Management Accounting Research 64 (2024) 100876

Contents lists available at ScienceDirect

Management Accounting Research

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

Employee benefits and company performance: Evidence from a high-dimensional machine learning model

Mikko Ranta [*], Mika Ylinen

School of Accounting and Finance, University of Vaasa, Vaasa, Finland

A R T I C L E I N F O

Keywords:
Organizational control
Employee benefits
Machine learning
Explainable AI
Social media

A B S T R A C T

By incorporating novel social media data, we analyze in detail how US companies offer different employee benefits and how they are associated with several company performance measures. Benefits such as 401(k), employee discounts, parking, and vision/dental healthcare are the most commonly provided, while free food -related benefits and family-related benefits are the most scarcely offered. Furthermore, with the aid of efficient machine learning -based models and tools from explainable artificial intelligence, we discover that family-related benefits are often associated with the most satisfied employees and best-performing companies. Our findings indicate that high-growth companies tend to provide a broad array of benefits to their employees. In contrast, highly profitable companies often concentrate on delivering a more limited and specialized set of benefits. We argue that companies offer rare and highly sought benefits to keep and recruit high-performers.

1. Introduction

Employees are essential company stakeholders in enabling strategy execution and, thus, play a vital role in the success of companies (Becker et al., 2001). Therefore, companies adopt many employee-friendly practices and offer numerous benefits to recruit and retain the best employees. In many ways, these benefits have become indispensable for companies to keep their know-how and enable long-term sustainability (Fauver et al., 2018) and strategy implementation (Garel & Petit-Romec, 2020). Thus, this role of employees as vital stakeholders has sparked substantial interest among scholars, making employee treatment and its effect on different aspects of companies a well-established field of study. This study contributes to the existing literature by providing additional specificity to the measurement of employee treatment. Previous research has primarily focused on a limited number of employee treatment types or an all-encompassing measure of overall employee treatment. In contrast, our investigation entails an analysis of 26 distinct employee benefits and explores the association of their widespread provision to rank-and-file personnel on organizational performance. We find that family-related benefits as well as stock and equity options, are associated with good company performance, while benefits like 401(k), employee discounts, and vision/dental healthcare are associated with bad company performance.

A substantial amount of literature has studied the effects of employee treatment. More specifically, previous literature has studied its effect on innovation (Chen et al., 2016; Mao & Weathers, 2019), capital structure (Bae et al., 2011), cost of loans (Francis et al., 2019), productivity (Darrough et al., 2019), dividend policy (Saeed, 2021; Benlemlih, 2019; Cheung et al. 2018; Yu, 2011) and financial performance (Gupta & Krishnamurti, 2020; Fauver et al., 2018; Flammer, 2015; Ding et al., 2009). Moreover, from the perspective of management control, emerging literature has examined the effectiveness of different reward types on employee performance (e.g., see Heninger et al., 2019; Kelly et al., 2017, Presslee et al., 2013), while more general business literature has studied the effect of employee treatment on other employee outcomes, like loyalty (Roehling et al., 2001), turnover intention (Guthrie, 2001), and employee satisfaction (Whitener, 2001).

The employee treatment research analyzing company performance and associated topics, like dividend policy, most closely relates to our research. Prior research has yielded mixed findings regarding the effectiveness of employee treatment on company performance. While some studies have revealed positive outcomes from good employee treatment, such as improved financial and operational performance for companies, others have reported negative effects (Ben-Nasr & Ghouma, 2018). The potential reasons underlying these conflicting results are multifarious. We suggest that these divergent results may also stem from differences between individual benefits. Some employee benefits may entail high costs that outweigh their benefits. Additionally, some

* Correspondence to: Mikko Ranta, University of Vaasa, Wolffintie 34, 65200 Vaasa, Finland.
E-mail address: mran@uwasa.fi (M. Ranta).

https://doi.org/10.1016/j.mar.2023.100876
Received 4 March 2022; Received in revised form 8 December 2023; Accepted 21 December 2023
Available online 27 December 2023

benefits may be highly esteemed by employees and are therefore more efficient at improving the financial performance of companies. These perspectives are corroborated by recent management control literature, which indicates notable differences in employee performance across various reward types (e.g., Heninger et al., 2019; Kelly et al., 2017).

Our research contributes to the literature by being one of the first papers that is able to investigate employee treatment in more detail by analyzing a large selection of individual benefits offered to employees and their association with the performance of companies. Previous studies use different proxies to estimate employee treatment, like the employee welfare index provided by the KLD STATS database (Francis et al., 2019) or the treatment level information in the Thomson Reuters ASSET4 database (Saeed, 2021), and have studied the association between employee treatment and company performance only at the aggregate level. Due to the applied machine learning (ML) approach, we can simultaneously analyze all the individual benefits offered to employees and analyze their association with company performance. Our results reveal interesting differences between benefits.

Furthermore, this paper makes methodological contributions to the literature. Our paper is one of the first papers in accounting research that uses tools from explainable AI (artificial intelligence) for model interpretation. The gradient boosting model applied in this research exhibits superior predictive power when compared to traditional OLS regression. It offers a data-driven approach for finding the optimal functional form for the model that minimizes out-of-sample prediction error. Furthermore, the gradient-boosting model can handle multicollinearity, which is crucial with high-dimensional models (Bertomeu, 2020; Jones, 2017). However, the lack of inference and interpretation has been the Achilles' heel of most ML models. We employ the SHAP method (Lundberg et al., 2020) to explain the results of the ML model. It is based on game theory and is considered one of the most consistent and accurate methods to explain features' local and global importance (Lundberg et al., 2020). We implement the approach suggested by Ranta et al. (2022) and use SHAP values for detailed association analysis. Moreover, we introduce an ML model designed to estimate causal effects from data by inferring the average treatment effects (ATE) of employee benefits on company performance using the double ML algorithm proposed by Chernozhukov et al. (2017a, 2017b).

Finally, our research contributes to the literature that exploits social media data in research. Recent research suggests that these alternative data sources offer a promising new source of information for management accounting research (Mahlendorf et al., 2023; Ranta et al., 2022; Bhimani, 2020). Our assessment of benefits policies employs an innovative dataset comprising a quarter of a million employee evaluations sourced from a job recruitment site. A number of research studies have established the value of these online platforms, which serve as venues for employees to voice their personal perspectives on different aspects of their employers. Such platforms have been recognized for offering insights that are pertinent to the forecast of future company outcomes (refer to studies by Huang et al., 2020; Green et al., 2019; Hales et al., 2018). These employee reviews provide a spyglass that researchers can use to acquire objective information about company practices previously only available to company insiders.

The remainder of the paper is organized as follows. Section 2 discusses the previous literature and motivates our research questions. Section 3 describes the sample and methodology. The section also reviews the previous business research literature exploiting social media data and implementing ML methods. Empirical findings are presented in Section 4. Section 5 discusses the findings and concludes.

2. Literature review

The social exchange theory, as proposed by Blau (1964) and Levinson (1965), provides a background explanation for the association between employee benefits and company performance. The central tenet of this theory, the norm of reciprocity (Gouldner, 1960), asserts that individuals are motivated to reciprocate positive actions that they receive (Eisenberger et al., 1997). As a result, employment can be viewed as a phenomenon where effort, dedication, and loyalty are exchanged for tangible and intangible rewards, like different benefits (van Knippenberg et al., 2007). The norm of reciprocity obliges employees to repay the positive treatment they receive from the organization (Eisenberger et al., 1997; Eisenberger et al., 2001; van Knippenberg et al., 2007). In line with the social exchange theory, the stakeholder theory (Donaldson & Preston, 1995) supports the view that good employee treatment improves employee loyalty and should appear as improved company performance. Well-treated employees support corporate plans and are less prone to exploit their organizations (Whitener, 2001). Good employment treatment creates a moral capital that provides firms protection against external shocks (Godfrey et al., 2009). Furthermore, good employee treatment can signal outside that the firm treats its stakeholders fairly (Faleye & Trahan, 2011).

There is a well-established empirical research base supporting these theories. Gupta & Krishnamurti (2020) found that employee-friendly procedures increase firm value. This connection is most substantial in countries with weak employment protection and with better infrastructure, productivity, and incentives. Chen et al. (2016) demonstrated that better employee treatment policies improve innovation, boost market valuation, and enable better future operating performance. Similarly, Mao & Weathers (2019) identified firm innovation as a path through which good employee practices flow to improved firm value. Darrough et al. (2019) found evidence that good employee treatment improves productivity by mitigating the adverse moral effects of strong unemployment insurance benefits. Benlemlih (2019) and Saeed (2021) found a positive relationship between employee treatment and dividend payments. Finally, Fauver et al. (2018) demonstrated that good employee treatment practices in companies are associated with improved market value and better operational performance.

However, several studies also document negative company outcomes from employee-friendly policies. Yu (2011) found that good employee treatment is negatively associated with the dividend payout decisions of companies. Similarly, Cheung et al. (2018) did not find any positive effects of good employee treatment on dividend payments. Most notably, Ben-Nasr and Ghouma (2018) found that high levels of employee well-being standards increase stock price crash risk. They argue that managers use good employee treatment to team up with employees to conceal the true status of a company and withhold bad news (Pagano & Volpin, 2005). Thus, good employee treatment can signal that the management is trying to hide the company's impaired future market or operational performance. Managers might treat employees kindly to avoid conflicts with employee unions, achieve improved social relations with employees (Cronqvist et al., 2009), and strive for a quiet life (Bertrand & Mullainathan, 2003). Ben-Nasr and Ghouma (2018) also found that good employee treatment is connected to the earnings management of companies.

In conclusion, previous research shows contradictory results on how employee treatment affects companies' performance. We argue that the discrepancies observed in the literature may stem from variations in offered employee benefits and their respective value as perceived by the employees. Furthermore, there are significant differences in costs associated with offering different employee benefits. To get more theoretical support, we also draw insight from the mental accounting theory (Thaler, 1985, 1999) to argue that a rare and highly sought benefit within an industry that is offered widely to employees boosts company performance. Furthermore, these benefits can also improve companies' recruitment process. The theory, applied to our setting, predicts that rare benefits suffer less from the diminishing marginal value, and employees are more motivated to earn these benefits and are more satisfied when having them (Choi and Presslee, 2023). This overall setting probably stems from the fact that rare benefits are difficult for companies to offer. However, as some companies still offer these benefits, they are probably highly sought after by employees, otherwise, all companies would cease offering these benefits.

An emerging narrative in management control literature supports the idea of strong variation between benefits at the company and employee levels. Henninger et al. (2019) examined organizational control in the background of wellness programs and found that tangible rewards are more efficient at improving employee performance than cash rewards. Kelly et al. (2017) made similar conclusions by analyzing the effects of tangible versus cash rewards in a repeated tournament setting. However, Presslee et al. (2013) found that cash rewards lead to better employee performance by conducting a quasi-experiment at five call centers of a financial services company. Thus, previous research documents some contradictory results and big variations between benefits on how they improve performance. We believe that an analysis of benefit-level differences might shed more light on this setting. Therefore, set against this backdrop, our goal is to find more knowledge on the following research questions:

RQ1: What are the most common benefits offered to employees?

RQ2: Which benefits are associated with company performance, and what is the nature of these associations?

3. Data and methodology

3.1. Employee benefits sample

We estimate employee treatment with a social media -based dataset that contains approximately 2,50,000 employee reviews from Kununu.com, a website similar to Glassdoor.com that offers employees a portal to rate their employer and review their work conditions. We use data from US companies, which is not publicly available anymore as Kununu has retired from the US market. Our choice to concentrate on US companies was due to the relatively homogenous business environment in the country, which effectively eliminates the potential distortions caused by transnational variations in benefits provision. Though Kununu also makes available analogous data for several Central European nations - a prospect that holds considerable promise for future research - one must approach this European dataset differently, as it is necessary to account for the potential influence of cross-country disparities. The dataset provided by Kununu is well-suited for our research context, given its ability to provide in-depth insight into employee experiences within their respective corporations. This depth of information enables a more reliable quantification of employee satisfaction. Moreover, it supplies an extensive account of the temporal evolution of employee benefits.

Previous research has shown that these platforms, where employees can share information about their employers, provide valuable information when estimating corporate conditions, like employee treatment (Huang et al., 2020). Thus, employee-provided social media information is increasingly used in academic research. Papers using Glassdoor.com or similar social-media data include Huang et al. (2020), Sheng et al. (2019), Green et al. (2018), Hales et al. (2018), and Huang et al. (2015), to name a few.

In 2019, Kununu reported hosting almost 4 million individual employer reviews of over 900,000 companies. The website provides information about open positions, employer reviews, and employee treatment in the form of 26 benefits offered to employees. The data allows us to analyze the connection between employee treatment and company performance in greater detail than previously. Furthermore, the large size of our dataset provides considerable cross-sectional and time-series variations allowing comprehensive analysis of the associations. From the data, we collect a sample representing S&P 1500 firms. To reduce noise, we include only those firms in the sample with at least 20 reviews per year during 2014–2019. The employee benefits data is combined with the financial data of the companies collected from the Refinitiv EIKON database. As the Kununu portal does not contain any common identifiers like CIK, the matching with the financial data is done using the company names in both databases. For matching, we use the Fuzzywuzzy library (https://pypi.org/project/fuzzywuzzy/) in Python and manual checking. The final sample consists of 493 companies with at least 20 reviews per year.[1]

Each review contains a grade between 1–5 for different company characteristics and a list of benefits that the employee has. Each review also includes information about the respondent's status within a company. This allows us to focus only on rank-and-file employees. The respondents were asked about 26 benefits with the question "What benefits and perks are offered to you?". For all the benefits, we calculate yearly averages by dividing the total number of employees having a benefit by the total number of reviews for that year. Thus, our measure is a proportion of reviews having a benefit, and we use it as a proxy for how widely the benefit is offered within a company. This estimate may be biased as the respondents do not necessarily represent all rank-and-file employees of a specific company. For example, blue-collar employees may be underrepresented in the data. Furthermore, how accurately respondents answer might be a source of another bias.

However, we consider this specific aggregation to be useful for our research setting as recent research suggests it might be beneficial to offer some benefits selectively and others more widely to employees (Opitz et al., 2022). To diminish the bias, we are including only companies with at least 20 reviews per year in the sample. We anticipate that there are substantial differences between industries in how they offer benefits. For example, flexible working arrangements are easier to offer in some industries. Furthermore, we assume that employees compare their benefits to the benefits offered by companies in the same industry. Therefore, the variables are transformed to be estimations for the deviation from the industry average by calculating means from the sample of companies belonging to the same 2-digit SIC group and transforming observations to be differences from these mean values.

Based on previous research (e.g., Huang et al., 2020; Hales et al., 2018; Huang et al., 2015), we decided to use seven different measures of company performance in our analysis: *Tobin's Q, ROA, Altman's Z* (Altman, 1968)*, Sales growth, Gross margin, Income before extraordinary items,* and *Employee grade*. The details of these measures can be found in the appendix. We consider a wide selection of measures that estimate company performance from different angles to be essential in our research setting. Our set of measures includes market performance measure (Tobin's Q), pure growth measure (Sales growth), performance in the eyes of a company's employees (*Employee grade*), and several measures that consider costs and benefits within a company (*ROA, Altman's Z, Gross margin, Income before extraordinary items*). We anticipate that measures considering the costs and benefits are important in determining if a benefit is helpful or not for a company (Busse et al., 2016). The control variables are total assets of a company, market-to-book ratio, leverage, implied volatility, R&D intensity, and firm age. Furthermore, we control employee satisfaction with compensation by adding a review question Overall compensation for work to the model. Using satisfaction to compensation, instead of absolute salary, has several advantages. Satisfaction with compensation provides a more nuanced and accurate representation of how employees perceive their pay. Absolute salary levels can vary widely based on factors such as industry, location, and job role. Furthermore, compensation satisfaction captures the subjective aspect of how employees feel about their pay. It takes into account not only the monetary value but also the perceived fairness, equity, and alignment with their contributions and market standards. The details of the control variables can be found in the

[1] To verify that our sample is an acceptable representation of S&P 1500 companies, we provide in the Supplementary material the kernel density estimation plots of Tobin's Q, ROA and Altman's Z for our sample and all S&P 1500 companies. The distributions are very similar, indicating that our dataset is an acceptable sample. The only minor difference is that companies in our dataset have slightly higher ROA when compared to all S&P 1500 companies of the same period.

appendix. To ensure a consistent and normally distributed set of feature values, transformations using logarithmic and cubic roots are applied when needed. We control year-fixed effects as well as industry effects using 2-digit SIC dummies. All variables are winsorized at the 1st and 99th percentiles to reduce the potential impact of outliers.

3.2. Descriptive statistics

Table 1 provides descriptive statistics for the employee benefits data, ordered according to how widely they are offered to employees (mean). To the best of our knowledge, descriptive information on employee benefits has not been previously documented by related research, at least at this level of detail. The final sample consists of 2507 firm-year observations. Benefits like *401(k), Employee discounts, Parking,* and *Vision and dental healthcare* are the most widely offered benefits followed by *Paid time off* and *Life insurance*. The most scarcely offered benefits include free food -related benefits and family-related benefits, like *On-site daycare facility* and *Reimbursed daycare*, as well as benefits like *Relocation allowance* and *Pet-friendly*. Although *Paid parental leave* is offered more widely, it is notable that 25% of the companies do not offer this benefit at all. Our arguments based on the mental accounting theory suggest that these family-related benefits would be more highly valued by employees, which can potentially lead to improved company performance through improved employee performance and recruitment of talented employees from the job market. The majority of our control variables are well-represented in the data, with the notable exceptions of R&D intensity and implied volatility, both of which have a significant number of missing values. Consequently, we have presented two versions of the reference OLS model in the appendix. The first version includes all control variables, while the second omits R&D intensity and implied volatility due to their substantial impact on the number of observations in the first version.

Table 2 provides correlations between the benefits variables. There are several quite high correlations between the variables (for example, 0.81 between *Vision and dental healthcare* and *Health and wellness programs*), indicating a multicollinearity problem. Therefore, a high-dimensional model including these variables should not be a traditional linear model but something that can handle multicollinearity. Gradient boosting has been shown to be robust against it due to the use of decision trees as weak estimators in the algorithm (Jones, 2017). Moreover, the hyperparameters of the models are finetuned to be such that the interpretations from the explainable AI model suffer from multicollinearity as little as possible.

3.3. Methodology

This paper proposes a novel ML approach that combines gradient boosting with state-of-the-art tools from explainable AI. We consider a nonlinear model to be crucial for our analysis, as the effectiveness of employee benefits is an interplay between costs and gains. The resulting difference between advantages and costs is linear only in special cases (Busse et al., 2016). For most of the benefits, the marginal value of a benefit decreases as it is offered more widely to employees (if it is even positive), and the marginal cost potentially increases. This naturally creates a nonlinear association with company performance. Thus, we think that a model that is able to capture all nonlinearities is a natural choice for our methodology. We use the SHAP method to interpret how different employee benefits affect company performance. SHAP uses game-theoretic Shapley values to rate the local importance of features. It estimates each feature's contribution to individual predictions and is considered to be one of the most consistent methods to explain the local importance of variables in an ML model (Lundberg et al., 2020; Lundberg & Lee, 2017). We implement the approach suggested by Ranta et al. (2022) to use SHAP values for (Gu et al., 2020) regression analysis.

Table 1
Descriptive statistics for the variables.

Variable	Count	Mean	St.dev.	Min	25%	50%	75%	Max
401(k)	2507	0.663	0.271	0.000	0.500	0.690	0.889	1.000
Employee discounts	2507	0.570	0.278	0.000	0.372	0.577	0.778	1.000
Parking	2507	0.506	0.268	0.000	0.330	0.500	0.678	1.000
Vision and dental healthcare	2507	0.460	0.368	0.000	0.000	0.500	0.785	1.000
Paid time off	2507	0.441	0.363	0.000	0.000	0.488	0.750	1.000
Life insurance	2507	0.397	0.345	0.000	0.000	0.400	0.695	1.000
401(k) Company match	2507	0.366	0.340	0.000	0.000	0.333	0.660	1.000
On-site cafeteria	2507	0.334	0.294	0.000	0.077	0.273	0.500	1.000
Flexible working hours	2507	0.310	0.248	0.000	0.125	0.267	0.476	1.000
Employee events	2507	0.266	0.275	0.000	0.000	0.207	0.447	1.000
Health and wellness programs	2507	0.235	0.242	0.000	0.000	0.188	0.395	1.000
Stock or equity options	2507	0.231	0.267	0.000	0.000	0.135	0.399	1.000
Tuition assistance	2507	0.230	0.261	0.000	0.000	0.150	0.400	1.000
Flexible spending account	2507	0.229	0.264	0.000	0.000	0.134	0.400	1.000
Paid parental leave	2507	0.161	0.215	0.000	0.000	0.083	0.250	1.000
Vehicle allowance	2507	0.153	0.117	0.000	0.029	0.160	0.246	0.614
Desirable office location	2507	0.151	0.199	0.000	0.000	0.083	0.229	1.000
Ability to work remotely	2507	0.149	0.213	0.000	0.000	0.050	0.229	1.000
On-site fitness center	2507	0.139	0.224	0.000	0.000	0.006	0.199	1.000
Easy access to public transportation	2507	0.131	0.177	0.000	0.000	0.063	0.213	1.000
Free meals	2507	0.108	0.170	0.000	0.000	0.034	0.148	1.000
Free snacks and drinks	2507	0.098	0.132	0.000	0.000	0.056	0.140	1.000
Relocation allowance	2507	0.039	0.106	0.000	0.000	0.000	0.014	1.000
On-site daycare facility	2507	0.036	0.110	0.000	0.000	0.000	0.008	1.000
Reimbursed daycare	2507	0.023	0.070	0.000	0.000	0.000	0.013	1.000
Pet-friendly	2507	0.016	0.063	0.000	0.000	0.000	0.000	1.000
Total assets (1000$)	2672	55,400	1,35,000	425	4020	12,100	39,200	9,07,000
Leverage	2661	1.197	4.755	-21.32	0.346	0.751	1.489	29.273
R&D intensity	1503	2.266	3.676	0	0	0.692	2.854	19.021
Return on assets	2672	5.482	6.455	-18.975	1.843	5.084	8.882	23.09
Implied volatility	1588	29.234	9.401	14.4	22.7	27.5	34.1	58.765
Age (months)	2672	404	226	12	228	372	576	828
Overall compensation for your work	2505	3.413	0.786	1.000	2.894	3.473	4.000	5.000

Note: The observations for the benefits are proportions for individual firm-years about how many of the reviews had a specific benefit

Furthermore, we add robustness to our analysis by using the double ML model introduced by Chernozhukov et al. (2017a, 2017b) to estimate the average treatment effects of benefits to company performance.

3.3.1. Gradient boosting

The gradient boosting algorithm (Friedman, 2001) applied in this research belongs to a group of ML methods called “ensemble methods”, other examples being random forests and bagging. Overall, ensemble methods have been applied quite extensively in general business research. For example, they have been applied to predict accounting fraud (Bao et al., 2020), bankruptcy (Jones, 2017), stock index returns (Kumar & Thenmozhi, 2014), financial fraud (Liu et al., 2015), and stock market prices (Khaidem et al., 2016). Furthermore, they have been used to detect misstatements in annual reports (Bertomeu et al., 2020) and improve managerial estimates (Ding et al., 2020). Barboza et al. (2017) compared several ML algorithms for predicting financial distress and found ensemble methods to be the most capable option. Similarly, Gu et al. (2020) discovered that for asset pricing, decision tree ensembles and neural networks were superior due to their efficiency in modeling nonlinear data interactions.

This study employs the gradient boosting algorithm to analyze the associations of various employee benefits with company performance. We use this approach to examine the importance of different predictors and the nonlinear associations between predictors and output variables. In our model, the employee benefit features are the predictors, and different measures of company performance are the predicted outcomes. The selected approach offers many advantages. It allows data-driven model selection because we do not restrict the functional form between the predicted variables and covariates. Thus, it is more efficient to allow any functional form and control for overfitting than to restrict the type of model and assume that it does not overfit (Beck, King, & Zeng, 2004). The associations can be nonlinear, and they can be learned from the data as the model searches for an optimal functional form using data without subjective (linearity) restrictions from the researcher (Bertomeu, 2020).

The high dimensional gradient boosting model can also automatically take into account important interaction effects. Decision trees in the model are constructed by sequentially picking the variables that improve the prediction the most. The best predictors are first used in the splits, and typically variables that have the strongest interactions with these are used in the following splits. Moreover, the approach also includes regularization, as the number of branches controls how many variables are used in the regression model (Friedman, 2001; Bertomeu, 2020). Additionally, because our model is based on decision trees, we can deal with the possible issue of multicollinearity between features that prevents using traditional linear methods. This follows from the building algorithm of decision trees. As branches are chosen based on their ability to decrease the loss function, highly correlating variables are rarely added to the same tree Storm et al. (2020). Thus, ensemble methods based on decision trees are considered relatively immune to multicollinearity (Jones, 2017), making the chosen ML approach crucial in our research setting.

Finally, our approach avoids losing observations with missing values, which is essential because we have a large dataset with many predictors and controls, where almost half of the observations have missing values. Tree-based gradient-boosting algorithms use the information from variables that have value for a specific observation, even if some other variables do not. This enables the dataset to be larger because the algorithm avoids dropping observations that have missing values.

In the current study, we have adopted the advanced version of the gradient boosting method known as the extreme gradient boosting (XGBoost) algorithm (Chen & Guestrin, 2016). In practice, it has been a highly successful prediction tool for structured data, winning numerous ML competitions. The algorithm uses highly optimized code, is scalable to large datasets, and is more accurate compared to the basic gradient-boosting approach. Moreover, it incorporates key elements from other aggregate methods, such as the random forest algorithm’s strategy for feature subsampling.

We briefly describe the theoretical foundations of XGBoost. The model’s prediction is improved iteratively so that a new weak predictor is trained with reweighted data, where the importance of those observations where the current phase is performing the worst is increased. Our approach can be defined with the model

$$\hat{y} = \sum_{j=1}^{J} f_j(\boldsymbol{x}_i), f_j \in \mathscr{F},$$

where $\mathscr{F}$ is the space of all possible decision trees. In our setting, the trees contain a continuous score on the leaves. The model is trained by minimizing the function

$$\mathscr{L} = \sum_i l(\hat{y}_i, y_i) + \sum_j \Omega(f_j),$$

where l is the model loss. The second term Ω penalizes the complexity of the model. The specific details of the algorithm are omitted and can be found in Chen and Guestrin (2016) and Friedman (2001).

We search for the optimal model structure by fine-tuning the following hyperparameters: the number of trees, the shrinkage parameter, the column subsampling parameter, the row subsampling parameter, the depth of the trees, and the gamma parameter. The shrinkage method decreases the importance of existing trees in the model to allow added trees to improve the model (Friedman, 2002). The column subsampling selects a subsample of features for the newly added tree, while the row subsampling does the same for observations. The gamma parameter sets a minimum loss reduction limit for new branches. The search for optimal parameters is implemented with a grid-search and cross-validation approach in two phases. The grid-search approach is efficient but computationally intensive, as we are trying every possible combination of the parameters in the grid. In the first phase, we search for optimal the optimal tree structure and regularization parameters. In the second phase, the number of decision trees is optimized. Table 3 provides the final hyperparameters for the models.[2]

3.3.2. SHAP values

Interpreting nonlinear ML models is challenging due to their complex nonlinear relationships. While there exist interpretation metrics that help analyze ML ensemble models’ features, like weight, gain, and cover metrics, they are problematic and may yield inconsistent results (Lundberg & Lee, 2017). TreeSHAP, a SHAP method variant used here, estimates feature importance for every observation and provides flexible tools for understanding the relationship between predictors and output variables (Lundberg et al., 2020). This flexibility is crucial for interpreting individual employee benefits’ associations on corporate performance. The benefits offered to employees are entangled in a complicated way for different companies. Thus, we consider SHAP to be essential for the correct interpretation of the results. More details of the methods and their applications in business research can be found in Valizade et al. (2022) and Ranta et al. (2022), among others. Papers in business research incorporating SHAP values for interpretation include Ranta & Ylinen (2023), Ylinen & Ranta (2023), and Bussmann et al. (2021).

3.3.3. Double ML

Our methods so far are extremely powerful at finding nonlinear associations between the examined variables. However, the issue of endogeneity is all the time present when analyzing the findings. Is it the benefits affecting company performance or the opposite? Or is there

[2] The Python code used in this research can be found in https://github.com/ML-for-Accounting/Explainable-AI-and-management-accounting

Table 2
Correlation matrix of the benefits variables.

	1	2	3	4	5	6	7	8	9	10	11	12
1 - 401(k)	1.0***	0.2***	0.22***	0.15***	0.05*	0.19***	0.17***	0.09***	0.12***	0.16***	0.32***	0.08***
2 - 401(k) Company match	0.2***	1.0***	0.74***	0.84***	0.28***	0.85***	0.6***	0.54***	0.00	0.83***	0.22***	0.28***
3 - Flexible spending account	0.22***	0.74***	1.0***	0.71***	0.32***	0.75***	0.56***	0.5***	0.05*	0.7***	0.22***	0.32***
4 - Vision and dental healthcare	0.15***	0.84***	0.71***	1.0***	0.26***	0.89***	0.61***	0.56***	-0.05*	0.91***	0.2***	0.26***
5 - Relocation allowance	0.05*	0.28***	0.32***	0.26***	1.0***	0.27***	0.3***	0.18***	0.1***	0.26***	0.1***	0.13***
6 - Life insurance	0.19***	0.85***	0.75***	0.89***	0.27***	1.0***	0.62***	0.55***	-0.02	0.88***	0.21***	0.28***
7 - Paid parental leave	0.17***	0.6***	0.56***	0.61***	0.3***	0.62***	1.0***	0.55***	0.13***	0.63***	0.22***	0.38***
8 - Desirable office location	0.09***	0.54***	0.5***	0.56***	0.18***	0.55***	0.55***	1.0***	0.06**	0.59***	0.17***	0.33***
9 - On-site daycare facility	0.12***	0.00	0.05*	-0.05*	0.1***	-0.02	0.13***	0.06**	1.0***	-0.02	0.12***	0.16***
10 - Paid time off	0.16***	0.83***	0.7***	0.91***	0.26***	0.88***	0.63***	0.59***	-0.02	1.0***	0.2***	0.29***
11 - Parking	0.32***	0.22***	0.22***	0.2***	0.1***	0.21***	0.22***	0.17***	0.12***	0.2***	1.0***	0.15***
12 - Reimbursed daycare	0.08***	0.28***	0.32***	0.26***	0.13***	0.28***	0.38***	0.33***	0.16***	0.29***	0.15***	1.0***
13 - Stock or equity options	0.17***	0.67***	0.58***	0.69***	0.28***	0.66***	0.54***	0.46***	0.00	0.67***	0.21***	0.27***
14 - Ability to work remotely	0.1***	0.53***	0.52***	0.53***	0.27***	0.56***	0.52***	0.55***	0.1***	0.55***	0.16***	0.28***
15- Easy access to public transportation	0.05*	0.42***	0.4***	0.47***	0.17***	0.46***	0.46***	0.5***	0.06**	0.49***	0.13***	0.2***
16 -Tuition assistance	0.17***	0.73***	0.66***	0.71***	0.31***	0.7***	0.63***	0.49***	0.07***	0.7***	0.25***	0.32***
17 - Flexible working hours	0.08***	0.03	0.05*	0.03	0.14***	0.01	0.15***	0.16***	0.14***	0.02	0.15***	0.08***
18 - On-site cafeteria	0.32***	0.18***	0.22***	0.12***	0.11***	0.17***	0.29***	0.18***	0.3***	0.14***	0.46***	0.17***
19 - Vehicle allowance	0.18***	0.79***	0.7***	0.92***	0.3***	0.83***	0.64***	0.58***	0.27***	0.85***	0.25***	0.44***
20 - Employee discounts	0.26***	-0.11***	-0.07***	-0.13***	0.00	-0.12***	0.04*	-0.08***	0.07***	-0.13***	0.17***	0.02
21 - Pet-friendly	-0.04*	0.07***	0.06**	0.1***	0.04	0.09***	0.12***	0.19***	0.11***	0.11***	0.05*	0.19***
22 - Employee events	0.11***	0.69***	0.56***	0.71***	0.27***	0.69***	0.61***	0.59***	0.04*	0.71***	0.27***	0.28***
23 - Health and wellness programs	0.18***	0.76***	0.69***	0.76***	0.33***	0.77***	0.67***	0.55***	0.12***	0.76***	0.3***	0.34***
24 - On-site fitness center	0.14***	0.52***	0.49***	0.49***	0.27***	0.51***	0.53***	0.42***	0.2***	0.5***	0.28***	0.29***
25 - Free snacks and drinks	0.00	0.26***	0.22***	0.31***	0.12***	0.3***	0.32***	0.35***	0.1***	0.31***	0.16***	0.18***
26 - Free meals	0.00	0.35***	0.29***	0.41***	0.14***	0.39***	0.4***	0.42***	0.06**	0.4***	0.14***	0.2***

Note: The asterisks indicate statistical significance in 10% (*), 5% (**), and 1% (***) levels.

some unknown variable in the background affecting both the output and the input variables?

To improve the robustness of our analysis, we implement a method from the causal ML literature (Hartford et al., 2017; Syrgkanis et al., 2019; Wager and Athey, 2018), namely the double ML method (Chernozhukov et al., 2017a, 2017b). The benefit of using an ML-based causal model lies in the power of ML models as predictors and in their flexibility to estimate many functional forms, reducing the risk of model misspecification. Furthermore, ML-based causal models are shown to be especially powerful when dealing with high-dimensional data, like we have here (Chernozhukov et al., 2017a). In fact, one of the goals of the double-ML algorithm is to provide a powerful estimate of the *average treatment effect* (ATE) for high-dimensional models (Chernozhukov et al., 2017b). The name of the method comes from two predictive steps in the process. An ML model is used to predict the outcome using the controls, and a second ML model is used to predict the treatment using the controls. In the next step, these two predictions are combined to estimate a heterogeneous treatment effect that can be used to calculate ATE. We omit the mathematical details of the double-ML algorithm and refer the reader to the two aforementioned papers by the inventors of the method (Chernozhukov et al., 2017a, 2017b).

3.3.4. Limitations of the ML approach

Although our ML approach, and ML methods in general, provide many advantages compared to traditional approaches, there are some limitations. There is no mature mathematical theory related to the statistical properties of most ML methods. In that respect, they are still somewhat black boxes. Thus, for statistical inference, we have to rely on numerical approaches, like bootstrapping and Monte Carlo -methods. These approaches lack reliability and very often need extensive computational capacities. For example, adding statistical inference for our SHAP methods, we need to estimate the gradient boosting model hundreds of times for different bootstrap samples, calculate the SHAP values for every model, and then estimate statistical significance from these samples of SHAP values. However, for some ML methods, we may have mathematical theory in the future, which allows statistical inferences without relying on numerical approaches. And even if numerical methods are the only option, the computing power of computers is increasing all the time, which diminishes this shortcoming in the future, at least for moderate-size datasets.

The double ML model also has some limitations in our research setting. As the model does not use external instrumental variables, we have to be careful with the interpretations. The authors of the method describe in their paper how the model can estimate ATE efficiently only when no relevant variables are missing from the model (Chernozhukov et al., 2017a). It is obvious that there are possibly many relevant variables missing from our model, and it is practically impossible to include all of them. To obtain more definitive answers regarding causality, we must adopt an alternative research methodology, such as conducting field experiments. However, we still think that by adding this analysis, we get more verification to the findings of the baseline models and more support for the arguments we draw from the background theories. Moreover, we want to include this model to demonstrate to management accounting scholars the opportunities causal ML is offering for academic research now and in the future.

4. Results

4.1. Model validation

We initiate our analysis by demonstrating that employee benefits possess meaningful predictive power, and as such, their association with company performance is valid to study. Furthermore, we demonstrate that, due to potential nonlinearities, XGBoost has higher predictive accuracy than LASSO or OLS regression and is the most appropriate method for our study. We compare the performance by calculating the coefficient of determination (R^2) for models that consist only of the benefits variables. The data is split into training and test sets using an 80/20% split. The training is implemented using 5-fold cross-validation (for the LASSO and the boosting model). *K*-fold cross-validation splits the dataset into *k* equally sized subsets and then iteratively combines $k - 1$ subsets into training samples to predict the remaining subset (Arlot & Celisse, 2010). Common choices for *k* are 3,5,10, and even the number of observations (Burman, 1989).

13	14	15	16	17	18	19	20	21	22	23	24	25	26
0.17***	0.1***	0.05*	0.17***	0.08***	0.32***	0.18***	0.26***	-0.04*	0.11***	0.18***	0.14***	0.00	0.00
0.67***	0.53***	0.42***	0.73***	0.03	0.18***	0.79***	-0.11***	0.07***	0.69***	0.76***	0.52***	0.26***	0.35***
0.58***	0.52***	0.4***	0.66***	0.05*	0.22***	0.7***	-0.07***	0.06**	0.56***	0.69***	0.49***	0.22***	0.29***
0.69***	0.53***	0.47***	0.71***	0.03	0.12***	0.92***	-0.13***	0.1***	0.71***	0.76***	0.49***	0.31***	0.41***
0.28***	0.27***	0.17***	0.31***	0.14***	0.11***	0.3***	0.00	0.04	0.27***	0.33***	0.27***	0.12***	0.14***
0.66***	0.56***	0.46***	0.7***	0.01	0.17***	0.83***	-0.12***	0.09***	0.69***	0.77***	0.51***	0.3***	0.39***
0.54***	0.52***	0.46***	0.63***	0.15***	0.29***	0.64***	0.04*	0.12***	0.61***	0.67***	0.53***	0.32***	0.4***
0.46***	0.55***	0.5***	0.49***	0.16***	0.18***	0.58***	-0.08***	0.19***	0.59***	0.55***	0.42***	0.35***	0.42***
0.00	0.1***	0.06**	0.07***	0.14***	0.3***	0.27***	0.07***	0.11***	0.04*	0.12***	0.2***	0.1***	0.06**
0.67***	0.55***	0.49***	0.7***	0.02	0.14***	0.85***	-0.13***	0.11***	0.71***	0.76***	0.5***	0.31***	0.4***
0.21***	0.16***	0.13***	0.25***	0.15***	0.46***	0.25***	0.17***	0.05*	0.27***	0.3***	0.28***	0.16***	0.14***
0.27***	0.28***	0.2***	0.32***	0.08***	0.17***	0.44***	0.02	0.19***	0.28***	0.34***	0.29***	0.18***	0.2***
1.0***	0.45***	0.38***	0.61***	0.08***	0.16***	0.67***	-0.03	0.1***	0.61***	0.65***	0.48***	0.27***	0.35***
0.45***	1.0***	0.37***	0.51***	0.24***	0.23***	0.57***	-0.08***	0.11***	0.53***	0.59***	0.47***	0.26***	0.33***
0.38***	0.37***	1.0***	0.4***	0.17***	0.13***	0.47***	0.02	0.18***	0.49***	0.48***	0.36***	0.32***	0.39***
0.61***	0.51***	0.4***	1.0***	0.1***	0.25***	0.7***	-0.04	0.1***	0.63***	0.74***	0.57***	0.26***	0.32***
0.08***	0.24***	0.17***	0.1***	1.0***	0.19***	0.09***	0.2***	0.09***	0.18***	0.12***	0.13***	0.15***	0.15***
0.16***	0.23***	0.13***	0.25***	0.19***	1.0***	0.23***	0.14***	0.02	0.24***	0.36***	0.42***	0.11***	0.1***
0.67***	0.57***	0.47***	0.7***	0.09***	0.23***	1.0***	-0.09***	0.17***	0.7***	0.77***	0.55***	0.37***	0.43***
-0.03	-0.08***	0.02	-0.04	0.2***	0.14***	-0.09***	1.0***	0.00	0.01	-0.04*	-0.04*	0.01	0.00
0.1***	0.11***	0.18***	0.1***	0.09***	0.02	0.17***	0.00	1.0***	0.18***	0.1***	0.11***	0.17***	0.12***
0.61***	0.53***	0.49***	0.63***	0.18***	0.24***	0.7***	0.01	0.18***	1.0***	0.74***	0.54***	0.43***	0.5***
0.65***	0.59***	0.48***	0.74***	0.12***	0.36***	0.77***	-0.04*	0.1***	0.74***	1.0***	0.84***	0.35***	0.43***
0.48***	0.47***	0.36***	0.57***	0.13***	0.42***	0.55***	-0.04*	0.11***	0.54***	0.84***	1.0***	0.28***	0.34***
0.27***	0.26***	0.32***	0.26***	0.15***	0.11***	0.37***	0.01	0.17***	0.43***	0.35***	0.28***	1.0***	0.82***
0.35***	0.33***	0.39***	0.32***	0.15***	0.1***	0.43***	0.00	0.12***	0.5***	0.43***	0.34***	0.82***	1.0***

Table 4 provides the results. They verify that the benefits variables have predictive power to explain company performance, and the boosting model achieves significantly better performance for *Tobin's Q, ROA, Altman's Z,* and *Gross margin.* Even though the gradient boosting model is performing better also for *Sales growth* and *Income before extraordinary items*, it is interesting that the coefficient of determination is significantly smaller for these two performance measures. In summary, XGBoost outperforms its counterparts by a discernible margin, owing to its ability to capture nonlinear relationships among variables that are completely overlooked by linear models. The explanative powers of LASSO and OLS regression models are in line with previous studies (e.g., Fauver et al., 2018; Hales et al., 2018; Huang et al., 2020). However, probably due to nonlinearities, the coefficient of determination for the boosting model is significantly higher than what has been achieved in previous studies using similar predictor variables.

4.2. Benefits associated with best-performing companies

We proceed with an analysis of the distinct features of the benefits offered, comparing how these diverge between top-tier companies and their counterparts. To do this, we initially categorize companies into two separate groups, the top 25% best-performing companies and the remaining majority, based on six financial performance indicators. The next step involves calculating the industry-adjusted prevalence of specific benefits within each of these two groups and contrasting them to highlight differences. The resulting comparisons are illustrated in Fig. 1.

While some disparities exist among the performance measures, there is a clear pattern with regard to certain benefits. They are either predominantly more widely offered by the top-performing companies or the opposite. For instance, when we examine all six performance measures, it is evident that *Stock and equity options* is invariably more widely offered by the top performers. Other benefits more frequently provided by these high-achieving companies include *Paid parental leave, Pet-friendly*, and complimentary food and snack options. Conversely, certain benefits tend to be scarcely offered by the top performers. These include *Flexible spending account, Vision and dental healthcare, Life insurance, and Health and wellness programs*. The pattern suggests a systematic tendency among top performers to offer some benefits more widely and others less so.

It's intriguing to observe that top-performing companies tend to offer a broader range of benefits more widely to employees when evaluated based on *Tobin's Q,* indicative of market performance, and *Sales growth*, a measure of pure growth excluding cost factors. This suggests that high-growth companies prioritize employee incentives. However, the tendency shifts noticeably when cost-incorporating performance measures, *ROA, Altman's Z, Income before extraordinary items,* and *Gross margin*, are taken into consideration. According to these measures, more profitable companies appear to offer a considerably smaller proportion of benefits more widely to employees. For instance, benefits such as *Vision and dental healthcare, Life insurance,* and *Health and wellness programs* initially seem to be more prevalent among the best-performing companies when assessed solely on *Sales growth*. However, a broader analysis

Table 3
Hyperparameters for the models.

Hyperparameter	Tobin's Q	ROA	Altman's Z	Sales growth	Gross margin	Income ex. ext.	Employee grade
Depth of trees	6	6	6	4	6	8	4
Shrinkage parameter	0.05	0.05	0.05	0.05	0.05	0.05	0.05
Row subsampling	0.8	0.8	0.8	0.8	0.8	0.8	0.8
Column subsampling	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Gamma	0.1	0.1	0.1	0.05	0.1	0.1	0.1

Note: The shrinkage method reduces weights by a constant factor to allow future trees to improve the model. The column subsampling selects a subsample of features for the newly added tree, while the row subsampling selects a subsample of observations for the tree. The gamma parameter sets a minimum loss reduction required to divide a leaf node of the tree further.

Table 4
Out-of-sample coefficient of determinations R^2 for the models.

Model	Tobin's Q	ROA	Altman's Z	Sales growth	Gross margin	Income ex. ext.	Employee grade
Boosting model	0.165	0.083	0.156	0.019	0.308	0.030	0.230
Lasso regression	0.027	0.020	0.009	0.010	0.009	0.007	0.188
OLS regression	0.026	0.016	0.003	0.006	0.002	0.001	0.188

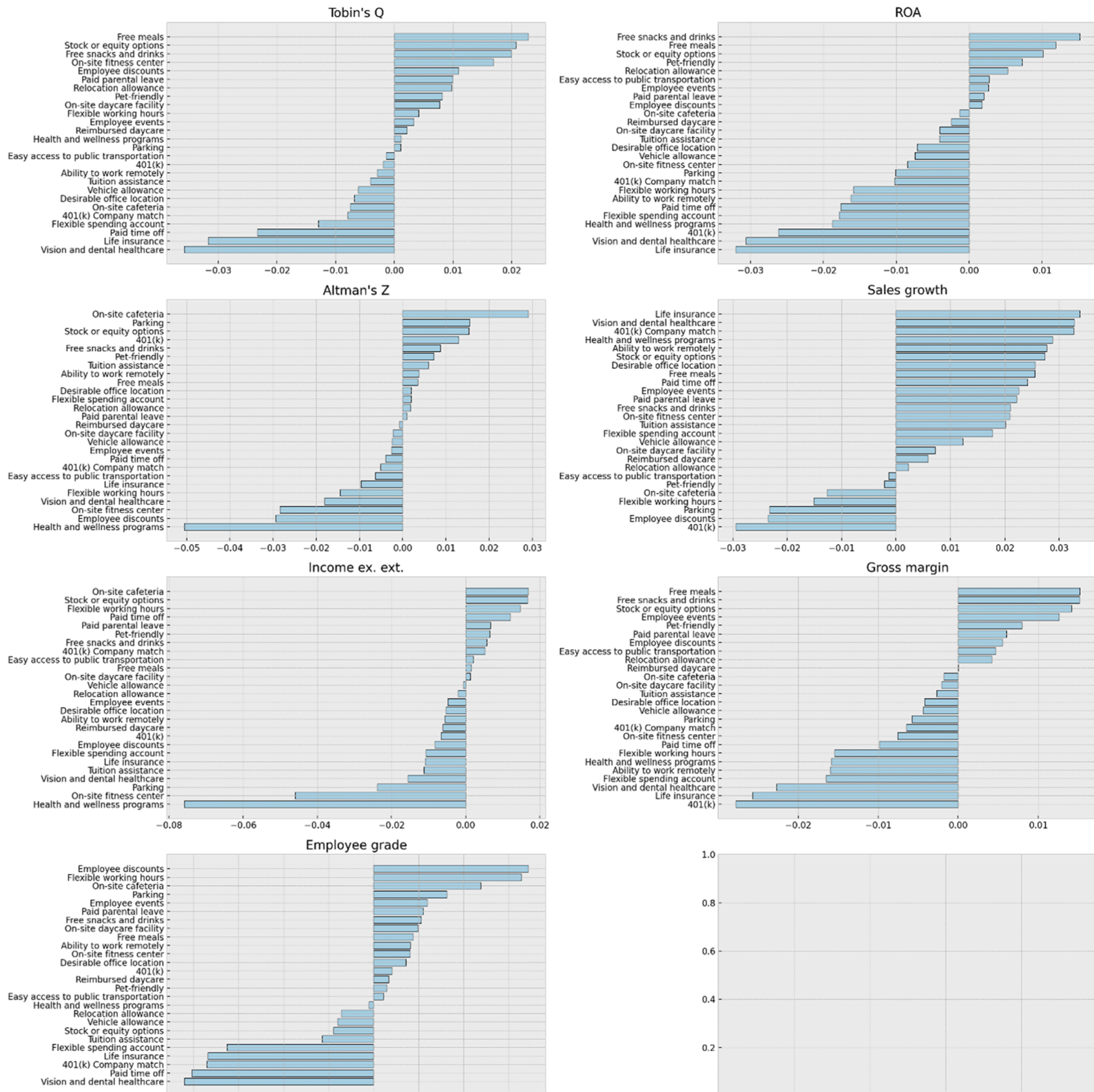

Fig. 1. The differences (industry-adjusted yearly means) in how widely benefits are offered between high-performing companies (top 25%) and the rest. The x-axis describes the difference between the high-performing companies and the rest in offering a specific benefit. Income ex. ext. is an abbreviation for Income excluding extraordinary items.

Table 5
Differences between the associations of the benefits and the performance measures for above- and below-sample industry average companies. Most consistent associations are marked with **bold** font.

	Tobin's Q	ROA	Altman's Z	Sales growth	Gross margin	Income ex. ext.	Employee grade
401(k)	0.0129	-0.0394	0.0286	-0.2892	0.0737	-0.0114	-0.0457***
401(k) Company match	0.0148	0.1818**	-0.0039	1.0568***	0.0371	0.1287**	-0.0214**
Flexible spending account	-0.0357***	-0.1932**	-0.0422**	-0.1673	-0.0044	-0.2121***	-0.1221***
Vision and dental healthcare	-0.0292***	-0.0901	-0.0176	-0.1171	-0.6756***	-0.0913*	-0.0818***
Relocation allowance	0.0007	0.0566	0.0202	-0.1354	0.3024*	-0.0164	-0.1877***
Life insurance	-0.0085	-0.1136	-0.0152	-0.4036*	-0.4361***	-0.0618	0.0038
Paid parental leave	0.0182*	0.0275	0.0118	0.4811**	0.5117***	0.0356	0.0715***
Desirable office location	0.0106	0.0221	0.0079	0.0246	0.2481	0.0259	0.0543***
On-site daycare facility	0.0475***	0.1401*	0.1074***	-0.1619	1.6031***	0.1368	0.0439***
Paid time off	0.0097	0.0066	0.0314*	-0.2929*	0.2465*	0.0104	-0.0606***
Parking	-0.0125	-0.0441	0.0226	-0.9669***	-0.7093****	-0.0136	0.0226*
Reimbursed daycare	0.0244*	0.1209*	0.0431*	0.2677	0.6826***	0.0794	0.0316**
Stock or equity options	0.0354**	0.1201	0.061**	0.4438**	0.1096	0.0889	0.0014
Ability to work remotely	0.0081	-0.2916***	0.0041	0.7174***	-0.1824	-0.206***	0.043***
Easy access to public transportation	0.0269*	0.0869	0.0234	-0.6876**	-0.2446	0.0391	-0.0173*
Tuition assistance	0.0061	0.1072	-0.0047	-0.1481	-0.7494***	0.0214	0.0106
Flexible working hours	-0.0058	-0.2473***	-0.0499**	-0.3981*	0.0204	-0.1693***	0.1136***
On-site cafeteria	-0.0331**	-0.1582*	-0.0179	0.0813	-0.0337	-0.0931	0.0415***
Vehicle allowance	-0.0117	-0.1122*	-0.0218	-0.0032	-0.3506***	-0.0861**	-0.0596***
Employee discounts	0.0059	0.0956	-0.0232	-0.2254	0.8214***	0.1062*	0.0454***
Pet-friendly	0.0535**	0.2236*	0.0372	-0.1442	0.2323	0.1211	0.0314**
Employee events	0.0097	0.1816**	0.0492**	-0.1928	0.3864*	0.1659***	0.1126***
Health and wellness programs	-0.0116	0.0031	-0.0513**	0.0264	-0.1272	-0.0107	0.0021
On-site fitness center	0.0124	0.0048	-0.0165	-0.1636	-0.1903	0.0146	-0.0004
Free snacks and drinks	0.0609***	0.1526**	0.0382	0.4845*	-0.5505*	0.0998	0.0507***
Free meals	0.0109	0.005	-0.0217	0.7022**	0.0056	0.0145	0.0023
Year-FE	yes	yes	yes	yes	yes	yes	yes
Industry-FE	yes	yes	yes	yes	yes	yes	yes
Controls	yes	yes	yes	yes	yes	yes	yes
Observations	2672	2672	2315	2657	2283	2670	2589

Note: The asterisks indicate statistical significance in 10% (*), 5% (**), and 1% (***) levels. We calculate the difference in contribution to the prediction (SHAP values) for companies with benefit values below the industry sample mean and companies with benefit values above the mean and test the significance of this difference. The values used are industry-adjusted. The R-square values are provided in Table 4.

incorporating all performance measures suggests that these benefits are infrequently offered by more profitable companies, implying that these companies often concentrate on a more limited and specialized set of benefits. This might be an indication of a selective approach in offering the benefits, as recent research suggests optimizing benefits packages at the employee level (Optiz et al., 2022). Moreover, this discrepancy highlights the complex interplay between financial performance, cost management strategies, and employee benefit programs. It suggests that while some companies prioritize extensive employee benefits as part of their growth strategy, others may limit such benefits to manage costs more effectively, thereby influencing their performance in different financial dimensions.

Employee satisfaction grade provides a different view of company performance. With a measure that estimates employee satisfaction, we want to analyze one viable channel on how employee benefits might profit companies by improving employee satisfaction, which in turn, could lead to improved financial performance (Huang et al., 2015; Edmans et al., 2011, 2012). Overall, the differences are similar to financial performance measures. For example, scarcely offering *Flexible spending account, Vision and dental healthcare,* and *Life insurance* are associated with companies that have the highest employee satisfaction. The case of *Flexible working hours* is interesting. Its wide offering is associated with companies that have the most satisfied employees. However, it was scarcely offered by companies with good financial performance. This could be due to decreased informal interaction and communication between employees, which can hurt productivity and innovativeness, and subsequently affect financial performance.

4.3. Predictive importance of benefits

We proceed by analyzing the importance of the benefits variables in predicting company performance using the XGBoost model. As endogeneity issues are an obvious concern with our research setting, we use all the performance measures to identify benefits that consistently show similar associations for all the performance measures. Using many variables that view company performance from different angles helps us find the most important benefits that demonstrate consistent associations with company performance, thus diminishing endogeneity concerns.

As SHAP values measure each feature's additive contribution to a prediction, they are useful in assessing the importance of a variable. In this study, we examine the significance of benefit variables by conducting a comparative analysis at a yearly level between companies that provide a particular benefit extensively and those that provide it sparingly. We calculate the difference in contribution to the prediction (SHAP values) for companies with benefit values below the industry sample mean and companies with benefit values above the mean and test the significance of this difference.[3] To infer statistical significance, we use bootstrapping with 1000 samples Table 5 provides the results [4] As the differences are computed using SHAP values, they can be straightforwardly interpreted as the impacts of each variable on the prediction. For example, for *Flexible spending account* the difference is −0.0357. In other words, companies who invest in good employee treatment by offering *Flexible spending account* more widely than the

[3] As a robustness test, we provide in the Supplementary material the results for a model that uses raw values of the benefits variables that are not industry-adjusted. The results for Tobin's Q are very similar to the results using the industry-adjusted variables.

[4] For comparison purposes, we provide in Appendix B the results for the OLS models predicting the performance measures. They can be compared to the results provided in Table 5 and Figs. 1 and 2. Notice that the difference in the number of observations arises from the ML model's ability to include observations that have missing values for some of the variables.

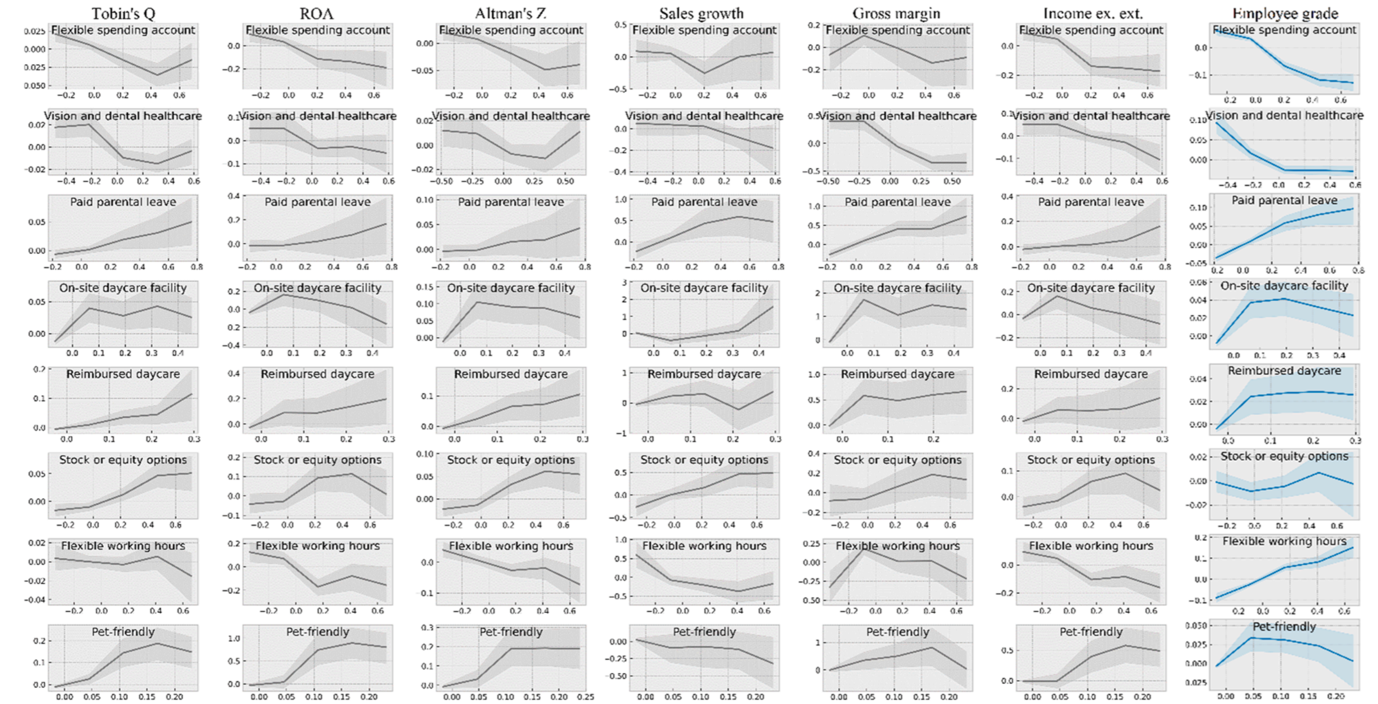

Fig. 2. The nonlinear associations of the financial performance measures from the SHAP values. The gray areas indicate 95% confidence intervals. The x-axis has the proportion of employees indicating having the benefit. The y-axis has SHAP values, and the curve is an average of the SHAP values calculated for six intervals. Thus, the y-axis estimates the effect on the prediction for different values of a benefit variable.

sample average have, on average, 0.0357 smaller Tobin's Q.[5]

To find the most important benefits with consistent associations, we use the following subjective criteria. We pick up for more detailed analysis the variables that 1) consistently have a positive or negative association for all the financial performance variables, and 2) have a statistically significant association for at least four variables. We excluded the employee satisfaction grade variable from the selection process as we consider it an inherently different performance measure that can be expected to show different associations to benefits than the other performance variables. This selection process reveals that *Flexible spending account, Vision and dental healthcare,* and *Flexible working hours* have a consistent negative association with the performance measures, while *Paid parental leave, On-site daycare facility, Reimbursed daycare, Stock or equity options,* and *Pet-friendly* have a consistent positive association with the performance measures. These are, by nature, quite different benefits and especially *Stock or equity options* benefit is often not even considered as a typical benefit. However, as it was considered such in the Kununu portal, we decided to keep it in our analysis despite its different nature.

4.4. Nonlinear associations between employee benefits and company performance

We proceed by analyzing more carefully how the eight benefits are associated with company performance. To analyze the nonlinear association between the performance measures and the benefits, we use individual SHAP values for all the observations and estimate an average contribution to the output variable (performance measures) for different values of the benefits variables. For 95% confidence intervals, we use bootstrapping with 1000 samples. The results are provided in Fig. 2.[6] The x-axis shows the deviations from the industry mean (calculated from a group of companies belonging to the same 2-digit SIC) of how widely a specific benefit is offered to the employees of a company, and the y-axis shows the estimates of the contribution a benefit has for the prediction of a performance measure, calculated as an average of the SHAP values for 6 intervals. The colored areas around the lines are 95% confidence intervals.

As can be seen from the results, *Flexible spending account* is negatively associated with all types of company performance, both financial and employee grade.[7] Furthermore, this benefit is relatively common and based on our arguments from the mental accounting theory, indicating

[5] The analytical outcomes presented in Table 5 and Fig. 1 address distinct aspects of our study and, consequently, are not directly comparable. Specifically, Fig. 1 delineates a comparative analysis between high-performing companies and other firms, with an emphasis on the range and extent of benefits offered, in relation to a defined performance metric. In contrast, Table 5 employs SHAP values to delve into the impact of various benefits on the prediction of company performance. This approach is centered on quantifying the extent to which each benefit contributes to deviations from the average performance prediction, expecting a balance between positive and negative effects. Therefore, while both Fig. 1 and Table 5 contribute to our understanding of the role of benefits in corporate performance, they do so from different analytical perspectives, rendering a direct comparison inappropriate.

[6] In the Supplementary material, we provide an analysis that uses an alternative approach for defining the benefits variables. In the analysis, we make the variables dichotomous so that they get value one if at least one of the respondents mentions having the benefit. Thus, instead of using a variable that measures how widely a benefit is used, we use a variable that measures if the benefit is used at all in the company. The results of this model are very similar to the main model.

[7] In the Supplementary material, we provide the main results for two additional employee satisfaction measures, in particular, review questions from the Kununu data that ask questions about employee satisfaction to overall compensation and pride in company. The results for those models are very similar to the results used in the main analysis. The only notable difference is that the negative associations for 401(k) and 401(k) Company match disappear for the output variable estimating satisfaction to compensation.

that it might be less attractive for employees. Similar conclusions can be made for *Vision and dental healthcare*, although the associations are not statistically significant for *Tobin's Q, ROA,* and *Altman's Z*.

For family-related benefits, firms that provide the benefit more widespread than other firms in the same industry have a higher performance. However, for *On-site daycare facility,* there are some indications for an inverted U-shaped association with *Tobin's Q, ROA,* and *Altman's Z*, although the decrease at high benefits values is not statistically significant. *Paid parental leave* and *Reimbursed daycare* are also associated with companies that have the highest employee satisfaction. Moreover, the overall commonness of these benefits is relatively low (see Table 1) giving support for our arguments that these benefits are highly sought after and therefore associated with the most satisfied employees. This can boost employee performance and the recruitment process, leading to improved financial performance (Huang et al., 2015; Edmans et al., 2011, 2012).

Stock or equity options has mainly a positive association with the financial performance measures. However, there is no positive association with the employee grade. This is interesting as our descriptive analysis indicated this benefit to be most strongly associated with the best-performing companies. *Flexible working hours* is also noteworthy. It is strongly associated with a high employee grade. However, this benefit is also negatively associated with all the financial performance measures. Thus, companies with this kind of flexibility are not among the best performing. As we argued earlier, this might indicate that the benefit is not necessarily improving employee performance. On the contrary, its use might negatively affect employee performance through decreased informal interaction and communication between employees and cause undesired consequences for productivity and innovativeness.

Pet-friendliness in workplaces has been a rising trend recently, and the benefit's role in employee well-being has also raised academia's interest (Foreman et al., 2017; Wilkin et al., 2016). This benefit is rarely offered to employees, and most companies do not offer this benefit at all (see Table 1). Companies offering this benefit often have the most satisfied employees and the benefit is also associated with positive financial performance. Thus, the findings support our arguments concerning rare, highly sought benefits.

4.5. Double ML

We proceed with a robustness test using a double ML model. We analyze only the variables for which we found a consistent association for company performance in the previous section and use the double ML algorithm to get more support for the findings. In the following models, we measure ATE (coefficient) when the value of a treatment variable (the offered benefits) changes from value 0 to value 0.5. The results, provided in Table 6, support the findings of our previous analyses, although with less statistically significant associations. *Flexible spending account* and *Vision and dental healthcare* are negatively associated with *Employee grade*, supporting our previous findings. Also, family-related benefits (*Paid parental leave, On-site daycare facility, Reimbursed daycare*) are positively associated with financial performance measures. *Stock and equity options* has similar associations as our baseline model. It is positively associated with financial performance measures and negatively associated with the employee grade. There is some disagreement with the benefits *Flexible working hours* and *Pet-friendly*, but the associations of the double ML model for these variables are statistically weak and mostly insignificant.

As the double ML analyses these associations using a completely different methodology, the findings give more support for the conclusions we make using the background theories. Because it is practically impossible to remove all the endogeneity concerns, we think that our wide range of analyses viewing the phenomena from different angles, with different outputs and several robustness models, diminishes the endogeneity concerns as much as what is achievable. Thus, we can draw several interesting conclusions from our results that we discuss in the next section.

5. Discussion and conclusions

There is strong theoretical support that good employee treatment should improve employee performance and, subsequently, company performance. Furthermore, previous empirical literature has shown that employee treatment affects many company outcomes. Most of it finds that good employee treatment improves company performance (see, e.g., Gupta & Krishnamurti, 2020; Mao & Weathers, 2019; Fauver et al., 2018; Chen et al., 2016). However, a small but significant part of the literature also finds a negative association (see, e.g., Ben-Nasr and Ghouma, 2018; Cheung et al., 2018; Yu, 2011). Furthermore, prior management control literature documents differences between reward types and employee performance (e.g., Heninger et al., 2019; Kelly et al., 2017, Presslee et al., 2013. As companies always aim to optimize their performance as they operate in a competitive environment, one could assume that their benefits packages offered to employees are continuously monitored, evaluated, and optimized to survive and keep up with the competitors. Companies using better benefits packages succeed while those using sub-optimal ones struggle and fall behind.

Descriptive analysis revealed that *401(k), Employee discounts, Parking,* and *Vision and dental healthcare* are the most widely offered benefits followed by *Paid time off* and *Life insurance*. On the other hand, free food -related benefits and family-related benefits, like *On-site daycare facility, Paid parental leave,* and *Reimbursed daycare*, as well as benefits like *Relocation allowance* and *Pet-friendly*, are most scarcely offered benefits. We also observed a trend where companies experiencing high growth are inclined to provide a diverse range of benefits to their employees, while companies with high profitability tend to focus on offering a more limited and specialized portfolio of benefits. Our subsequent analysis also revealed that more scarcely offered benefits, like family-related benefits, are often associated with the best-performing companies, while more widely offered benefits are associated with poorly performing companies, supporting our arguments drawn from the mental accounting theory.

Why rare benefits would be more efficient in improving performance? For various reasons, companies might see offering them as complicated, for example, from the cost perspective. Or company management might be hesitant to offer them due to possible detrimental effects on productivity. If they were easy to offer or did not create productivity concerns, all companies would start to offer them widely. However, as some companies still offer them, these rare benefits are probably highly sought by employees, otherwise all companies would cease to offer them (Mas & Pallais, 2017; Maestas et al., 2023). Companies want to indicate that they are good employers by offering these highly sought benefits to keep and attract high-performers. Employees see these benefits and form opinions based on their preconceived notions. These opinions then influence their decisions to join and remain in the company.

We consider our results as the first step in uncovering the fine details of associations between different employee benefits and company performance. Although several theories indicate a causal direction from employee benefits to company performance, it is equally possible that good company performance improves benefits packages offered to employees. Due to good performance, companies have more resources, and the management team tries to optimize the use of these resources to succeed in the future. At least a part of these extra resources is likely channeled to improve benefits offered to the employees. Thus, in this case, a found association between benefits and performance would be because of a causal effect from the improved performance to the benefits. In the end, the causal structure is probably even more complex, including bi-directional effects or causal recursive structure (Erhart et al., 2017; Luft & Shields, 2003), where employee benefits improve company performance and good performance improves employee benefits, generating a cyclical causal structure.

Table 6
Average treatment effects of the double ML model.

Name	Tobin's Q	ROA	Altman's Z	Sales growth	Gross margin	Income ex. ext.	Employee Grade
Flexible spending account	0.006	0.0	-0.119	0.010	-0.057	-0.0	-0.755***
	(−0.015)	(−0.011)	(−0.146)	(−0.249)	(−0.760)	(−0.017)	(−4.370)
Vision and dental healthcare	-0.016	0.0	-0.327	-0.033	-0.062	-0.001	-0.591***
	(−0.064)	(−0.007)	(−0.627)	(−1.215)	(−1.280)	(−0.110)	(−5.298)
Paid parental leave	0.813	0.049*	1.240	0.020	0.064	0.043*	-0.151
	(−1.641)	(−1.895)	(−1.235)	(−0.381)	(−0.694)	(−1.664)	(−0.689)
On-site daycare facility	0.267	-0.040	4.655**	0.125	0.464**	-0.51	0.768
	(−0.229)	(−0.655)	(−1.965)	(−1.035)	(−2.123)	(−0.835)	(−1.488)
Reimbursed daycare	4.446**	0.239**	4.979	0.056	0.179	0.218**	-0.278
	(−2.25)	(−2.274)	(−1.205)	(−0.27)	(−0.476)	(−2.093)	(−0.316)
Stock or equity options	0.602*	0.046**	1.100	-0.025	0.036	0.040**	-0.309**
	(−1.71)	(−2.454)	(−1.558)	(−0.686)	(−0.537)	(−2.177)	(−1.965)
Flexible working hours	-0.048	0.008	-0.529	0.041	0.130*	0.005	0.286*
	(−0.129)	(−0.394)	(−0.701)	(−1.061)	(−1.877)	(−0.276)	(−1.758)
Pet-friendly	-5.552**	-0.092	-4.098	-0.072	-0.360	-0.113	-1.240
	(−2.545)	(−0.795)	(−0.944)	(−0.316)	(−0.863)	(−0.982)	(−1.269)
Year-FE	yes	yes	yes	yes	yes	yes	yes
Industry-FE	yes	yes	yes	yes	yes	yes	yes
Controls	yes	yes	yes	yes	yes	yes	yes
Observations	2672	2672	2315	2657	2283	2670	2589

Note: The asterisks indicate statistical significance in 10% (*), 5% (**), and 1% (***) levels. Z-values are provided in parentheses. The values used are industry-adjusted.

Future research could build on these initial findings and collect more evidence on how the advantages of the benefits lead to better company performance. For example, Fig. 2 offers compelling insights into the effects of equitable treatment of employees with respect to various benefits. The data indicate that uniform implementation of certain benefits, such as *Paid parental leave*, is positively correlated with enhanced organizational performance. In contrast, for benefits like *Flexible spending account*, a selective approach appears to yield more favorable outcomes. This preliminary observation invites further research to elucidate the underlying mechanisms that determine when a benefit should be universally applied versus when a selective approach is more advantageous.

Furthermore, subsequent research, particularly those employing robust causal inference methodologies like field experiments, could investigate what kind of mechanism often connects more scarcely offered benefits with enhanced employee motivation. The findings of *Stock and equity options* and *Flexible working hours* are interesting, which also calls for more research. Why do they have differing associations, when comparing financial performance measures and employee satisfaction grades? For *Stock and equity options*, a study by Chang et al. (2018) provides some insight. They offer evidence that the positive association of this benefit with employee grade can be delayed due to a lag in the advantages of these benefits. Thus, future research could analyze the temporal association between benefits and company performance. Another interesting line of research would be to repeat this research in other geographical areas, like Europe. Differences in legislation between countries add another dimension to the research setting. Kununu operates mainly in Europe and provides a much more extensive database for Central European countries, which opens an exciting research opportunity.

Data availability

The code is provided in an online repository mentioned in the manuscript.

Acknowledgments

We would like to thank the editor and the reviewers of our paper for their invaluable contributions and guidance regarding our paper. Their insights and feedback have been instrumental in enhancing the quality and depth of this research. Furthermore, we would like to thank the participants of the Monforma 2022 Conference for the perceptive feedback and constructive discussions regarding our paper.

Appendix A. Description of the variables

Performance measures	
Tobin's Q	(Market capitalization + Liabilities) / (Total assets)
ROA	(Net income) / (Total assets)
Altman's Z	3.3 * (EBIT / Total assets) + (Sales / Total assets) + 0.6 * (Market value / Total liabilities) + 1.2 * (Working capital / Total assets) + 1.4 * (Retained earnings / Total assets)
Sales growth	(Sales(t) – Sales(t-1)) / Sales(t-1)
Gross margin	(Sales – Cost of goods sold) / (Sales)
Income ex. ext.	(Income before extraordinary items) / (Total assets)
Employee grade	An average grade of 18 different questions related to company practices, graded between 1-5. The employee ratings are described below.
Control variables	
Total Assets	Total assets of a company.
Leverage	(Total long-term debt + Current liabilities) / (Stockholder's equity)
Market-to-Book	(Market value) / (Book value of equity)
R&D Intensity	Research and development expenses divided by the total assets of a company.
Implied Volatility	Volatility used in the fair value calculation for stock options.
Firm Age	Firm age computed from the first observation in the database

(*continued on next page*)

(*continued*)

Performance measures	
Overall compensation for work	Yearly average of the grades (1-5) to the question: "Overall, do you feel that you are fairly compensated for your work?"
Employees' rating of their employer ranked on a five-point scale, with five (one) being the most favourable (least favourable):	
Company culture	How is the overall company culture?
Internal communication	How is the internal communication? To what extent are you informed about company results, successes, and challenges through regular communication?
Teamwork	How are co-workers at working together and interacting in an honest, direct manner?
Work–life balance	How does the company value work–life balance? Are families considered? Is there pressure to work long hours?
Support from management	Does the leadership set realistic expectations, communicate clear goals, and involve employees in the decision-making process?
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?
Office/work environment	Is the work environment comfortable and suited to do the work you are doing? Is there proper ventilation, lighting, temperature control, and technology available?
Environmental friendliness	To what extent does the company demonstrate concern for or awareness of the environment (e.g., having recycling programme in place)?
Accessible to people with disabilities	Does the company have facilities that are handicapped accessible to support people with disabilities?
Workplace safety	Does the company maintain a safe and compliant working environment?
Overall compensation for work	Overall, do you feel that you are fairly compensated for your work?
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?
Challenging work	How challenging is your work? Are you proud of the work that you produce?

Appendix B. Results of the OLS models

Here we provide two OLS models for comparison purposes. The first one uses the same controls as the ML models while the second one does not have R&D intensity and implied volatility as controls, as these variables cause almost all the missing observations in the first model.

	Tobin's Q	ROA	Altman's Z	Sales growth	Gross margin	Income ex. ext.	Employee grade
401(k)	0.2978*	0.8436	1.1986***	0.5285	0.7974	0.8498	0.0421
401(k) Company match	0.2181	-0.6744	0.1026	8.1793***	2.7374	-0.6789	-0.1101
Flexible spending account	-0.2200	-0.2842	-0.2932	-1.4983	-4.8829	-0.4144	-0.4858***
Vision and dental healthcare	-0.7501	-5.0708*	-1.7369	-5.6491	-25.4592**	-5.4287*	-0.3376
Relocation allowance	-0.6793	-0.5217	-0.3722	3.2546	-6.2797	-1.4423	-0.8201***
Life insurance	-0.4319	-2.9080*	-0.0543	-3.7124	-3.9355	-2.4	0.0678
Paid parental leave	0.3582	1.2344	0.4229	3.1982	10.2456*	1.1111	0.4689***
Desirable office location	0.1793	-0.7081	1.1017*	-5.9924	3.0288	-1.0504	0.4203***
On-site daycare facility	-0.9255	-8.6360**	-0.824	-8.8817	4.3297	-8.6087**	0.5048
Paid time off	-0.208	1.3	0.1236	-4.8031	8.7370*	1.5564	-0.1205
Parking	-0.1234	-0.2342	-0.0336	-4.4556**	-8.2586**	-0.2461	0.1006
Reimbursed daycare	-0.1754	-4.5191	-2.1043	-4.5642	-32.5587*	-5.4591	0.3763
Stock or equity options	0.4788**	1.6114	0.8385*	-0.7881	5.9539	1.9159*	-0.0629
Ability to work remotely	-0.0301	-2.2464	0.1153	7.8220**	6.6235	-2.5280*	0.3367***
Easy access to public transportation	0.4435	0.4334	0.6305	-8.4162**	7.8791	-0.0768	-0.108
Tuition assistance	0.0082	0.7215	0.197	0.1792	-2.9593	1.0375	0.2028
Flexible working hours	-0.0967	-1.3789	-0.7299**	-0.1682	2.2155	-1.3757	0.0981
On-site cafeteria	-0.3918**	-0.985	-0.6414*	0.6308	-1.676	-0.8407	0.0162
Vehicle allowance	3.7502**	19.9563**	6.0788*	32.6284	57.6928*	20.8238**	-0.7091
Employee discounts	-0.1388	0.0228	-0.2162	-4.7276**	2.9756	0.0244	0.2650***
Pet-friendly	1.1527	9.5141*	0.4092	6.0419	32.5932*	9.6941**	-0.1285
Employee events	-0.3673	0.64	-0.5709	3.8219	-7.436	0.4405	0.4374***
Health and wellness programs	-0.1734	-0.7014	-1.7533*	-4.7516	-2.4583	-1.6198	0.1056
On-site fitness center	0.2625	2.8567	1.1518	-1.111	5.5265	3.4425*	-0.1883
Free snacks and drinks	1.1646**	-3.217	-0.5117	-4.864	-23.7986**	-4.7442*	0.6128***
Free meals	-0.3525	2.5879	0.0508	2.4405	11.4738	3.7774*	-0.3264*
Year-FE	x	x	x	x	x	x	x
Industry-FE	x	x	x	x	x	x	x
Controls	x	x	x	x	x	x	x
Observations	905	905	866	900	870	905	905
R-squared	0.354	0.303	0.298	0.087	0.351	0.320	0.262
	Tobin's Q	ROA	Altman's Z	Sales growth	Gross margin	Income ex. ext.	Employee grade
401(k)	-0.0418	0.1699	0.0994	-0.8559	-3.1769	0.0336	0.0308
401(k) Company match	0.2759*	2.0465**	0.2032	5.2218***	3.5795	1.7774**	-0.1584**
Flexible spending account	-0.3140**	-1.3703*	-0.8054**	-0.7988	-4.9620*	-1.3692*	-0.4192***
Vision and dental healthcare	-0.7490**	-2.6486	-1.5181*	-4.892	-15.2905**	-3.0489*	-0.1996
Relocation allowance	0.3314	3.4324**	1.2277**	-1.0117	6.0106	2.7609*	-0.8471***

(*continued on next page*)

(*continued*)

Life insurance	-0.4316**	-2.3267**	-0.3683	-3.0717	-4.5284	-1.8256*	0.0284
Paid parental leave	0.4967***	1.8129**	0.453	3.1406*	8.2751**	1.8032**	0.4132***
Desirable office location	-0.0431	-0.1429	0.3941	0.2958	-0.1477	-0.1053	0.2351***
On-site daycare facility	-0.7209	-0.9738	-0.2541	-3.9687	7.7006	-1.0851	0.2157
Paid time off	-0.1535	-0.8134	0.0465	-2.4269	7.3545**	-0.6674	-0.2208***
Parking	-0.2722**	-1.5439***	-0.4281*	-3.9257***	-9.0219***	-1.4574***	0.1191**
Reimbursed daycare	0.3809	3.0493	0.7639	-3.3469	-3.2691	2.7798	0.3475
Stock or equity options	0.4825***	1.5450**	0.9897***	2.2841	4.6663*	1.6347**	0.1158*
Ability to work remotely	0.0246	-1.6142*	-0.3115	3.1981*	-0.7082	-1.7383**	0.3204***
Easy access to public transportation	-0.1906	-0.6994	-0.112	-4.0869**	0.8756	-0.8405	-0.0991
Tuition assistance	0.1025	0.8026	-0.2781	-1.8148	-6.9559**	0.5596	0.1112
Flexible working hours	-0.1529	-0.8344	-0.4393*	-2.7759**	2.414	-1.0138*	0.2504***
On-site cafeteria	0.1567	0.0487	0.6273**	1.4205	4.5687**	0.0406	0.1391***
Vehicle allowance	2.9445***	6.4509	5.5227**	20.0071*	31.272	7.0514	-0.7122
Employee discounts	-0.1196	-0.1734	-0.6921***	-1.7633	3.2242*	-0.0723	0.1617***
Pet-friendly	1.2286**	10.1836***	2.0448	-3.6752	19.0203	10.1649***	0.0129
Employee events	-0.0658	0.9669	0.2757	-0.5193	0.891	1.2446	0.4748***
Health and wellness programs	-0.3975	-0.4405	-1.8081**	2.3463	-9.4193	-0.6224	-0.0362
On-site fitness center	0.3076	1.6156	1.2382**	-2.2621	8.3428*	1.8611	-0.0411
Free snacks and drinks	1.1993***	2.7145	1.8811**	3.8313	-9.3594	1.7338	0.5730***
Free meals	-0.156	-0.3242	-1.2150**	1.6507	1.0893	0.2174	-0.2905**
Year-FE	x	x	x	x	x	x	x
Industry-FE	x	x	x	x	x	x	x
Controls (no R&D int., Volatility)	x	x	x	x	x	x	x
Observations	2493	2493	2160	2480	2123	2492	2495
R-squared	0.116	0.099	0.102	0.076	0.177	0.100	0.218

Appendix C. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.mar.2023.100876.

References

Altman, E.I., 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. J. Financ. 23 (4), 589–609. https://doi.org/10.2307/2978933.

Arlot, S., Celisse, A., 2010. A survey of cross-validation procedures for model selection. Stat. Surv. 4 (none), 40–79. https://doi.org/10.1214/09-SS054.

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.

Beck, N., King, G., Zeng, L., 2004. Theory and evidence in international conflict: a response to de marchi, gelpi, and grynaviski. Am. Political Sci. Rev. 98 (2), 379–389. https://doi.org/10.1017/S0003055404001212.

Becker, B.E., Huselid, M.A., Huselid, M.A., Ulrich, D., 2001. The HR Scorecard: Linking People, Strategy, and Performance. Harvard Business Press.

Benlemlih, M., 2019. Corporate social responsibility and dividend policy. Res. Int. Bus. Financ. 47, 114–138. https://doi.org/10.1016/j.ribaf.2018.07.005.

Ben-Nasr, H., Ghouma, H., 2018. Employee welfare and stock price crash risk. J. Corp. Financ. 48, 700–725. https://doi.org/10.1016/j.jcorpfin.2017.12.007.

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., 2020. Using machine learning to detect misstatements. Rev. Account. Stud. https://doi.org/10.1007/s11142-020-09563-8.

Bertrand, M., Mullainathan, S., 2003. Enjoying the quiet life? corporate governance and managerial preferences. J. Political Econ. 111 (5), 1043–1075. https://doi.org/10.1086/376950.

Bhimani, A., 2020. Digital data and management accounting: Why we need to rethink research methods. J. Manag. Control 31 (1), 9–23. https://doi.org/10.1007/s00187-020-00295-z.

Burman, P., 1989. A comparative study of ordinary cross-validation, v-fold cross-validation and the repeated learning-testing methods. Biometrika 76 (3), 503–514. https://doi.org/10.1093/biomet/76.3.503.

Busse, C., Mahlendorf, M.D., Bode, C., 2016. The ABC for studying the too-much-of-a-good-thing effect: a competitive mediation framework linking antecedents, benefits, and costs. Organ. Res. Methods 19 (1), 131–153. https://doi.org/10.1177/1094428115579699.

Bussmann, N., Giudici, P., Marinelli, D., Papenbrock, J., 2021. Explainable machine learning in credit risk management. Comput. Econ. 57 (1), 203–216. https://doi.org/10.1007/s10614-020-10042-0.

Chang, Y.-C., Chang, K.-H., Wu, G.-J., 2018. Application of eXtreme gradient boosting trees in the construction of credit risk assessment models for financial institutions. Appl. Soft Comput. 73, 914–920. https://doi.org/10.1016/j.asoc.2018.09.029.

Chen, C., Chen, Y., Hsu, P.-H., Podolski, E.J., 2016. Be nice to your innovators: employee treatment and corporate innovation performance. J. Corp. Financ. 39, 78–98. https://doi.org/10.1016/j.jcorpfin.2016.06.001.

Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system ([Cs]). ArXiv 1603, *02754*. https://doi.org/10.1145/2939672.2939785.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., Robins, J., 2017a. Double/debiased machine learning for treatment and causal parameters. ArXiv 1608, *00060 [Econ, Stat]*. http://arxiv.org/abs/1608.00060.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., 2017. Double/Debiased/Neyman machine learning of treatment effects. ArXiv 1701, *08687 [Stat]*. http://arxiv.org/abs/1701.08687.

Cheung, A. (Waikong, Hu, M., Schwiebert, J., 2018. Corporate social responsibility and dividend policy. Account. Financ. 58 (3), 787–816. https://doi.org/10.1111/acfi.12238.

Choi, J. (Willie, Presslee, A., 2023. When and why tangible rewards can motivate greater effort than cash rewards: an analysis of four attribute differences. Account., Organ. Soc. 104, 101389 https://doi.org/10.1016/j.aos.2022.101389.

Cronqvist, H., Heyman, F., Nilsson, M., Svaleryd, H., Vlachos, J., 2009. Do entrenched managers pay their workers more? J. Financ. 64 (1), 309–339. https://doi.org/10.1111/j.1540-6261.2008.01435.x.

Darrough, M., Kim, H., Zur, E., 2019. The impact of corporate welfare policy on firm-level productivity: evidence from unemployment insurance. J. Bus. Ethics 159 (3), 795–815. https://doi.org/10.1007/s10551-018-3817-2.

Ding, D.Z., Akhtar, S., Ge, G.L., 2009. Effects of inter- and intra-hierarchy wage dispersions on firm performance in Chinese enterprises. Int. J. Hum. Resour. Manag. 20 (11), 2370–2381. https://doi.org/10.1080/09585190903239716.

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.

Edmans, A., 2011. Does the stock market fully value intangibles? employee satisfaction and equity prices. J. Financ. Econ. 101 (3), 621–640. 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), 1–19. https://doi.org/10.5465/amp.2012.0046.

Eisenberger, R., Armeli, S., Rexwinkel, B., Lynch, P.D., Rhoades, L., 2001. Reciprocation of perceived organizational support. J. Appl. Psychol. 86 (1), 42–51. https://doi.org/10.1037/0021-9010.86.1.42.

Eisenberger, R., Cummings, J., Armeli, S., Lynch, P., 1997. Perceived organizational support, discretionary treatment, and job satisfaction. J. Appl. Psychol. 82 (5), 812–820. https://doi.org/10.1037/0021-9010.82.5.812.

Erhart, R., Mahlendorf, M.D., Reimer, M., Schäffer, U., 2017. Theorizing and testing bidirectional effects: the relationship between strategy formation and involvement of controllers. Account., Organ. Soc. 61, 36–52. https://doi.org/10.1016/j.aos.2017.07.004.

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. Financ. 50, 84–108. https://doi.org/10.1016/j.jcorpfin.2018.02.003.

Flammer, C., 2015. Does corporate social responsibility lead to superior financial performance? a regression discontinuity approach. Manag. Sci. 61 (11), 2549–2568. https://doi.org/10.1287/mnsc.2014.2038.

Foreman, A.M., Glenn, M.K., Meade, B.J., Wirth, O., 2017. Dogs in the workplace: a review of the benefits and potential challenges. Int. J. Environ. Res. Public Health, 14(5), Artic. 5. https://doi.org/10.3390/ijerph14050498.

Francis, B., Hasan, I., Liu, L., Wang, H., 2019. Employee treatment and contracting with bank lenders: an instrumental approach for stakeholder management. J. Bus. Ethics 158 (4), 1029–1046. https://doi.org/10.1007/s10551-017-3722-0.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29 (5), 1189–1232.

Garel, A., Petit-Romec, A., 2020. Engaging employees for the long run: long-term investors and employee-related CSR. J. Bus. Ethics. https://doi.org/10.1007/s10551-020-04572-8.

Godfrey, P.C., Merrill, C.B., Hansen, J.M., 2009. The relationship between corporate social responsibility and shareholder value: An empirical test of the risk management hypothesis. Strateg. Manag. J. 30 (4), 425–445. https://doi.org/10.1002/smj.750.

Gouldner, A.W., 1960. The norm of reciprocity: a preliminary statement. Am. Sociol. Rev. 25 (2), 161–178. https://doi.org/10.2307/2092623.

Green, T.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.

Gupta, K., Krishnamurti, C., 2020. Do countries matter more in determining the relationship between employee welfare and financial performance? Int. Rev. Financ. 20 (2), 415–450. https://doi.org/10.1111/irfi.12231.

Guthrie, J.P., 2001. High-involvement work practices, turnover, and productivity: evidence from New Zealand. Acad. Manag. J. 44 (1), 180–190. https://doi.org/10.5465/3069345.

Hales, J., Moon, J.R., Swenson, L.A., 2018. A new era of voluntary disclosure? empirical evidence on how employee postings on social media relate to future corporate disclosures. Account. Organ. Soc. 68–69, 88–108. https://doi.org/10.1016/j.aos.2018.04.004.

Hartford, J., Lewis, G., Leyton-Brown, K., & Taddy, M., 2017. Deep IV: A Flexible Approach for Counterfactual Prediction. In D. Precup & Y. W. Teh (Eds.), Proceedings of the 34th International Conference on Machine Learning (Vol. 70, pp. 1414–1423). PMLR. https://proceedings.mlr.press/v70/hartford17a.html.

Heninger, W.G., Smith, S.D., Wood, D.A., 2019. Reward type and performance: an examination of organizational wellness programs. Manag. Account. Res. 44, 1–11. https://doi.org/10.1016/j.mar.2019.02.001.

Huang, K., Li, Meng, 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. Financ. 34, 108–127. https://doi.org/10.1016/j.jcorpfin.2015.08.002.

Jones, S., 2017. Corporate bankruptcy prediction: a high dimensional analysis. Rev. Account. Stud. 22 (3), 1366–1422. https://doi.org/10.1007/s11142-017-9407-1.

Khaidem, L., Saha, S., & Dey, S.R., 2016. Predicting the direction of stock market prices using random forest. ArXiv:1605.00003 [Cs]. http://arxiv.org/abs/1605.00003.

Kelly, K., Presslee, A., Webb, R.A., 2017. The effects of tangible rewards versus cash rewards in consecutive sales tournaments: a field experiment. Account. Rev. 92 (6), 165–185. https://doi.org/10.2308/accr-51709.

van Knippenberg, D., van Dick, R., Tavares, S., 2007. Social identity and social exchange: identification, support, and withdrawal from the job. J. Appl. Soc. Psychol. 37, 457–477. https://doi.org/10.1111/j.1559-1816.2007.00168.x.

Kumar, M., Thenmozhi, M., 2014. Forecasting stock index returns using ARIMA-SVM, ARIMA-ANN, and ARIMA-random forest hybrid models. Int. J. Bank., Account. Financ. 5 (3), 284–308. https://doi.org/10.1504/IJBAAF.2014.064307.

Levinson, H., 1965. Reciprocation: the relationship between man and organization. Adm. Sci. Q 9 (4), 370–390. https://doi.org/10.2307/2391032.

Liu, C., Chan, Y., Kazmi, S.H.A., Fu, H., 2015. Financial fraud detection model: based on random forest. Int. J. Econ. Financ. 7 (7), p178. https://doi.org/10.5539/ijef.v7n7p178.

Luft, J., Shields, M.D., 2003. Mapping management accounting: graphics and guidelines for theory-consistent empirical research. Account., Organ. Soc. 28 (2), 169–249. https://doi.org/10.1016/S0361-3682(02)00026-0.

Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.-I., 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. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), Advances in Neural Information Processing Systems 30 (pp. 4765–4774). Curran Associates, Inc. http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf.

Mahlendorf, M.D., Martin, M., Smith, D.A., 2023. Innovative data – use-cases in management accounting research and practice. Eur. Account. Rev. https://doi.org/10.1080/09638180.2023.2213258.

Maestas, N., Mullen, K.J., Powell, D., von Wachter, T., Wenger, J.B., 2023. The value of working conditions in the united states and the implications for the structure of wages. Am. Econ. Rev. 113 (7), 2007–2047. https://doi.org/10.1257/aer.20190846.

Mao, C.X., Weathers, J., 2019. Employee treatment and firm innovation. J. Bus. Financ. Account. 46 (7–8), 977–1002. https://doi.org/10.1111/jbfa.12393.

Mas, A., Pallais, A., 2017. Valuing alternative work arrangements. Am. Econ. Rev. 107 (12), 3722–3759. https://doi.org/10.1257/aer.20161500.

Opitz, S., Sliwka, D., Vogelsang, T., & Zimmermann, T. 2022. The Algorithmic Assignment of Incentive Schemes (SSRN Scholarly Paper 4077778). https://doi.org/10.2139/ssrn.4077778.

Pagano, M., Volpin, P.F., 2005. Managers, workers, and corporate control. J. Financ. 60 (2), 841–868. https://doi.org/10.1111/j.1540-6261.2005.00748.x.

Presslee, A., Vance, T.W., Webb, R.A., 2013. The effects of reward type on employee goal setting, goal commitment, and performance. Account. Rev. 88 (5), 1805–1831. https://doi.org/10.2308/accr-50480.

Ranta, M., & Ylinen, M., 2023. Board gender diversity and workplace diversity: A machine learning approach. Corporate Governance: The International Journal of Business in Society, ahead-of-print(ahead-of-print). https://doi.org/10.1108/CG-01-2022-0048.

Ranta, M., Ylinen, M., Järvenpää, M., 2022. Machine learning in management accounting research: literature review and pathways for the future. Eur. Account. Rev. 0 (0), 1–30. https://doi.org/10.1080/09638180.2022.2137221.

Roehling, P.V., Roehling, M.V., Moen, P., 2001. The relationship between work-life policies and practices and employee loyalty:a life course perspective. J. Fam. Econ. Issues 22 (2), 141–170. https://doi.org/10.1023/A:1016630229628.

Saeed, A., 2021. The impact of employee friendly practices on dividend payments: evidence from emerging economies. J. Bus. Res. 135, 592–605. https://doi.org/10.1016/j.jbusres.2021.07.009.

Storm, H., Baylis, K., & Heckelei, T. (n.d.)., 2020. Machine learning in agricultural and applied economics. European Review of Agricultural Economics. https://doi.org/10.1093/erae/jbz033.

Syrgkanis, V., Lei, V., Oprescu, M., Hei, M., Battocchi, K., & Lewis, G., 2019. Machine Learning Estimation of Heterogeneous Treatment Effects with Instruments. ArXiv:1905.10176 [Cs, Econ, Stat]. http://arxiv.org/abs/1905.10176.

Thaler, R., 1985. Mental accounting and consumer choice. Mark. Sci. 4 (3), 199–214.

Thaler, R.H., 1999. Mental accounting matters. J. Behav. Decis. Mak. 12 (3), 183–206.

Valizade, D., Schulz, F., & Nicoara, C. (2022). Towards a Paradigm Shift: How Can Machine Learning Extend the Boundaries of Quantitative Management Scholarship? British Journal of Management, n/a(n/a). https://doi.org/10.1111/1467-8551.12678.

Wager, S., Athey, S., 2018. Estimation and inference of heterogeneous treatment effects using random forests. J. Am. Stat. Assoc. 113 (523), 1228–1242. https://doi.org/10.1080/01621459.2017.1319839.

Whitener, E.M., 2001. Do "high commitment" human resource practices affect employee commitment?: a cross-level analysis using hierarchical linear modeling. J. Manag. 27 (5), 515–535. https://doi.org/10.1177/014920630102700502.

Wilkin, C.L., Fairlie, P., Ezzedeen, S.R., 2016. Who let the dogs in? a look at pet-friendly workplaces. Int. J. Workplace Health Manag. 9 (1), 96–109. https://doi.org/10.1108/IJWHM-04-2015-0021.

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

Yu, B., 2011. Employee rights and dividend policy around the world (SSRN Scholarly Paper ID 1786598). Soc. Sci. Res. Netw. https://doi.org/10.2139/ssrn.1786598.