero otherwise, $NITA$ is net income (NI) over total assets (AT), $FUTL$ is funds provided by operations[30] (PI plus DP) over total liabilities (LT), $INTWO$ is a dummy that takes the value of one if net income (NI) is negative for the past two years and zero otherwise, $CHIN$ is the change in net income (NI) and O is the Ohlson-score (overall index). $WCTA$ and $CLCA$ measure liquidity. $NITA$, $FUTL$, $INTWO$, and $CHIN$ capture the different aspects of profitability. $TLTA$ and $OENEG$ describe the capital structure and $SIZE$ is a measure of firm size.

Shumway (2001) bankruptcy prediction model

In addition to selected financial ratios used by Ohlson, Shumway (2001) adds two market variables, the excess return and its standard deviation:

$$S = \beta_0 + \beta_1 RSIZE + \beta_2 TLTA + \beta_3 NITA + \beta_4 ER + \beta_5 STDER,$$

where $RSIZE$ is the logarithm of market equity divided by the value-weighted market equity of the index, $TLTA$ is total liabilities (LT) over total assets (AT), $NITA$ is net income (NI) over total assets (AT), ER is excess returns calculated as the difference between the previous year's returns and risk-free rate, $STDER$ is the standard deviation of the returns, and S is the S-score (overall index). $TLTA$ measures solvency and describes the capital structure and profitability is captured by $NITA$. ER measures the profit of an investment, where $STDER$ determines the variability of excess returns. Returns are

[30] Since funds provided by operations are no longer reported, we perform an approximation by summing "pretax income" and "depreciations and amortization".

calculated as the present year's adjusted stock price (PRCCD multiplied by TRFD and divided by AJEXDI) divided by the previous year's adjusted stock price minus one. $RSIZE$ is a measure of firm size.

Bharath and Shumway (2008) bankruptcy prediction model

Bharath and Shumway (2008) extend the distance-to-default models that Vassalou and Xing (2004) and Hillegeist et al. (2004) construct by applying Merton's (1974) option pricing theory. Merton's probability of bankruptcy is calculated as

$$PD - Merton \ = N\left(-\frac{ln\left(\frac{V}{F}\right) + (\mu - 0.5\sigma_V^2)}{\sigma_V}\right),$$

where V is the market value of a firm's assets, σ_V is its standard deviation, μ is the expected return on assets, F is the market value of firms' debt, and $N(\cdot)$ is the cumulative standard normal distribution function. Vassalou and Xing (2004) numerically compute V and σ_V by applying an iterative procedure. Bharath and Shumway, however, propose a naïve approach. They approximate the market value of debt using the book value of debt and, thus, calculate F as debt in current liabilities plus one half of long-term debt. Furthermore, the volatility of a firm's debt is approximated by

$$\sigma_F = 0.05 + 0.25 \cdot \sigma_E,$$

where σ_E is the volatility of market equity. Market equity is denoted by E and calculated as the product of share price at the end of the month and the number of outstanding shares. Accordingly, an approximation for the volatility of the firm's assets is denoted by

$$\sigma_V = \frac{E}{E+F}\sigma_E + \frac{F}{E+F}\sigma_F.$$

The expected return on assets μ is approximated using the previous year's return on assets. In addition, the market value of assets is approximated by the sum of the market value of equity and book value of debt.

The most effective model in Bharath and Shumway (2008) includes PD-Merton as constructed above, the logarithm of market equity E (PRCC_F multiplied by CSHO), the logarithm of the book value of debt F calculated as current debt (DLC) plus one half of long-term debt (DLTT), the inverse of market equity volatility, excess returns calculated as the difference between the previous year's returns and the risk-free rate measured by the return on a one-year Treasury Bill from the Board of Governors of the Federal Reserve system.

Hess and Huettemann (2018) bankruptcy prediction model

The key idea of the Hess and Huettemann (2018) model is that a firm becomes bankrupt if its book equity becomes negative. Thus, the key bankruptcy predictor is the probability that the sum of a firm's current book equity and earnings forecast for the subsequent month is negative. This probability for firm *i* at time *t* can be expressed as

$$PNBE_{i,t} = 1 - \Phi\left(\frac{\widehat{Earn_{\iota,t+1}} + \mathrm{BkEq_{i,t}}}{\sigma\left(\widehat{Earn_{\iota,t+1}}\right)}\right)$$

where $BkEq_{i,t}$ denotes the current book equity for the previous quarterly or yearly report, $Earn_{i,t+1}$ is the expected earnings for the subsequent month, $\sigma\left(\widehat{Earn_{\iota,t+1}}\right)$ is the corresponding volatility of the individual earnings forecast, and $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. $BkEq_{i,t}$ equals stockholder's equity (SEQ). If SEQ is missing, they include common equity (CEQ) plus the value of preferred stocks (PSTK). If CEQ or PSTK are missing, book equity is evaluated as total assets (AT) minus total liabilities (LT), minus minority interest (MIB). Earnings are the change in retained earnings, which equals net income (NI) minus dividend payments (DVT).

To calculate the earnings forecasts and their volatilities, Hess and Huettemann (2018) use cross-sectional models. More specifically, they adopt a rolling regression technique with accounting data from the past five years to estimate the parameters they use for forecasting.

Hess and Huettemann's (2018) market model contains $PNBE_{i,t}$, the probability that book equity becomes negative, $NegBkEq_{i,t}$, a dummy that equals one if book equity is negative and zero otherwise, $NegEarnFrc_{i,t}$, a dummy that equals one if the earnings forecast is negative and zero otherwise; $CAPXTA$ as capital expenditures (CAPX) divided by total assets (AT); TXT as paid taxes (TXT); $EBITTA$ as profitability calculated as earnings before interest and taxes (EBIT) over total assets (AT); $SIZE$ measured by the logarithm of total assets (AT); and the market leverage ratio (MLR) calculated as the sum of long-term debt (DLTT) and current debt (DLC) divided by the sum of long-term debt, current debt, and market equity. Market equity is the fiscal year-end equity price (PRCC_F) multiplied by the number of common outstanding shares (CSHO). It adds excess returns calculated as the difference between the previous year's returns and the risk-free rate (ER), i.e., excess return, and the standard deviation of the excess return ($STDER$).

Appendix B: List of German words related to bankruptcy

1	Insolvenaverwalter	51	Insolvenzeröffnungsbilanz	101	Insolvenzplansanierung
2	Insolvencies	52	Insolvenzeröffnungsgutachten	102	Insolvenzplanteilnehmer
3	Insolvency	53	Insolvenzeröffnungsverfahren	103	Insolvenzplanverfahren
4	insolvency	54	Insolvenzeröffnungsverfahrens	104	Insolvenzplanverfahrens
5	Insolvent	55	Insolvenzerwaltung	105	Insolvenzprozesses
6	insolvent	56	Insolvenzexpertin	106	Insolvenzquote
7	Insolvente	57	Insolvenzfälle	107	Insolvenzrecht
8	insolvente	58	Insolvenzfällen	108	Insolvenzrechtes
9	insolventen	59	Insolvenzfordcrungen	109	Insolvenzrechtlich
10	insolventer	60	Insolvenzforderung	110	insolvenzrechtlich
11	Insolvenverfahrens	61	Insolvenzforderungen	111	insolvenzrechtliche
12	insolvenz	62	Insolvenzfrist	112	insolvenzrechtlichen
13	Insolvenz	63	Insolvenzgefahr	113	Insolvenzrechts
14	insolvenzabwendenden	64	insolvenzgefährdet	114	Insolvenzreife
15	Insolvenzabwicklung	65	insolvenzgefährdeten	115	Insolvenzrisiken
16	insolvenzähnliche	66	Insolvenzgefährdung	116	Insolvenzrisiko
17	Insolvenzanfechtungs	67	Insolvenzgeld	117	Insolvenzsanierungsplan
18	Insolvenzankündigung	68	Insolvenzgeldes	118	Insolvenzschuldnerin
19	Insolvenzanmeldung	69	Insolvenzgeldvorfinanzierung	119	Insolvenzschutz
20	insolvenzantrag	70	Insolvenzgericht	120	insolvenzsichere
21	Insolvenzantrag	71	Insolvenzgerichten	121	Insolvenzsituation
22	Insolvenzanträge	72	Insolvenzgerichtes	122	Insolvenzsituationen
23	Insolvenzanträgen	73	Insolvenzgerichts	123	Insolvenzspezialisten
24	Insolvenzantrages	74	Insolvenzgeschäft	124	Insolvenzstatus
25	Insolvenzantrags	75	Insolvenzgeschichte	125	Insolvenzszenario
26	Insolvenzantragsgründe	76	Insolvenzgesetzes	126	Insolvenztabelle
27	Insolvenzantragsgründen	77	Insolvenzgläubiger	127	Insolvenztatbestände
28	Insolvenzantragspflicht	78	Insolvenzgläubigern	128	insolvenztypische
29	Insolvenzantragspflichten	79	Insolvenzgläubigerversammlung	129	Insolvenzursachen
30	Insolvenzantragsprüfung	80	Insolvenzgrund	130	Insolvenzverewalter
31	Insolvenzantragstellung	81	Insolvenzgründe	131	insolvenzverfahren
32	Insolvenzantragsverfahren	82	Insolvenzgutachten	132	Insolvenzverfahren
33	Insolvenzantragsverfahrens	83	Insolvenzgutachtens	133	Insolvenzverfähren
34	Insolvenzantragverfahrens	84	Insolvenzjahr	134	Insolvenzverfahrens
35	Insolvenzausfallgeld	85	Insolvenzkanzlei	135	insolvenzverfahrens
36	Insolvenzausgleichsfonds	86	Insolvenzkapitel	136	insolvenzverfahrensgestützten
37	Insolvenzbedingte	87	Insolvenzlage	137	Insolvenzverordnung
38	insolvenzbedingte	88	insolvenzlichen	138	Insolvenzverschleppung
39	insolvenzbedingten	89	insolvenzliches	139	Insolvenzvertreter
40	Insolvenzbedingter	90	Insolvenzmanagement	140	insolvenzverwalter
41	insolvenzbefangene	91	Insolvenzmasse	141	Insolvenzverwalter
42	insolvenzbekanntmachungen	92	Insolvenzmassen	142	Insolvenzverwalterin
43	Insolvenzbüro	93	Insolvenzordnung	143	Insolvenzverwaltern
44	insolvenzen	94	insolvenzphase	144	Insolvenzverwalters
45	Insolvenzen	95	Insolvenzplan	145	Insolvenzverwaltung

46	Insolvenzentwicklung	96	Insolvenzpläne	146	Insolvenzwelle
47	Insolvenzer	97	Insolvenzplänen	147	Insolvenzwirtschaft
48	Insolvenzeröffhung	98	Insolvenzplanes	148	Insolvenzzahlen
49	Insolvenzeröffnung	99	Insolvenzplans	149	lnsolvenz
50	Insolvenzeröffnungs	100	insolvenzplans	150	lnsolvenzverfahrens

Note: This list of German words that are connected to news releases capturing insolvency proceedings is created using a PYTHON script. We parse each news release within the data archives of DGAP and APA OTS into individual words and use dynamic regular expressions to test if the root of each word contains insolvency wordings. This procedure yields a list of words that are connected to textual content of insolvency proceedings.