1 Introduction

Measurement, the act of translating real-world phenomena into numbers, is at the heart of accounting (Bae et al., 2023). Governed by a set of standards and rules, accounting measurements are concerned with recording, quantifying, and verifying firms' transactions (Andreicovici et al., 2023). They commonly involve expert judgment and entail extensive disclosures of accounting information to share- and stakeholders. In recent years, several factors contributed to an increase in their volume, scope, and complexity, including the growing demand for non-financial information (Stolowy & Paugam, 2018; Christensen et al., 2021; Bourveau et al., 2023; Cohen et al., 2023; Rouen et al., 2023), the proliferation of big data (Stice-Lawrence, 2022), and the prevalence of managerial estimates and human judgment (Seidel et al., 2020; Chen et al., 2023).

One potential solution to these challenges, which garners increasing attention in the contemporaneous accounting literature, is the use of machine learning. For example, research finds that machine learning models outperform human experts in specific, clearly delineated tasks, such as summarizing earnings news and narrative disclosures (Cardinaels et al., 2019; Kim et al., 2023a), performing audit interviews (Pickard et al., 2020), taking stock of inventories (Christ et al., 2021), generating accounting estimates (Ding et al., 2020), and providing investment recommendations (Coleman et al., 2022). These studies echo broader labor market trends that project jobs in accounting, auditing, and financial analysis to be largely computerizable and heavily exposed to novel machine learning technologies (Frey & Osborne, 2017; Eloundou et al., 2023). The goal of this cumulative dissertation is to extend this line of research and demonstrate the application of machine learning to address select measurement problems in financial reporting, voluntary disclosure, and management accounting.

Figure 1 distinguishes between the four types of machine learning applications commonly encountered in the literature and illustrates their role in solving measurement problems faced by both accounting practitioners and researchers. To begin with, many measurement problems require the accurate prediction of future accounting outcomes. Machine learning models are naturally suited for this task because they are designed and trained to minimize (maximize) the out-of-sample prediction error (accuracy) (Hastie et al., 2009). Accordingly, previous studies have used machine learning to predict accounting fraud and irregularities (Perols et al., 2017; Bertomeu et al., 2021), accounting estimates (Ding et al., 2020), earnings (Chen et al., 2022; Hunt et al., 2022), credit risk (Donovan et al., 2021), firm valuations (Geertsema & Lu, 2023), bankruptcies (Jones, 2017), effective tax rates (Guenther et al., 2023), and profitability (Jones et al., 2023). Collectively, these studies demonstrate that machine learning models are consistently more accurate in predicting accounting outcomes relative to existing benchmarks, such as linear models from the literature or forecasts by human experts.

In addition, measurement problems in accounting frequently deal with the extraction of complex information from high-dimensional, unstructured data. Machine learning models accomplish this task via dimensionality reduction, i.e., by projecting high-dimensional data into a lower-dimensional space and accounting for non-linear relationships (Bochkay et al., 2023). Bertomeu et al. (2023) exploits neural networks to account for non-linearities between a set of financial variables and produces a novel measure of accounting conservatism. Turning to unstructured data, recent works leverage machine learning to extract parsimonious variables from visual data, such as facial expressions from images of executives (Hsieh et al., 2020), management postures from videos of entrepreneurial pitches (Dávila & Guasch, 2022), and retailers' store performance from satellite imagery of parking lots (Kang et al., 2021). Other studies use machine learning and textual data to derive empirical proxies of firm-level constructs, such as corporate culture (Li et al., 2021), investment opportunities (Basu et al., 2022), and management accounting practices (Qiu et al., 2023). Moreover, the rise of large language models (LLMs) offers new opportunities for researchers and practitioners to summarize and parse corporate narratives (Jha et al., 2023; Kim et al., 2023b; Kim et al., 2023a; Wu et al., 2023), allowing for the extraction of complex information in a highly efficient manner with little to no training data (de Kok, 2023).[1,2]

Figure 1 further accounts for two auxiliary use cases: the use of machine learning for interpretability and for identification and inference. Machine learning for interpretability, commonly referred to as 'interpretable machine learning', involves a toolbox of "methods and models that make the behavior and predictions of machine learning systems understandable to humans" (Molnar, 2022, p. 13). Interpretability techniques can complement prediction and dimensionality reduction and reveal insights into the most important predictors and key sources of variation in high-dimensional datasets. Accounting scholars rely on these tools to assess the relative importance of individual explanatory variables (Ding et al., 2020; Chen et al., 2022; Geertsema & Lu, 2023; Guenther et al., 2023; Jones et al., 2023), gauge the functional form of the relationship between predictor variables and the outcome (Chen et al., 2022; Krupa & Minutti-Meza, 2022; Jones et al.,

---

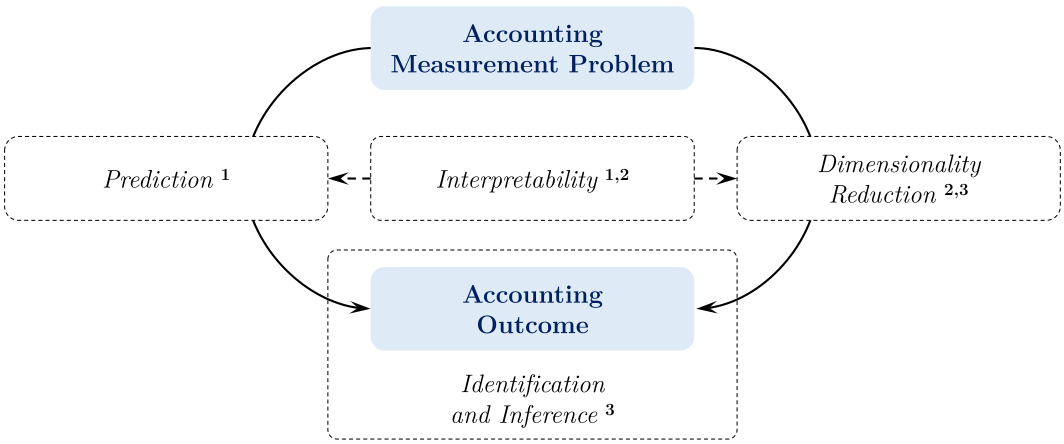
[1] For an overview of machine learning applications that involve textual data and techniques from natural language processing see Loughran & McDonald (2016), Gentzkow et al. (2019), El-Haj et al. (2019), Bochkay et al. (2023), and Webersinke (2023).

[2] Note that some measurement problems may involve both, dimensionality reduction and prediction. For example, Donovan et al. (2021) use machine learning to derive a text-based measure of credit risk from firms' 10-K filings (machine learning for dimensionality reduction) to predict future credit events, such as bankruptcies and rating downgrades (machine learning for prediction).

2023), extract interpretable decision rules (Bertomeu et al., 2021), and obtain word importance scores when the input data is text (Huang et al., 2023).

Machine learning for identification and inference integrates machine learning into the estimation of an unknown, plausibly causal estimand (Mullainathan & Spiess, 2017; Athey & Imbens, 2019; Huber, 2023). This integration is particularly promising when some component of the estimation can be framed as a prediction task. For example, the recent econometrics literature introduces machine learning applications for fitting the first stage in instrumental variable estimation (Belloni et al., 2012), estimating heterogeneous treatment effects (Athey & Imbens, 2016; Wager & Athey, 2018), and controlling for non-linear relationships and prohibitively large sets of control variables (Chernozhukov et al., 2018). Roberts et al. (2020) proposes text matching, a method that uses machine learning to extract low-dimensional variables from textual data and conditions on these 'latent' variables via matching.

Figure 1 Overview of Machine Learning Applications in Accounting



The figure presents an overview of machine learning applications in accounting that complements the reviews by Krupa & Minutti-Meza (2022) and Ranta et al. (2023). Dashed boxes indicate machine learning applications. Superscripts enumerate the three papers of this dissertation and indicate the role that machine learning plays in each of them.

[1] *'Beating Best Estimates': Machine Learning Estimates, Human Estimation Errors, and the Warranty Provision (see section 2.1)*
[2] *Transformation Talk: Drivers and Economic Consequences (see section 2.2)*
[3] *Value-Based Management Sophistication, Credit Risk, and Shareholder Value: Evidence from Credit Ratings (see section 2.3)*

In line with recent calls for future research (Krupa & Minutti-Meza, 2022; Ranta et al., 2023), the contributions of this cumulative dissertation are twofold. First, it introduces new machine learning applications to the accounting literature to address long-standing measurement problems and enable novel measurements. Second, the dissertation generates insights that are contributions to the accounting literature in their own right. The first paper adds to the financial reporting literature and investigates the ability of machine learning to generate decision-useful accounting estimates in the context of firms' warranty provisions. It demonstrates that machine learning estimates are consistently more accurate than estimates by human experts and discusses the potential causes of human estimation errors. The second paper of this dissertation studies the prevalence of transformation talk, measured as the share of managers' voluntary disclosures on firm transformation in quarterly earnings calls, and provides evidence on its economic consequences (future business model changes and capital market responses) and drivers (CEOs, peer firms, and regulatory change). The third paper revisits the cost-benefit-debate in management accounting research on implementing value-based management (VBM) systems and establishes a link between VBM sophistication, risk management, and shareholder value in the context of credit rating changes.

Section 2 discusses the specific machine learning applications employed in each paper of the dissertation, summarizes the papers' key motivations and results, and lays out in more detail their individual contributions to the literature. Section 3 concludes the central theme and presents an outlook on future research at the intersection of accounting and machine learning.

2 Summary and Contributions of the Dissertation

2.1 *Beating 'Best Estimates': Machine Learning Predictions, Human Estimation Errors, and the Warranty Provision*

The first paper of this dissertation, "Beating 'Best Estimates': Machine Learning Predictions, Human Estimation Errors, and the Warranty Provision" (joint work with Martin Becker), sheds light on the accuracy of accounting estimates disclosed by human experts and the potential reasons for human estimation errors related to lost information, a misspecified estimation model, and misaligned incentives (Bertomeu, 2020). Therefore, we use machine learning for prediction and construct machine-based estimates, which serve as a benchmark for accounting estimates disclosed by human experts. Further, we employ interpretability techniques to unravel the inner workings of the trained machine learning estimators and facilitate comparisons with the managerial estimation model.

Current financial reporting rules produce accounting estimates that are widely criticized for their deteriorating decision-usefulness (Dechow & Dichev, 2002; Lev et al., 2010; Balachandran & Mohanram, 2011; Lev & Gu, 2016). Managerial discretion is frequently regarded as a key culprit since it permits (un)intentional human biases to enter the estimation (Lev et al., 2010; Bell & Griffin, 2012; Christensen et al., 2012; Bratten et al., 2013; Szerwo et al., 2022). Building on recent works that advocate the use of machine learning to bypass human discretion, facilitate decision-making, and improve the accuracy of managerial estimates (Bertomeu, 2020; Cho et al., 2020; Ding et al., 2020; Ranta et al., 2023), the paper studies the accuracy of human experts vis-à-vis machine learning models in estimating a pervasive and economically relevant accounting estimate in the manufacturing sector: the warranty provision.

We obtain proprietary data on statutory warranty obligations from a large European manufacturing firm that accounts for more than 55,000 product sales. We train various machine learning models to generate point-in-time predictions of future (i.e., next-year) warranty cost and benchmark these predictions with historical managerial estimates of this cost. Importantly, we can then compare both the machine and human estimate with the actual realization of the warranty cost as observed in the data.

We find that warranty provision estimates by the machine are consistently more accurate than estimates provided by human experts. We document this result under two regimes: when we compare machine learning and human estimates on the basis of individual warranty obligations (i.e., per product) and after aggregating obligations across homogeneous classes of products (i.e., per product type). Notably, the observed performance differences are more pronounced under the second regime, which is prescribed by the regulator and represents the preferred level of aggregation of the accounting expert. Drawing upon institutional insights about the mechanism by which the firm's experts generate estimates, we attribute the larger human errors to misspecifications of the managerial estimation model, which produces frequent and more severe overstatements. In addition, we discuss the role of misaligned managerial incentives.

Our results replicate and extend the previous findings of Ding et al. (2020). While the authors focus on loss reserve estimates by U.S. property and casualty insurers, which are strictly regulated in accordance with solvency rules that apply to all U.S. insurance firms, we demonstrate that machine learning models outperform human experts in estimating the warranty reserve under IFRS. Drawing on institutional knowledge about the process through which the firm's accounting experts generate their estimates, i.e., the managerial estimation model, we discuss potential reasons for human errors due to lost information and a misspecified estimation model (Bertomeu, 2020). However, we acknowledge that our single-firm dataset prevents us from testing the presence of misaligned incentives in the cross-section (e.g., Ding et al., 2020). In summary, the paper highlights the role of machine learning as a versatile tool for filers, recipients, and auditors of financial statements to facilitate decision-useful disclosures and enable the independent assessment of managerial estimates.

2.2 Transformation Talk: Drivers and Economic Consequences

The second paper of this dissertation, "Transformation Talk: Drivers and Economic Consequences" (joint work with Martin Artz and Hannes Doering), proposes an empirical measure of transformation talk that serves as a leading indicator of firm transformation[3]. The measure can be intuitively interpreted as the share of executives' discussions on firm transformation during the question-and-answer section of quarterly earnings calls. We identify transformation talk using a BERT[4]-style language model (Devlin et al., 2019; Liu et al., 2019), which operates on entire text sequences and considers long-range dependencies between individual words. Importantly, the model enables us to capture both explicit (i.e., individual signal words) and implicit (i.e., 'reading-between-the-lines') discussions on firm transformation by accounting for the different contexts in which executives discuss fundamental changes to the business model, e.g., in terms of its capital structure, product market positioning, or organizational design.

[3] In the paper, we define firm transformation as a large-scale, multi-year process that affects the business model and fundamentally changes how a firm operates (Kotter, 1995; Fiedler et al., 2023).

[4] BERT is short for "bidirectional encoder representations from transformers" and describes a class of neural network models that is widely used in natural language processing for solving various types of discriminative tasks, such as text classification or named entity recognition.

Transformation is a timely and relevant topic in business research (Fiedler et al., 2023) and mentions of the term “transformation” in business news have more than tripled over the past decade (see Figure 2, panel B in the paper). While firms set out to install chief transformation officers to manage transformation-related initiatives (Gorter et al., 2016), obtaining detailed information on future firm transformations and their economic consequences remains challenging for outside stakeholders. While they can refer to a variety of different data sources, such as financial statements, narrative disclosures, news articles, and ad hoc filings, these information mainly account for past changes and/or a limited set of economic events that are sufficiently salient to warrant an official press release or media coverage. This research aims to close this gap and shed light on the prevalence, economic consequences, and drivers of talk about future firm transformation.

We identify transformation talk in over 170,000 earnings calls using a pretrained RoBERTa language model (Liu et al., 2019). We compile a large dataset of manually labeled sentences and fine-tune the model to distinguish sentences that contain discussions on firm transformation from non-transformation-related sentences. The model outputs a continuous score between zero and one, indicating its confidence that a sentence contains transformation talk. We then obtain a firm-quarter-level measure of transformation talk by aggregating all sentences in a given call and scaling by the length of the call. Finally, we assess the measure’s construct validity (Bochkay et al., 2023) through a set of qualitative and quantitative tests and demonstrate its superiority over a wordlist-based approach that we employ as a benchmark.

We conduct three series of analyses to attest the economic relevance of transformation talk. First, we demonstrate that our measure is positively associated with future business model changes, as signaled by changes in property, plant, and equipment, operating expenditures, research and development expenditures, and transitions in the firm life cycle stage, up to six quarters into the future. Second, we highlight the information content of transformation talk for capital market participants and show that investors' trading activities around quarterly earnings calls increase in the level of transformation talk. This reaction occurs with a delay and coincides with a decrease in the speed of price discovery, consistent with the idea that transformation talk is complex and raises information processing cost for investors. Third, we conduct three tests to attest the relevance of the CEO, the firm's peers, and regulatory change, specifically the COP21 Paris Agreement, as major drivers of transformation talk at the firm, industry, and regulatory level. Collectively, we predict and find that transformation talk increases in response to endogenous changes in the CEO position, to talk about transformation in a peer's preceding earnings call, and to the adoption of the COP21 Paris Agreement.

The paper contributes to the contemporary accounting literature in three ways. First, we leverage text as data (Bae et al., 2023) and extend recent works that propose firm-level measures derived from textual data (e.g., Hassan et al., 2019; Donovan et al., 2021; Li et al., 2021; Balakrishnan et al., 2023; Sautner et al., 2023). Specifically, we focus on talk about firm transformation and provide large-sample evidence that such talk is highly prevalent, accounting for more than 12% of the average earnings call, and informative to investors. Second, we contribute to the scarce but growing literature on the information content of earnings calls for predicting future firm behavior, especially corporate investments (Jha et al., 2023; Kim et al., 2023b; Li et al., 2023). Our findings are consistent with transformation talk predicting business model changes in future periods, as proxied by incremental changes in four important firm fundamentals (property, plant, and equipment, operating expenditures, research and development expenditures, and the firm's life cycle stage). Third, we add to the literature on textual analysis and demonstrate that subtle cues and linguistic context hold valuable information for construct measurement. We are by no means the first to attest the ability of (large) language models to extract complex information from corporate disclosures (Siano & Wysocki, 2021; Huang et al., 2023; Kim et al., 2023b; Webersinke, 2023). However, our findings highlight that the accurate measurement of transformation talk hinges on the use of language models and, more importantly, fails if we resort to simpler methods of textual analysis, such as wordlist-based approaches. Hence, our approach should be of interest to future research that is concerned with deriving economically meaningful measures of latent constructs from textual data.

2.3 Value-Based Management Sophistication, Risk Management, and Shareholder Value: Evidence from Credit Ratings

The third paper, titled "Value-Based Management Sophistication, Risk Management, and Shareholder Value: Evidence from Credit Ratings" (joint work with Friedrich Sommer), investigates the consequences of value-based management (VBM) sophistication, i.e., the extent of firms' VBM practices (Firk et al., 2019b), for risk management and shareholder value. Focusing on credit risk as a key dimension of firm-level risk, our findings suggest that VBM sophistication encourages management decisions that reduce credit risk and create shareholder value. The paper includes two machine learning applications to assess the robustness of the main findings. The first application leverages double/debiased machine learning (Chernozhukov et al., 2018), which accounts for non-linear relationships between modeling variables and thus facilitates identification and inference. The second applications harnesses the capabilities GPT-4 (OpenAI, 2023) and identifies confounding, stock-price relevant events in a large corpus of business news. The model's output is then used to construct control variables that correct for the effect of concurrent news in an event study setting.

There is an ongoing debate among management accounting academics and practitioners on the cost and benefits of VBM, its influence on managerial decision-making, and implications for shareholder value. Empirical evidence shows that (sophisticated) VBM practices are associated with many managerial decisions, including investment activities (Wallace, 1997; Hogan & Lewis, 2005; Ryan & Trahan, 2007; Mavropulo et al., 2021), share repurchases and dividend payouts (Wallace, 1997; Kister et al., 2022), earnings management (Hörner & Sommer, 2023), and enterprise risk management (Mikes, 2009). Moreover, numerous studies have researched the average effect of VBM adoption and VBM sophistication on different dimensions of firm performance, with mixed results (Lueg & Schäffer, 2010). In contrast, more recent papers demonstrate that VBM is positively associated with firm performance in specific settings, such as M&A (Knauer et al., 2018; Firk et al., 2019a) and divestitures (Firk et al., 2021), or when conditioning on certain environmental or firm characteristics (Firk et al., 2019b; Mavropulo et al., 2021; Kister et al., 2023). This paper is the first to examine the relationship between sophisticated VBM, risk management, and shareholder value changes in the context of decreases and increases in credit risk for a large number of firms.

We focus on 200 large, non-financial European companies during the 2005–2020 period and gather, for each firm, the full history of credit ratings issued by Moody's. We replicate the VBM sophistication measure proposed by Firk et al. (2019b), using hand-collected data on firms' VBM practices as disclosed in their annual reports. First, to test the risk management consequences of VBM, we construct a firm-year panel and correlate firms' VBM sophistication with proxies for decreases and increases in credit risk. Second, we assess the shareholder value consequences of VBM using a cross-sectional event study research design and measure the short-window capital market reactions to risk decreases and increases, as proxied by credit rating upgrade and downgrade announcements.

We find that sophisticated VBM promotes decreases while it cannot prevent increases in credit risk. Considering shareholder value effects, we document a statistically robust, positive relationship between VBM sophistication and cumulative abnormal returns following rating downgrades. We do not find a significant positive effect for rating upgrades. In the paper, we term these two results the risk management effect and the performance effect of VBM and interpret our results as follows: While sophisticated VBM users manage, on average, for decreases in credit risk, their shareholder value benefits even as credit risk increases. That is, investors potentially attribute risk increases in those firms to value-increasing management decisions in the past or, alternatively, trust their management to take countermeasures that promote risk decreases in the future.

The paper contributes to the VBM and credit ratings literature. First, we provide evidence on the implications of VBM sophistication for risk management and shareholder value. Our findings extend the literature on the consequences of VBM for managerial decision-making (e.g., Hogan & Lewis, 2005; Mikes, 2009; Schultze et al., 2018; Mavropulo et al., 2021) and firm performance (e.g., Knauer et al., 2018; Firk et al., 2019a; Firk et al., 2021). Second, we add to the scarce body of research on the information contents of external credit rating changes in Europe (Calderoni et al., 2009; Hu et al., 2016) and show that sophisticated VBM practices positively influence the post-announcement effect to rating downgrade announcements.

3 Conclusion

This cumulative dissertation presents novel machine learning applications in accounting and, in so doing, contributes to the financial reporting, voluntary disclosure, and management accounting literature. First, it demonstrates the accuracy of machine learning models vis-à-vis human experts in generating estimates of the warranty provision and discusses potential reasons for human estimation errors. Second, it provides evidence on the prevalence, drivers, and economic consequences of executives' talk about firm transformation. Our measurement of transformation talk is enabled by the use of language models that improve over prior approaches to textual analysis and account for the context, structure, and semantics in corporate disclosures. Third, it presents theory and evidence on the benefits of sophisticated VBM practices in the context of risk management and illustrates how accounting scholars can harness machine learning to strengthen identification and inference and reduce human labor by efficiently analyzing thousands of text documents. In concert, the three empirical papers of this dissertation demonstrate the applicability and versatility of machine learning methods for accounting research.

Recent reviews of the literature agree that state-of-the-art machine learning and natural language processing methods offer many opportunities and avenues for future accounting research (El-Haj et al., 2019; Ranta et al., 2023; Webersinke, 2023). In fact, Bertomeu (2020, p. 1151) argues that "it is uncontroversial that the tools of machine learning are revolutionizing empirical research." In the following, I revisit the four different machine learning applications discussed in section 1 and provide an outlook on future research at the intersection of accounting and machine learning.[5]

[5] Note that the three papers in this dissertation provide more specific directions for future research in financial reporting, voluntary disclosure, and management accounting, respectively.

Construct Measurement. First, machine learning can improve the measurement of existing constructs and facilitate the development of new empirical measures. For example, Bertomeu et al. (2023) propose a non-linear model of differential timeliness based on neural networks that improves traditional measures of accounting conservatism. In the context of text data, Webersinke (2023) demonstrate that language models produce accounting variables with higher accuracy (i.e., lower measurement error) compared to simpler machine learning methods and wordlist-based approaches. Furthermore, they can perform mundane measurement tasks, which would otherwise require a large group of research assistants, in a cost-effective manner and without sacrificing accuracy (de Kok, 2023). Other contemporary works adopt machine learning to derive entirely new measures (e.g., Basu et al., 2022; Kim et al., 2023a; Qiu et al., 2023), as does the second paper of this dissertation. These examples illustrate how machine learning can be used to extract economically meaningful representations from high-dimensional data and construct accounting variables for use in empirical analyses. I expect that machine learning applications will prove particularly valuable in measurement problems that involve complex theoretical constructs and datasets that exhibit relatively low signal-to-noise ratios.

Interpretability. Second, machine learning applications can help accounting researchers and practitioners gain insights into the true data-generating process underlying the focal measurement problem, learn about relevant non-linearities, and identify important variables. Previously, advanced machine learning models were often referred to as ‘black box’ models, owing to their inherent lack of interpretability as compared with a linear regression model (Krupa & Minutti-Meza, 2022; Stice-Lawrence, 2022; Bochkay et al., 2023; de Kok, 2023). Given the plethora of interpretability techniques that have been developed in recent years (see Molnar (2022) for a review), these arguments become less and less relevant. For example, in paper one and two of this dissertation, we employ interpretability techniques to unravel the inner workings of our models and illustrate the most important drivers (words) for predicting future warranty cost (identifying transformation talk). While simplicity is a desirable principle in empirical accounting research (Bae et al., 2023), large datasets and complex theoretical constructs may necessitate the use of more sophisticated measurement approaches. I expect that interpretability methods will be the default in future machine learning applications in accounting, to understand and demonstrate the inner workings of the implemented model, be transparent about what the model can and cannot do, and reconcile (non-linear) variable relationships and important predictors with prior research.

Identification and Inference. Third, machine learning can promote the identification of empirical associations and causal effects. Accounting researchers commonly select control variables based on theory, institutional knowledge, and findings from prior work, and presume linearity among the main independent variable, dependent variable, and covariates. Future research can rely on machine learning to systematically evaluate a large set of potential control variables and allow for flexible relationships between variables in the empirical model, particularly in the context of big data and selection-on-observables (Bertomeu, 2020; Huber, 2023). In the third paper of this dissertation, we employ double/debiased machine learning (Chernozhukov et al., 2018) to assess the robustness of the main findings to the linearity assumption of the linear regression model and we use GPT-4 (OpenAI, 2023) to derive additional controls to allay omitted variable bias. I expect future accounting research to resort to machine learning methods to complement statistical analyses and improve robustness tests.

References

Andreicovici, I., L. van Lent, V. V. Nikolaev, and R. Zhang. (2023). *Accounting Measurement Intensity*. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3853951. Accessed November 13, 2023.

Athey, S., and G. Imbens. (2016). Recursive Partitioning for Heterogeneous Causal Effects. *Proceedings of the National Academy of Sciences of the United States of America, 113*(27), pp. 7353–7360. doi: https://doi.org/10.1073/pnas.1510489113.

Athey, S., and G. W. Imbens. (2019). Machine Learning Methods That Economists Should Know About. *Annual Review of Economics, 11*(1), pp. 685–725. doi: https://doi.org/10.1146/annurev-economics-080217-053433.

Bae, J., C. Y. Hung, and L. van Lent. (2023). Mobilizing Text As Data. *European Accounting Review, Forthcoming*. doi: https://doi.org/10.1080/09638180.2023.2218423.

Balachandran, S., and P. Mohanram. (2011). Is the Decline in the Value Relevance of Accounting Driven by Increased Conservatism? *Review of Accounting Studies, 16*(2), pp. 272–301. doi: https://doi.org/10.1007/s11142-010-9137-0.

Balakrishnan, K., R. Copat, D. de La Parra, and K. Ramesh. (2023). Racial Diversity Exposure and Firm Responses Following the Murder of George Floyd. *Journal of Accounting Research, 61*(3), pp. 737–804. doi: https://doi.org/10.1111/1475-679X.12484.

Basu, S., X. Ma, and H. Briscoe-Tran. (2022). Measuring Multidimensional Investment Opportunity Sets with 10-K Text. *The Accounting Review, 97*(1), pp. 51–73. doi: https://doi.org/10.2308/TAR-2019-0110.

Bell, T. B., and J. B. Griffin. (2012). Commentary on Auditing High-Uncertainty Fair Value Estimates. *Auditing: A Journal of Practice & Theory, 31*(1), pp. 147–155. doi: https://doi.org/10.2308/ajpt-10172.

Belloni, A., D. Chen, V. Chernozhukov, and C. Hansen. (2012). Sparse Models and Methods for Optimal Instruments With an Application to Eminent Domain. *Econometrica, 80*(6), pp. 2369–2429. doi: https://doi.org/10.3982/ECTA9626.

Bertomeu, J. (2020). Machine Learning Improves Accounting: Discussion, Implementation and Research Opportunities. *Review of Accounting Studies, 25*(3), pp. 1135–1155. doi: https://doi.org/10.1007/s11142-020-09554-9.

Bertomeu, J., E. Cheynel, E. Floyd, and W. Pan. (2021). Using Machine Learning to Detect Misstatements. *Review of Accounting Studies, 26*(2), pp. 468–519. doi: https://doi.org/10.1007/s11142-020-09563-8.

Bertomeu, J., E. Cheynel, Y. Liao, and M. Milone. (2023). *Using Machine Learning to Measure Conservatism*. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3924961. Accessed November 13, 2023.

Bochkay, K., S. V. Brown, A. J. Leone, and J. W. Tucker. (2023). Textual Analysis in Accounting: What's Next?*. *Contemporary Accounting Research, 40*(2), pp. 765–805. doi: https://doi.org/10.1111/1911-3846.12825.

Bourveau, T., M. Chowdhury, A. Le, and E. Rouen. (2023). *Human Capital Disclosures*. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4138543. Accessed November 13, 2023.

Bratten, B., L. M. Gaynor, L. McDaniel, N. R. Montague, and G. E. Sierra. (2013). The Audit of Fair Values and Other Estimates: The Effects of Underlying Environmental, Task, and Auditor-Specific

Factors. *Auditing: A Journal of Practice & Theory, 32*(Supplement 1), pp. 7–44. doi: https://doi.org/10.2308/ajpt-50316.

Calderoni, F., P. Colla, and S. Gatti. (2009). *Rating Changes across Europe*. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1515988. Accessed November 13, 2023.

Cardinaels, E., S. Hollander, and B. J. White. (2019). Automatic Summarization of Earnings Releases: Attributes and Effects on Investors' Judgments. *Review of Accounting Studies, 24*(3), pp. 860–890. doi: https://doi.org/10.1007/s11142-019-9488-0.

Chen, H., J. V. Chen, and F. Li. (2023). The Number of Estimates in Footnotes and Accruals. *Management Science, Forthcoming*. doi: https://doi.org/10.1287/mnsc.2022.4659.

Chen, X., Y. H. Cho, Y. Dou, and B. Lev. (2022). Predicting Future Earnings Changes Using Machine Learning and Detailed Financial Data. *Journal of Accounting Research, 60*(2), pp. 467–515. doi: https://doi.org/10.1111/1475-679X.12429.

Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, et al. (2018). Double/Debiased Machine Learning for Treatment and Structural Parameters. *The Econometrics Journal, 21*(1), C1-C68. doi: https://doi.org/10.1111/ectj.12097.

Cho, S., M. A. Vasarhelyi, T. Sun, and C. Zhang. (2020). Learning from Machine Learning in Accounting and Assurance. *Journal of Emerging Technologies in Accounting, 17*(1), pp. 1–10. doi: https://doi.org/10.2308/jeta-10718.

Christ, M. H., S. A. Emett, S. L. Summers, and D. A. Wood. (2021). Prepare for Takeoff: Improving Asset Measurement and Audit Quality with Drone-Enabled Inventory Audit Procedures. *Review of Accounting Studies, 26*(4), pp. 1323–1343. doi: https://doi.org/10.1007/s11142-020-09574-5.

Christensen, B. E., S. M. Glover, and D. A. Wood. (2012). Extreme Estimation Uncertainty in Fair Value Estimates: Implications for Audit Assurance. *Auditing: A Journal of Practice & Theory, 31*(1), pp. 127–146. doi: https://doi.org/10.2308/ajpt-10191.

Christensen, H. B., L. Hail, and C. Leuz. (2021). Mandatory CSR and Sustainability Reporting: Economic Analysis and Literature Review. *Review of Accounting Studies, 26*(3), pp. 1176–1248. doi: https://doi.org/10.1007/s11142-021-09609-5.

Cohen, S., I. Kadach, and G. Ormazabal. (2023). Institutional Investors, Climate Disclosure, and Carbon Emissions. *Journal of Accounting and Economics, Forthcoming*. doi: https://doi.org/10.1016/j.jacceco.2023.101640.

Coleman, B., K. Merkley, and J. Pacelli. (2022). Human Versus Machine: A Comparison of Robo-Analyst and Traditional Research Analyst Investment Recommendations. *The Accounting Review, 97*(5), pp. 221–244. doi: https://doi.org/10.2308/TAR-2020-0096.

Dávila, A., and M. Guasch. (2022). Managers' Body Expansiveness, Investor Perceptions, and Firm Forecast Errors and Valuation. *Journal of Accounting Research, 60*(2), pp. 517–563. doi: https://doi.org/10.1111/1475-679X.12426.

de Kok, T. (2023). *Generative LLMs and Textual Analysis in Accounting: (Chat)GPT as Research Assistant?*. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4429658. Accessed November 15, 2023.

Dechow, P. M., and I. D. Dichev. (2002). The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors. *The Accounting Review, 77*(Supplement: Quality of Earnings Conference), pp. 35–59. doi: https://doi.org/10.2308/accr.2002.77.s-1.35.

Devlin, J., M.-W. Chang, K. Lee, and K. Toutanova. (2019). *BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding*. arXiv. https://arxiv.org/abs/1810.04805v2. Accessed November 13, 2023.

Ding, K., B. Lev, X. Peng, T. Sun, and M. A. Vasarhelyi. (2020). Machine Learning Improves Accounting Estimates: Evidence from Insurance Payments. *Review of Accounting Studies, 25*(3), pp. 1098–1134. doi: https://doi.org/10.1007/s11142-020-09546-9.

Donovan, J., J. Jennings, K. Koharki, and J. Lee. (2021). Measuring Credit Risk Using Qualitative Disclosure. *Review of Accounting Studies, 26*(2), pp. 815–863. doi: https://doi.org/10.1007/s11142-020-09575-4.

El-Haj, M., P. Rayson, M. Walker, S. Young, and V. Simaki. (2019). In Search of Meaning: Lessons, Resources and Next Steps for Computational Analysis of Financial Discourse. *Journal of Business Finance & Accounting, 46*(3-4), pp. 265–306. doi: https://doi.org/10.1111/jbfa.12378.

Eloundou, T., S. Manning, P. Mishkin, and D. Rock. (2023). *GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models*. arXiv. https://arxiv.org/abs/2303.10130v5. Accessed November 13, 2023.

Fiedler, M., T. Hutzschenreuter, M. Klarmann, and B. E. Weißenberger. (2023). Transformation: Challenges, Impact, and Consequences. *Schmalenbach Journal of Business Research, 75*(3), pp. 271–279. doi: https://doi.org/10.1007/s41471-023-00172-0.

Firk, S., F. Maybuechen, J. Oehmichen, and M. Wolff. (2019a). Value-Based Management and Merger & Acquisition Returns: A Multi-level Contingency Model. *European Accounting Review, 28*(3), pp. 451–482. doi: https://doi.org/10.1080/09638180.2018.1492947.

Firk, S., S. Richter, and M. Wolff. (2021). Does Value-Based Management Facilitate Managerial Decision-Making? An Analysis of Divestiture Decisions. *Management Accounting Research, 51*, p. 100736. doi: https://doi.org/10.1016/j.mar.2021.100736.

Firk, S., T. Schmidt, and M. Wolff. (2019b). Exploring Value-Based Management Sophistication: The Role of Potential Economic Benefits and Institutional Influence. *Contemporary Accounting Research, 36*(1), pp. 418–450. doi: https://doi.org/10.1111/1911-3846.12402.

Frey, C. B., and M. A. Osborne. (2017). The Future of Employment: How Susceptible are Jobs to Computerisation? *Technological Forecasting and Social Change, 114*, pp. 254–280. doi: https://doi.org/10.1016/j.techfore.2016.08.019.

Geertsema, P., and H. Lu. (2023). Relative Valuation with Machine Learning. *Journal of Accounting Research, 61*(1), pp. 329–376. doi: https://doi.org/10.1111/1475-679X.12464.

Gentzkow, M., B. Kelly, and M. Taddy. (2019). Text as Data. *Journal of Economic Literature, 57*(3), pp. 535–574. doi: https://doi.org/10.1257/jel.20181020.

Gorter, O., R. Hudson, and J. Scott. (2016). The Role of the Chief Transformation Officer. McKinsey & Company. https://www.mckinsey.com/capabilities/rts/our-insights/the-role-of-the-chief-transformation-officer. Accessed November 13, 2023.

Guenther, D. A., K. Peterson, J. Searcy, and B. M. Williams. (2023). How Useful Are Tax Disclosures in Predicting Effective Tax Rates? A Machine Learning Approach. *The Accounting Review, 98*(5), pp. 1–26. doi: https://doi.org/10.2308/TAR-2021-0398.

Hassan, T. A., S. Hollander, L. van Lent, and A. Tahoun. (2019). Firm-Level Political Risk: Measurement and Effects. *The Quarterly Journal of Economics, 134*(4), pp. 2135–2202. doi: https://doi.org/10.1093/qje/qjz021.

Hastie, T., R. Tibshirani, and J. H. Friedman. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). New York: Springer.

Hogan, C. E., and C. M. Lewis. (2005). Long-Run Investment Decisions, Operating Performance, and Shareholder Value Creation of Firms Adopting Compensation Plans Based on Economic Profits. *The Journal of Financial and Quantitative Analysis, 40*(4), pp. 721–745. doi: https://doi.org/10.1017/S0022109000001952.

Hörner, S., and F. Sommer. (2023). *Value-Based Management Sophistication and Earnings Management*. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4498375. Accessed November 13, 2023.

Hsieh, T.-S., J.-B. Kim, R. R. Wang, and Z. Wang. (2020). Seeing is Believing? Executives' Facial Trustworthiness, Auditor Tenure, and Audit Fees. *Journal of Accounting and Economics, 69*(1), p. 101260. doi: https://doi.org/10.1016/j.jacceco.2019.101260.

Hu, H., T. Kaspereit, and J. Prokop. (2016). The Information Content of Issuer Rating Changes: Evidence for the G7 Stock Markets. *International Review of Financial Analysis, 47*, pp. 99–108. doi: https://doi.org/10.1016/j.irfa.2016.06.012.

Huang, A. H., H. Wang, and Y. Yang. (2023). FinBERT: A Large Language Model for Extracting Information from Financial Text. *Contemporary Accounting Research, 40*(2), pp. 806–841. doi: https://doi.org/10.1111/1911-3846.12832.

Huber, M. (2023). *Causal Analysis: Impact Evaluation and Causal Machine Learning with Applications in R*. Cambridge: The MIT Press.

Hunt, J. O. S., J. N. Myers, and L. A. Myers. (2022). Improving Earnings Predictions and Abnormal Returns with Machine Learning. *Accounting Horizons, 36*(1), pp. 131–149. doi: https://doi.org/10.2308/HORIZONS-19-125.

Jha, M., J. Qian, M. Weber, and B. Yang. (2023). *ChatGPT and Corporate Policies*. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4521096. Accessed November 13, 2023.

Jones, S. (2017). Corporate Bankruptcy Prediction: A High Dimensional Analysis. *Review of Accounting Studies, 22*(3), pp. 1366–1422. doi: https://doi.org/10.1007/s11142-017-9407-1.

Jones, S., W. J. Moser, and M. M. Wieland. (2023). Machine Learning and the Prediction of Changes in Profitability. *Contemporary Accounting Research, Forthcoming*. doi: https://doi.org/10.1111/1911-3846.12888.

Kang, J. K., L. Stice-Lawrence, and Y. T. F. Wong. (2021). The Firm Next Door: Using Satellite Images to Study Local Information Advantage. *Journal of Accounting Research, 59*(2), pp. 713–750. doi: https://doi.org/10.1111/1475-679X.12360.

Kim, A. G., M. Muhn, and V. Nikolaev. (2023a). *Bloated Disclosures: Can ChatGPT Help Investors Process Information?*. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4425527. Accessed November 13, 2023.

Kim, A. G., M. Muhn, and V. V. Nikolaev. (2023b). *From Transcripts to Insights: Uncovering Corporate Risks Using Generative AI*. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4593660. Accessed November 13, 2023.

Kister, N., T. Knauer, F. Sommer, and M. Wiegerling. (2022). *Analyzing the Relation between Value-Based Management and Dividend Payouts*. EAA Annual Congress 2022.

Kister, N., T. Knauer, F. Sommer, and M. Wiegerling. (2023). *Shareholders vs. Stakeholders? Value-Based Management Sophistication, Corporate Sustainability, and Financial Performance*. EAA Annual Congress 2023.

Knauer, T., L. Silge, and F. Sommer. (2018). The Shareholder Value Effects of Using Value-Based Performance Measures: Evidence from Acquisitions and Divestments. *Management Accounting Research, 41*, pp. 43–61. doi: https://doi.org/10.1016/j.mar.2018.02.001.

Kotter, J. P. (1995). Leading Change: Why Transformations Fail. *Harvard Business Review, 73*(2), pp. 59–67.

Krupa, J., and M. Minutti-Meza. (2022). Regression and Machine Learning Methods to Predict Discrete Outcomes in Accounting Research. *Journal of Financial Reporting, 7*(2), pp. 131–178. doi: https://doi.org/10.2308/JFR-2021-010.

Lev, B., and F. Gu. (2016). *The End of Accounting and the Path Forward for Investors and Managers*. Hoboken: Wiley.

Lev, B., S. Li, and T. Sougiannis. (2010). The Usefulness of Accounting Estimates for Predicting Cash Flows and Earnings. *Review of Accounting Studies, 15*(4), pp. 779–807. doi: https://doi.org/10.1007/s11142-009-9107-6.

Li, K., F. Mai, R. Shen, and X. Yan. (2021). Measuring Corporate Culture Using Machine Learning. *The Review of Financial Studies, 34*(7), pp. 3265–3315. doi: https://doi.org/10.1093/rfs/hhaa079.

Li, Q., H. Shan, Y. Tang, and V. Yao. (2023). Corporate Climate Risk: Measurements and Responses. *Review of Financial Studies, Forthcoming*.

Liu, Y., M. Ott, N. Goyal, Du Jingfei, M. Joshi, D. Chen, et al. (2019). *RoBERTa: A Robustly Optimized BERT Pretraining Approach*. arXiv. https://arxiv.org/abs/1907.11692. Accessed November 13, 2023.

Loughran, T. I., and B. McDonald. (2016). Textual Analysis in Accounting and Finance: A Survey. *Journal of Accounting Research, 54*(4), pp. 1187–1230. doi: https://doi.org/10.1111/1475-679X.12123.

Lueg, R., and U. Schäffer. (2010). Assessing Empirical Research on Value-Based Management: Guidelines for Improved Hypothesis Testing. *Journal für Betriebswirtschaft, 60*(1), pp. 1–47. doi: https://doi.org/10.1007/s11301-009-0055-9.

Mavropulo, O., M. S. Rapp, and I. A. Udoieva. (2021). Value-Based Management Control Systems and the Dynamics of Working Capital: Empirical Evidence. *Management Accounting Research, 52*, p. 100740. doi: https://doi.org/10.1016/j.mar.2021.100740.

Mikes, A. (2009). Risk Management and Calculative Cultures. *Management Accounting Research, 20*(1), pp. 18–40. doi: https://doi.org/10.1016/j.mar.2008.10.005.

Molnar, C. (2022). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable* (2nd ed.). Online (https://christophm.github.io/interpretable-ml-book): Independently Published.

Mullainathan, S., and J. Spiess. (2017). Machine Learning: An Applied Econometric Approach. *Journal of Economic Perspectives, 31*(2), pp. 87–106. doi: https://doi.org/10.1257/jep.31.2.87.

OpenAI. (2023). *GPT-4 Technical Report*. arXiv. https://arxiv.org/abs/2303.08774. Accessed November 13, 2023.

Perols, J. L., R. M. Bowen, C. Zimmermann, and B. Samba. (2017). Finding Needles in a Haystack: Using Data Analytics to Improve Fraud Prediction. *The Accounting Review, 92*(2), pp. 221–245. doi: https://doi.org/10.2308/accr-51562.

Pickard, M. D., R. Schuetzler, J. S. Valacich, and D. A. Wood. (2020). Innovative Accounting Interviewing: A Comparison of Real and Virtual Accounting Interviewers. *The Accounting Review, 95*(6), pp. 339–366. doi: https://doi.org/10.2308/tar-2017-0235.

Qiu, F., N. Hu, P. Liang, and K. Dow. (2023). Measuring Management Accounting Practices Using Textual Analysis. *Management Accounting Research, 58*, p. 100818. doi: https://doi.org/10.1016/j.mar.2022.100818.

Ranta, M., M. Ylinen, and M. Järvenpää. (2023). Machine Learning in Management Accounting Research: Literature Review and Pathways for the Future. *European Accounting Review, 32*(3), pp. 607–636. doi: https://doi.org/10.1080/09638180.2022.2137221.

Roberts, M. E., B. M. Stewart, and R. A. Nielsen. (2020). Adjusting for Confounding with Text Matching. *American Journal of Political Science, 64*(4), pp. 887–903. doi: https://doi.org/10.1111/ajps.12526.

Rouen, E., K. Sachdeva, and A. Yoon. (2023). *The Evolution of ESG Reports and the Role of Voluntary Standards*. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4261044. Accessed November 13, 2023.

Ryan, H. E., and E. A. Trahan. (2007). Corporate Financial Control Mechanisms and Firm Performance: The Case of Value-Based Management Systems. *Journal of Business Finance & Accounting, 34*(1-2), pp. 111–138. doi: https://doi.org/10.1111/j.1468-5957.2006.00660.x.

Sautner, Z., L. van Lent, G. Vilkov, and R. Zhang. (2023). Firm-level Climate Change Exposure. *The Journal of Finance, 78*(3), pp. 1449–1498. doi: https://doi.org/10.1111/jofi.13219.

Schultze, W., T. List, B. Schabert, and T. Dinh. (2018). Economic Consequences of Implementing and Communicating Value Based Management Systems. *Journal of Business Finance & Accounting, 45*(5-6), pp. 511–543. doi: https://doi.org/10.1111/jbfa.12297.

Seidel, T. A., C. A. Simon, and N. M. Stephens. (2020). Management Bias across Multiple Accounting Estimates. *Review of Accounting Studies, 25*(1), pp. 1–53. doi: https://doi.org/10.1007/s11142-019-09518-8.

Siano, F., and P. Wysocki. (2021). Transfer Learning and Textual Analysis of Accounting Disclosures: Applying Big Data Methods to Small(er) Datasets. *Accounting Horizons, 35*(3), pp. 217–244. doi: https://doi.org/10.2308/HORIZONS-19-161.

Stice-Lawrence, L. (2022). Practical Issues to Consider When Working with Big Data. *Review of Accounting Studies, 27*(3), pp. 1117–1124. doi: https://doi.org/10.1007/s11142-022-09708-x.

Stolowy, H., and L. Paugam. (2018). The Expansion of Non-Financial Reporting: An Exploratory Study. *Accounting and Business Research, 48*(5), pp. 525–548. doi: https://doi.org/10.1080/00014788.2018.1470141.

Szerwo, B., Z. Axelton, and J. Gramlich. (2022). *Auditor Responses to Estimation Uncertainty in Financial Statements*. Hawai'i Accounting Research Conference (HARC) 2022. http://hdl.handle.net/10125/77003. Accessed November 13, 2023.

Wager, S., and S. Athey. (2018). Estimation and Inference of Heterogeneous Treatment Effects Using Random Forests. *Journal of the American Statistical Association, 113*(523), pp. 1228–1242. doi: https://doi.org/10.1080/01621459.2017.1319839.

Wallace, J. S. (1997). Adopting Residual Income-Based Compensation Plans: Do You Get What You Pay for? *Journal of Accounting and Economics, 24*(3), pp. 275–300. doi: https://doi.org/10.1016/S0165-4101(98)00009-3.

Webersinke, N. (2023). *Natural Language Processing Meets Accounting and Finance: Review and Performance Comparison of Textual Analysis Approaches*. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4527724. Accessed November 13, 2023.

Wu, S., O. Irsoy, S. Lu, V. Dabravolski, M. Dredze, S. Gehrmann, et al. (2023). *BloombergGPT: A Large Language Model for Finance*. arXiv. https://arxiv.org/abs/2303.17564. Accessed November 13, 2023.