Technovation 148 (2025) 103322

Contents lists available at ScienceDirect

Technovation

journal homepage: www.elsevier.com/locate/technovation

# Predicting corporate innovation using machine learning and social media data

Mika Ylinen *, Mikko Ranta

*University of Vaasa, School of Accounting and Finance, Wolffintie 34, FI-65200, Vaasa, Finland*

ARTICLE INFO

*JEL classification:*
G3
J24
M14
O31

*Keywords:*
Social media analytics
Human resource policies
Corporate innovation
Machine learning
Textual analysis

ABSTRACT

This study explores the potential of employee reviews on social media to predict corporate innovation performance. We investigate these relationships using a novel social media dataset and an explainable machine learning approach to assess the predictive value and importance of various employee treatment policies in driving corporate innovation. In addition to traditional patent-based innovation measures, we employ a text-based innovation metric derived from 10-K filings. Our findings reveal that several employee ratings on social media provide valuable insights for predicting corporate innovation. Specifically, we highlight the importance of flexible working hours and employee stock or equity options in predicting patent counts, patent citations, and text-based innovation. Other significant predictors of patent-based innovation include employees' career growth prospects and pride in the company. Furthermore, we find that the ability to work remotely is a strong predictor of text-based innovation but is less significant for patent counts and citations. Our findings reveal notable differences in the key determinants of various types of innovation, contributing to a deeper understanding of how employee experiences associate corporate innovation outcomes.

## 1. Introduction

Corporate innovation is among the most important drivers in boosting long-term growth and the competitiveness of a firm (e.g., Holmstrom, 1989; Bellstam et al., 2020; Chang et al., 2015). Innovative accomplishments of a company are highly dependent on human capital because employees' active behaviour, engagement, and motivation drive the generation of new ideas and their implementation into new services, products, processes, and business model innovations (Zingales, 2000; Chang et al., 2015; Chen et al., 2016a). As such, to efficiently fuel corporate innovation, organizations need to harness all the potential of rank-and-file employees as innovators of a company (Chang et al., 2015). Therefore, it is critical for organizations to understand which employee treatment aspects are the most influential drivers of corporate innovation. In this study, we try to advance our understanding of this important topic.

This study is explicitly positioned as exploratory and predictive, aimed at uncovering associations between employee treatment practices and corporate innovation outcomes using advanced machine learning (ML) methods. By leveraging ML models, we identify key predictors of innovation without making causal claims, focusing instead on detecting patterns that can guide future research and theory development. To complement this predictive analysis, we incorporate explanatory elements through the use of traditional statistical models and state-of-the-art explainable AI tools. This dual approach allows us to balance predictive accuracy with theoretical depth, offering both practical insights and a foundation for future causal investigations. We seek to identify which disclosed human resource policies are most influential in predicting various forms of corporate innovation. To achieve this, we employ a novel social media analytics approach to innovation management, as detailed in previous research by Geissinger et al. (2023). This methodology enables us to utilize the abundance of publicly available employee feedback to gain a deeper understanding of how workplace practices affect innovation outcomes.

We respond to recent calls for the integration of machine learning (ML) techniques in management research (e.g., Choudhury et al., 2021; Doornenbal et al., 2021; Valizade et al., 2024), we advocate for a more comprehensive and systematic application of ML principles within quantitative innovation management studies. This includes using ML methods alongside traditional statistical modeling where they complement each other, and as a substitute where ML offers distinct advantages. Unlike many previous studies, we employ a nonlinear and

* Corresponding author.
*E-mail addresses:* mika.ylinen@uwasa.fi (M. Ylinen), mikko.ranta@uwasa.fi (M. Ranta).

https://doi.org/10.1016/j.technovation.2025.103322
Received 23 August 2024; Received in revised form 14 June 2025; Accepted 19 July 2025
Available online 9 August 2025

explainable ML-driven methodology to identify practically relevant patterns in social media data that most effectively predict future corporate innovation. This algorithmic approach to quantitative data analysis aims to advance theory-generating research and uncover surprising, counterintuitive insights that challenge existing norms and expectations. It achieves this by enabling the analysis of complex, non-monotonic, and nonlinear effects that traditional statistical models often overlook (e.g., Choudhury et al., 2021; Doornenbal et al., 2021; Shrestha et al., 2021; Valizade et al., 2024). Our ML-based research approach enables us to identify key features from employee reviews that significantly predict corporate innovation.

Social media platforms have transformed corporate disclosure channels, facilitating the dissemination of information by non-professionals. Recently, these platforms have simplified the process for individual employees to share and access opinions about their firms. Prior studies have demonstrated that these platforms offer information relevant to forecasting future corporate performance (e.g., Huang et al., 2020; Green et al., 2019; Hales et al., 2018). Consequently, information from social media presents a valuable and potential avenue for uncovering new insights about a company and its management. However, prior innovation management literature has largely neglected this promising source of information and the role of employee reviews in revealing information valuable in predicting corporate innovation. Following the recent call for research in utilizing social media analytics in innovation management research (see Geissinger et al., 2023), our goal is to study the practical relevance (see Sarstedt and Danks, 2022) of social media information related to corporate innovation. Our analysis of social media data highlights potential drivers of innovation, including diversity, equity, and inclusion; employee compensation through stock or equity options; career development opportunities; work-life balance; freedom and autonomy; job security; and company pride.

Long-standing research has shown evidence for the crucial role of employee treatment in enhancing company performance (e.g. Cao and Rees, 2020; Edmans, 2011; Edmans et al., 2014; Jiao, 2010; Faleye and Trahan, 2011; Huang et al., 2015; Fauver et al., 2018; Ranta and Ylinen, 2023a; Ylinen and Ranta, 2023). However, considerably less attention has been paid to corporate innovation and to identifying which specific HR policies are the most important drivers of corporate innovation. There is a gap in solid evidence regarding the effectiveness of HR policies in fostering innovation within organizations and the specific factors that influence this dynamic (e.g. Krammer, 2022). Prior research has either aggregated various dimensions of employee treatment or focused solely on a single practice, potentially leading to an incomplete understanding of how specific practices and benefits can promote innovation. The existing literature (e.g., Acar et al., 2019; Amabile, 1988; Amabile et al., 1996; Anderson et al., 2014; Shalley et al., 2004; Shipton et al., 2006) indicates that a variety of employee treatment practices are available for companies to implement, each of which is a theoretically significant driver of corporate innovation. However, it is likely that not all employee treatment practices are equally effective in enhancing different types of innovation.

As such, previous studies do not provide clear evidence or guidance as to which particular HR policies are most important for companies to focus on and whether there are any differences in stimulating various types of innovation. To achieve meaningful, practical outcomes, research must advance and guide innovation policies by identifying the most effective employee treatment practices that enhance corporate innovation. We aim to provide evidence about the practical relevance of employee treatment practices by adopting a primarily predictive approach. Using ML methods, we identify key predictors of corporate innovation and supplement this with explanatory analyses to deepen theoretical insights into these relationships (see e.g. Sarstedt and Danks, 2022). This helps companies develop and select employee treatment practices, policies, and strategies tailored to various types of innovation.

This research differs from earlier studies in several ways. First, in this study, we utilize social media data analytics to uncover detailed insights into the relationships between various HR policies and corporate innovation. As such, by capturing these complex patterns in data, our ML-driven approach can provide advancements to current explanatory models (e.g. Choudhury et al., 2021; Doornenbal et al., 2021; Valizade et al., 2024). Future research can utilize these identified patterns as observations to enhance theories using exploratory inductive or abductive research methods (Choudhury et al., 2021; Ranta et al., 2023). Additionally, these findings can aid researchers in forming more robust hypotheses based on data, which can subsequently be tested deductively with traditional econometric tools. This ML-led approach to pattern discovery, though atheoretical by nature, can unearth significant insights that enable researchers to identify key variables for inclusion in theoretical causal models (Bertomeu et al., 2021; Ranta et al., 2023; Valizade et al., 2024). Therefore, it provides fresh perspectives and insights into which theories might effectively explain a sample (Bertomeu, 2020). Furthermore, the pattern discovery and interpretation capabilities of ML demonstrated in this study can enhance post hoc analyses of conventional models, revealing previously unnoticed patterns.

Secondly, besides widely used innovation proxies of patent counts and citations, we employ a novel procedure to create a text-based corporate innovation measure suggested by Bellstam et al. (2020). In particular, we use latent Dirichlet allocation (LDA), a topic modelling method, to employ a text-based innovation measure from 10-K fillings. These filings are annual reports required by the U.S. Securities and Exchange Commission (SEC) that provide a detailed overview of a company's financial performance, risk factors, and other material information. They serve as a critical source of structured and comprehensive data for understanding corporate activities, including innovation-related efforts. By adhering to the established framework Bellstam et al. (2020), we address concerns regarding the validity of our initial approach while maintaining the ability to capture a broader spectrum of innovation activities. This represents a significant difference, as many innovation measures used in prior studies rely solely on either cross-sectional survey instruments or, in the case of longitudinal data, on patenting and R&D expenditures. Innovation extends far beyond the creation and introduction of new products and patents.[1] As such, a text-based measure of innovation can reflect a broader spectrum of a firm's innovative activities beyond product launches, encompassing new production methods, novel supply sources, the exploitation of new markets, and new organizational structures (Bellstam et al., 2020). By considering a more comprehensive view of measuring innovation, our ML approach provides new aspects and enables an understanding of the associations between employee treatment and corporate innovation performance in a more detailed manner. This is a significant advantage because a text-based innovation measure can be applied to all firms, including those that do not generate patents or invest in R&D.[2]

While we propose two guiding premises informed by prior literature, our methodological approach remains primarily exploratory. Prior studies suggest that employee-generated reviews on social media contain meaningful insights about workplace practices and culture (e.g., Green et al., 2019; Hales et al., 2018; Huang et al., 2020) that could relate to innovation outcomes. We build on this work by exploring whether such reviews are indeed predictive of firm-level innovation. Additionally, we examine whether particular dimensions of employee experience, such as autonomy, flexible work arrangements, or equity-based compensation, tend to emerge as stronger predictors.

[1] For example, a study by Hall et al. (2013) argues that only about 4 % of firms that innovate are creating patents. Similarly, Bellstam et al. (2020) found that within their sample of 703 firms from the S&P 500 during the period 1990–2010, 219 firms did not have any patents and 329 firms had no R&D.

[2] Bellstam et al.'s (2020) study shows clear evidence that text-based measures capture innovative accomplishments more broadly beyond the narrowly focused view of product innovations that have used traditionally measured proxies based on patent counts and citations.

However, we do not assume in advance which specific practices are most relevant. Instead, we allow the data-driven analysis to surface the most salient associations, acknowledging that other, less-emphasized practices could also play an important role.

Rather than relying on pre-specified functional forms or selecting variables based strictly on theoretical assumptions, we use ML methods to explore complex, potentially nonlinear associations. Our approach is especially well suited to this context, where the existing literature does not provide definitive guidance on which aspects of employee treatment are most critical to innovation. To complement the predictive modeling, we incorporate explanatory tools such as SHAP values and follow up with instrumental variable analysis to cautiously explore potential causal mechanisms. Throughout, we interpret our findings primarily as associations, and we clearly distinguish between predictive insights and causal interpretations.

This approach enables us to address several interrelated questions. Can social media based employee reviews predict corporate innovation? What types of HR practices most strongly relate to innovation outcomes? Are the relationships linear or more complex in form? And how much each employee treatment factor contributes to the model's predictive power?

Building on this approach, we also acknowledge the distinction between prediction and causality. ML techniques such as the ones employed in this study are powerful tools for detecting complex, nonlinear associations but do not inherently support causal claims. To probe potential causal relationships, we supplement our analysis with instrumental variable techniques. Nonetheless, we interpret these findings with appropriate caution and consistently frame our results as associative rather than definitively causal unless clearly justified. This distinction is important for accurately communicating the scope and implications of our findings.

The rest of the paper is structured as follows: Section 2 reviews the existing literature and outlines the motivation for our research questions. Section 3 describes the employer review sample, presents the measures for research variables and descriptive statistics, and outlines the research methods. The empirical findings are presented in Section 4. The discussion of the findings and the conclusion are presented in Section 5.

## 2. Related research literature

### 2.1. Social media data and firm performance

Our paper adds to a growing line of research that uses various websites and social media platforms as possible sources of company-specific information (for extensive reviews, see Geissinger et al., 2023; Lei et al., 2019). Previous research has demonstrated the predictive power of social media on various corporate outcomes. Bartov et al. (2018) and Tang (2017) highlighted Twitter's role in forecasting earnings and sales, respectively, while Saura et al. (2023) applied social media analytics to gain insights into open innovation. Elliott et al. (2018) and Jung et al. (2018) discussed the impact of CEO Twitter usage and corporate communications on investor trust and the likelihood of sharing negative news. Cade (2018) noted that Twitter criticism affects investor perceptions based on retweet frequency. Ozcan et al. (2021) mined tweets to identify trends and extract ideas related to product development, technology, and sustainability. They developed a classification model to pinpoint innovative ideas in tweets. Their findings show that text mining combined with supervised or semi-supervised classification techniques effectively retrieves ideas from social media.

Moreover, studies have shown that social media platforms like Seeking Alpha and Estimize offer valuable forecasts on stock returns and earnings expectations (Campbell et al., 2019; Chen et al., 2014; Jame et al., 2016; Kelley and Tetlock, 2013) and reviews on platforms like Amazon also predict firm performance (Zhu and Zhang, 2010; Huang, 2018). Similarly, employee-generated content on platforms such as Glassdoor provides insights into company performance and investor returns. Studies by Huang et al. (2015) and Hales et al. (2018) have found that employee reviews correlate with firm valuation and future corporate disclosures. Further, research by Sheng (2019) and Green et al. (2019) has shown that employee outlook and satisfaction levels can predict stock returns and earnings surprises, underscoring the value of employee sentiment in forecasting firm performance.

Overall, these findings support the argument that employee ratings in social media enclose vital information. These recent studies conclude that employee reviews contain new and valuable information about firms' fundamentals, and employees as company insiders possess valuable information about their firms. However, these studies show less clear and mixed results in terms of the value of other types of employee review ratings, such as information related to various HR policies like employee assessments of career opportunities, organizational culture, work-life balance, compensation, job security, autonomy, and flexibility in working conditions. Additionally, prior research has not explored whether information from employee reviews on social media can predict corporate innovation.

Building on this literature, our study explores whether employee-generated content on social media, particularly reviews of workplace experiences, can offer predictive insights into firms' innovation outcomes. While prior research has demonstrated the informational value of such content for financial performance, its role in forecasting innovation remains underexplored. We therefore propose the following hypothesis:

**H1**. *Employee-generated social media reviews contain valuable information on corporate culture and HR practices that is predictive of corporate innovation outcomes.*

### 2.2. Employee treatment and corporate innovation

Secondly, our paper adds to the recent and growing literature on the relationship between employee-friendly workplaces and firm performance. In particular, our paper contributes to recent studies that examine the role of good employee treatment and various HR policies in stimulating corporate innovation.

Several prior studies have shown evidence that good employee treatment is positively associated with better future financial performance (Jiao, 2010; Bae et al., 2011; Edmans, 2011; Faleye and Trahan, 2011; Ertugrul, 2013; Huang et al., 2015; Fauver et al., 2018) and firm innovation (e.g., Chen et al., 2016a; Chen et al., 2016b; Mao and Weathers, 2019). The fundamental theoretical rationale for the positive impact of good employee treatment on company performance lies in its potential to boost employee satisfaction, commitment, loyalty, and productivity. This, in turn, supports lower turnover, reduced absenteeism, and ultimately leads to increased profitability and higher firm value (Edmans, 2011, 2012; Faleye and Trahan, 2011; Huang et al., 2015; Fauver et al., 2018).

Although the fast-growing body of literature has provided clear empirical evidence supporting the effects of employee treatment practices on financial performance, their associations with corporate innovation are less studied, even if corporate innovation can be a channel through which positive employee treatment affects firm performance by enhancing and creating competitive advantages (e.g., Chen et al., 2016b; Mao and Weathers, 2019).

Chen et al. (2016a) highlighted that favourable employee treatment is positively associated with patent counts and citations, fostering a failure-tolerant environment that boosts experimentation and innovation, particularly in high-risk industries. Contrary to the notion that such treatment primarily attracts or retains top talent, their findings emphasize the role of positive workplace practices in enhancing innovation sustainability against shocks, based on data from the KLD database and Best Companies list. Furthering this perspective, Chen et al. (2016b) identified additional pathways through which positive employee treatment impacts innovation, suggesting that it not only attracts and retains talented employees but also contributes to higher patent value and improved operating performance through efficient patent implementation and enhanced innovation strategies.

Krammer (2022) provides substantial empirical evidence supporting the impact of job autonomy and performance-based pay policies in driving firm innovation. Additionally, there is some evidence suggesting that the influence of institutional and competitive environments may moderate this relationship. Research including studies by Manso (2011) and Ederer and Manso (2013) underline job security and incentive schemes that balance short-term failure tolerance with long-term rewards as critical for motivating innovation. Chang et al. (2015) found that stock options for non-executive employees significantly encourage corporate innovation, evidenced by increased patent output. Conversely, Bradley et al. (2017) showed that labour unionization and reduced R&D spending negatively affect innovation productivity and could lead firms to relocate innovation activities, underscoring the complex impact of employee treatment practices on corporate innovation.

Our study builds on the existing literature by focusing on a small number of key HR practices identified through ML methods, such as flexible working hours and stock options, to predict corporate innovation. We then analyze these predictors in greater depth to develop stronger theoretical insights into their role in driving innovation. We employ a novel social media dataset to examine firms' employee treatment practices using both traditional patent-based measures and a text-based innovation measure to capture a broader array of innovation achievements. Like Green et al. (2019) and Huang et al. (2020), we aim to assess the predictive value of this social media information. Our research contributes further by adopting an explainable, ML-driven approach to uncover the most impactful employee treatment practices that enhance various types of corporate innovation. We detail these associations by analyzing their individual contributions and the dynamics of these relationships in the context of innovation prediction.

Building on prior literature, our study examines various theoretical perspectives to identify employee-friendly HR practices that promote corporate innovation. Our ML models include the following theoretically significant drivers of innovation based on prior research: job security (Manso, 2011; Acharya et al., 2014; Tian and Wang, 2014); diversity, equity, and inclusivity (DEI) values within company culture (Østergaard et al., 2011); freedom, autonomy, and flexibility in work arrangements (Amabile et al., 1996; Anderson et al., 2014; Krammer, 2022); and compensation features like employee stock or equity options (Manso, 2011; Chang et al., 2015). Another theoretical perspective suggests that good employee treatment can enhance innovation by attracting and retaining talented innovators (Chen et al., 2016b). In line with this, our research model also considers less-studied factors, such as work-life balance, organizational pride (company image), and employee career opportunities (Anderson et al., 2014), as potential drivers of innovation.

While prior studies suggest that employee treatment practices can foster innovation, it remains unclear which specific HR-related dimensions are most closely linked to innovation outcomes. Our research seeks to clarify this by identifying which workplace factors, as captured through employee reviews, most strongly predict corporate innovation. Accordingly, we propose a second hypothesis:

**H2.** *There is a positive association between employee-friendly HR practices and corporate innovation, with certain practices (*e.g.*, flexible work arrangements, autonomy, equity compensation) having a stronger predictive impact than others.*

To empirically evaluate these hypotheses, we adopt a methodological framework that balances data-driven discovery with explanatory analysis. While prior research on the relationship between employee treatment and innovation has predominantly relied on parametric methods such as OLS regression, our study takes a different approach by employing machine learning techniques. Rather than specifying a predefined functional form or selecting variables based on strict theoretical assumptions, we allow the data to guide the selection of models and relationships. This approach is particularly suited to contexts such as ours where the theoretical and empirical literature does not yet provide clear guidance on which aspects of employee treatment are most strongly associated with innovation outcomes (Karolyi and Van Nieuwerburgh, 2020). As Bertomeu et al. (2021) note, machine learning methods are especially valuable in settings where complex, nonlinear, and possibly interacting relationships may exist but are not well captured by conventional linear models. Accordingly, our framework is designed to identify relevant predictors from a broad set of employee review variables, while later stages of analysis such as SHAP interpretation and two-stage least squares estimation allow for further explanatory insight.

## 3. Data and methodology

### 3.1. Measuring employee treatment

We measure employee treatment by using novel social media data. In particular, we collect information from Kununu, a social media-based recruiting website, which data has been used recently in prior research (e.g., Ranta and Ylinen, 2023a; Cai et al., 2024). Detailed definitions for all the variables collected from the website and other research variables are reported in Appendix Table A. As of 2019, Kununu hosted over 3.96 million individual employer reviews across more than 933,000 companies. From this dataset, we compiled a sample of 250,000 reviews representing 493 firms from the S&P 1500. For robustness purposes, we included only those firms with at least 20 observations per year from 2014 to 2019.[3] The sample includes heterogenous group of companies well distributed across industries, as the data was collected according to these two criteria (S&P1500 and at least 20 reviews per year). This ensures a more representative dataset for analyzing employee treatment practices and corporate innovation. However, we acknowledge the potential bias due to the reliance on social media data, as firms with more active employee engagement on platforms like Kununu may be overrepresented. Although our sample is skewed toward larger, publicly listed firms, Appendix D shows that differences in employee review scores between S&P 1500 companies and the broader firm population are modest. We therefore consider it unlikely that this selection introduces substantial bias into our results, though it may limit the generalizability of our findings to smaller or privately held firms.

In our research model, we incorporate the following characteristics related to employee treatment: *work-life balance, freedom to work independently, inclusive/diverse values, gender equality, attitudes towards older colleagues, handicapped accessibility, job security, pride in the company, career development*. The following employee benefits are included in ML models: *stock or equity options, flexible working hours, ability to work remotely, paid parental leave,* and *paid time off*.

Our measure of employee treatment has four clear advantages. First, Kununu data is based on employees' perceptions of their companies and management, rather than on values proclaimed by management. This distinction is crucial for reliable measures of employee treatment, as previous studies have shown that employee perceptions on social media are significant for firm performance (e.g., Huang et al., 2015; Green et al., 2019; Huang et al., 2020), whereas proclaimed values are often inconsequential (Guiso et al., 2015). Second, our dataset includes over 250,000 employee reviews from a wide range of firms, offering substantial time and cross-sectional variations. This breadth allows for a

[3] We used data from Kununu US, which includes reviews from employees who are working in the US. This data is not anymore publicly available as Kununu has retaired US markets.

comprehensive examination of employee treatment characteristics. Third, compared to previously used data (e.g., Landier et al., 2009; Bae et al., 2011; Faleye and Trahan, 2011; Ghaly et al., 2015; Chen et al., 2016a; Chen et al., 2016b; Fauver et al., 2018; Mao and Weathers, 2019) on employee treatment from sources such as the KLD Socrates and ASSET4 databases or employee reviews from Glassdoor, which include only nine review attributes, our dataset provides a more detailed and comprehensive range of features capturing various aspects of employee treatment practices. The employee opinion variables are averages of observations for the whole year and lagged by one year relative to independent variables in all ML models. All variables are winsorized at the 1st and 99th percentiles to reduce the potential impact of outliers.

### 3.2. Innovation measures

#### 3.2.1. Text-based innovation

We use three different measures for corporate innovation. First, we follow the established procedure introduced by Bellstam et al. (2020) to construct a text-based innovation measure using the LDA method of Blei et al. (2003). By adopting this methodology, we leverage a framework that has undergone extensive testing and validation, thereby enhancing the validity and robustness of our analysis. We use 10-K filings to form the corpus for our LDA analysis. These filings, required by the SEC, provide a structured and comprehensive view of a firm's financial performance and other material disclosures, making them a valuable resource for analyzing corporate innovation. We assume that the textual content of these filings reveals valuable data on firms' existing innovation activities and accomplishments. This approach ensures consistency with prior research while capturing a broad spectrum of innovative activities, extending beyond traditional patent-based metrics. Furthermore, recent research suggests that information distilled by third parties might be biased and most unbiased and timely information on strategic decisions comes directly from the companies themselves (Kimbrough et al., 2022). Several prior studies have already used ML methods to study 10-K filings, and the number of new papers is increasing (e.g., Lehavy et al., 2011; Hoberg and Maksimovic, 2015; Kim et al., 2019; Frankel et al., 2016; Buehlmaier and Whited, 2018; Basu et al., 2021; Donovan et al., 2021). For example, a recent study by Lu and Chesbrough (2022) used LDA topic modelling and 10-K filings to measure open innovation practices. Whereas, Nousiainen et al. (2024) used 10-K filings to measure corporate innovation.

When applied to our corpus of 10-K filings, the LDA model assumes the filings are composed of a fixed number of topics represented as a multinomial distribution over a predetermined vocabulary. The model assumes that the texts are formed by sampling a mixture of the topics, and the words are sampled from the topic distributions. This approach follows an intuition that texts include multiple topics in different proportions. The model is trained using the Bayesian approach that searches for such parameters that best explain the data (the text of 10-K filings). First, the algorithm chooses a random distribution over the topics, then picks one of these topics and randomly chooses a word from that topic. While doing this iteratively, the model parameters are optimized continuously until the topic and word distributions explain the data adequately. This optimization is usually done using variational inference (Jordan et al., 1999). For a more thorough review of the mathematical details of LDA, see Blei et al. (2003) and Hoffman et al. (2010), and for the details of creating our text-based innovation measure, see Bellstam et al. (2020).

To construct our primary innovation measure, we adopt the LDA-based text innovation score proposed by Bellstam et al. (2020), which identifies firm-level innovation content from 10-K filings. To ensure consistency with prior work, we adopted a similar text preprocessing pipeline. Each 10-K document was cleaned through tokenization, lemmatization, and removal of common English stop words (the stop words list of the NLTK library in Python). We retained financial and domain-relevant terms.

The number of topics in the LDA model is an important methodological decision that affects both the level of detail and the interpretability of the results. To select the optimal number of topics, we evaluated models ranging from 10 to 60 topics using two standard evaluation metrics: perplexity and coherence scores (Röder et al., 2015). As shown in Appendix B, Table B1, the model with 40 topics provided the best balance between lower perplexity and higher coherence. This model also produced the most interpretable output based on manual inspection. Models with fewer topics, such as 10 or 20, tended to combine distinct themes into broad categories, while models with more than 40 topics often split meaningful themes into overly narrow subtopics. Based on these quantitative and qualitative criteria, we selected the 40-topic model for our main analysis.

To address concerns about the sensitivity of topic modeling results to the number of latent topics specified, we conducted a robustness check comparing LDA models with 30, 40, and 50 topics. As detailed in Appendix C, the core semantic content of the identified innovation topic remains highly stable across these specifications. Key terms such as "solution," "device," and "patent" consistently appear across models, supporting the robustness of our text-based innovation measure to reasonable variations in topic number.

We follow the methodology developed by Bellstam et al. (2020) and compute the Kullback-Leibler (KL) divergence between the word distributions of 40 LDA-generated topics and a reference text on innovation. This reference, consistent with Bellstam et al. (2020), is the book *Managing Innovation* by Tidd et al. (2005), a comprehensive source on innovation theory and practices. The topic with the lowest KL divergence to this text is selected as our innovation topic. The topic's document-level loading is then used as our firm-year text-based innovation score.

To assess the validity of the LDA-derived innovation measure, we conduct multiple robustness checks and benchmark comparisons. First, we evaluate its association with conventional innovation outputs, specifically patent filings and citations, using baseline OLS regressions. The results, shown in Appendix E Table E1, indicate that the Bellstam innovation score is positively and significantly related to both patent metrics, even after controlling for key firm characteristics such as research and development intensity, firm size, profitability, and capital intensity.

To further benchmark the effectiveness of our text-based innovation measure, we construct an alternative metric using a word embedding approach following Li et al. (2020). This method calculates the semantic proximity of each firm's annual filing to a vector space defined by innovation-related seed phrases. While this embedding-based measure shows positive associations with patent outcomes, our results indicate that the Bellstam LDA-based innovation score demonstrates consistently stronger statistical relationships, as reported in Appendix E Table E1.

In addition to regression-based validation, we evaluate the predictive performance of both text-based innovation measures using machine learning models. We implement gradient boosting models to predict patent filings and citations and apply SHAP value analysis to interpret variable importance. As shown in Figures F1 and F2 in Appendix F, the Bellstam innovation score provides superior predictive power compared to the embedding-based metric and ranks higher in feature importance across both innovation outcomes. This provides further support for the validity and utility of the LDA-derived innovation measure.

#### 3.2.2. Patent-based measures

We extract patent information from the dataset of Kogan et al. (2017). This database is created from Google patent data and complemented with the hand-collected reference data of Nicholas (2008). This database has information for 1,928,123 matched patents, and 523, 301 (27 %) of these are not included in the NBER data (Kogan et al., 2017). From this patent data, we create two measures for corporate innovation. First, we construct a measure for innovation quantity by using a measure of a firm's total number of patent applications filled in a given year that are eventually granted. Following previous studies (e.g., Chang et al., 2015; Chen et al., 2016a; Mao and Weathers, 2019), we create our measure using a patent application year instead of its grant year to capture the realistic time of innovation. As a second measure of firm innovation, we measure the total number of non-self patent citations that the patent receives in the following years. This measures patent quality by taking into account the patent's impact and distinguishes more fundamental, radical innovations from incremental discoveries. As a patent approval process typically takes years, there is, on average, a two-year time lag between patent applications and patent grants (Bradley et al., 2017). This will lead to a decrease in the total number of patent applications (that will be later granted) in our sample, especially during the last couple of years at the end of the sample period. We solve this truncation problem by following guidelines given in existing research (e.g., Bradley et al., 2017; Chang et al., 2015; Chen et al., 2016a; Chen et al., 2016b; Mao and Weathers, 2019) by using correction weights counted from the empirical distribution of application-grant numbers as suggested by Hall et al. (2001). Similarly, we use the shape of the distribution related to citation lag to correct the second truncation problem caused by the time lag in accumulated patent citation numbers at the end of the sample period.

### 3.3. Control variables

Following prior literature, we include several control variables to isolate employee treatment practices' effect on innovation and examine their relative importance and predictive power. These firm-level control variables have been reported in the extant literature as important characteristics that may be related to future firm innovation (Chang et al., 2015; Bradley et al., 2017; Chen et al. 2016a, 2016b; Mao and Weathers, 2019; Bellstam et al., 2020). The control variables include firm size measured by total assets, as larger firms often have more resources to invest in R&D and innovation, which could influence innovation outcomes. The next control variable is leverage calculated as the ratio of long-term debt to total assets. The ratio of long-term debt to total assets accounts for a firm's capital structure and has potential impact on innovation investments. We also control for R&D intensity assessed as the ratio of R&D expenditure to total assets. This variable captures the firm's direct investment in innovation activities. We also control profitability using return on asset (ROA) as the ratio of operating income over book assets. ROA reflects the firm's operational efficiency, which could affect its ability to fund and sustain innovation.

Tobin's Q is also included as a control variable, as it captures growth opportunities and the firm's ability to attract external funding for innovation. We also control company risk through implied volatility, as risk may influence a firm's innovation strategy and outcomes. To control for the role of physical assets in supporting innovation, we control capital intensity, calculated as the ratio of net property, plant, and equipment over the number of employees, Finally, we control sales growth and firm age. Sales growth works as a proxy for market performance. Regarding firm age, older firms may have more established processes and resources but may also face challenges related to organizational inertia, which could affect innovation. To account for state-specific factors such as the wrongful discharge laws introduced by Acharya et al. (2014), we include state controls in our models. The model also controls for year and industry effects, with industry effects managed using two-digit SIC code dummies. The control variables are the end-of-the-fiscal-year values. All control variables are lagged by one year, relative to the dependent variables. A detailed definition of these control variables is provided in Appendix 1.

### 3.4. Employer review descriptive statistics

Table 1 provides descriptive statistics and Table 2 correlation coefficients for the employee opinion variables. To improve clarity and ease of interpretation, we have also included a heatmap to visually highlight the most relevant correlations between employee treatment practices, control variables, and innovation outcomes. To improve the heatmap, the diagonal elements of the correlation matrix (correlation with the variable itself) were manually changed to zero. The total number of employee reviews is approximately 250,000, and the number of features is 14. The reviews are utilized to compute annual averages for each firm. The total number of observations is 2,584, except for the feature *inclusive/diverse,* where the number is 2,170, because this question was introduced to the web portal later than the others. No significant anomalies are present in the statistics. The mean values are 3.295–3.688 for the variables with domain 1–5 and 0.165–0.322 for the variables with domain 0–1.

### 3.5. A tree-based gradient-boosting model

Our methodological approach follows a structured three-stage framework. First, we employ an extreme gradient boosting (XGBoost) machine learning model to predict firm-level innovation outcomes based on a wide range of employee experience and workplace-related variables. Second, we use SHAP (SHapley Additive exPlanations) values to interpret the predictive model and identify which features contribute most significantly to the model's predictions. This stage provides insight into the relative importance of various factors, without implying causality. Finally, to investigate potential causal relationships, we conduct two-stage least squares (2SLS) regression analyses using an external instrument. These regression-based analyses are clearly distinguished from the prediction and interpretation stages and are interpreted with appropriate caution to avoid overstating causal claims.

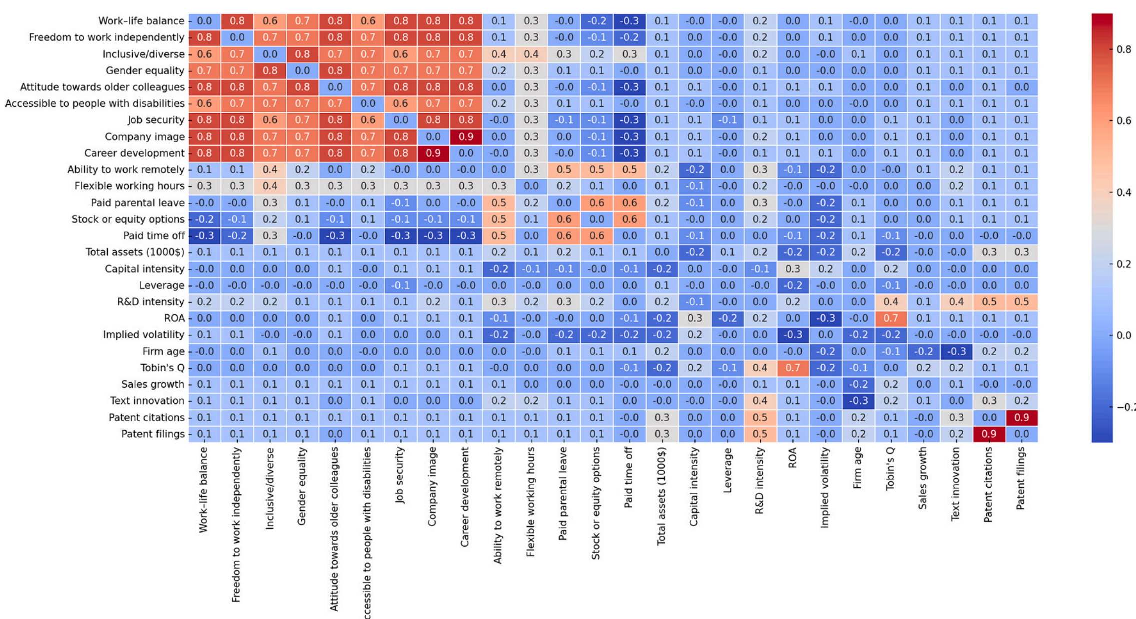
Ensemble methods refer to a category of machine learning approaches that rely on the combined predictions of multiple simple predictors. Common ensemble methods include bagging, random forests, and boosting. The gradient-boosting algorithm (Friedman et al., 2000; Friedman, 2001, 2002) applies a gradient descent technique within the boosting framework, significantly improving its efficiency for applications. Previous research has found ensemble methods often to be the most efficient option for prediction setting (Barboza et al., 2017; Gu et al., 2020), and they have been applied in various contexts, such as identifying key innovation determinants (Xu et al., 2024), predicting financial failure (Jones, 2017; Jiang and Jones, 2018), determining factors influencing citations in academic papers (Jones and Alam, 2019), detecting fraud (Bao et al., 2020; Gepp et al., 2021), enhancing managerial estimations in corporate reporting (Ding et al., 2020), forecasting corporate bond defaults (Lu and Zhuo, 2021), discovering the most important employee-friendly corporate culture values in predicting firm value and profitability (Ylinen and Ranta, 2023), and examining the value and relative importance of various employee benefits in predicting company performance (Ranta and Ylinen, 2023a).

This study employs a predictive research design, utilizing gradient boosting to identify key predictors of corporate innovation. Our

**Table 1**
Descriptive statistics for the variables.

| | count | mean | std | min | 25 % | 50 % | 75 % | max |
|---|---|---|---|---|---|---|---|---|
| Work–life balance | 2584 | 3.360 | 0.682 | 1.000 | 2.929 | 3.369 | 3.818 | 5.000 |
| Freedom to work independently | 2583 | 3.523 | 0.664 | 1.000 | 3.125 | 3.560 | 4.000 | 5.000 |
| Inclusive/diverse | 2170 | 3.418 | 0.751 | 1.000 | 3.000 | 3.500 | 4.000 | 5.000 |
| Gender equality | 2584 | 3.627 | 0.677 | 1.000 | 3.200 | 3.700 | 4.107 | 5.000 |
| Attitude towards older colleagues | 2584 | 3.620 | 0.673 | 1.000 | 3.213 | 3.667 | 4.068 | 5.000 |
| Accessible to people with disabilities | 2584 | 3.688 | 0.652 | 1.000 | 3.286 | 3.761 | 4.130 | 5.000 |
| Job security | 2584 | 3.295 | 0.738 | 1.000 | 2.788 | 3.333 | 3.831 | 5.000 |
| Company image | 2584 | 3.592 | 0.691 | 1.000 | 3.143 | 3.627 | 4.053 | 5.000 |
| Career development | 2584 | 3.303 | 0.724 | 1.000 | 2.818 | 3.333 | 3.814 | 5.000 |
| Ability to work remotely | 2584 | 0.165 | 0.194 | 0.000 | 0.000 | 0.100 | 0.261 | 1.000 |
| Flexible working hours | 2584 | 0.322 | 0.207 | 0.000 | 0.182 | 0.308 | 0.444 | 1.000 |
| Paid parental leave | 2584 | 0.194 | 0.189 | 0.000 | 0.006 | 0.158 | 0.300 | 1.000 |
| Stock or equity options | 2584 | 0.257 | 0.230 | 0.000 | 0.042 | 0.221 | 0.407 | 1.000 |
| Paid time off | 2584 | 0.490 | 0.304 | 0.000 | 0.255 | 0.541 | 0.721 | 1.000 |
| Total assets (1000$) | 2666 | 40224 | 64780 | 1032 | 4009 | 12047 | 39266 | 250165 |
| Capital intensity | 2664 | 0.037 | 0.031 | 0.001 | 0.013 | 0.030 | 0.054 | 0.109 |
| Leverage | 2655 | 1.123 | 1.432 | −1.552 | 0.345 | 0.753 | 1.493 | 5.142 |
| R&D intensity | 1503 | 0.021 | 0.030 | 0.000 | 0.000 | 0.007 | 0.029 | 0.106 |
| ROA | 2666 | 0.057 | 0.051 | −0.029 | 0.018 | 0.051 | 0.089 | 0.164 |
| Implied volatility | 1584 | 29.031 | 8.612 | 17.000 | 22.700 | 27.500 | 34.100 | 49.270 |
| Firm age | 2666 | 403.845 | 226.235 | 0.000 | 228.000 | 372.000 | 576.000 | 837.000 |
| Tobin’s Q | 2666 | 2.033 | 1.063 | 0.981 | 1.219 | 1.692 | 2.438 | 4.810 |
| Sales growth | 2651 | 0.052 | 0.099 | −0.123 | −0.008 | 0.041 | 0.095 | 0.292 |
| Text innovation | 2065 | 0.047 | 0.077 | 0.000 | 0.000 | 0.011 | 0.053 | 0.284 |
| Patent citations | 2332 | 295.486 | 751.231 | 0.000 | 0.000 | 0.000 | 51.340 | 2933.353 |
| Patent filings | 2666 | 43.679 | 112.288 | 0.000 | 0.000 | 0.000 | 6.772 | 436.809 |

**Table 2**
Correlation matrix.




| | Work–life balance | Freedom to work independently | Inclusive/diverse | Gender equality | Attitude towards older colleagues | Accessible to people with disabilities | Job security | Company image | Career development | Ability to work remotely | Flexible working hours | Paid parental leave | Stock or equity options | Paid time off | Total assets (1000$) | Capital intensity | Leverage | R&D intensity | ROA | Implied volatility | Firm age | Tobin's Q | Sales growth | Text innovation | Patent citations | Patent filings |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Work–life balance | 0.0 | 0.8 | 0.6 | 0.7 | 0.8 | 0.6 | 0.8 | 0.8 | 0.8 | 0.1 | 0.3 | -0.0 | -0.2 | -0.3 | 0.1 | -0.0 | -0.0 | 0.2 | 0.0 | 0.1 | -0.0 | 0.0 | 0.1 | 0.1 | 0.1 | 0.1 |
| Freedom to work independently | 0.8 | 0.0 | 0.7 | 0.7 | 0.8 | 0.7 | 0.8 | 0.8 | 0.8 | 0.1 | 0.3 | -0.0 | -0.1 | -0.2 | 0.1 | 0.0 | -0.0 | 0.2 | 0.0 | 0.1 | 0.0 | 0.0 | 0.1 | 0.1 | 0.1 | 0.1 |
| Inclusive/diverse | 0.6 | 0.7 | 0.0 | 0.8 | 0.7 | 0.7 | 0.6 | 0.7 | 0.7 | 0.4 | 0.4 | 0.3 | 0.2 | 0.3 | 0.1 | 0.0 | -0.0 | 0.2 | 0.0 | -0.0 | 0.1 | 0.0 | 0.1 | 0.1 | 0.1 | 0.1 |
| Gender equality | 0.7 | 0.7 | 0.8 | 0.0 | 0.8 | 0.7 | 0.7 | 0.7 | 0.7 | 0.2 | 0.3 | 0.1 | 0.1 | -0.0 | 0.1 | 0.0 | -0.0 | 0.1 | 0.0 | -0.0 | 0.0 | 0.0 | 0.1 | 0.1 | 0.1 | 0.1 |
| Attitude towards older colleagues | 0.8 | 0.8 | 0.7 | 0.8 | 0.0 | 0.7 | 0.8 | 0.8 | 0.8 | 0.0 | 0.3 | -0.0 | -0.1 | -0.3 | 0.1 | 0.1 | -0.0 | 0.1 | 0.1 | 0.1 | 0.0 | 0.0 | 0.1 | 0.0 | 0.1 | 0.0 |
| Accessible to people with disabilities | 0.6 | 0.7 | 0.7 | 0.7 | 0.7 | 0.0 | 0.6 | 0.7 | 0.7 | 0.2 | 0.3 | 0.1 | 0.1 | -0.0 | 0.1 | -0.0 | -0.0 | 0.1 | 0.0 | 0.0 | -0.0 | 0.0 | 0.1 | 0.1 | 0.1 | 0.1 |
| Job security | 0.8 | 0.8 | 0.6 | 0.7 | 0.8 | 0.6 | 0.0 | 0.8 | 0.8 | -0.0 | 0.3 | -0.1 | -0.1 | -0.3 | 0.1 | 0.1 | -0.1 | 0.1 | 0.1 | 0.0 | 0.0 | 0.1 | 0.1 | 0.0 | 0.1 | 0.1 |
| Company image | 0.8 | 0.8 | 0.7 | 0.7 | 0.8 | 0.7 | 0.8 | 0.0 | 0.9 | 0.0 | 0.3 | 0.0 | -0.1 | -0.3 | 0.1 | 0.1 | -0.0 | 0.2 | 0.1 | 0.0 | 0.0 | 0.1 | 0.1 | 0.0 | 0.1 | 0.1 |
| Career development | 0.8 | 0.8 | 0.7 | 0.7 | 0.8 | 0.7 | 0.8 | 0.9 | 0.0 | -0.0 | 0.3 | -0.0 | -0.1 | -0.3 | 0.1 | 0.1 | -0.0 | 0.1 | 0.1 | 0.1 | 0.0 | 0.1 | 0.1 | 0.0 | 0.1 | 0.1 |
| Ability to work remotely | 0.1 | 0.1 | 0.4 | 0.2 | 0.0 | 0.2 | -0.0 | 0.0 | -0.0 | 0.0 | 0.3 | 0.5 | 0.5 | 0.5 | 0.2 | -0.2 | 0.0 | 0.3 | -0.1 | -0.2 | 0.0 | -0.0 | 0.1 | 0.2 | 0.1 | 0.1 |
| Flexible working hours | 0.3 | 0.3 | 0.4 | 0.3 | 0.3 | 0.3 | 0.3 | 0.3 | 0.3 | 0.3 | 0.0 | 0.2 | 0.1 | 0.0 | 0.1 | -0.1 | -0.0 | 0.2 | -0.0 | -0.0 | -0.0 | 0.0 | 0.0 | 0.2 | 0.1 | 0.1 |
| Paid parental leave | -0.0 | -0.0 | 0.3 | 0.1 | -0.0 | 0.1 | -0.1 | 0.0 | -0.0 | 0.5 | 0.2 | 0.0 | 0.6 | 0.6 | 0.2 | -0.1 | 0.0 | 0.3 | -0.0 | -0.2 | 0.1 | 0.0 | 0.0 | 0.1 | 0.1 | 0.1 |
| Stock or equity options | -0.2 | -0.1 | 0.2 | 0.1 | -0.1 | 0.1 | -0.1 | -0.1 | -0.1 | 0.5 | 0.1 | 0.6 | 0.0 | 0.6 | 0.1 | -0.0 | 0.0 | 0.2 | 0.0 | -0.2 | 0.1 | 0.0 | 0.0 | 0.1 | 0.1 | 0.1 |
| Paid time off | -0.3 | -0.2 | 0.3 | -0.0 | -0.3 | -0.0 | -0.3 | -0.3 | -0.3 | 0.5 | 0.0 | 0.6 | 0.6 | 0.0 | 0.1 | -0.1 | 0.0 | 0.0 | -0.1 | -0.2 | 0.1 | -0.1 | -0.0 | 0.0 | -0.0 | -0.0 |
| Total assets (1000$) | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.2 | 0.1 | 0.2 | 0.1 | 0.1 | 0.0 | -0.2 | 0.1 | 0.2 | -0.2 | -0.2 | 0.2 | -0.2 | -0.0 | -0.0 | 0.3 | 0.3 |
| Capital intensity | -0.0 | 0.0 | 0.0 | 0.0 | 0.1 | -0.0 | 0.1 | 0.1 | 0.1 | -0.2 | -0.1 | -0.1 | -0.0 | -0.1 | -0.2 | 0.0 | -0.0 | -0.1 | 0.3 | 0.2 | 0.0 | 0.2 | 0.0 | -0.0 | 0.0 | 0.0 |
| Leverage | -0.0 | -0.0 | -0.0 | -0.0 | -0.0 | -0.0 | -0.1 | -0.0 | -0.0 | 0.0 | -0.0 | 0.0 | 0.0 | 0.0 | 0.1 | -0.0 | 0.0 | -0.0 | -0.2 | -0.0 | 0.0 | -0.1 | -0.0 | -0.0 | 0.0 | 0.0 |
| R&D intensity | 0.2 | 0.2 | 0.2 | 0.1 | 0.1 | 0.1 | 0.1 | 0.2 | 0.1 | 0.3 | 0.2 | 0.3 | 0.2 | 0.0 | 0.2 | -0.1 | -0.0 | 0.0 | 0.2 | 0.0 | 0.0 | 0.4 | 0.1 | 0.4 | 0.5 | 0.5 |
| ROA | 0.0 | 0.0 | 0.0 | 0.0 | 0.1 | 0.0 | 0.1 | 0.1 | 0.1 | -0.1 | -0.0 | -0.0 | 0.0 | -0.1 | -0.2 | 0.3 | -0.2 | 0.2 | 0.0 | -0.3 | -0.0 | 0.7 | 0.1 | 0.1 | 0.1 | 0.1 |
| Implied volatility | 0.1 | 0.1 | -0.0 | -0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 0.1 | -0.2 | -0.0 | -0.2 | -0.2 | -0.2 | -0.2 | 0.2 | -0.0 | 0.0 | -0.3 | 0.0 | -0.2 | -0.2 | -0.0 | -0.0 | -0.0 | -0.0 |
| Firm age | -0.0 | 0.0 | 0.1 | 0.0 | 0.0 | -0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.0 | 0.1 | 0.1 | 0.1 | 0.2 | 0.0 | 0.0 | 0.0 | -0.0 | -0.2 | 0.0 | -0.1 | -0.2 | -0.3 | 0.2 | 0.2 |
| Tobin's Q | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.1 | 0.1 | 0.1 | -0.0 | 0.0 | 0.0 | 0.0 | -0.1 | -0.2 | 0.2 | -0.1 | 0.4 | 0.7 | -0.2 | -0.1 | 0.0 | 0.2 | 0.2 | 0.1 | 0.1 |
| Sales growth | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.0 | 0.0 | 0.0 | -0.0 | -0.0 | 0.0 | -0.0 | 0.1 | 0.1 | -0.0 | -0.2 | 0.2 | 0.0 | 0.1 | -0.0 | -0.0 |
| Text innovation | 0.1 | 0.1 | 0.1 | 0.1 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 0.2 | 0.2 | 0.1 | 0.1 | 0.0 | -0.0 | -0.0 | -0.0 | 0.4 | 0.1 | -0.0 | -0.3 | 0.2 | 0.1 | 0.0 | 0.3 | 0.2 |
| Patent citations | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | -0.0 | 0.3 | 0.0 | 0.0 | 0.5 | 0.1 | -0.0 | 0.2 | 0.1 | -0.0 | 0.3 | 0.0 | 0.9 |
| Patent filings | 0.1 | 0.1 | 0.1 | 0.1 | 0.0 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | -0.0 | 0.3 | 0.0 | 0.0 | 0.5 | 0.1 | -0.0 | 0.2 | 0.1 | -0.0 | 0.2 | 0.9 | 0.0 |

approach is exploratory and focuses on uncovering patterns and associations within the data, rather than testing causal hypotheses. To enhance interpretability, we complement the predictive analysis with linear regression models that include all key variables alongside controls and fixed effects for state, year, and industry. This parsimonious approach enhances the interpretability and rigor of our findings. This approach differs from the more common research approach, as we do not restrict the functional form of our model (Karolyi and Van Nieuwerburgh, 2020; Valizade et al., 2024).). Instead, by employing a tree-based ensemble method with efficient regularization, the model identifies the optimal functional form from the data, free from subjective linearity constraints imposed by the researcher (Agrawal et al., 2019). Compared to linear models, ML-driven approaches are particularly advantageous in research contexts where the relationships among variables, their interactions, and their connections to outcomes are not easily drawn from theory (Bertomeu et al., 2021). This approach is advantageous for us because the prior literature does not offer robust evidence, which could guide us to reason about the relevance of one employee treatment practice over another.[4] For our research approach, it is more efficient to allow any functional form and control for overfitting than to restrict the form and assume that it does not overfit (Beck et al., 2004).

With this ML approach, we can examine the true shape of associations between individual employee treatment practices and corporate innovation. ML methods model the joint probability function of variables using highly nonlinear functions. This approach also offers us other benefits. Since our model is based on decision trees, multicollinearity concerns are mitigated. This reduction arises from the tree construction process. As decision trees develop, new nodes are introduced by selecting features that significantly reduce prediction error, which typically prevents highly correlated covariates from being included in the same tree (Storm et al., 2020). However, the multicollinearity issue is not completely removed, and one still needs to use some care when selecting covariates for the model (Bertomeu, 2020). Our approach also preserves observations with missing values, which is crucial given our large dataset containing many predictors and controls, where most observations lack values for at least one of the variables. Decision tree based ensemble methods handle missing data by utilizing information only from non-missing values when making splits. This method enables the use of a larger dataset by retaining observations with missing values instead of discarding them.

We employ an advanced version of the gradient-boosting algorithm, known as the extreme gradient-boosting algorithm (Chen and Guestrin, 2016), to predict corporate innovation outcomes based on employee treatment practices. This algorithm enhances the original gradient-boosting approach through optimized code, improved scalability, and greater accuracy. Additionally, it incorporates feature subsampling from the random forest algorithm. Key predictors identified through this approach are further analysed using traditional statistical models to provide explanatory insights into their relationships with innovation outcomes.

Boosting models typically use standard decision trees or regression trees as base predictors. The model's accuracy is enhanced iteratively by feeding the next decision tree with reweighted data, increasing the weight of misclassified observations. Friedman (2001) introduced an efficient method for training tree ensemble models using the gradient descent algorithm (Cauchy, 1847). When using tree ensemble models for regression, the trees are often classification and regression trees (CART), which provide continuous scores on each leaf, unlike ordinary decision trees that display class labels. Through the gradient descent algorithm, both the optimal leaf weights and the optimal split points for tree branches are estimated for each added CART. The exact mathematical details are omitted and can be found in Friedman (2001) and Chen and Guestrin (2016).

When optimizing the model, several hyperparameters can be adjusted: the number of CARTs, a shrinkage parameter, column subsampling, row subsampling, the depth of the CARTs, and a gamma parameter. The shrinkage method reduces the influence of newly added weights by a predefined factor, allowing future trees to enhance the model. Column subsampling, borrowed from random forest algorithms, involves selecting a subset of features for each new tree, while row subsampling selects a subset of data points. The gamma parameter dictates the minimum loss reduction necessary to further split a leaf node. These methods help mitigate overfitting, and column subsampling also enhances computational efficiency, especially with parallel computing.

Finding optimal hyperparameters is very data-dependent. Therefore, the common practice is to use either a grid search method or "commonly accepted values" for the parameter optimization. The grid search method is efficient for parameter optimization but computationally intensive, as it tests all possible parameter combinations. In this study, we optimize the model in two steps using five-fold cross-validation. First, we conduct a comprehensive grid search for the tree structure and regularization parameters. Subsequently, we perform a separate cross-validation run to determine the optimal number of decision trees. Table 3 provides the hyperparameters used in the models.

### 3.6. SHAP method

In this paper, we employ the Shapley additive explanations (SHAP) model interpretation method (Lundberg and Lee, 2017; Covert et al., 2020) to explain our "black box" ML model. With the help of SHAP values, we can examine each feature's contribution to the prediction of corporate innovation for any particular value of the predictor (Lundberg et al., 2020; Covert et al., 2020). In addition, we employ bootstrapping to define the statistical significance of SHAP values, which allows us to move the interpretation of our ML models on par with traditional linear approaches (Ranta et al., 2023). This procedure also enables us to examine whether certain associations are nonlinear and statistically significant only at specific sub-intervals of a specific employee treatment feature. As such, this study aims not to find the best possible model for prediction, but rather to help advance what we know about the associations between employee treatment practices and corporate innovation.

The SHAP method, derived from game-theoretically optimal Shapley values (Vajda et al., 1954), is one of the most effective techniques for explaining tree-based machine learning models (Covert et al., 2020; Lundberg et al., 2020). This method offers several desirable properties for importance metrics, including consistency, local accuracy, symmetry, efficiency, and additivity (Lundberg and Lee, 2017). Moreover, SHAP values offer a highly intuitive interpretation. In a regression context, they estimate the impact of an individual feature on a specific prediction. Thus, SHAP values explain each observation by showing how each feature influences the corresponding prediction. The sum of the effects adds to the final prediction for a given company's corporate innovation prediction. By employing the SHAP method, we can explore the shape of the relationships in detail and define their statistical significance via bootstrap. Thus, the SHAP method not only enhances the interpretability of our ML models but also provides a bridge between predictive and explanatory approaches, allowing us to uncover the theoretical relevance of key predictors.

**Table 3**
Hyperparameters for the models.

| Hyperparameter | Text innovation model | Paten counts model | Patent citations model |
|---|---|---|---|
| Depth of trees | 7 | 5 | 4 |
| Shrinkage parameter | 0.05 | 0.1 | 0.05 |
| Row subsampling | 0.8/1.0 | 0.8/1.0 | 0.8/1.0 |
| Column subsampling | 0.6 | 0.6 | 0.6 |
| Gamma | 0.05 | 0.1 | 0.1 |

[4] Shmueli (2010) stated that, despite the dominant role of causal-explanatory statistical models in social sciences, predictive modelling is an essential component of scientific research aiming to create new theories, develop existing theories and compare competing theories, as well as the relevance assessment of different theories, evaluation of the predictability of empirical phenomena, and generation of new measures. Furthermore, as outlined by Shmueli (2010), despite common belief, explanatory power does not imply predictive power. Instead, predictive power should be assessed out-of-sample because measures computed from the training dataset tend to overestimate predictive accuracy (Shmueli, 2010). In this study, we use 80 % of the observations for training and 20 % for testing, and the model's predictive performance is analysed with the test data.

So far, only a few prior research papers in management and related fields have employed SHAP values to explain the results of ML models. Valizade et al. (2024) demonstrate the use of SHAP in predicting product innovation. Jabeur et al. (2021) use ML and SHAP values to reveal the most important features in predicting oil prices during the COVID-19 pandemic. Ylinen and Ranta (2023) employ SHAP values to examine how various aspects of an employee-friendly corporate culture predict company performance, utilizing bootstrap methods (Efron, 1979) for the statistical significance analysis of SHAP values in their interpretations. Similarly, Futagami et al. (2021) demonstrate the applicability of SHAP values to identify key features and interpret their contributions to predicting mergers and acquisitions. Additionally, Ranta and Ylinen (2023a) apply SHAP values to assess the practical relevance of different employee benefits on company performance. In a related study, Ranta and Ylinen (2023b) use SHAP values and bootstrap methods to explore the relationship between firm board characteristics and workplace diversity. Lin and Bai (2021) also utilize SHAP values to identify the most significant predictors of debt financing in heavily polluting enterprises.

## 4. Results

### *4.1. Model validation*

We proceed by demonstrating that employee treatment features possess significant predictive power, validating the importance of these independent variables for our study. Given potential nonlinear relationships in the data, we compare our model to several models: Random Forest, Multilayer Perceptron, as well as two linear benchmarks—LASSO and OLS regression. Model performance is measured using the out-of-sample mean squared error (MSE) and the coefficient of determination ($R^2$). We employ five-fold cross-validation on 80 % of the data for tuning, and we reserve the remaining 20 % for testing. Table 4 summarizes the results, reporting predictive power across our three innovation measures.

While all models demonstrate some capacity to predict innovation outcomes, XGBoost achieves the best performance, consistently delivering the highest $R^2$ and lowest MSE. This indicates that XGBoost effectively captures nonlinear interactions between employee treatment practices and corporate innovation, resulting in a superior fit compared to the other techniques. Random Forest, being a similar ML method, generally ranks second in predictive accuracy but exhibits more variance when applied to patent-based innovation measures. The Multilayer Perceptron, though capturing some nonlinear patterns, is more prone to variability given its sensitivity to hyperparameter tuning and the structure of the data. Finally, both LASSO and OLS regressions yield the weakest performance, underscoring how linear methods can overlook important nonlinear relationships present in the data. Overall, these findings justify our use of the XGBoost model. Its consistent outperformance across all innovation measures underscores the importance of employing flexible, nonlinear modeling when investigating the link between employee treatment features and corporate innovation outcomes.

**Table 4**
The out-of-sample coefficient of determination $R^2$ and mean square error for the models.

| Model | Mean squared error (MSE) | Coefficient of determination ($R^2$) |
|---|---|---|
| **Text innovation** | | |
| XGBoost model | 1.642 | 0.566 |
| Random forest | 2.460 | 0.314 |
| Multilayer perceptron | 2.950 | 0.177 |
| LASSO regression | 3.130 | 0.128 |
| OLS regression | 3.140 | 0.125 |
| **Patent counts** | | |
| XGBoost model | 9815 | 0.716 |
| Random forest | 9853 | 0.253 |
| Multilayer perceptron | 12490 | 0.053 |
| LASSO regression | 12490 | 0.052 |
| OLS regression | 14520 | 0.010 |
| **Patent citations** | | |
| XGBoost model | 761000 | 0.737 |
| Random forest | 1088000 | 0.553 |
| Multilayer perceptron | 2318000 | 0.047 |
| LASSO regression | 2357000 | 0.031 |
| OLS regression | 2360000 | 0.029 |

Note: The predictive power is calculated using an out-of-sample test set (20 % of the total) that is not used to train the data. The parameters of the models are optimized using five-fold cross-validation within the training set (80 % of the total).

### *4.2. Text-based innovation*

SHAP values determine how much each feature contributes to individual prediction. As such, we employ the SHAP values to explore the nonlinear associations between individual predictors and corporate innovation. We proceed by analysing the SHAP values of each feature in detail and estimate the statistical significance of the associations using 1000 bootstrap samples.

Table 5 reports the significance analysis for each feature. We use SHAP values to estimate the average effect for four equal-length subdomains for each feature and calculate the difference between the minimum and maximum effects of these subdomains. We use this difference to assess the feature importance; that is, we analyze which employee treatment practices are most strongly associated with corporate innovation. The results show that control variables *patent citations*, *R&D intensity*, *patent counts*, *firm age*, and *Tobin's Q* are the most important predictors of text-based innovation performance. The most important employee treatment feature is *flexible working hours*. The next five other important features are *paid parental leave, ability to work remotely, stock or equity options*, *attitude towards older colleagues*, and *freedom to work independently*. The least important variable is *accessible to people with disabilities*.[5]

Fig. 1 provides the nonlinear associations with 95 % confidence intervals estimated from the SHAP values for every feature. Of the most important control variables, R&D intensity has a positive association, but mainly for low R&D intensity values. Patent variables also have a positive association, with some nonlinear variation in the upper end of *patent counts*. *Firm age* and *total assets* have a negative association with text-based innovation, suggesting that text-based innovation is stronger in younger and smaller companies. These are conditions where high innovation, creativity, and entrepreneurship are common. This is a finding opposite to prior research where firm size and firm age have been found to be positively associated with patent counts and citations and supports the suggestions of Bellstam et al. (2020) that the text-based innovation measure can capture a wider range of innovation types and activities that patent proxies cannot reach. However, the results verify that our text-based innovation measure also captures traditional

[5] In our additional analyses, we complement our findings on state, industry, and year effects in all models by utilizing the averages of absolute SHAP values. It is not meaningful to calculate SHAP values in different intervals for these variables because the change between values lacks interpretation. The average SHAP values provide alternative estimates for the global importance of features based on the SHAP values in tree-based ensemble models. Our results confirm the importance of control variables such as R&D intensity, industry, total assets, firm age, capital intensity, and patent citations as the most important predictors of text-based innovation performance. Moreover, our findings suggest that the most important predictors of patent counts include control variables such as total assets, R&D intensity, text-based innovations, industry, capital intensity, and firm age. In the patent citations model, the most influential predictors are total assets, R&D intensity, industry, text-based innovation, and Tobin's Q.

**Table 5**
Significance analysis of the features for the text-based innovation model.

| | 1 | 2 | 3 | 4 | Max effect |
|---|---|---|---|---|---|
| Patent citations | −0.0026** | 0.037*** | 0.0512*** | 0.0794*** | 0.082 |
| R&D intensity | −0.0053*** | 0.0434*** | 0.0461*** | 0.0531*** | 0.0584 |
| Patent counts | −0.0012** | 0.0153*** | 0.0289*** | 0.0525*** | 0.0538 |
| Firm age | 0.0181*** | 1.00E-04 | −0.0088*** | −0.0123*** | 0.0304 |
| Tobin's Q | −0.0022** | 0.0049** | 0.0079** | 0.0136** | 0.0158 |
| Leverage | 0.0128*** | −0.0009*** | 0.0075* | 0.0144** | 0.0153 |
| Flexible working hours | −0.0027*** | 0.0004 | 0.0049*** | 0.0115*** | 0.0142 |
| Paid parental leave | 0 | −0.0002 | 0.0012 | 0.0082 | 0.0084 |
| Ability to work remotely | −0.0015*** | 0.0061*** | 0.0065*** | 0.0059*** | 0.0079 |
| Total assets (1000$) | 0.0002 | 1.00E-04 | −0.0034 | −0.0063** | 0.0065 |
| Stock or equity options | −0.0004 | 0 | 0.002** | 0.0054** | 0.0058 |
| ROA | −0.0028 | −0.0011 | 0.0003 | 0.003** | 0.0058 |
| Attitude towards older colleagues | 0.0037*** | 0.001* | −0.0002 | −0.0009* | 0.0047 |
| Sales growth | −0.0015* | −0.0006** | 0.0022** | 0.0029* | 0.0044 |
| Capital intensity | 0.0008 | −0.0004 | −0.0011 | −0.0031 | 0.0039 |
| Freedom to work independently | −0.0021* | −0.0011** | 0.0005* | 0.0009 | 0.003 |
| Work–life balance | −0.0025* | −0.0004 | 0.0005* | 0.0004 | 0.003 |
| Job security | 0.002* | −1.00E-04 | −0.0002 | −0.0005 | 0.0025 |
| Gender equality | −0.001 | −0.0008* | −1.00E-04 | 0.0013* | 0.0023 |
| Inclusive/diverse | 0.0017* | 0.0002 | 0.0003 | 0.0024* | 0.0022 |
| Career development | 0.0015 | −0.0004 | −1.00E-04 | 0.0005 | 0.0019 |
| Paid time off | −0.0005 | −0.0004 | 0.0006* | 0.001* | 0.0015 |
| Company image | −0.0002 | 0.0009* | −0.0004* | 0.0002 | 0.0013 |
| Accessible to people with disabilities | 0.0003 | 0 | 0 | −1.00E-04 | 0.0004 |
| Implied volatility | 0.0003 | 0.0004 | 0.0002 | 0.0005 | 0.0004 |

*Notes:* The SHAP values are the mean effects of the bootstrap samples for the four intervals. The domain of each variable is divided into four intervals, and the average contribution to the prediction is calculated for each variable and interval. Approximately, this means dividing each sub-figure of Fig. 1 into four intervals and taking the average from those intervals. Then, the average effect and the statistical significance are estimated using a bootstrap sample of 1000 subsamples. The asterisks indicate features with a statistically significant effect at 1 %, 5 %, and 10 % levels. All independent variables are lagged by one year, relative to the dependent variables in all three ML models. Industry, year, and state effects are controlled in the model.

patent-based innovation, as patent citations and patent counts are positively associated with the measure.

From the most important employee treatment variables, *flexible working hours*, *paid parental leave*, *ability to work remotely,* and *stock or equity options* have a positive association with text-based innovation. Interestingly, the association is negative for *job security*. However, the study of Bradley et al. (2017) also indicates that job security, in the form of labour unionization, has a negative effect on corporate innovation. This also suggests, as our above-presented results show, that high text-based innovations seem to be related to firm characteristics, where highly stable work conditions are unlikely.

To estimate the importance of the nonlinear structures, we divide the domain of each variable into four intervals and estimate the mean effect for each interval for all variables. Bootstrapping with 1000 samples is used to estimate statistical significance. For employee benefit variables (*ability to work remotely*, *flexible working hours*, *paid parental leave*, *paid time off*, and *stock or equity options*), each interval accounts for 25 % of company employees (in our dataset) who are entitled to have that particular benefit. For employee reviews concerning employee treatment practices, the SHAP values are divided into four bins (grades 1–2, grades 2–3, grades 3–4, and grades 4–5). The results provided in Table 5 verify the statistical significances identified in Fig. 1. In addition, the max effect column shows that employee treatment practices' predictive power varies considerably and that particular employee treatment features are clearly more important predictors of text-based innovation than others.

### *4.3. Patent counts and citations*

#### *4.3.1. Patent counts*

We repeat a similar analysis for patent counts and citations, which are commonly used proxies to measure innovation performance. Fig. 2 provides estimates for nonlinear associations estimated from the SHAP values of the patent counts model. Table 6 provides again estimates for the mean effects in four subdomains and the maximum difference in these effects (last column). This information can be used in determining the significance of different features when predicting patent counts.

The strongest associations are mainly with the control variables, *R&D intensity, total assets,* and *text innovation* being the three strongest predictors of patent counts by a clear margin. The four most important employee treatment features are *stock or equity options*, *flexible working hours*, *company image,* and *career development*. Compared to text-based innovation, the role of *stock or equity options* is now more important relative to other employee treatment features. Accordingly, the effects of *ability to work remotely* and *freedom to work independently* are less important in predicting patent counts than text-based innovation. Results also indicate that the relative importance of employee treatment features compared to control values is less important in predicting patent counts than text-based innovations. This makes sense, as most of the patents are more likely to be created by research groups of highly talented engineers and scientists than by rank-and-file innovators in organizations. In those situations, the role of innovation inputs in the form of sufficient R&D resources seems to be the most important requirement for reaching high levels of corporate innovation measured by patent counts.

Of the most important control variables, *R&D intensity* and *text-based innovation* have very strong positive associations. Firm size has substantial predictive power and an inverse U-shaped association with patent counts, the turning point being around the level of 600,000 (1000 $). Firm age and capital intensity have positive associations with patent counts, in accordance with prior research findings. For example, Hall and Ziedonis (2001) suggest that a firm's size and capital intensity are positively related to the number of patents and citations. In addition, ROA exhibits an interesting U-shaped relationship with patent counts, whereas Tobin's Q has a more linear, inverted U-shaped effect.

Although the control variables in our ML model have a major role in predicting patent counts, employee treatment features also have a significant contributing role to the prediction. Of the most important employee treatment variables, employee *stock or equity options* has a relatively strong positive effect on patent counts. The SHAP values show

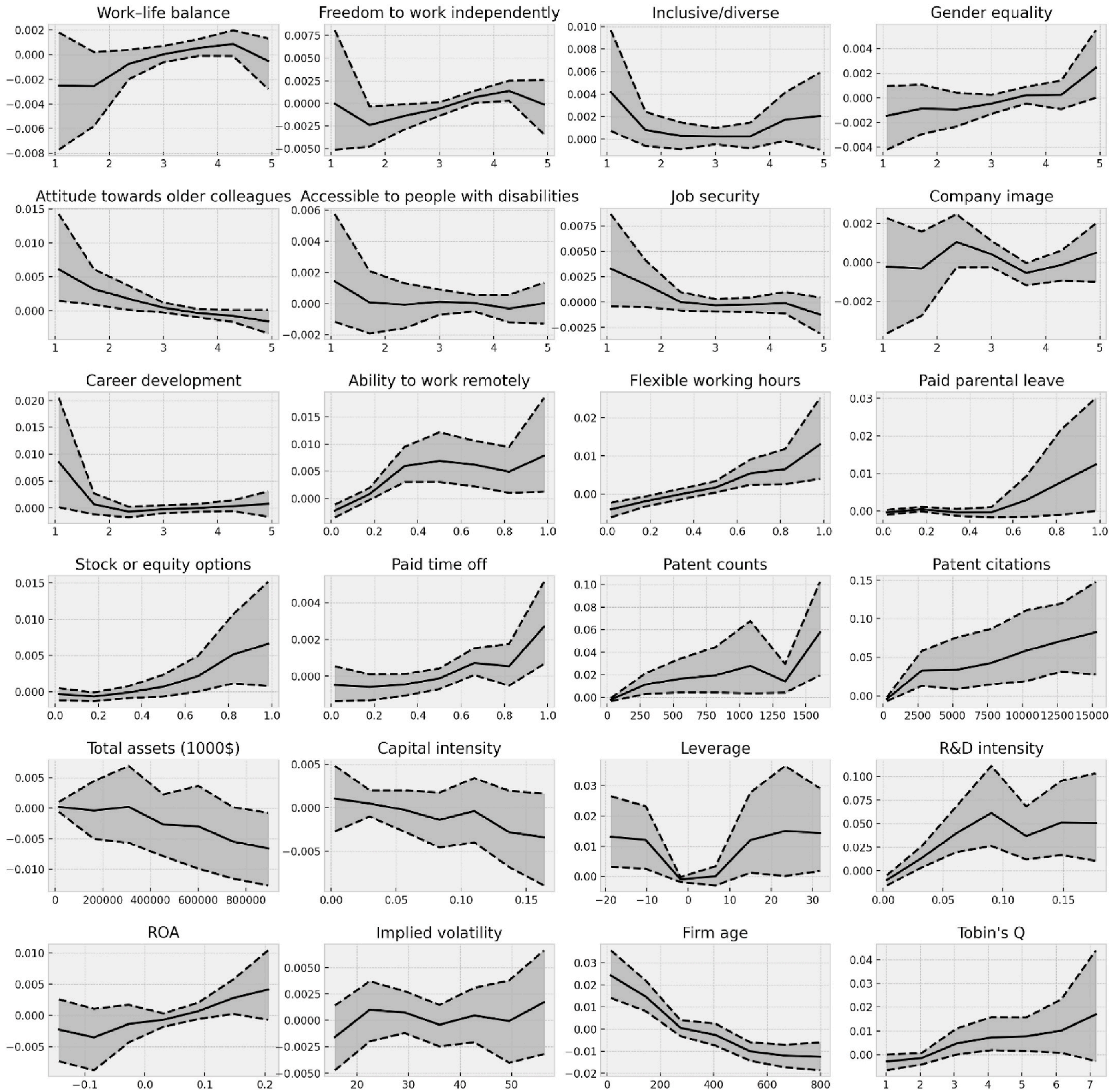



**Fig. 1.** *The SHAP values of the text-based innovation model for the whole bootstrap sample*
*Notes:* The black line is the mean effect, and the dashed lines are the 95 % confidence lines calculated from the individual SHAP values. One point in the figure gives the contribution of that feature to the prediction of a single observation. The horizontal axis represents the domain of each variable, and the vertical axis the effect of that variable on the prediction (patent counts). Thus, the charts show the average contribution of each variable to the prediction of the model. The sum of these contributions is the final prediction (averaged here over 1000 bootstrap samples). The x-axis defines the average of employee reviews (scale 1–5) or the percentage of employees with certain benefits. The y-axis defines the total amount of patents. Individual charts are scaled to a suitable level to visualize the associations to the reader. All independent variables are lagged by one year, relative to the dependent variables in all three ML models. Industry, year, and state effects are controlled in the model.

that the contribution to the prediction of patent counts increases by more than 10 patents when the benefit is offered to more than 75 % of the employees, compared to the situation where the benefit is not offered. This supports the findings of Chang et al. (2015), who found that non-executive stock options had a positive impact on driving corporate innovation, demonstrating the significance of long-term incentives in boosting innovation performance.

*Flexible working hours* has a slightly nonlinear association, where the positive effect increases until approximately 65 % of employees are offered that particular benefit, and after that point, the effect starts to drop slightly. *Career development* has a positive and statistically significant trend. The effect is negative for the first three intervals and positive for the last interval (employee ratings 4–5). However, only the effect for the last interval is statistically significant. The effect of *company image* is statistically significant at three intervals. The association is strongly positive and, in particular, the low grades for this benefit have a

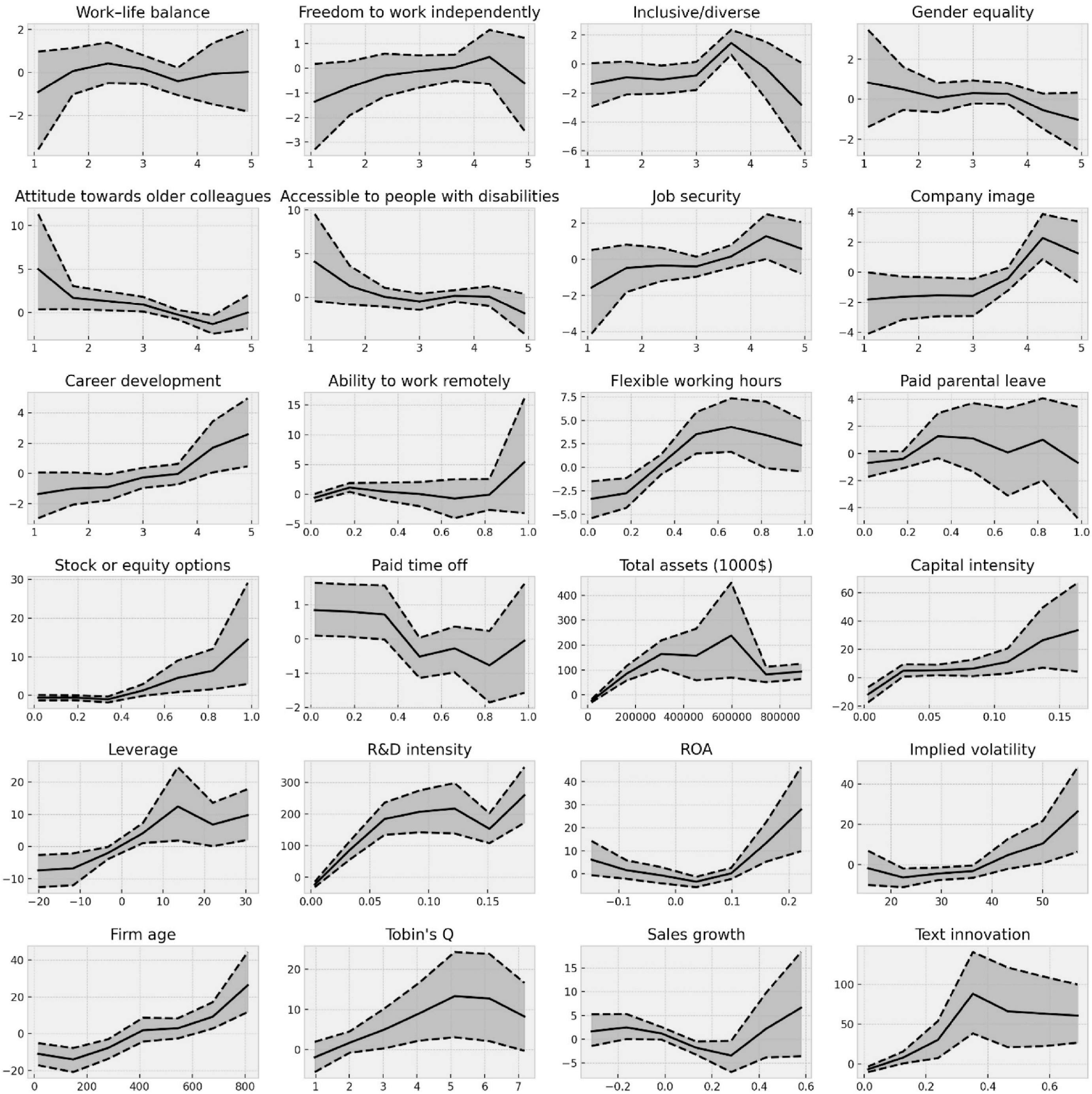



**Fig. 2.** *The SHAP values of the patent counts model for the whole bootstrap sample*
*Notes:* The black line is the mean effect, and the dashed lines are the 95 % confidence lines calculated from the individual SHAP values. One point in the figure gives the contribution of that feature to the prediction of a single observation. The horizontal axis represents the domain of each variable, and the vertical axis the effect of that variable on the prediction (patent citations). Thus, the charts show the average contribution of each variable to the prediction of the model. The sum of these contributions is the final prediction (averaged here over 1000 bootstrap samples). The x-axis defines the average of the employee reviews (scale 1–5) or percentage of employees in the company who have certain benefits. The y-axis defines the total amount of patent citations. Individual charts are scaled to a reasonable level to visualize the associations to the reader. All independent variables are lagged by one year, relative to the dependent variables in all three ML models. Industry, year, and state effects are controlled in the model.

significant negative effect on patent counts. The positive and statistically significant effect is present for the last interval (employee ratings 4–5). Our findings reveal that freedom, specifically in the form of flexible working hours, is among the two most significant employee treatment practices influencing both patent counts and text-based innovation. However, other forms of work-related freedom, such as the ability to work remotely and the autonomy to work independently, show a notably weaker association with patent counts compared to text-based innovation. This intriguing result suggests the need for further in-depth research to better understand how different dimensions of workplace freedom and flexibility influence various types of innovation outcomes. Exploring these distinctions could provide valuable insights into optimizing work arrangements to foster specific forms of innovation.

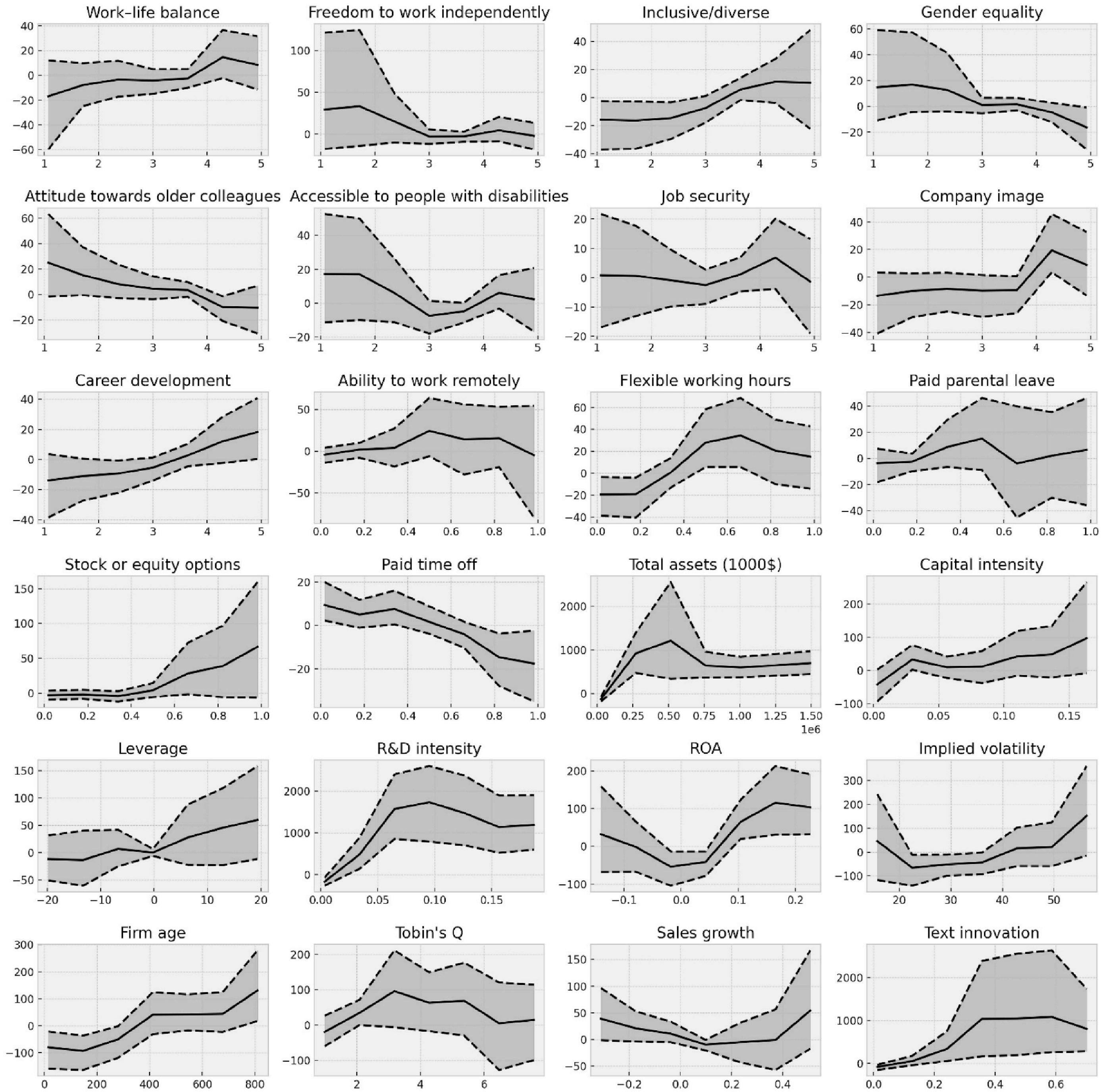



**Fig. 3.** *The SHAP values of the patent citations model for the whole bootstrap sample.*
*Notes:* The black line is the mean effect, and the dashed lines are the 95 % confidence lines calculated from the individual SHAP values. One point in the figure gives the contribution of that feature to the prediction of a single observation. The horizontal axis represents the domain of each variable, and the vertical axis the effect of that variable on the prediction (patent counts). Thus, the charts show the average contribution of each variable to the prediction of the model. The sum of these contributions is the final prediction (averaged here over 1000 bootstrap samples). The x-axis defines the average of employee reviews (scale 1–5) or the percentage of employees who have a certain benefit. The y-axis defines the total amount of patents. Individual charts are scaled to a suitable level to visualize the associations to the reader. All independent variables are lagged by one year, relative to the dependent variables in all three ML models. Industry, year, and state effects are controlled in the model.

#### 4.3.2. *Patent citations*

We continue by repeating a similar analysis for patent citations. Table 7 provides the effect estimates for our predictors across four intervals. The results for patent citations are similar to those for patent counts. The control variables continue to exhibit the highest predictive power. Similar to the patent count model, *R&D intensity, text-based innovation,* and *total assets* are the most important predictors of patent citations. The associations with *total assets* and *R&D intensity* are nonlinear, resembling an inverse U-shape; however, the SHAP values for total assets peak slightly earlier at approximately 500,000 (1000 $) compared to the patent counts model. *R&D intensity* has a turning point around the 10 % level, differing from the patent counts model, where the association was more positive. This finding supports the argument by Hall and Ziedonis (2001) that large firms generate more patents and

**Table 6**
Significance analysis of the features for the patent counts innovation model.

| Feature | 1 | 2 | 3 | 4 | Max effect |
|---|---|---|---|---|---|
| R&D intensity | −0.6896 | 192.1651*** | 214.1146*** | 226.276*** | 226.9656 |
| Total assets (1000$) | −14.3022*** | 166.0055*** | 159.4888*** | 91.0834*** | 180.3077 |
| Text innovation | −4.9526*** | 38.0942*** | 81.9675*** | 63.0293*** | 86.9201 |
| Capital intensity | −3.2428** | 5.6093** | 9.1816** | 30.1378*** | 33.3805 |
| Firm age | −13.2638*** | −5.2988* | 3.1628 | 19.6664*** | 32.9303 |
| Implied volatility | −5.3404* | −4.4704** | 1.6286 | 17.0895** | 22.4299 |
| ROA | 3.9126 | −2.22 | −1.5733 | 17.7473*** | 19.9673 |
| Leverage | −7.4153** | 0.2033 | 8.4187** | 9.0474* | 16.4627 |
| Tobin's Q | −0.1814 | 5.5927* | 12.6258** | 10.2985* | 12.8072 |
| Stock or equity options | −0.616 | −0.6169 | 4.0744** | 7.6749** | 8.2918 |
| Flexible working hours | −2.9536*** | 1.0787* | 4.4909*** | 3.0567 | 7.4445 |
| Sales growth | 2.3436 | 0.3052 | −2.919* | 4.2042 | 7.1232 |
| Company image | −1.6326** | −1.5473** | −0.4494 | 2.2481** | 3.8807 |
| Career development | −1.0282 | −0.7501 | 0.0437 | 2.2782* | 3.3063 |
| Attitude towards older colleagues | 2.2611** | 1.147* | −0.188 | −0.9135 | 3.1746 |
| Ability to work remotely | −0.0449 | 0.395 | −0.5472 | 1.8726 | 2.4197 |
| Inclusive/diverse | −1.0493 | −0.9643 | 0.9709** | −1.281 | 2.2519 |
| Accessible to people with disabilities | 1.7491 | −0.2295 | 0.0938 | −0.4869 | 2.236 |
| Paid parental leave | −0.5955 | 1.2457 | 0.1748 | 0.2501 | 1.8412 |
| Job security | −0.6271 | −0.3886 | 0.1174 | 1.2026 | 1.8297 |
| Paid time off | 0.8338* | 0.1642 | −0.373 | −0.6439 | 1.4777 |
| Gender equality | 0.5435 | 0.1032 | 0.2951 | −0.7594 | 1.3029 |
| Freedom to work independently | −0.8381 | −0.3338 | 0.1058 | 0.1508 | 0.9889 |
| Work–life balance | 0.0201 | 0.4901 | −0.3084 | −0.0653 | 0.7985 |

*Notes:* The SHAP values are the mean effects of the bootstrap samples for the four intervals. The domain of each variable is divided into four intervals, and the average contribution to the prediction is calculated for each variable and interval. This means approximately dividing each sub-figure of Fig. 2 into four intervals and taking the average from those intervals. Then, the average effect and the statistical significance are estimated using a bootstrap sample of 1000 subsamples. The bolded effects indicate significance at least at the 10 % level. The asterisks indicate features with a statistically significant effect at 1 %, 5 %, and 10 % levels. All independent variables are lagged by one year, relative to the dependent variables in all three ML models. Industry, year, and state effects are controlled in the model.

**Table 7**
Significance analysis of the features for the patent citations innovation model.

| Feature | 1 | 2 | 3 | 4 | Max effect |
|---|---|---|---|---|---|
| R&D intensity | −30.5796 | 1603.7499*** | 1574.6627*** | 1219.4364*** | 1634.33 |
| Text innovation | −63.48*** | 403.7844*** | 1111.177*** | 928.1888*** | 1174.657 |
| Total assets (1000$) | −73.2034*** | 991.7973*** | 591.4821*** | 693.586*** | 1065.001 |
| Firm age | −89.8488*** | −25.0046 | 44.9439* | 97.2267** | 187.0755 |
| ROA | 17.9926 | −53.3842*** | 21.0145* | 114.3379*** | 167.7221 |
| Implied volatility | −32.4298 | −51.3575** | −9.825 | 79.8644 | 131.222 |
| Tobin's Q | 4.4668 | 93.2901* | 63.0812 | 12.394 | 88.8234 |
| Capital intensity | −2.4457 | 11.1175 | 27.7277 | 69.0242 | 71.4699 |
| Leverage | −12.0263 | 5.9231 | 2.2918 | 54.8005* | 66.8268 |
| Flexible working hours | −19.4474** | 7.229* | 35.1824*** | 17.8562 | 54.6298 |
| Stock or equity options | −2.5487 | −2.9667 | 19.9467* | 51.2198** | 54.1864 |
| Sales growth | 29.0199* | 2.3626 | −12.9678 | 25.9408 | 41.9877 |
| Freedom to work independently | 32.9049 | 5.0073 | −2.5914 | 4.4887 | 35.4964 |
| Company image | −10.2789 | −8.7001 | −8.832* | 21.9692** | 32.2482 |
| Inclusive/diverse | −16.0818** | −11.1982** | 2.4879 | 12.9673 | 29.0491 |
| Ability to work remotely | −2.409 | 7.3691 | 25.9984 | 8.9156 | 28.4075 |
| Attitude towards older colleagues | 16.9189** | 6.7971 | 2.0034 | −10.4448** | 27.3637 |
| Career development | −11.2389* | −8.6199** | 1.2127 | 15.3238* | 26.5627 |
| Paid time off | 8.2163** | 4.5062* | −3.1619 | −15.8732*** | 24.0896 |
| Gender equality | 16.4726 | 5.193 | 1.4201 | −7.3444* | 23.817 |
| Work–life balance | −8.155 | −3.0209 | −2.0381 | 15.2568* | 23.4118 |
| Accessible to people with disabilities | 17.0193 | −1.3859 | −4.5787* | 6.0616 | 21.5979 |
| Paid parental leave | −3.3813 | 9.6735 | 1.8759 | 4.2975 | 13.0548 |
| Job security | 0.7612 | −1.5893 | 0.6513 | 4.7213 | 6.3106 |

*Notes:* The SHAP values are the mean effects of the bootstrap samples for the four intervals. The domain of each variable is divided into four intervals, and the average contribution to the prediction is calculated for each variable and interval. Approximately, this means dividing each sub-figure of Fig. 3 into four intervals and taking the average from those intervals. Then, the average effect and the statistical significance are estimated using a bootstrap sample of 1000 subsamples. The bolded effects indicate significance at least at the 10 % level. The asterisks indicate features with a statistically significant effect at 1 %, 5 %, and 10 % levels. All independent variables are lagged by one year, relative to the dependent variables in all three ML models. Industry, year, and state effects are controlled in the model.

citations. Our results suggest that increasing resources and inputs to R&D can lead to more patents (innovation quantity), but not necessarily higher innovation quality (patent citations) after a certain level of inputs is achieved.

*Text innovation, firm age*, *capital intensity*, and *implied volatility* have somewhat linear positive associations with patent citations. However, the importance (SHAP values) for *capital intensity*, relative to *R&D intensity* and *total assets*, seems to be less important when compared to the patent counts model. One possible explanation is that capital-intensive firms tend to generate more incremental or process-oriented innovations, thus filing more patents without necessarily garnering many citations. As a result, capital intensity emerges as a stronger predictor for patent counts than for citations or text-based measures, which likely capture more intangible, high-impact innovations.

In addition, *text innovation* seems to be a strong predictor for both patent counts and patent citations. The most important employee treatment features with positive and statistically significant associations include *flexible working hours, stock or equity options,* and *company image*.

Additionally, *career development* and *inclusive/diverse* company values show a statistically significant positive association. Surprisingly, variables related to *work-life balance* and *job security* are weak predictors of patent citations. Meanwhile, *attitudes toward older colleagues*, *gender equality*, and *paid time off* exhibit statistically significant negative associations with patent citations. These findings are intriguing, and we explore their theoretical implications and practical significance more deeply in our discussion and conclusions section.

### 4.4. Additional analyses - evidence supporting the validity of text-based innovation measure

Our analyses so far have shown that the text-based innovation metric is strongly linked to patenting and R&D efforts, indicating an ability to measure innovation. Thus, the evidence so far is in line with Bellstam et al. (2020). To further validate that our text-based innovation measure derived from 10-K filings accurately captures innovation, we examine the association between our text-based innovation measure and the research variables separately on patenting and non-patenting firms. To facilitate straightforward interpretation, we restricted the models to linear relationships. Fig. 4 provides the results where the variables are plotted against our text-based innovation measure to determine their predictive importance.

Our findings indicate no systematic differences between patenting

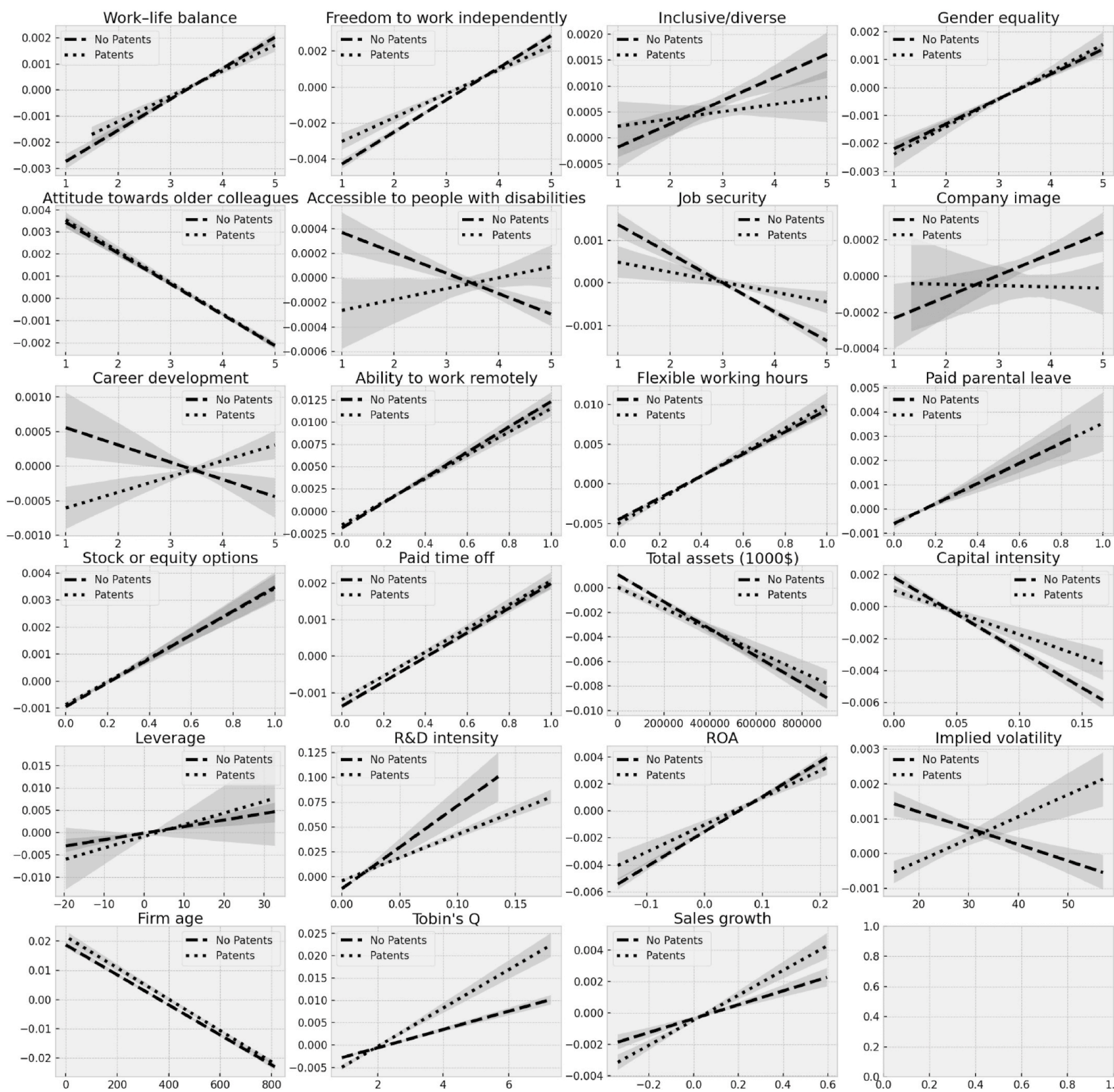



**Fig. 4.** *SHAP values indicating feature importance in predicting text-based innovation for patenting vs. non-patenting firms.*

and non-patenting firms when using our text-based measure (note that the subplots have varying scales on the y-axis). For example, variables such as *R&D intensity*, *firm age*, *capital intensity*, and *total assets* display consistent SHAP values across both patenting and non-patenting firms, suggesting that our measure effectively captures innovation irrespective of patenting activity. Interestingly, the association between *R&D intensity* and text-based innovation is slightly stronger in non-patenting firms. Additionally, employee-related factors such as *flexible working hours*, *stock or equity options*, *freedom to work independently, gender equality*, and *career development* show similar patterns of importance in both types of firms. This consistency reinforces the robustness of our text-based innovation measure and its capability to assess innovation beyond traditional patent metrics. Our analysis is in line with Bellstam et al. (2020) and provides strong evidence that our text-based innovation measure is a reliable and comprehensive tool for assessing corporate innovation in both patenting and non-patenting firms. This finding is significant as it underscores the versatility of text-based analysis in capturing a broader spectrum of innovative activities that might not be reflected solely through patent counts.

### 4.5. Endogeneity concerns

We acknowledge that reverse causality could pose a significant issue if our goal were to assert causal relationships based on our ML analyses. For instance, it could be argued that more innovative and successful companies are better positioned to invest in superior employee treatment practices. However, this issue is less critical in our context, as our research approach is descriptive, aiming to uncover novel patterns in complex social media data, assess the predictive value of various employee treatment practices, and determine their relative importance as predictors of corporate innovation. Our ML-led research approach is intended to generate discoveries by identifying interesting associations that warrant inclusion in more explanatory causal models in future research (Bertomeu et al., 2021).

To further address endogeneity concerns, we complement our machine learning (ML) analyses with a two-stage least squares (2SLS) regression approach. This method allows us to isolate the exogenous component of employee treatment practices and examine its effect on corporate innovation outcomes. We focus our analyses on four employee treatment practices identified as strong and statistically significant predictors in the ML analyses, which also remain statistically significant in a linearly restricted ordinary least squares (OLS) regression model for both text-based and patent innovation (Fig. 4). These practices include *flexible working hours* (an employee benefit), *employee stock or equity options* (an employee benefit), *remote work* (an employee benefit), and *company image* (a measure of how proud employees are to work for their company). Notably, these variables are key predictors of 10-K text-based innovation and are also highly significant predictors of patent counts and patent citations in non-linear models, except that remote work is primarily relevant to the text-based innovation measure. Given the challenge of identifying appropriate instruments for all our employee treatment practices, we also construct two new aggregate variables from our employee review data: one representing employee treatment features and another representing employee benefits.

**Table 8**
Employee treatment practices and innovative output — baseline results.

| Dependent Variable | Text-inno (t) (1) | Text-inno (t+1) (2) | Ln(1+Pat) (3) | Ln(1+Cit) (4) | Text-inno (t) (5) | Text-inno (t+1) (6) | Ln(1+Pat) (7) | Ln(1+Cit) (8) |
|---|---|---|---|---|---|---|---|---|
| Independent Variables | | | | | | | | |
| Employee stock or equity options | −0.0003 | 0.0064 | 0.6382** | 0.8576** | – | – | – | – |
| | (-0.02) | (0.49) | (2.87) | (2.58) | | | | |
| Flexible working hours | 0.0443*** | 0.0460*** | 0.4097** | 0.7424*** | – | – | – | – |
| | (3.99) | (4.03) | (2.26) | (2.82) | | | | |
| Company image | 0.0110*** | 0.0078** | 0.1942*** | 0.1476* | – | – | – | – |
| | (3.14) | (2.04) | (3.09) | (1.66) | | | | |
| Remote work | 0.0724*** | 0.0726*** | 0.0138 | 0.2451 | – | – | – | – |
| | (4.07) | (3.39) | (0.04) | (0.51) | | | | |
| Employee treatment | – | – | – | – | 0.0185*** | 0.0145*** | 0.1616** | 0.1364 |
| | | | | | (4.39) | (3.16) | (2.20) | (1.30) |
| Employee benefits | – | – | – | – | 0.1070*** | 0.1231*** | 1.4944*** | 2.3797*** |
| | | | | | (3.83) | (4.06) | (3.53) | (3.92) |
| Text-inno | – | – | 3.8865*** | 6.5806*** | – | – | 3.9112*** | 6.6644*** |
| | | | (7.24) | (9.34) | | | (7.31) | (9.58) |
| Tobin's Q | 0.0034 | 0.0040 | −0.0308 | 0.0105 | 0.0035 | 0.0039 | −0.0302 | 0.0067 |
| | (0.96) | (1.00) | (-0.65) | (0.16) | (0.97) | (0.96) | (-0.64) | (0.10) |
| Firm age | −0.00007*** | −0.00007*** | 0.0017*** | 0.0013*** | −0.00007*** | −0.00007*** | 0.0017*** | 0.0013*** |
| | (-7.17) | (-6.70) | (9.01) | (4.84) | (-7.17) | (−6.69) | (9.09) | (4.82) |
| Capital intensity | 0.2005** | 0.1039 | 0.4027 | 2.5007 | 0.1882** | 0.0945 | 0.4448 | 2.4480 |
| | (2.27) | (1.03) | (0.28) | (1.29) | (2.06) | (0.92) | (0.30) | (1.27) |
| R&D intensity | 1.4345*** | 1.5085*** | 22.4918*** | 25.2200*** | 1.4507*** | 1.5233*** | 22.7521*** | 25.5252*** |
| | (8.59) | (7.98) | (8.89) | (7.19) | (8.75) | (8.14) | (8.95) | (7.24) |
| Leverage | −0.0001 | −0.0001 | 0.0165** | 0.0169 | −0.0002 | −0.0001 | 0.0163** | 0.0164 |
| | (-0.43) | (-0.21) | (2.09) | (1.27) | (-0.52) | (-0.25) | (2.05) | (1.24) |
| ROA | 0.0340 | 0.0548 | 2.5620*** | 2.6835** | 0.0255 | 0.0495 | 2.7092*** | 2.8377** |
| | (0.67) | (0.92) | (2.94) | (2.14) | (0.50) | (0.82) | (3.09) | (2.27) |
| Total assets | 2.00e-08 | 1.47e-08 | 2.90e-06*** | 3.94e-06*** | 2.61e-08* | 2.06e-08 | 2.90e-06*** | 3.95e-06*** |
| | (1.35) | (0.91) | (7.54) | (7.15) | (1.76) | (1.28) | (7.58) | (7.20) |
| Industry FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1985 | 1578 | 1965 | 1773 | 1985 | 1578 | 1965 | 1773 |
| Adj. R-squared | 0.5606 | 0.5736 | 0.5681 | 0.5816 | 0.5556 | 0.569 | 0.5659 | 0.5802 |

*Notes:* This table reports the results of the linear relationship between employee treatment variables and various types of innovative outputs. Robust t-statistics are in parentheses. ***, **, and * indicate significance at the 1 %, 5 %, and 10 % levels, respectively. All independent variables are lagged by one year relative to the dependent variables in all OLS models, except for the Text-inno (t) model. Controls include firm size, measured by total assets; leverage, calculated as the ratio of long-term debt to total assets; R&D intensity, measured as the ratio of R&D expenditure to total assets; return on assets (ROA), measured as the ratio of operating income to book assets; Tobin's Q; capital intensity, calculated as the ratio of net property, plant, and equipment to the number of employees; and firm age. Additionally, text-based innovation is included as a control in the patent counts and patent citations models. All regressions include year and industry (SIC2) fixed effects.

We begin by constraining our non-linear ML models to linear OLS regressions to examine the associations between employee treatment practices and various corporate innovation measures. Our results, presented in Table 8, indicate that most employee treatment coefficient estimates are positive and statistically significant. While reverse causality remains a potential concern, our predictive-driven approach, supplemented by the 2SLS regression method, enhances the robustness of our findings and underscores the importance of employee treatment practices in fostering corporate innovation.[6]

In our 2SLS instrumental variable regressions, we use one instrument for each endogenous regressor. Following Deng et al. (2013) and Mao and Weathers (2019), we instrument employee treatment practices, specifically *flexible working hours* and the aggregate *employee treatment index*, using a Democratic state dummy. This variable equals one if a firm's headquarters is located in a state that predominantly voted for Democratic presidential candidates during the relevant time period, and zero otherwise.[7] Unlike state-level averages of firm behavior, which may raise concerns of endogeneity, this instrument is based solely on historical voting patterns and is independent of individual firm actions. Both Deng et al. (2013) and Mao and Weathers (2019) demonstrate the relevance of this measure. Prior research by Rubin (2008) shows that firms located in Democratic-leaning states are more likely to exhibit stronger corporate social responsibility practices, including employee-related policies. We assume that political alignment affects innovation outcomes only indirectly through its influence on internal social practices such as employee treatment, thereby satisfying the exclusion restriction.

We acknowledge the inherent difficulty of finding suitable instrumental variables for all our employee treatment and benefit variables that are unrelated to innovation performance while satisfying the exclusion restriction. To mitigate this challenge, we employ the IV technique proposed by Lewbel (2012), which has been extensively utilized in recent economics and finance literature (see, e.g., Chen et al., 2021; Dimic et al., 2024; Emran and Hou, 2013; Gong et al., 2018; Hasan et al., 2021; Mavis et al., 2020). Formally, Lewbel's (2012) internal instrumental variables, based on a heteroscedastic covariance restriction, are constructed by multiplying the mean-centered forms of existing exogenous variables with the residuals from the first-stage regression of the instrumented independent variable. We use Lewbel's instrument for *employee stock and equity options*, *remote work*, *company image*, and *aggregate employee benefits* variables.

The estimates from the two-stage IV regressions are reported in Table 9. Overall, the IV regressions indicate that firms with strong employee treatment and benefit practices exhibit higher innovation levels even after controlling for potential endogeneity. Specifically, the coefficient estimates for the instrumented variables are positive, statistically significant, and comparable to our ML analyses. In general, the IV regressions, which incorporate heteroscedasticity-based augmentations of external instruments, support the argument that positive employee treatment policies enhance firm innovation.

To further validate our second-stage findings related to Lewbel's instruments, we also use firm option volatility as an alternative instrument for employee stock and equity options. Prior research suggests several theoretical explanations for firms' use of stock option grants in compensation plans. For example, agency theory predicts a negative relationship between risk and incentives, implying that option plans should be less common at high-volatility firms (Oyer and Schaefer, 2005). Our first-stage regression results support this theoretical argument. However, while option volatility was not a strong instrument for employee stock option benefits, the second-stage regressions still demonstrated positive and statistically significant associations with text- and patent-based measures of firm innovation, thereby supporting the findings obtained using Lewbel's instrument.

We report the diagnostic statistics for the first-stage regressions in Panel B of Table 9 to validate the IV estimates based on Lewbel's (2012) approach and the Democratic-state dummy. In the first-stage regressions (coefficients not reported due to space constraints), all instruments, including the Democratic-state dummy and Lewbel's variables, are strongly related to employee benefits and treatment variables across all models at the 0.01 significance level. The Kleibergen-Paap rk Wald F statistics are high and comfortably exceed the critical value suggested by Stock and Yogo (2005) (except for the blue-state dummy for the aggregate employee treatment variable), thereby rejecting the null hypothesis of weak instruments. Furthermore, the Kleibergen-Paap rk LM statistics are high and statistically significant at the 1 % level, rejecting the null hypothesis of underidentification. Collectively, the diagnostics in Panel B confirm the validity of Lewbel's (2012) internal instruments and the Democratic-state dummy in our two-stage IV regressions.

The results remain robust after accounting for fixed effects and potential confounding factors. This approach strengthens the validity of our findings by providing a more holistic view of the relationships between variables and innovation outcomes. These findings align with previous research (e.g., Chen et al., 2016b; Chang et al., 2015; Mao and Weathers, 2019), demonstrating that positive employee treatment has a causal effect on corporate innovation. This effect is evident not only in patent-based innovations but also in other forms of innovation activities.

## 5. Discussion and conclusions

In this paper, we showcased the application and strengths of an explainable ML research approach. Utilizing nonlinear ML predictor and SHAP values, we analysed the global importance of several variables in predicting corporate innovation, examined the shape of their relationships, and estimated effect sizes. Our study is primarily predictive, leveraging ML methods to identify key employee treatment practices that drive corporate innovation. This approach revealed the impact of specific features on predicting corporate innovation and allowed us to assess whether effects were significant at certain value intervals of employee treatment features. We suggest that many research questions in innovation management could benefit from this explainable ML approach. By narrowing our focus to key predictors and analyzing them through explanatory models, we enhance the theoretical depth and practical relevance of our findings. This dual approach allows us to balance predictive accuracy with interpretability.

Our study extends the existing innovation management research by integrating ML methods with a refined empirical strategy that addresses potential endogeneity concerns (e.g., Chiarello et al., 2024; Lu and Chesbrough, 2022; Xu et al., 2024; Validaze, 2024). Furthermore, utilizing social media data analysis brings novelty to our research (e.g., Saura et al., 2023; Ozcan et al., 2021; Petruzzelli et al., 2024). We demonstrate that employer reviews on social media platforms are

[6] Our ML analysis (gradient boosting with SHAP) finds a strong positive association between employee shares and firm innovation, suggesting that stock-based incentives play a crucial role in motivating employees toward innovative outcomes. However, in OLS regression, this association weakens when controlling for other employee treatment variables (remote work, company image, and flexible working hours). This difference likely arises because ML captures complex interactions and non-linear effects, while OLS assumes a linear and additive relationship, potentially underestimating the independent effect of employee shares in the presence of correlated employee benefits. Further supporting this explanation, we find that in OLS regression, the association between employee shares and text-based innovation remains statistically significant when other employee treatment variables are removed from the model. This suggests that multicollinearity among employee benefits may obscure the independent contribution of stock-based incentives to innovation when multiple highly correlated variables are included simultaneously. These findings highlight the need for careful interpretation of OLS results when analyzing interrelated workplace policies and their impact on firm innovation.

[7] The classification of states is based on publicly available data, such as those summarized at http://en.wikipedia.org/wiki/Red_states_and_blue_states. This classification was also used in Deng et al. (2013) and Mao and Weathers (2019).

**Table 9**
Employee treatment practices and innovative output — 2SLS analysis.

Panel A: Second-stage IV regressions

| | (1) Ln(1+Pat) | (2) Ln(1+Cit) | (3) Text-inno | (4) Ln(1+Pat) | (5) Ln(1+Cit) | (6) Text-inno | (7) Ln(1+Pat) | (8) Ln(1+Cit) |
|---|---|---|---|---|---|---|---|---|
| Workhours (IV) | 12.2059*** | 11.3994*** | 0.3044*** | | | | | |
| | (4.17) | (3.61) | (3.03) | | | | | |
| Emp. stock opt (IV) | | | | 3.2655** | 4.7281** | 0.2591** | | |
| | | | | (1.98) | (2.06) | (2.06) | | |
| Company image (IV) | | | | | | | 0.3839 | 0.6846** |
| | | | | | | | (1.48) | (2.18) |
| Tobin's Q | 0.0657 | 0.0939 | 0.0073* | −0.0190 | 0.0159 | 0.0042 | −0.0326 | −0.0068 |
| | (0.85) | (1.08) | (1.77) | (−0.36) | (0.20) | (0.94) | (−0.68) | (−0.09) |
| Firm age | 0.00136*** | 0.00081** | −0.00008*** | 0.00153*** | 0.00101*** | −0.00009*** | 0.00165*** | 0.00115*** |
| | (4.58) | (2.23) | (−6.23) | (7.88) | (3.57) | (−6.38) | (9.29) | (4.31) |
| Capital intensity | −0.385 | 2.9082 | 0.0938 | 0.4600 | 2.5161 | 0.1058 | 0.3536 | 2.2865 |
| | (−0.15) | (1.03) | (0.89) | (0.32) | (1.26) | (0.96) | (0.25) | (1.14) |
| R&D intensity | 16.5182*** | 24.7405*** | 1.2613*** | 27.000*** | 32.9178*** | 1.2705*** | 29.3217*** | 36.0697*** |
| | (3.65) | (4.60) | (5.45) | (9.07) | (7.87) | (5.62) | (11.48) | (10.20) |
| Leverage | 0.0238* | 0.0287** | 0.00011 | 0.01137 | 0.01566 | −0.00029 | 0.0146** | 0.0184 |
| | (1.91) | (1.96) | (0.24) | (1.53) | (1.24) | (−0.71) | (2.12) | (1.55) |
| ROA | 2.0239 | 2.8224 | 0.0078 | 2.5191** | 2.5019 | 0.0050 | 3.0835*** | 3.3266** |
| | (1.50) | (1.75) | (0.12) | (2.57) | (1.64) | (0.07) | (3.45) | (2.42) |
| Total assets | 2.09e–06*** | 3.08e–06*** | 5.94e–09 | 2.82e–06*** | 3.71e–06*** | 2.72e–08 | 2.70e–06*** | 3.57e–06*** |
| | (4.00) | (4.64) | (0.32) | (7.15) | (6.54) | (1.57) | (6.64) | (6.31) |
| Industry FE (SIC2) | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 2108 | 1846 | 1578 | 2108 | 1846 | 1578 | 2108 | 1846 |
| RMSE (2nd stage) | 2.563 | 2.729 | 0.0913 | 1.527 | 2.088 | 0.087 | 1.473 | 2.019 |

**Panel B: IV first stage diagnostics**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Kleibergen-Paap rk LM statistic | 22.43 | 22.21 | 20.42 | 21.22 | 20.68 | 17.14 | 45.05 | 51.35 |
| Underidentification test (p-value) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Kleibergen-Paap rk Wald F statistic | 21.97 | 21.68 | 19.89 | 41.65 | 41.46 | 32.74 | 61.13 | 136.63 |
| Overidentification (Sargan) | 0.000 (exact ID) | 0.000 (exact ID) | 0.000 (exact ID) | 0.000 (exact ID) | 0.000 (exact ID) | 0.000 (exact ID) | 0.000 (exact ID) | 0.000 (exact ID) |

Panel A: Second-stage IV regressions

| | (9) Text-inno | (10) Text-inno | (11) Ln(1+Pat) | (12) Ln(1+Cit) | (13) Text-inno | (14) Ln(1+Pat) | (15) Ln(1+Cit) | (16) Text-inno |
|---|---|---|---|---|---|---|---|---|
| Company image (IV) | 0.0407** | | | | | | | |
| | (2.21) | | | | | | | |
| Remote work (IV) | | 0.2210*** | | | | | | |
| | | (2.15) | | | | | | |
| Emp.benefits (IV) | | | 3.8158** | 5.3909** | 0.2664** | | | |
| | | | (2.07) | (2.14) | (2.12) | | | |
| Emp.treatment (IV) | | | | | | 8.07** | 8.77** | 0.20** |
| | | | | | | (2.70) | (2.22) | (2.23) |
| Tobin's Q | 0.0022 | 0.0033 | −0.0225 | 0.0091 | 0.0039 | −0.0794 | −0.0205 | 0.0016 |
| | (0.54) | (0.85) | (−0.45) | (0.12) | (0.99) | (−0.66) | (−0.15) | (0.37) |
| Firm age | −0.00007*** | −0.00007*** | 0.00157*** | 0.00100*** | −0.000077*** | 0.00148*** | 0.00114** | −0.000066*** |
| | (−6.39) | (−6.70) | (8.69) | (3.71) | (−6.93) | (3.19) | (1.97) | (−4.10) |
| Capital intensity | 0.0777 | 0.1159 | 0.4378 | 2.8337 | 0.0981 | −6.7347 | −6.2404 | −0.0413 |
| | (0.75) | (1.16) | (0.31) | (1.49) | (0.97) | (−1.47) | (−1.00) | (−0.28) |
| R&D intensity | 1.5281*** | 1.4860*** | 27.9427*** | 34.3721*** | 1.4575*** | 7.715 | 15.152 | 1.0833*** |
| | (8.36) | (7.53) | (10.42) | (9.18) | (7.43) | (0.81) | (1.30) | (3.10) |
| Leverage | −0.00004 | −0.00018 | 0.0133* | 0.0168 | −0.00015 | 0.0230 | 0.0319 | 0.00015 |
| | (−0.11) | (−0.46) | (1.89) | (1.42) | (−0.39) | (1.22) | (1.28) | (0.23) |
| ROA | 0.0499 | 0.0700 | 3.1347*** | 3.4230** | 0.0467 | −1.7829 | −2.0171 | −0.0597 |
| | (0.82) | (1.18) | (3.58) | (2.55) | (0.78) | (−0.66) | (−0.57) | (−0.62) |
| Total assets | 1.68e–08 | 2.84e–09 | 2.71e–06*** | 3.59e–06*** | 1.59e–08 | 8.47e–07 | 2.26e–06** | −6.61e–09 |
| | (0.95) | (0.13) | (6.92) | (6.34) | (0.91) | (0.85) | (2.08) | (−0.26) |
| Industry FE (SIC2) | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1578 | 1578 | 2108 | 1846 | 1578 | 2108 | 1846 | 1578 |
| RMSE (2nd stage) | 0.0798 | 0.07967 | 1.481 | 2.01 | 0.0788 | 3.853 | 4.342 | 0.1165 |

**Panel B: IV first stage diagnostics**

| | (9) | (10) | (11) | (12) | (13) | (14) | (15) | (16) |
|---|---|---|---|---|---|---|---|---|
| Kleibergen-Paap rk LM statistic | 35.22 | 29.65 | 31.86 | 32.09 | 27.58 | 7.84 | 5.75 | 7.32 |
| Underidentification test (p-value) | 0 | 0 | 0 | 0 | 0 | 0.0051 | 0.0165 | 0.0068 |
| Kleibergen-Paap rk Wald F statistic | 40.61 | 50.69 | 94.33 | 97.5 | 79.41 | 7.61 | 5.54 | 7.05 |
| Overidentification (Sargan) | 0.000 (exact ID) | 0.000 (exact ID) | 0.000 (exact ID) | 0.000 (exact ID) | 0.000 (exact ID) | 0.000 (exact ID) | 0.000 (exact ID) | 0.000 (exact ID) |

*Notes:* This table reports the results on the relationship between employee treatment practices and various innovative outputs using a two-stage least squares regression. The democratic state (blue state) dummy is employed as an instrument for employee treatment schemes in the first stage for flexible working hours and aggregate employee treatment variables. For employee stock and equity options, the ability to work remotely, company image (pride in the company), and aggregate employee benefit variables, this table presents the results of instrumental variable regressions using heteroskedasticity-based instruments following Lewbel's (2012) IV regressions. Panel A reports the results of the second-stage regression for the dependent variables: patent counts, patent citations, and text-based innovation, respectively. Panel B presents the first-stage diagnostics for each endogenous regressor. Control variables include firm size (measured by total assets), leverage (calculated as the ratio of long-term debt to total assets), R&D intensity (measured as the ratio of R&D expenditure to total assets), return on assets (ROA) (calculated as the ratio of operating income to book assets), Tobin's Q, capital intensity (measured as the ratio of net property, plant, and equipment to the number of employees), firm age, and SIC2 industry-year fixed effects. All independent variables are lagged by one year relative to the dependent variables. Robust t-statistics are shown in parentheses. ***, **, and * indicate significance at the 1 %, 5 %, and 10 % levels, respectively.

instrumental in predicting corporate innovation by revealing key insights into employee treatment. Our findings extend prior research (e.g., Chang et al., 2015; Bradley et al., 2017; Chen et al., 2016a; Chen et al., 2016b; Mao and Weathers, 2019) by identifying key employee treatment practices that show strong associations with innovation outcomes. This research provides managers with actionable insights and offers a foundation for scholars to develop new theories and hypotheses regarding corporate innovation drivers.

These findings collectively support our two proposed hypotheses. First, the predictive relevance of employee-generated reviews for innovation outcomes provides strong empirical support for H1. This confirms the value of social media-based employee feedback as a signal of corporate innovation potential. Second, the SHAP analysis and regression results show that specific human resource practices, particularly flexible working hours and equity-based compensation, have stronger associations with innovation than others. This affirms H2. By confirming both hypotheses, our results demonstrate that employee treatment information extracted from social media captures meaningful firm-level human resource characteristics. It also reveals significant patterns linked to innovation outputs across multiple measures.

Additionally, our study distinguishes itself by employing the text innovation metric of Bellstam et al. (2020) on 10-K filings to develop an additional innovation measure, proving its predictive strength for patent outcomes and revealing dimensions of innovation not captured by traditional patent metrics. This supports the broader potential of text-based measures to capture diverse innovative activities across organizational functions, a finding corroborated by Bellstam et al. (2020). By aligning our methodology with the framework proposed by Bellstam et al. (2020), we address concerns regarding the validity of our text-based innovation measure and leverage their extensive testing and validation to ensure a more robust and reliable analysis. Additionally, this alignment allows us to contribute to the literature by applying a well-established framework to examine innovation across a broader spectrum of firm activities. We observe a robust positive link between innovation texts and patent outputs. However, firm characteristics like size, age, and capital intensity showed negative associations with text-based innovation, contrasting with patent data, and indicative of traits in a modern, innovative organization (e.g., Zingales, 2000; Edmans, 2011). Furthermore, our results show that *R&D intensity* and *total assets* are critical predictors for patent metrics, indicating that extensive resources play a more vital role in patent-based than text-based innovation. This contrasts with the broader impact of positive employee treatment on generating diverse innovations across the organization, a factor more significantly captured by text-based measures. Our results align with Bellstam et al. (2020), indicating that text-based metrics better reflect the innovation contributions of rank-and-file employees across various organizational functions, highlighting their role as innovators more distinctly than patent-focused measures.

Despite the significant predictive power of control variables in the ML models, our findings emphasize the importance of employee treatment practices. Several employee treatment features show statistically significant contributions to predicting patent counts, patent citations, and text-based innovation. Our analysis reveals that specific HR policies and practices are notably stronger predictors of corporate innovation than others. This contributes to prior research by identifying the most effective types of employee treatment practices for predicting innovation metrics. Notably, *flexible working hours* and *stock or equity options* emerge as key predictors across all three innovation measures, demonstrating strong and positive statistical associations. A higher proportion of employees with access to these benefits correlates with increased innovation output.

Prior research (e.g., Oyer and Schaefer, 2005) has provided several theoretical explanations for why firms should or should not offer stock options to all employees. Our results indicate that this benefit is among the most significant in supporting firm innovation—not only patent-based innovation but also other forms of innovativeness. This finding suggests that firms should consider offering stock options more broadly to all employees. Moreover, in innovation-driven contexts, attractive benefits such as stock options play a crucial role in recruiting top talent, as highlighted by studies such as Chang et al. (2015) and Chen et al. (2016b).

Our study also contributes to the growing body of research on flexible work arrangements and their implications for firm innovation. We provide empirical evidence that flexible working hours are positively associated with both patent-based and text-based innovation, highlighting their role as a valuable employee benefit that enhances motivation and fosters innovation. This finding supports existing literature emphasizing the importance of freedom, autonomy, and flexibility in work arrangements as key drivers of creativity and innovation (Amabile et al., 1996; Anderson et al., 2014; Krammer, 2022). However, our results also reveal that not all types of flexibility contribute equally to different forms of innovation. While the ability to work remotely is strongly associated with text-based innovation, we do not observe a similar effect for patent-based innovation. In addition, freedom to work independently is statistically insignificant or among the least important employee benefits in relation to patent-based innovation. However, in the case of text-based innovation freedom to work independently is a highly important predictor with a statistically significant positive association.

This suggests that different forms of innovation may require distinct work environments. Text-based innovation benefits from the autonomy and concentration facilitated by remote work, while patent-based innovation, which often involves collaboration and hands-on experimentation, may rely more on in-person interactions. This nuanced insight aligns with prior research on remote work, which has identified both advantages (e.g., productivity, work-life balance, knowledge sharing, creativity, employee retention) and challenges (e.g., reduced collaboration, isolation, communication barriers, and work-life balance tensions) (e.g. Allen et al., 2015; Battisti et al., 2022; Becker et al., 2024; Donnelly and Johns, 2021; Kelliher et al., 2019; Kelliher and de Menezes, 2019; Pérez et al., 2002; Vyas and Butakhieo, 2020).

Our findings also contribute to the broader discussion on the effectiveness of flexible work arrangements. While previous research has explored the implications of flexitime and flexispace, the results remain mixed, with studies documenting both positive and negative effects (Allen et al., 2013; Azar et al., 2018; Becker et al., 2024; Edwards and Rothbard, 2000; Haines et al., 2024; Masuda et al., 2012). Our study adds to this debate by demonstrating that the impact of flexible work policies is context-dependent, with different types of flexibility fostering distinct innovation outcomes.

Moreover, the literature has pointed out potential downsides of remote working, such as productivity concerns, social isolation, reduced collaboration, and stress related to digital work environments (e.g. Atkin et al., 2023; Soga et al., 2022; Biron et al., 2021; Fuller and Hirsh, 2019;

Jacobs and Padavic, 2015; Yang et al., 2022). Our results suggest that these challenges may be particularly relevant for patent-based innovation, where collaboration and direct engagement with colleagues play a crucial role.

Our findings have several practical implications for firms seeking to optimize their innovation outcomes through flexible work policies. First, they suggest that offering flexible working hours can be a powerful tool for fostering innovation across multiple dimensions. Second, remote work policies should be designed with a nuanced understanding of their impact on different innovation types. While they are beneficial for text-based innovation, they may not be as effective for patent-based innovation. From a theoretical perspective, our study highlights the need to further explore the mechanisms through which different types of autonomy and flexibility influence innovation outcomes. Future research should examine how specific job roles, team dynamics, and firm cultures mediate these relationships. Additionally, as more firms transition to hybrid work models, future studies should investigate how the balance between remote and in-person work affects innovation performance over the long term.

In addition, our results demonstrate a statistically significant positive association between company image (employee pride in the company) and both patent counts and citations, aligning with social identity theory and prior research on organizational identification (Ashforth and Mael, 1989; Jones, 2010; Riketta, 2005). Employees who feel proud to work for their company are more likely to internalise organisational goals and align their efforts with the firm's strategic objectives, fostering greater innovation (Farooq et al., 2017; Jones, 2010). Organisational pride has been shown to enhance employee motivation, collaboration, and commitment—key factors that contribute to knowledge creation and the successful development of new technologies (Dukerich et al., 2002; Schuh et al., 2016; van Knippenberg et al., 2007). Furthermore, firms that cultivate positive company image might indicate about an environment where employees are more engaged in discretionary innovative activities, leading to higher patent output and greater technological impact. Our findings suggest that fostering a strong sense of identification and pride among employees can be an important driver of firm-level innovation outcomes, reinforcing the role of good employee treatment and corporate culture in shaping innovation performance.

Employees' pride in their company (company image), appears to capture fundamental and enduring aspects of corporate culture, values, and reputation, which can have a meaningful impact on firm innovation. Unlike more immediate indicators such as job security, which may reflect short-term operational performance, company image reflects employees' deeper connection to their organisation—its mission, stability, and long-term prospects. Employees who take pride in their company are more likely to internalise its goals, engage in knowledge-sharing, and contribute to innovative efforts that drive long-term success. This aligns with research suggesting that changes in employee satisfaction reflect fundamental shifts within firms, which markets may be slow to recognise (Green et al., 2019). A strong company image, as a proxy for employee identification and organisational pride, may indicate a firm's ability to sustain innovative capabilities over time. Rather than reacting to short-term financial fluctuations, company image reflects employees' broader perceptions of their firm's strategic direction, long-term stability, and capacity for continuous improvement. Our findings suggest that company image plays a crucial role in fostering an environment conducive to innovation, where employees are motivated to contribute to technological advancement and long-term value creation.

These findings provide novel insights into the relationship between employee treatment practices and innovation outcomes. The robustness of these predictors across both predictive and explanatory models underscores their importance and validates their theoretical significance. By focusing on these underexplored variables, our study contributes to a more nuanced understanding of the factors that drive corporate innovation. Future research can build on these findings to explore these key variables in more detail.

This study's conclusions should be viewed considering its limitations. First, while the extreme gradient-boosting algorithm is effective, it complicates interpretation due to the inherent trade-off between flexibility and interpretability. We utilized advanced nonlinear methods, as a linear model approach would have missed the nonlinear relationships uncovered here. To overcome interpretability challenges, we utilized efficient tools from explainable AI, SHAP values. Second, the voluntary nature of employee ratings on social media raises concerns about omitted variable bias, particularly if specific events influence these reviews. While our sample includes both successful and less successful firms, and we incorporate controls to address these issues, the reliance on social media data may introduce non-random sampling, as firms with higher employee engagement on platforms like Kununu are more likely to be represented. These limitations should be considered when interpreting the findings. Third, the possibility of positive bias in 10-K filings due to impression management by companies may distort the portrayal of their innovativeness. Future research could compare text-based and patent-based innovation measures under various conditions to assess the strength of their association. Moreover, our method's ability to detect innovation may be limited if companies withhold disclosing innovative ideas as trade secrets.

Our study paves the way for explanatory research with causal models to explore deeper into workplace diversity values and their impact on innovation. Crowdsourced social media data could provide a novel resource for investors, employees, and companies to access and interpret information regarding employee treatment and diversity. We find that the relationship between flexible work arrangements and firm innovation performance is complex, and employee ratings on social media can provide fresh insights into company policies that may not be captured elsewhere. While we leverage firm-level data to analyze the impact of flexible work arrangements on innovation, more granular data at the individual or team level could provide deeper insights into how employees perceive and utilize flexibility for creative work. Additionally, as remote work and hybrid work models continue to evolve, future research should track how firms adjust their innovation strategies and collaboration practices in response to changing work environments. Furthermore, future research could employ the SHAP method for a detailed analysis of how employee treatment values interact with other critical firm characteristics. The ML approach we introduced allows for estimating the contributions of different variable interactions at both global and local scales, presenting opportunities to expand this line of inquiry. Focusing solely on linear effects or simple interactions might overlook crucial factors that influence whether employee treatment practices drive innovation (Doornenbal et al., 2021).

## CRediT authorship contribution statement

**Mika Ylinen:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **Mikko Ranta:** Writing – review & editing, Methodology, Formal analysis, Data curation.

## Declaration of conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Definitions of Research Variables

| Variable | Definition |
|---|---|
| *Independent variables* | Employee's rating of employer ranked on a 5-point scale, with 5 (1) being most favourable (unfavourable) |
| Work-life balance | How does the company value work-life balance? Are families considered? Is there pressure to work long hours? |
| Freedom to work independently | To what extent are you trusted to work independently? |
| Inclusive/diverse | To what extent does the company value diversity in the workplace? Are diverse ideas and opinions supported? |
| Gender equality | Are women treated equally and given the same career opportunities? |
| Attitude towards older colleagues | Does the company hire older workers? Are senior colleagues appreciated, supported, and given equal opportunities? |
| Handicapped accessibility | Does the company have facilities that are handicapped accessible to support people with disabilities? |
| Job security | How stable do you feel your job is? |
| Company image | Are you proud to work for your company? |
| Career development | How are your career prospects for growth and professional development? |
| Ability to work remotely | Percentage of respondents of a company that are entitled to benefit. |
| Flexible working hours | Percentage of respondents of a company that are entitled to benefit. |
| Paid parental leave | Percentage of respondents of a company that are entitled to benefit. |
| Stock or equity options | Percentage of respondents of a company that are entitled to benefit. |
| Paid time off | Percentage of respondents of a company that are entitled to benefit. |
| *Dependent variables* | |
| Patent counts (innovation quantity) | Total number of patents filed (and eventually granted) by a firm in a year. |
| Patent citations (innovation quality) | Total number of citations. |
| Text-based innovation | Latent Dirichlet allocation (LDA), a topic modelling method, to employ a text-based innovation measure from 10-K fillings in line with Bellstam et al. (2020). |
| *Control variables* | |
| Total assets | Book value of assets at the end of fiscal year. |
| Leverage | Ratio of the sum of short-term and long-term debts to total book value of assets. |
| Sales growth | The growth rate in sales from year t-1 to year t. |
| Year | |
| State | |
| Industry | SIC-2 code |
| R&D Intensity | Research and development expenditure scaled by total assets of a company. |
| ROA | Net income divided by total assets of a company. |
| Implied volatility | Volatility assumption percentage. |
| Capital intensity | Property, plant & equipment divided by total number of employees. |
| Firm age | Firm age is computed by the difference between the firm's first year appeared in Compustat and the current year. |
| Tobin's Q | Computed as (Total assets – common ordinary equity + (common shares outstanding × annual fiscal price close))/total assets. |

## Appendix B. Sensitivity Analysis of the LDA Model

The choice of the number of topics in the LDA model significantly impacts the granularity and interpretability of the output. To determine the optimal number of topics, we evaluated models with topic numbers ranging from 10 to 60, using perplexity and coherence scores as evaluation metrics. As shown in Table B1, the model with 40 topics achieved the best balance between minimizing perplexity and maximizing coherence scores.

In addition to these metrics, we conducted a manual inspection of the topic outputs to ensure interpretability and alignment with the study's objectives. Models with fewer topics (e.g., 10 or 20) grouped distinct themes into overly broad categories, while models with more topics (e.g., 50 or 60) produced fragmented themes that were less meaningful for the analysis.

To assess the robustness of our findings, we compared the results of models with different topic numbers. While the number of topics slightly influenced the granularity of the themes identified, the main conclusions regarding the key predictors of corporate innovation and their relevance remained consistent across models.

**Table B1**
Perplexity and coherence scores for different topic numbers

| No. topics | Coherence | Perplexity |
|---|---|---|
| 10 | 0.51 | −0.45 |
| 20 | 0.48 | −0.61 |
| 30 | 0.50 | −0.68 |
| 40 | 0.52 | −0.69 |
| 50 | 0.51 | −0.69 |
| 60 | 0.51 | −0.66 |
| 70 | 0.51 | −0.64 |

Note: Table plots the changes in perplexity and coherence as the number of topics increases, identifying 40 topics as the optimal balance point.

## Appendix C. Topic Stability Across LDA Models

**Table C1**
Most frequent words in innovation topics across topic models

| 30-topic model | 40-topic model | 50-topic model |
|---|---|---|
| [('solution', 0.017823877), | [('patent', 0.020066546), | [('solution', 0.017285082), |
| ('care', 0.015486031), | ('device', 0.019253021), | ('device', 0.015527119), |
| ('device', 0.014159026), | ('patient', 0.014454303), | ('patent', 0.011134772), |
| ('software', 0.012840858), | ('tcs', 0.014210451), | ('software', 0.010133649), |
| ('medical', 0.01264999), | ('medical', 0.011496481), | ('patient', 0.0100174565), |
| ('patient', 0.009894571), | ('confidential', 0.009973075), | ('medical', 0.009558127), |
| ('patent', 0.008871885), | ('licensee', 0.008213935), | ('care', 0.008093627), |
| ('goodwill', 0.0075140344), | ('care', 0.0068054446), | ('goodwill', 0.0076180873), |
| ('healthcare', 0.006815474), | ('clinical', 0.006490707), | ('supplier', 0.0063377125), |
| ('intellectual', 0.0051179286), | ('intellectual', 0.005877296), | ('intellectual', 0.005586089), |
| ('supplier', 0.004960969), | ('reimbursement', 0.0057979887), | ('manufacturer', 0.0053414796), |
| ('hospital', 0.0047146687), | ('physician', 0.005682688), | ('thousand', 0.0051795086), |
| ('thousand', 0.004545686), | ('healthcare', 0.0051219305), | ('manufacture', 0.005067814), |
| ('manufacturer', 0.00422764), | ('manufacture', 0.0048371665), | ('manufacturing', 0.004682697), |
| ('reimbursement', 0.00407512)] | ('convertible', 0.004398373)] | ('distributor', 0.004442076)] |

Note: Semantic stability across topic models suggests consistent identification of innovation-related themes.

The comparison of the most salient words associated with innovation-related topics across the 30, 40, and 50 topic LDA models reveals a notable consistency in the semantic content of these topics, despite minor variations in word rankings and weights. In each of the three models, key terms such as "solution," "device," "medical," "patient," and "care" consistently appear among the top-ranked words. This consistency suggests that the core thematic focus of the innovation topics remains stable regardless of the number of topics specified in the LDA model. In addition, terms related to intellectual property and business operations, including "patent," "intellectual," "licensee," and "goodwill," appear across all models, underscoring the importance of proprietary knowledge and organizational assets in discussions of innovation within annual reports. Operational terms such as "supplier," "manufacturer," and "manufacture" also occur frequently, reflecting the broader context of innovation in relation to production and supply chain considerations.

The strong lexical overlap and the stability of the core vocabulary across the three models indicate that, even as the level of topic granularity increases, the fundamental contours of innovation discourse in annual reports remain robust. This consistency suggests that the number of topics, within the tested range, does not materially influence the identification of central innovation themes, which continue to reflect a mix of technological and organizational concepts.

## Appendix D. Descriptive Comparison of S&P 1500 vs. Full Sample

**Table D1**
Average employee review scores for S&P 1500 vs. full sample

| | S&P1500 | Full sample |
|---|---|---|
| Work–life balance | 3.360 | 3.260 |
| Freedom to work independently | 3.523 | 3.470 |
| Inclusive/diverse | 3.418 | 3.380 |
| Gender equality | 3.627 | 3.660 |
| Attitude towards older colleagues | 3.620 | 3.580 |
| Accessible to people with disabilities | 3.688 | 3.690 |
| Job security | 3.295 | 3.230 |
| Company image | 3.592 | 3.500 |
| Career development | 3.303 | 3.220 |
| Ability to work remotely | 0.165 | 0.180 |
| Flexible working hours | 0.322 | 0.320 |
| Paid parental leave | 0.194 | 0.210 |
| Stock or equity options | 0.257 | 0.270 |
| Paid time off | 0.490 | 0.550 |

Table D1 compares average employee review scores across several workplace dimensions for firms in the S&P 1500 index and for the full sample of companies on the Kununu platform. While S&P 1500 firms tend to have slightly higher scores in areas such as work–life balance, job security, company image, and career development, the differences are generally small. In contrast, the full sample shows marginally higher ratings in specific benefits including gender equality, remote work, and paid time off.

## Appendix E. Regression Validation of Innovation Measures

In Table E1, we compare our text-based innovation measure to an alternative augmented dictionary approach. We construct our dictionary measure using an ML technique, specifically a word embedding model. This approach quantifies the degree of innovation-related disclosure in firms' 10-K filings. We draw on Li et al. (2020) that leverage word embeddings to measure corporate concepts and begin with a set of innovation-related seed words. The word embedding model, trained on a large corpus of 10-K and 10-Q filings from the SEC's EDGAR database, identifies a broad set of words and phrases that are contextually similar to our innovation seeds. Using a neural network model (word2vec, Mikolov et al., 2013a, 2013b), we generate vector representations of words and expand our innovation dictionary with the most semantically similar terms. We then calculate the relative frequency of these innovation-related words and phrases in each firm's 10-K report, providing a data-driven, context-sensitive measure of innovation disclosure. This method addresses the limitations of traditional keyword approaches by capturing the nuanced language of innovation in corporate filings.

**Table E1**
Comparison of Bellstam text-based and word-embedding innovation measures in predicting patent outcomes

| Variable | Patent Counts (1) | Patent Counts (2) | Patent Counts (3) | Patent Citations (4) | Patent Citations (5) | Patent Citations (6) |
|---|---|---|---|---|---|---|
| Bellstam text-innovation | 5.51*** (15.68) | | 5.30*** (14.74) | 8.83*** (20.54) | | 8.48*** (19.53) |
| Word-embedding Innovation | | 0.00057*** (6.54) | 0.00029*** (3.38) | | 0.00093*** (7.09) | 0.00047*** (3.73) |
| R&D Intensity | 20.58*** (13.11) | 27.20*** (16.56) | 20.28*** (12.90) | 21.85*** (10.32) | 32.44*** (14.41) | 21.32*** (10.06) |
| Firm Age | 0.0018*** (14.37) | 0.0016*** (12.23) | 0.0019*** (14.49) | 0.0018*** (9.64) | 0.0014*** (7.41) | 0.0018*** (9.78) |
| Firm Size | 3.13e-06*** (12.58) | 2.73e-06*** (9.75) | 2.86e-06*** (10.77) | 4.10e-06*** (11.53) | 3.40e-06*** (8.38) | 3.66e-06*** (9.64) |
| Capital Intensity | −0.43 (-0.50) | 0.05 (0.06) | −0.51 (-0.59) | 1.54 (1.23) | 1.93 (1.49) | 1.43 (1.15) |
| Tobin's Q | −0.029 (-0.91) | 0.025 (0.68) | −0.017 (-0.54) | −0.041 (-0.94) | 0.046 (0.85) | −0.022 (-0.50) |
| ROA | 3.08*** (5.79) | 3.51*** (5.78) | 3.13*** (5.89) | 4.16*** (5.56) | 5.01*** (5.77) | 4.23*** (5.67) |
| Industry and Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| R-squared | 0.606 | 0.574 | 0.607 | 0.623 | 0.58 | 0.625 |
| Observations | 4277 | 4277 | 4277 | 3943 | 3943 | 3943 |

This table reports the results of linear regressions examining the relationship between the Bellstam text-based innovation measure and the word-embedding innovation measure with two types of innovation outputs: patent citations and patent counts. Robust t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels, respectively. All independent variables are lagged by one year relative to the dependent variable. Control variables include firm size (total assets), leverage (long-term debt to total assets), R&D intensity (R&D expenditure to total assets), return on assets (operating income to book assets), Tobin's Q, capital intensity (net property, plant, and equipment per employee), and firm age. All regressions include year and industry (SIC2) fixed effects. The sample covers the period 2008–2019. The $R^2$ for the patent counts model is 0.564 when neither text-innovation variable is included, and 0.544 when both are excluded from the patent citations model.

## Appendix F. SHAP Analysis of Innovation Predictors

*Patents filings:*

Fig. F1 highlights the average SHAP values for features included in the patent filings prediction model. The SHAP value quantifies the average magnitude of each feature's contribution to the model's output, thus providing a transparent measure of feature importance in the context of machine learning interpretability. Among the innovation-related variables, the Bellstam text-innovation measure demonstrates notably greater predictive importance than the word-embedding-based innovation measure, despite the latter's sophistication and popularity in prior textual analysis research. The Bellstam measure ranks just below traditional financial predictors such as R&D intensity and firm size, underscoring its strong explanatory contribution to patent activity. In contrast, the word-embedding innovation measure appears among the least influential predictors, suggesting that the KL-divergence-based method used in Bellstam et al. (2020) may capture more relevant language patterns related to innovation. These findings reinforce the value of carefully targeted domain-specific text metrics in complementing machine learning models of corporate innovation.

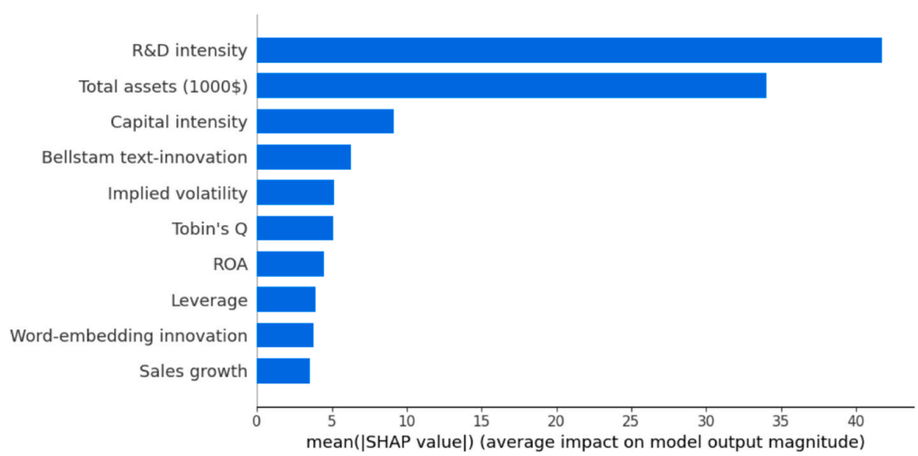



**Fig. F1.** SHAP Value importance for patent filings prediction model

*Patent citations:*

Fig. F2 displays the mean absolute SHAP values associated with each predictor in the model estimating patent citations. Consistent with the results from the patent filings model, the Bellstam text innovation measure demonstrates significantly greater predictive importance than the word embedding-based innovation metric. Despite the methodological sophistication and wide adoption of word embeddings in textual research, their contribution to explaining patent citation outcomes remains relatively limited in this context. In contrast, the Bellstam measure, constructed using topic modeling and aligned with innovation-related reference text, emerges as one of the more influential innovation-related predictors, following R&D intensity and firm size. These findings support the validity of the Bellstam measure and demonstrate its usefulness as a complementary text-based metric that adds value alongside traditional patent-based indicators.

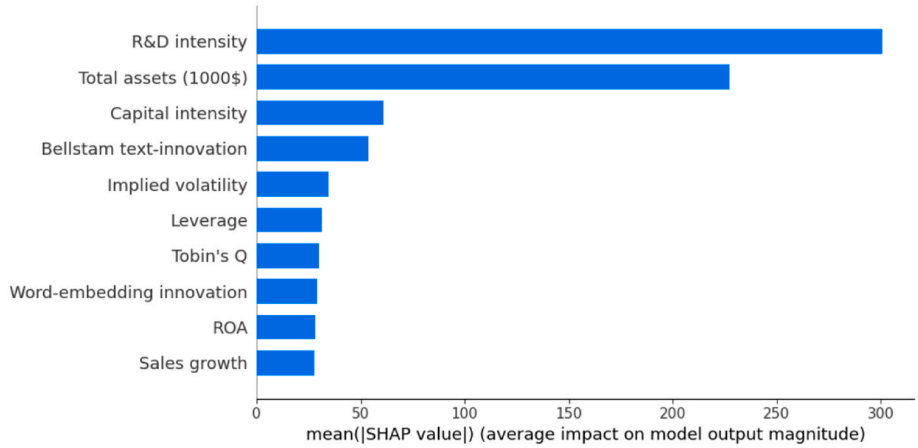



**Fig. F2.** SHAP Value importance for patent citations prediction model

## Data availability

The authors do not have permission to share data.

## References


Acar, O.A., Tarakci, M., van Knippenberg, D., 2019. Creativity and innovation under constraints: a cross-disciplinary integrative review. J. Manag. 45 (1), 96–121. https://doi.org/10.1177/0149206318805832.

Acharya, V., Baghai, R., Subramanian, K., 2014. Wrongful discharge laws and innovation. Rev. Financ. Stud. 27, 301–346.

Agrawal, A., Gans, J., Goldfarb, A., 2019. The impact of machine learning on economics. In: Agrawal, A., Gans, J., Goldfarb, A. (Eds.), The Economics of Artificial Intelligence: an Agenda. University of Chicago Press, Chicago, pp. 507–552. https://doi.org/10.7208/chicago/9780226613475.003.0021.

Allen, T.D., Golden, T.D., Shockley, K.M., 2015. How effective is telecommuting? Assessing the status of our scientific findings. Psychol. Sci. Publ. Interest 16 (2), 40–68. https://doi.org/10.1177/1529100615593273.

Allen, T.D., Johnson, R.C., Kiburz, K.M., Shockley, K.M., 2013. Work–family conflict and flexible work arrangements: deconstructing flexibility. Pers. Psychol. 66 (2), 345–376. https://doi.org/10.1111/peps.12012.

Anderson, N., Potočnik, K., Zhou, J., 2014. Innovation and creativity in organizations. J. Manag. 40 (5), 1297–1333. https://doi.org/10.1177/0149206314527128.

Amabile, T.M., 1988. A model of creativity and innovation in organizations. Res. Organ. Behav. 10, 123–167.

Amabile, T.M., Conti, R., Coon, H., Lazenby, J., Herron, M., 1996. Assessing the work environment for creativity. Acad. Manag. J. 39 (5), 1154–1184, 1996.

Ashforth, B.E., Mael, F., 1989. Social identity theory and the organisation. Acad. Manag. Rev. 14 (1), 20–39. https://doi.org/10.2307/258189.

Atkin, D., Schoar, A., Shinde, S., 2023. Working from home, worker sorting, and development. NBER Working Paper No., 31515. http://www.nber.org/papers/w31515.

Azar, S., Khan, A., Van Eerde, W., 2018. Modelling linkages between flexible work arrangements' use and organizational outcomes. J. Bus. Res. 91, 134–143. https://doi.org/10.1016/j.jbusres.2018.06.004.

Bae, K.H., Kang, J.K., Wang, J., 2011. Employee treatment and firm leverage: a test of the stakeholder theory of capital structure. J. Financ. Econ. 100 (1), 130–153. https://doi.org/10.1016/j.jfineco.2010.10.019.

Bao, Y., Ke, B., Li, B., Yu, Y.J., Zhang, J., 2020. Detecting accounting fraud in publicly traded US firms using a machine learning approach. J. Account. Res. 58 (1), 199–235. https://doi.org/10.1111/1475-679X.12292.

Barboza, F., Kimura, H., Altman, E., 2017. Machine learning models and bankruptcy prediction. Expert Syst. Appl. 83, 405–417. https://doi.org/10.1016/j.eswa.2017.04.006.

Bartov, E., Faurel, L., Mohanram, P.S., 2018. Can Twitter help predict firm-level earnings and stock returns? Account. Rev. 93 (3), 25e57. https://doi.org/10.2308/accr-51865.

Battisti, E., Alfiero, S., Leonidou, E., 2022. Remote working and digital transformation during the COVID-19 pandemic: economic–Financial impacts and psychological drivers for employees. J. Bus. Res. 150, 38–50. https://doi.org/10.1016/j.jbusres.2022.06.010.

Basu, S., Ma, X., Briscoe-Tran, H., 2021. Measuring multidimensional investment opportunity sets with 10-K text. Account. Rev. https://doi.org/10.2308/TAR-2019-0110.

Beck, N., King, G., Zeng, L., 2004. Theory and evidence in international conflict: a response to De Marchi, Gelpi, and Grynaviski. Am. Polit. Sci. Rev. 98, 379–389. https://doi.org/10.1017/S0003055404001212.

Becker, K., Ebbers, J.J., Engel, Y., 2024. Going online: peer entrepreneur networks in a startup accelerator before and during the COVID-19 pandemic. Technovation 130. https://doi.org/10.1016/j.technovation.2023.102917. Article 102917.

Biron, M., De Cieri, H., Fulmer, I., Lin, C.-H. (Veronica), Mayrhofer, W., Nyfoudi, M., Sanders, K., Shipton, H., Sun, J. M.James, 2021. Structuring for innovative responses to human resource challenges: a skunk works approach. Hum. Resour. Manag. Rev. 31 (2), 100768. https://doi.org/10.1016/j.hrmr.2020.100768.

Bellstam, G., Bhagat, S., Cookson, J.A., 2020. A text-based analysis of corporate innovation. Manag. Sci. 67 (7), 4004–4031.

Bertomeu, J., 2020. Machine learning improves accounting: discussion, implementation and research opportunities. Rev. Account. Stud. 25 (3), 1135–1155. https://doi.org/10.1007/s11142-020-09554-9.

Bertomeu, J., Cheynel, E., Floyd, E., Pan, W., 2021. Using machine learning to detect misstatements. Rev. Account. Stud. 26 (2), 468–519. https://doi.org/10.1007/s11142-020-09563-8.

Blei, D.M., Ng, A.Y., Jordan, M.I., Lafferty, J., 2003. Latent dirichlet allocation. J. Mach. Learn. Res. 3 (4/5), 993–1022.

Bradley, D., Kim, I., Tian, X., 2017. Do unions affect innovation? Manag. Sci. 63 (7), 2251–2271.

Buehlmaier, M.M.M., Whited, T.M., 2018. Are financial constraints priced? Evidence from textual analysis. Rev. Financ. Stud. 31 (7), 2693–2728. https://doi.org/10.1093/RFS/HHY007.

Cade, N.L., 2018. Corporate social media: how two-way disclosure channels influence investors. Account. Organisat. Soc. 68–69 (July), 63–79. https://doi.org/10.1016/j.aos.2018.03.004.

Cai, W., Mahlendorf, M.D., Wu, F., 2024. Supervisor Behavior in Times of External Performance Pressure. Available at: SSRN 4827928.

Campbell, J.L., DeAngelis, M.D., Moon, J.R., 2019. Skin in the game: personal stock holdings and investors' response to stock analysis on social media. Rev. Account. Stud. 24, 731–779. https://doi.org/10.1007/s11142-019-09498-9.

Cao, Z., Rees, W., 2020. Do employee-friendly firms invest more efficiently? Evidence from labor investment efficiency. J. Corp. Finance 65, 101744.

Cauchy, A., 1847. Méthode générale pour la résolution des systemes d'équations simultanées. C. R. Acad. Sci. 25, 536–538.

Chang, X., Fu, K., Low, A., Zhang, W., 2015. Non-executive employee stock options and corporate innovation. J. Financ. Econ. 115, 168–188. https://doi.org/10.1016/j.jfineco.2014.09.002.

Chen, Y., Fan, Q., Yang, X., Zolotoy, L., 2021. CEO early-life disaster experience and stock price crash risk. J. Corp. Finance 68, 101928. https://doi.org/10.1016/j.jcorpfin.2021.101928.

Chen, C., Chen, Y., Hsu, P.H., Podolski, E.J., 2016b. Be nice to your innovators: employee treatment and corporate innovation performance. J. Corp. Finance 39, 78–98.

Chen, H., De, P., Hu, Y., Hwang, B.-H., 2014. Wisdom of crowds: the value of stock opinions transmitted through social media. Rev. Financ. Stud. 27 (5), 1367e1403. https://doi.org/10.1093/rfs/hhu001.

Chen, J., Leung, W.S., Evans, K., 2016a. Are employee-friendly workplaces conducive to innovation? J. Corp. Finance 40, 61–79. https://doi.org/10.1016/j.jcorpfin.2016.07.011.

Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16). Association for Computing Machinery, pp. 785–794. New York, NY, USA. https://doi.org/10.1145/2939672.2939785.

Chiarello, F., Giordano, V., Spada, I., Barandoni, S., Fantoni, G., 2024. Future applications of generative large language models: a data-driven case study on

ChatGPT. Technovation 133, 103002. https://doi.org/10.1016/j.technovation.2024.103002.
Choudhury, P., Allen, R.T., Endres, M.G., 2021. Machine learning for pattern discovery in management research. Strateg. Manag. J. 42, 30–57. https://doi.org/10.1002/smj.3215.
Covert, I., Lundberg, S., Lee, S.-I., 2020. Understanding global feature contributions with additive importance measures. http://arXiv:2004.00668.
Deng, X., Kang, J.-K., Low, B., 2013. Corporate social responsibility and stakeholder value maximization: evidence from mergers. J. Financ. Econ. 110 (1), 87–109. https://doi.org/10.1016/j.jfineco.2013.04.014.
Dimic, N., Fatmy, V., Vähämaa, S., 2024. Religiosity and corporate social responsibility: a study of firm-level adherence to Christian values in the United States. Corp. Soc. Responsib. Environ. Manag. 31 (1), 396–413. https://doi.org/10.1002/csr.2576.
Ding, K., Lev, B., Peng, X., Sun, T., Vasarhelyi, M.A., 2020. Machine learning improves accounting estimates: evidence from insurance payments. Rev. Account. Stud. 25 (3), 1098–1134. https://doi.org/10.1007/s11142-020-09546-9.
Donnelly, R., Johns, J., 2021. Recontextualising remote working and its HRM in the digital economy: an integrated framework for theory and practice. The international journal of *human resource management*, 32(1), 84-105. https://doi.org/10.1080/09585192.2020.1737834.
Donovan, J., Jennings, J., Koharki, K., Lee, J., 2021. Measuring credit risk using qualitative disclosure. Rev. Account. Stud. 26 (2), 815–863. https://doi.org/10.1007/S11142-020-09575-4.
Doornenbal, B., Spisak, B., van der Laken, P., 2021. Opening the Black box: uncovering the leader trait paradigm through machine learning. Leadersh. Q. 33 (5), 101515. https://doi.org/10.1016/j.leaqua.2021.101515.
Dukerich, J.M., Golden, B.R., Shortell, S.M., 2002. Beauty is in the eye of the beholder: the impact of organisational identification, identity, and image on the cooperative behaviors of physicians. Adm. Sci. Q. 47 (3), 507–533. https://doi.org/10.2307/3094849.
Ederer, F., Manso, G., 2013. Is pay-for-performance detrimental to innovation? Manag. Sci. 59, 1498–1513.
Edmans, A., 2011. Does the stock market fully value intangibles? Employee satisfaction and equity prices. J. Financ. Econ. 101 (3), 621e640. https://doi.org/10.1016/j.jfineco.2011.03.021.
Edmans, A., 2012. The link between job satisfaction and firm value, with implications for corporate social responsibility. Acad. Manag. Perspect. 26 (4), 1e19. https://doi.org/10.5465/amp.2012.0046.
Edmans, A., Li, L., Zhang, C., 2014. Employee Satisfaction, Labor Market Flexibility, and Stock Returns Around the World. National Bureau of Economic Research. Available at::Working paper 20300 http://www.nber.org/papers/w20300.
Edwards, J.R., Rothbard, N.P., 2000. Mechanisms linking work and family: clarifying the relationship between work and family constructs. Acad. Manag. Rev. 25 (1), 178–199. https://doi.org/10.2307/259269.
Efron, B., 1979. Bootstrap methods: another look at the jackknife. Ann. Stat. 7 (1), 1–26. https://doi.org/10.1214/aos/1176344552.
Elliott, W.B., Grant, S.M., Hodge, F.D., 2018. Negative news and investor trust: the role of $firm and # CEO Twitter use. J. Account. Res. 56 (5), 1483–1519. https://doi.org/10.1111/1475-679X.12217.
Emran, M.S., Hou, Z., 2013. Access to markets and rural poverty: evidence from household consumption in China. Rev. Econ. Stat. 95 (2), 682–697. https://doi.org/10.1162/REST_a_00354.
Ertugrul, M., 2013. Employee-friendly acquirers and acquisition performance. J. Financ. Res. 36 (3), 347–370. https://doi.org/10.1111/j.1475-6803.2013.12014.x.
Faleye, O., Trahan, E.A., 2011. Labor-friendly corporate practices: is what is good for employees good for shareholders? J. Bus. Ethics 101 (1), 1–27. https://doi.org/10.1007/s10551-010-0705-9.
Farooq, O., Rupp, D.E., Farooq, M., 2017. The multiple pathways through which internal and external corporate social responsibility influence organisational identification and multifoci outcomes: the moderating role of cultural and social orientations. Acad. Manag. J. 60 (3), 954–985. https://doi.org/10.5465/amj.2014.0849.
Fauver, L., McDonald, M.B., Taboada, A.G., 2018. Does it pay to treat employees well? International evidence on the value of employee-friendly culture. J. Corp. Finance 50 (June), 84–108. https://doi.org/10.1016/j.jcorpfin.2018.02.003.
Frankel, R., Jennings, J., Lee, J., 2016. Using unstructured and qualitative disclosures to explain accruals. J. Account. Econ. 62 (2–3), 209–227. https://doi.org/10.1016/J.JACCECO.2016.07.003.
Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29 (5), 1189–1232. https://doi.org/10.2307/2699986.
Friedman, J.H., 2002. Stochastic gradient boosting. Comput. Stat. Data Anal. 38 (4), 367–378. https://doi.org/10.1016/S0167-9473(01)00065-2.
Friedman, J., Hastie, T., Tibshirani, R., 2000. Additive logistic regression: a statistical view of boosting. Ann. Stat. 28, 337–407. https://doi.org/10.1214/aos/1016218223.
Futagami, K., Fukazawa, Y., Kapoor, N., Kito, T., 2021. Pairwise acquisition prediction with SHAP value interpretation. J. Finance Data Sci. 7, 22–44. https://doi.org/10.1016/j.jfds.2021.02.001.
Geissinger, A., Laurell, C., Öberg, C., Sandström, C., 2023. Social media analytics for innovation management research: a systematic literature review and future research agenda. Technovation 123, 102712. https://doi.org/10.1016/j.technovation.2023.102712.
Gepp, A., Kumar, K., Bhattacharya, S., 2021. Lifting the numbers game: identifying key input variables and a best-performing model to detect financial statement fraud. Account. Finance 61 (3), 4601–4638. https://doi-org.proxy.uwasa.fi/10.1111/acfi.12742.
Ghaly, M., Dang, V.A., Stathopoulos, K., 2015. Cash holdings and employee welfare. J. Corp. Finance 33, 53–70.
Gong, G., Xu, S., Gong, X., 2018. On the value of corporate social responsibility disclosure: an empirical investigation of corporate bond issues in China. J. Bus. Ethics 150, 227–258. https://doi.org/10.1007/s10551-016-3193-8.
Green, C., Huang, R., Wen, Q., Zhou, D., 2019. Crowdsourced employer reviews and stock returns. J. Financ. Econ. 134 (1), 236–251. https://doi.org/10.1016/j.jfineco.2019.03.012.
Gu, S., Kelly, B., Xiu, D., 2020. Empirical asset pricing via machine learning. Rev. Financ. Stud. 33 (5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009.
Guiso, L., Sapienza, P., Zingales, L., 2015. The value of corporate culture. J. Financ. Econ. 117 (1), 60–76. https://doi.org/10.1016/j.jfineco.2014.05.010.
Hales, J., Moon, J., Swenson, L., 2018. A new era of voluntary disclosure? Empirical evidence on how employee postings on social media relate to future corporate disclosures. Account. Organisat. Soc. 68–69 (July), 88–108. https://doi.org/10.1016/j.aos.2018.04.004.
Hall, B., Helmers, C., Rogers, M., Sena, V., 2013. The importance (or not) of patents to UK firms. Oxf. Econ. Pap. 65 (3), 603–629. https://doi-org.proxy.uwasa.fi/10.1093/oep/gpt012.
Hall, B., Jaffe, A., Trajtenberg, M., 2001. The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. National Bureau of Economic Research. NBER working paper 8498.
Hall, B., Ziedonis, R., 2001. The patent paradox revisited: an empirical study of patenting in the U.S. semiconductor industry, 1979–1995. Rand J. Econ. 32, 101–128.
Hasan, M.M., Taylor, G., Richardson, G., 2021. Brand capital and stock price crash risk. Manag. Sci. 68 (10), 7221–7247. https://doi.org/10.1287/mnsc.2021.4197.
Hoberg, G., Maksimovic, V., 2015. Redefining financial constraints: a text-based analysis. Rev. Financ. Stud. 28 (5), 1312–1352. https://doi.org/10.1093/RFS/HHU089.
Hoffman, M., Bach, F., Blei, D., 2010. Online learning for latent dirichlet allocation. In: Lafferty, J., Williams, C., Shawe-Taylor, J., Zemel, R., Culotta, A. (Eds.), Advances in Neural Information Processing Systems, vol. 23. Curran Associates, Inc, pp. 856–864.
Holmstrom, B., 1989. Agency costs and innovation. J. Econ. Behav. Organ. 12 (3), 305–327. https://doi.org/10.1016/0167-2681(89)90025-5.
Huang, J., 2018. The customer knows best: the investment value of consumer opinions. J. Financ. Econ. 128 (1), 164–182. https://doi.org/10.1016/j.jfineco.2018.02.001.
Huang, K., Li, M., Markov, S., 2020. What do employees know? Evidence from a social media platform. Account. Rev. 95 (2), 199–226. https://doi.org/10.2308/accr-52519.
Huang, M., Li, P., Meschke, F., Guthrie, J.P., 2015. Family firms, employee satisfaction, and corporate performance. J. Corp. Finance 34, 108e127. https://doi.org/10.1016/j.jcorpfin.2015.08.002.
Jabeur, B., Khalfaoui, R., Ben Arfi, W., 2021. The effect of green energy, global environmental indexes, and stock markets in predicting oil price crashes: evidence from explainable machine learning. J. Environ. Manag. 298, 113511. https://doi.org/10.1016/j.jenvman.2021.113511.
Jacobs, A.W., Padavic, I., 2015. Hours, scheduling and flexibility for women in the US low-wage labour force. Gend. Work. Organ. 22 (1), 67–86.
Jame, R., Johnston, R., Markov, S., Wolfe, M.C., 2016. The value of crowdsourced earnings forecasts. J. Account. Res. 54 (4), 1077–1110. https://doi.org/10.1111/1475-679X.12121.
Jiang, Y., Jones, S., 2018. Corporate distress prediction in China: a machine learning approach. Account. Finance 58 (4), 1063–1109. https://doi-org.proxy.uwasa.fi/10.1111/acfi.12432.
Jiao, Y., 2010. Stakeholder welfare and firm value. J. Bank. Finance 34 (10), 2549–2561. https://doi.org/10.1016/j.jbankfin.2010.04.013.
Jones, D.A., 2010. Does serving the community also serve the company? Using organisational identification and social exchange theories to understand employee responses to a volunteerism programme. J. Occupat. Organisat. Psychol. 83 (4), 857–878. https://doi.org/10.1348/096317909X477495.
Jones, S., 2017. Corporate bankruptcy prediction: a high dimensional analysis. Rev. Account. Stud. 22, 1366–1422. https://doi.org/10.1007/s11142-017-9407-1.
Jones, S., Alam, N., 2019. A machine learning analysis of citation impact among selected Pacific basin journals. Account. Finance 59 (4), 2509–2552. https://doi-org.proxy.uwasa.fi/10.1111/acfi.12584.
Jordan, M., Ghahramani, Z., Jaakkola, T., Saul, L., 1999. An introduction to variational methods for graphical models. Mach. Learn. 37, 183–233. https://doi.org/10.1023/A:1007665907178.
Jung, M.J., Naughton, J.P., Tahoun, A., Wang, C., 2018. Do firms strategically disseminate? Evidence from corporate use of social media. Account. Rev. 93 (4), 225–252. https://doi.org/10.2308/accr-51906.
Karolyi, G.A., Van Nieuwerburgh, S., 2020. New methods for the cross-section of returns. Rev. Financ. Stud. 33 (5), 1879–1890. https://doi.org/10.1093/rfs/hhaa019.
Kelley, E., Tetlock, P., 2013. How wise are crowds? Insights from retail orders and stock returns. J. Finance 68 (3), 615–641. https://doi.org/10.1111/jofi.12028.
Kelliher, C., de Menezes, L.M., 2019. Flexible Working in Organisations: a Research Overview. Routledge/Taylor & Francis Group. https://doi.org/10.4324/9781351128346.
Kelliher, C., Richardson, J., Boiarintseva, G., 2019. All of work? All of life? Reconceptualising work-life balance for the 21st century. Hum. Resour. Manag. J. 29 (2), 97–112. https://doi.org/10.1111/1748-8583.12215.
Kim, C., Wang, K., Zhang, L., 2019. Readability of 10-K reports and stock price crash risk. Contemp. Account. Res. 36, 1184–1216. https://doi-org.proxy.uwasa.fi/10.1111/1911-3846.12452.
Kimbrough, M.D., Wang, X., Wei, S., Zhang, J., 2022. Does voluntary ESG reporting resolve disagreement among ESG rating agencies? Eur. Account. Rev. 1–33.

Kogan, L., Papanikolaou, D., Seru, A., Stoffman, N., 2017. Technological innovation, resource allocation, and growth. Q. J. Econ. 132 (2), 665–712.

Krammer, S.M.S., 2022. Human resource policies and firm innovation: the moderating effects of economic and institutional context. Technovation 110, 102366. https://doi.org/10.1016/j.technovation.2021.102366.

Landier, A., Nair, V., Wulf, J., 2009. Trade-offs in staying close: corporate decision making and geographic dispersion. Rev. Financ. Stud. 22 (3), 1119–1148. https://doi.org/10.1093/rfs/hhm042.

Lei, L.G., Li, Y., Luo, Y., 2019. Production and dissemination of corporate information in social media: a review. J. Account. Lit. 42, 29–43. https://doi.org/10.1016/j.acclit.2019.02.002.

Lehavy, R., Li, F., Merkley, K., 2011. The effect of annual report readability on analyst following and the properties of their earnings forecasts. Account. Rev. 86 (3), 1087–1115. https://doi.org/10.2308/accr.00000043.

Lewbel, A., 2012. Using heteroscedasticity to identify and estimate mis-measured and endogenous regressor models. J. Bus. Econ. Stat. 30 (1), 67–80. https://doi.org/10.1080/07350015.2012.643126.

Li, K., Mai, F., Shen, R., Yan, X., 2020. Measuring corporate culture using machine learning. Rev. Financ. Stud. 33 (8), 3591–3623. https://doi.org/10.1093/rfs/hhaa079.

Lin, B., Bai, R., 2021. Machine learning approaches for explaining determinants of the debt financing in heavy-polluting enterprises. Finance Res. Lett., 102094 https://doi.org/10.1016/j.frl.2021.102094.

Lu, Q., Chesbrough, H., 2022. Measuring open innovation practices through topic modelling: revisiting their impact on firm financial performance. Technovation 114, 102434. https://doi.org/10.1016/j.technovation.2021.102434.

Lu, Z., Zhuo, Z., 2021. Modelling of Chinese corporate bond default – a machine learning approach. Account. Finance. https://doi.org/10.1111/acfi.12846.

Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., et al., 2020. From local explanations to global understanding with explainable AI for trees. Nat. Mach. Intell. 2 (1), 56–67. https://doi.org/10.1038/s42256-019-0138-9.

Lundberg, S.M., Lee, S.I., 2017. A unified approach to interpreting model predictions. Adv. Neural Inf. Process. Syst. 4766–4775. Available at: https://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.

Mao, C., Weathers, J., 2019. Employee treatment and firm innovation. J. Bus. Finance Account. 46, 977–1002. https://doi.org/10.1111/jbfa.12393.

Manso, G., 2011. Motivating innovation. J. Finance 66 (5), 1823–1860.

Masuda, A.D., Poelmans, S.A., Allen, T.D., Spector, P.E., Lapierre, L.M., Cooper, C.L., et al., 2012. Flexible work arrangements availability and their relationship with work-to-family conflict, job satisfaction, and turnover intentions: a comparison of three country clusters. Appl. Psychol. 61 (1), 1–29.

Mavis, C.P., McNamee, N.P., Petmezas, D., Travlos, N.G., 2020. Selling to buy: asset sales and acquisitions. J. Corp. Finance 62, 101587. https://doi.org/10.1016/j.jcorpfin.2020.101587.

Mikolov, T., Chen, K., Corrado, G., Dean, J., 2013a. Efficient estimation of word representations in vector space. ArXiv 1301, 3781. https://arxiv.org/abs/1301.3781.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., Dean, J., 2013b. Distributed representations of words and phrases and their compositionality. ArXiv 1310, 4546. https://arxiv.org/abs/1310.4546.

Nicholas, T., 2008. Does innovation cause stock market runups? Evidence from the great crash. Am. Econ. Rev. 98, 1370–1396.

Nousiainen, E., Ranta, M., Ylinen, M., Järvenpää, M., 2024. Using machine learning and 10-K filings to measure innovation. Account. Finance 00, 1–29. https://doi.org/10.1111/acfi.13245.

Østergaard, C.R., Timmermans, B., Kristinsson, K., 2011. Does a different view create something new? The effect of employee diversity on innovation. Res. Pol. 40 (3), 500–509.

Oyer, P., Schaefer, S., 2005. Why do some firms give stock options to all employees? An empirical examination of alternative theories. J. Financ. Econ. 76, 99–133. https://doi.org/10.1016/j.jfineco.2004.03.004.

Ozcan, S., Suloglu, M., Sakar, C.O., Chatufale, S., 2021. Social media mining for ideation: identification of sustainable solutions and opinions. Technovation 107, 102322. https://doi.org/10.1016/j.technovation.2021.102322. ISSN 0166-4972.

Pérez, M.P., Sánchez, A.M., Carnicer, M.P., 2002. Benefits and barriers of telework: perception differences of human resources managers according to company's operations strategy. Technovation 22, 775–783.

Petruzzelli, A.M., Mora, L., Natalicchio, A., Platania, F., Toscano Hernandez, C., 2024. Consumers' reaction to Sci-Fi as a source of information for technological development: an empirical analysis. Technovation 132, 102970. https://doi.org/10.1016/j.technovation.2024.102970.

Ranta, M., Ylinen, M., 2023a. Employee benefits and company performance: evidence from a high-dimensional machine learning model. Manag. Account. Res., 100876 https://doi.org/10.1016/j.mar.2023.100876.

Ranta, M., Ylinen, M., 2023b. Board gender diversity and workplace diversity: a machine learning approach. Corp. Gov. 23, 995–1018. https://doi.org/10.1108/CG-01-2022-0048.

Ranta, M., Ylinen, M., Järvenpää, M., 2023. Machine learning in management accounting research: literature review and pathways for the future. Eur. Account. Rev. 32, 607–636. https://doi.org/10.1080/09638180.2022.2137221.

Riketta, M., 2005. Organisational identification: a meta-analysis. J. Vocat. Behav. 66 (2), 358–384. https://doi.org/10.1016/j.jvb.2004.05.005.

Röder, M., Both, A., Hinneburg, A., 2015. Exploring the space of topic coherence measures. Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, pp. 399–408. https://doi.org/10.1145/2684822.2685324.

Rubin, A., 2008. Political views and corporate decision making: the case of corporate social responsibility. Financ. Rev. 43 (3), 337–360. https://doi.org/10.1111/j.1540-6288.2008.00197.x.

Sarstedt, M., Danks, N.P., 2022. Prediction in HRM research–A gap between rhetoric and reality. Hum. Resour. Manag. J. 32 (2), 485–513. https://doi.org/10.1111/1748-8583.12400.

Saura, J.R., Palacios-Marqués, D., Ribeiro-Soriano, D., 2023. Exploring the boundaries of open innovation: evidence from social media mining. Technovation 119, 102447. https://doi.org/10.1016/j.technovation.2021.102447.

Schuh, S.C., Van Quaquebeke, N., Göritz, A.S., Xin, K.R., De Cremer, D., Van Dick, R., 2016. Mixed feelings, mixed blessing? How ambivalence in organisational identification relates to employees' regulatory focus and citizenship behaviors. Hum. Relat. 69 (12), 2224–2249. https://doi.org/10.1177/0018726716639117.

Shalley, C.E., Zhou, J., Oldham, G.R., 2004. The effects of personal and contextual characteristics on creativity: where should we go from here? J. Manag. 30 (6), 933–958, 2004.

Sheng, J., 2019. Asset pricing in the information age: employee expectations and stock returns. https://doi.org/10.2139/ssrn.3321275.

Shrestha, Y.R., He, V.F., Puranam, P., von Krogh, G., 2021. Algorithm supported induction for building theory: how can we use prediction models to theorize? Organ. Sci. 32 (3), 856–880. https://doi.org/10.1287/orsc.2020.1382.

Shipton, H., West, M.A., Dawson, J., Birdi, K., Patterson, M., 2006. HRM as a predictor of innovation. Hum. Resour. Manag. J. 16 (1), 3–27.

Shmueli, G., 2010. To explain or to predict? Stat. Sci. 25 (3), 289–310.

Soga, L., Bolade-Ogunfodun, O., Mariani, M., Laker, B., 2022. Unmasking the other face of flexible working practices: a systematic literature review. J. Bus. Res. 142, 648–662. https://doi.org/10.1016/j.jbusres.2022.01.024.

Stock, J.H., Yogo, M., 2005. Testing for weak instruments in linear IVregression. In: Stock, J.H., Andrews, D.W.K. (Eds.), Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg. Cambridge University Press, pp. 80–108. Chapter 5 in.

Storm, H., Baylis, K., Heckelei, T., 2020. Machine learning in agricultural and applied economics. Eur. Rev. Agric. Econ. 47 (3), 849–892. https://doi.org/10.1093/erae/jbz033.

Tang, V.W., 2017. Wisdom of crowds: cross-Sectional variation in the informativeness of third-party-generated product information on Twitter. J. Account. Res. 56 (3), 989–1034. https://doi.org/10.1111/1475-679X.12183.

Tian, X., Wang, T.Y., 2014. Tolerance for failure and corporate innovation. Rev. Financ. Stud. 27 (1), 211–255.

Tidd, J., Bessant, J., Pavitt, K., 2005. Managing Innovation: Integrating Technological, Market and Organizational Change. John Wiley & Sons, New York.

Vajda, S., Kuhn, H.W., Tucker, A.W., 1954. Contributions to the theory of games. Volume II. Math. Gaz. 38 (326), 319–320. https://doi.org/10.2307/3611199.

Valizade, D., Schulz, F., Nicoara, C., 2024. Towards a paradigm shift: how can machine learning extend the boundaries of quantitative management scholarship? Br. J. Manag. 35 (1), 99–114. https://doi.org/10.1111/1467-8551.12678.

Van Knippenberg, D., van Dick, R., Tavares, S., 2007. Social identity and social exchange: identification, organisational and supervisor support, and withdrawal from the job. Journal of *applied social psychology*, 37(3), 457–477. https://doi.org/10.1111/j.1559-1816.2007.00168.x.

Vyas, L., Butakhieo, N., 2020. The impact of working from home during COVID-19 on work and life domains: an exploratory study on Hong Kong. Policy Design Practice 4 (1), 59–76. https://doi.org/10.1080/25741292.2020.1863560.

Xu, Y., Fan, Y., Yeh, C.-H., Zhang, L., 2024. A machine learning-based approach to identifying key innovation determinants for industrial firms: the case of the pharmaceutical industry in China. IEEE Trans. Eng. Manag. 71, 6103–6115. https://doi.org/10.1109/TEM.2023.3281422.

Yang, L., Holtz, D., Jaffe, S., et al., 2022. The effects of RW on collaboration among information workers. Nat. Hum. Behav. 6, 43–54. https://doi.org/10.1038/s41562-021-01196-4.

Ylinen, M., Ranta, M., 2023. Employer ratings in social media and firm performance: evidence from an explainable machine learning approach. Account. Finance 64, 247–276. https://doi.org/10.1111/acfi.13146.

Zhu, F., Zhang, X., 2010. Impact of online consumer reviews on sales: the moderating role of product and consumer characteristics. J. Market. 74 (2), 133–148. https://doi.org/10.1509/jmkg.74.2.133.

Zingales, L., 2000. In search of new foundations. J. Finance 55 (4), 1623–1653.