Technology in Society 76 (2024) 102453

Contents lists available at ScienceDirect

Technology in Society

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

Corporate social responsibility in family business: Using machine learning to uncover who is doing good

Feng Liu [a], Wanying Huang [b], Jing Zhang [c], Mingjie Fang [d,*]

[a] Business School, Shandong University, Weihai, China
[b] Business School, Central University of Finance and Economics, Beijing, China
[c] Department of Electronic Business, South China University of Technology, Guangzhou, China
[d] Department of Logistics, Service & Operations Management, Korea University Business School, Seoul, South Korea

ARTICLE INFO

Keywords:
Corporate social responsibility
Family business
Firm characteristics
CEO characteristics
Machine learning

ABSTRACT

Rapid social and environmental changes continue to highlight the critical role of corporate social responsibility (CSR) in business operations. Research on the determinants of CSR has received widespread attention. However, there has been little focus on family businesses, which play an important function in the economy and differ from non-family firms. Synthesizing the resource-based view (RBV) with the behavioral agency theory (BAT), this study aimed to examine the factors that drive the CSR performance of family businesses from the perspective of the firm and chief executive officer (CEO) characteristics. By comparing the performance of a set of machine learning (ML) techniques, we found that the best predictive model with the lowest mean squared error is the random forest algorithm. The results from the random forest indicated that profitability was the most important attribute of CSR performance, followed by firm size, CEO education, leverage, and board ownership. This study is one of the first to adopt an ML approach to investigate the drivers of CSR performance in family businesses. The novel findings provide a deeper understanding of how the various aspects of a family business firm affect its CSR performance, which can facilitate future research.

1. Introduction

Along with the increasing interest in organizational prosocial behaviors, corporate social responsibility (CSR) has been a striking indicator influencing investors, stakeholders, and regulators' decision-making for many years [1–3]. Business analysts predict firms' future growth and performance based on CSR ratings to make investment recommendations [4], and governments adjust public policies to support CSR for economic growth [5]. Recently, the socially responsible behaviors of family businesses have attracted considerable attention, and scholars have emphasized their unique roles in promoting regional environmental and social welfare [6,7].

Following previous studies, CSR performance refers to a signal of management integrity and ethics, which entails achievements in environmental and social performance areas ([8,9,97]. The better a family business performs and fulfills these aspects, the higher its CSR performance. Previous articles have involved research on businesses in many different countries, while our research focuses on the CSR performance of Chinese family businesses for two main reasons. First, family businesses are one of the major contributors to China's rapid economic development, contributing half of tax revenue, more than 60 % of GDP, and 90 % of new jobs [10]. Moreover, family firms are the main component of China's private enterprises, accounting for approximately 90 % [11]. Second, most previous studies have focused on developed economies like the United States instead of emerging economies like China. However, China is a representative emerging economy because of its unique institutional environment [10,12]. Notably, the impact of Confucianism significantly shapes business ethics across East Asia, particularly within China [13]. Thus, Chinese companies exhibit a heightened emphasis on CSR performance. Hence, due to the uniqueness of the cultural and institutional environment, studying the CSR performance of Chinese family businesses is meaningful. Although existing literature has suggested several determinants of firms' CSR practices, such as social media [14], stakeholder pressure [15], and cultural values [16], there is a dearth of a systematic framework and family business-dedicated research to explain how the various aspects of a

* Corresponding author. Department of Logistics, Service & Operations Management, Korea University Business School, Seoul, 02841, South Korea.
E-mail addresses: liufeng@sdu.edu.cn (F. Liu), 2023211238@email.cufe.edu.cn (W. Huang), ebzhangjing@mail.scut.edu.cn (J. Zhang), mj_fang@kroea.ac.kr (M. Fang).

https://doi.org/10.1016/j.techsoc.2024.102453
Received 27 July 2022; Received in revised form 30 December 2023; Accepted 1 January 2024
Available online 2 January 2024

family business firm affect its CSR performance.

Research suggests that family firms perform better than non-family businesses regarding CSR [17]. On the one hand, family firms tend to exhibit ethical behavior because of their typical long-term orientation [18]. On the other hand, to forestall social concerns about the potential hazard to society, family firms devote themselves to assuming social responsibility in a multitude of ways [6,19,20]. Additionally, to safeguard the family image and transgenerational control, family businesses are more prone to participate in CSR activities [21]. Furthermore, family firms are more likely to pursue environmental responsibility than non-family firms because of their relatively concentrated dispersion of business operations [22,23]. As a result, family businesses attach importance to promoting communication with their external stakeholders, establishing foundations, and publishing CSR reports on their websites to preserve their reputation [6,24].

Research has established that a family firm's characteristics are crucial to its CSR. The board of directors, a core corporate governance tool, can reduce information asymmetries and conflicts of interest between internal managers and organizational stakeholders [25]. [26] found that board ownership is positively related to financial performance and, thus, leads to higher CSR performance. Regarding board characteristics and CSR [27], pointed out that board size is significantly negatively related to the adoption of CSR activities, while board independence is significantly positive. [28] found that when a family has a strong influence, which equals strong involvement, the board cannot promote CSR activities in its business. Moreover, several studies suggest that firm size and profitability positively impact CSR performance [28–30].

Chief executive officers (CEOs) in family firms also play an important role in explaining variances in CSR compared to non-family firms. For instance, an increasing number of Chinese executives promote active engagement in CSR initiatives in alignment with global reporting initiative standards trends. In China, the evolving socioeconomic landscape provides a flourishing environment for CSR practices, which requires local firms to integrate social and environmental concerns into business strategies. In family firms, when a family member assumes the CEO's office, their actions primarily focus on the family's interests [31]. Building on previous research (e.g. Refs. [25,32]), we can conclude that CEO characteristics are key influencing factors in CSR performance.

Although previous studies have argued for the decisive role of CEOs in business operations and CSR activities [32–34], it remains unknown about the permutation importance of various CEO characteristics in predicting the CSR performance of family firms and how CEO characteristics drive CSR performance in the context of family business. This study fills this research gap by using machine learning (ML) methods to develop a prediction model of CSR performance for uncovering how firm-level and CEO-level characteristics influence family firms' CSR performance. Drawing on a synthesized view of the resource-based view (RBV) and the behavioral agency theory (BAT), this study investigates whether and how firm and CEO characteristics influence family firms' CSR performance using an ML approach. First, RBV emphasizes the significance of resources in enabling organizations to gain a competitive advantage [35]. It underscores that a firm's tangible and intangible resources are critical for achieving competitive advantages. Moreover, RBV posits that as firms acquire unique and difficult-to-imitate resources, they achieve more enduring competitive advantages. Following this logic, we contend that the firm characteristics of family businesses can be a resource, and some of their inherent features, such as firm age and ownership structure, can assist them in gaining an advantage in CSR performance [36]. Second, while we emphasized firm characteristics as "heterogeneous resources" in fueling CSR performance, we utilized BAT to justify the importance of CEO characteristics in determining the CSR performance of family businesses. In the context of family businesses, BAT suggests that the family CEO plays a crucial role in CSR decisions as the key leader of a family firm's visions, strategies, and operations. In addition, concerning the heterogeneity of family firms, CEOs in family firms may pay more attention to the non-economic objectives linked with the family [37,38]. Since CEOs' strategic orientation and leadership differ across characteristics, we believe it is necessary to consider CEO characteristics as determinants of CSR performance in family firms.

This study contributes to the existing body of knowledge on CSR in family firms by investigating the determinants of the CSR of family businesses. The study approach is novel in that ML algorithms explore the determinants of family firms' CSRs. Despite much research using traditional estimation methods, few studies have used ML methods to investigate CSR in family businesses. This novel algorithm-supported approach compensates for the shortcomings of traditional statistical models (e.g., excessive model complexity and sample dependence, complex functional forms) [39]. ML algorithms facilitate the analysis of suitable complex function forms and alleviate the issue of overfitting, thereby providing a systematic determinant model for measuring the CSR performance of family firms [40–42].

In addition, this study provides a systematic framework for determining the drivers of family firms' CSR through family firm characteristics. Previous research on the impact of firm characteristics on CSR performance remains limited because the evidence is piecemeal and lacks a global perspective. The present study examines firm characteristics related to accounting, financial aspects, and board information (firm size, firm age, market value, leverage, profitability, board size, and board ownership). It determines their predictive power and permutation importance to CSR performance.

Furthermore, this study analyzes the CSR performance of family firms based on comprehensive CEO characteristics, which is thus far still lacking in existing research, with most studies focusing only on limited aspects of CEO characteristics, such as age and family membership [32,43]. This study bridges this literature gap by synthesizing CEO characteristics involving the CEO's gender, age, tenure, education, family membership, pay, and duality. To this end, we developed a comprehensive theoretical CSR performance determinant model by integrating RBV and BAT.

We organized the remainder of the paper as follows. The next section reviews related literature on CSR in family business. The third section illustrates the ML method used in the current research. The fourth section discusses the main findings, and the last section concludes with the implications and limitations, outlining future directions.

2. Literature review

2.1. Related literature on CSR in family business

CSR in family businesses can be defined as discretionary multidimensional activities, including social, political, environmental, economic, and ethical actions undertaken by family firms [32,44,45]. Family businesses play a critical role in local economies and global markets, making focusing on their CSR activities important because of their ubiquity and significance [46]. systematically reviewed the literature on CSR in family businesses. They summarized the driving factors of family firms' CSR, including firm size, family involvement, corporate governance, ethics, and religion. In addition [47], identified the value system that defines the social interaction of family businesses as an important aspect affecting the orientation of CSR. Similarly [7], state that CSR orientation enhances family business performance by highlighting the role of family values and CSR attitudes and promoting relevant connections with the surrounding community.

Researchers have investigated links between the board directors and family firms' CSR commitment. For example [48], pointed out that female directors are associated with higher levels of CSR commitment in non-family firms, while CSR commitment is not significantly related to the presence of female directors in family firms [27]. examined the relationship between board characteristics and CSR performance. They found that board size is significantly negatively related to the adoption of CSR activities, while board independence is significantly positively related. Additionally, researchers have studied the relationship between a family firm and its stakeholders. For instance Ref. [49], indicated that family involvement influences internal and external stakeholders, including employees, customers, and suppliers. Thus, existing literature emphasizes that family firms are inclined to promote CSR activities to improve their reputation with trade partners and the community.

Previous research on the determinants of CSR has hinted at the importance of two broad categories: family firm characteristics and CEO characteristics. The first category addresses the family firm characteristics that impact the CSR performance of family firms [23]. suggested that the number of directors on a board affects agency concerns and firms' prosocial decisions [27]. pointed out that board size is significantly negatively associated with adopting CSR activities, while board independence is positively related [36]. found that family firms' CSR performance positively correlates with firm age and growth opportunities. Other studies, such as [6,43]; stated that firm size impacts CSR performance, and the larger the enterprise scale, the higher the disclosure level of CSR with an increase in CSR investment due to improvement in firm performance.

Regarding the role of CEO characteristics, a recent study by Ref. [32] states that with increasing CEO age, family firms will mildly affect their decision reference, leading to a change in the priority of all stakeholders related to CSR performance. As the CEO ages, the anticipation of succession becomes the main concern. In a study on the role of a CEO's incentives in driving CSR [33], argued that as CEOs get older, they feel less pressure on career concerns, which motivates them to pay more attention to the concerns suggested by the diverse stakeholders who are not strongly tied to economic profits, even if such activities may damage corporate profits. Furthermore, CEO education is an important factor in CSR performance [34]. state that CEO education positively impacts CSR disclosure. If a family member assumes the CEO's role in the family business, firms may emphasize priorities mainly relevant to family members' organizational identification, binding social connections, and internal norms in the group [32,50]. Additionally, CEO duality (serving as CEO and chair of the board of directors) and CEO family membership (the CEO is a family member) influence the CSR performance of family businesses [32,43]. In addition, research has shown other CEO characteristics, such as gender, tenure, and pay, influence CSR performance in family businesses [32,51,52].

Based on the above discussion, despite previous studies having reached meaningful findings, there is currently no consensus on the determinants of CSR performance in family businesses, especially regarding CEO characteristics. Thus, this study fills an important literature gap by exploring how family firm characteristics and CEO characteristics affect CSR performance in Chinese family businesses by developing a prediction model using ML methods.

2.2. Theoretical background and research framework

This study primarily draws on the RBV and BAT perspectives. The RBV has its roots in the early works of [53,54]; which emphasize the critical role of firm-specific resources and capabilities in achieving and sustaining a competitive advantage [35]. seminal work raised an important point: the resources that can bring sustainable competitive advantage to the enterprise are valuable, rare, inimitable, and non-substitute. Therefore, the RBV can explain robustly how family firm-level characteristics are unique resources that affect CSR performance. As [55] noted, firms can leverage their unique resources, such as reputation, brand, and other resources, to fulfill CSR activities. Following this logic, the firm-level characteristics of family businesses can be a source of advantage for firms to improve their CSR performance. For instance, firm size is significantly and positively related to CSR performance; larger-sized family firms tend to disclose more CSR activities [6,46]) [30]. findings indicate that corporate profitability contributes to CSR performance. From the RBV perspective, the unique characteristics of family businesses provide potential competitive advantages and thus contribute to improving CSR performance [56].

Furthermore, the present study argues that agencies such as CEOs dominate firms' CSR performance; therefore, this study employs BAT as a theoretical guide [57]. Agency issues are common because managers may adopt business actions that are difficult to observe for the principals and take advantage of information asymmetries to deceive shareholders into acquiring private interests [58]. Thus, the impetus to practice CSR may not necessarily originate from CEOs. However, the heterogeneity of family firms may challenge this notion because family members occupy a dominant position in family firms. As a result, non-economic objectives linked with the family are particularly significant [37,38]. In addition, as [45] argue, in contrast to other shareholders, family members have a greater chance to participate in the daily operations of family firms. Hence, they have more access to the information they need. Accordingly, whether family members are top executives and/or board chairmen plays a crucial role in CSR decisions in the family business. Based on RBV and BAT, this study established an ML prediction model for the CSR performance of family businesses from the firm-level and CEO-level characteristics (Fig. 1).

3. Methodology

3.1. Data collection

The data used in this study came from the China Stock Market Accounting Research (CSMAR) and Hexun databases. The CSMAR is a reliable and high-accuracy research database covering various fields, such as securities, foreign exchange, macro-economy, industry, and other major economies and finance systems. It is also among the largest databases of listed companies in China and is a major source of authentic data on the background and financial records of Chinese listed companies, which is widely used in family business research [59,60]. We collected data on publicly listed Chinese family firms from 2010 to 2020, including the firm's basic description, accounting and financial indicators, and CEO information. Similar to previous research [40,42], we applied CSR ratings released by the Hexun database every year to measure CSR merits in family businesses. To obtain relatively accurate annual CSR scores for listed companies in China, we used Hexun's assessment system as the foundation of firms' CSR and annual reports. This CSR performance assessment system of the Hexun database comprises five dimensions: shareholders, employees, supplier-customer rights, preservation of the environment, and charity. The above five first-level dimensions include 13 secondary-level indicators and 37 third-level indicators.

3.2. Dataset preparation

In its original form, the dataset included all Chinese A-share listed firms' annual CSR performance data from 2010 to 2020 from the Hexun database and relevant data from CSMAR, including information related to family firms' and CEOs' characteristics. Following existing studies, we identify a Chinese firm as a family business if the ultimate controlling shareholder is a family or an individual or if the ultimate controlling shareholder has a blood or marital relationship [61–63]. We compiled the current dataset using the following procedure. First, we matched CSR ratings with family business data using stock codes, excluding observations with missing values. This process resulted in a sample of 1, 964 Chinese family businesses (13,756 firm-year observations).

For the dependent variable, we measured CSR performance using CSR ratings. Inspired by Ref. [64]; a CSR rating with five categories was used, with E-related firms assigned a value of 1 and firms with ratings of A assigned a value of 5. Lower values represented firms with lower CSR ratings. We divided the independent variables into two main categories: family firm and CEO characteristics. Family firm characteristics included firm size, the natural logarithm of total assets, firm age, number of years since the firm's establishment, market value, the ratio

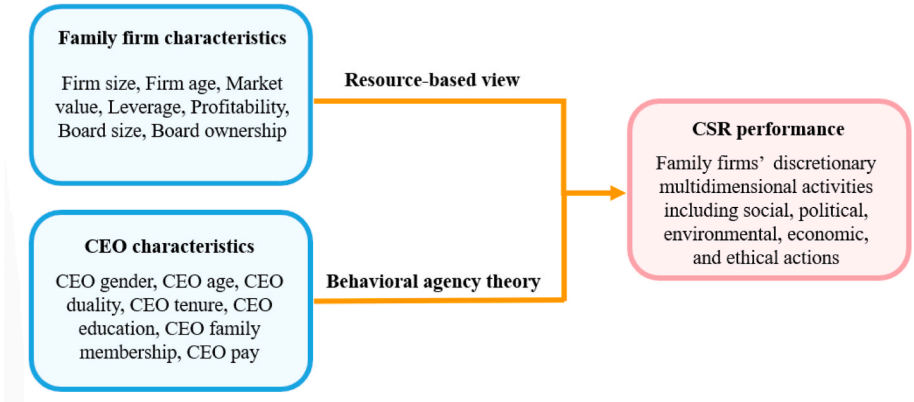

Fig. 1. Theoretical framework.

of the sum of equity market value and book value of debt to total assets (Tobin's q), leverage, the ratio of total debts to total assets, profitability, return on assets (ROA), board size, the number of directors on the board, board ownership, and the proportion of the firm's total outstanding equity held by board members.

Regarding CEO characteristics, we define CEO gender, CEO duality, CEO education, and CEO family membership as dummy variables. CEO age is the natural logarithm of the CEO's age in a given year; CEO tenure is the natural logarithm of the tenure time in years plus one; CEO pay is the natural logarithm of the CEO's total pay plus one. Detailed descriptions and sources of these variables are in Table 1.

3.3. Machine learning methods

The use of ML can not only contribute to real-world information identification and acquisition but also ameliorate the execution of tasks

Table 1
Definitions of variables.

Category	Variable	Definition	Source
CSR Performance	CSR Rating	1 = E, 2 = D, 3 = C, 4 = B, 5 = A	[64]
Family Firm Characteristics	Firm Size	The natural logarithm of total assets	[32]
	Firm Age	The number of years the firm has been established	[45]
	Market Value	Tobin's q (equity market value + book value of debt)/total assets	[65]
	Leverage	Total debt/total assets	[32]
	Profitability	ROA, net income/total assets	[51]
	Board Size	The number of directors on the board	[66]
	Board Ownership	The proportion of the firm's total outstanding equity held by the board members	[67]
CEO Characteristics	CEO Gender	Dummy variable, 0 if the CEO is male and 1 if the CEO is female	[68]
	CEO Age	The natural logarithm of CEO age in a given year	[32]
	CEO Duality	Dummy variable, 1 for both CEO and chairman and 0 otherwise	[32]
	CEO Tenure	The natural logarithm of the tenure time in years plus one.	[69]
	CEO Education	Dummy variable, 1 if a CEO holds a postgraduate degree (e.g., Masters and/or PhD) and 0 otherwise.	[70]
	CEO Family Membership	Dummy variable, which equals 1 if the CEO is a family member and 0 otherwise.	[43]
	CEO Pay	The natural logarithm of the CEO's total pay plus one	[51]

in light of the new information. This methodology is particularly useful for analyzing problems involving ambiguous functional forms [42]. Unlike traditional statistical models with assumed causality between factors (e.g., linear, non-linear), there is no need for assumptions or benchmarks in ML; it rather detects patterns of data in the real world [39,71]. This approach can help detect orders from a large amount of disordered information, especially high-dimensional datasets, and ML methods usually have a better performance than linear models [40,41]. Moreover, the various regulation and cross-validation functions in ML algorithms can help enhance inductive inference and guarantee the interpretability and replicability of the analytic results [72–74].

Based on the chosen target, researchers have used supervised learning models such as lasso regression (Lasso), ridge regression (Ridge), elastic net regression (ElasticNet), logistic regression (LR), multi-layer perceptron (MLP), k nearest neighbor (KNN), Bernoulli Naïve Bayes (BernoulliNB), support vector machine (SVM), decision tree (DT), random forest (RF), extreme gradient boosting (XGBoost), and light gradient boosting machine (LightGBM) to predict CSR performance.

Linear models make predictions by using a linear function of the input features. First, Lasso and Ridge, an extension of ordinary least squares (OLS), added a penalty term in the formulation to minimize the sum of the squared errors [75]. The difference between the two is in the regularization parameter. Specifically, the regularization parameter of Lasso is the L1 norm that penalizes the absolute values of the coefficients, whereas that of Ridge is the L2 norm that penalizes the squares of coefficients [76]. Second, ElasticNet is a linear model that integrates Lasso and Ridge, and its design reduces the loss functions [77]. The regularization parameter is calculated based on the hybrid parameter between the L1 and L2 norms. As such, by including L1 and L2 penalties, it is possible to determine their effects on the penalty term. Finally, logistic regression, or logit regression, is a linear machine model for classification [75]. This model fits well for data linearly related to the target variable requiring prediction. In this model, we employed a logistic function to calculate the probability of possible results per test. We applied these linear models to the dataset for comparison with other ML algorithms.

This study employed the MLP, one of the most commonly used supervised neural network algorithms, with input and output layers and hidden layers based on a feedforward backpropagation network [78]. The feedforward MLP unit admits only numerical data in the input layer through its nodes. Compared with other algorithms, MLP has various merits, such as the capability to learn non-linear models in real time. We tuned the hyperparameters to the most widely used set of parameters.

KNN is a non-parametric supervised learning approach used for classification based on the distance between the test point and points in the training data [79]. The basic process of KNN comprises five steps: 1) calculate the distance between the samples in the given training set and the samples to be classified; 2) rank in ascending sequence of distance; 3) select the optimal k samples that are closest to the samples to be classified; 4) obtain the appearance frequency of the category of the k samples; and 5) select the category with the highest appearance frequency of the k samples as the predicted category of the sample to be classified. Particularly in cases where the data lack uniform sampling and the object encompasses many class labels, KNN is a suitable choice.

The Naïve Bayes algorithm is a supervised learning algorithm based on the Bayesian theory, while "Naïve" has been derived from an important assumption that the features are independent [75]. Bayes is a classical method used in probability theory and the cornerstone of ML and statistical theory. There are three methods used in the Naïve Bayes algorithm: BernoulliNB, GaussianNB, and multinomialNB. This study employed the BernoulliNB to represent the Naïve Bayes algorithm for prediction. It has been popularly applied to binomial distribution problems and provided relatively better results with the current dataset.

SVM is effective for learning assignments involving multiple features; it is one of the best-performing ML algorithms in recent decades for its robust and accurate output [79]. The SVM establishes an n-dimensional separating hyperplane that can achieve optimal class separation. In addition, the characteristics of the SVM model depend on which kernel function the model uses. This study used an SVM with a polynomial kernel to transform the data points and create an optimal decision boundary because researchers widely use it in classification prediction.

DT, a hierarchical model based on a dendrogram, is an inductively discovered series of rules that visually resembles an inverted tree [75]. Recently, decision trees have been developed and improved with applications in various fields [80,81]. It is a non-parametric approach with no assumptions of normality or independence. It eliminates estimation bias caused by outliers and missing values, and the analysis results are intuitive as they are in an "if-then" structure. Moreover, researchers quickly extended DTs to analyze numeric variables and address regression issues.

RF is an important ensemble approach that employs forecasts from DT-based algorithms and relies on bagging (bootstrap aggregation; [82]. Each random tree in the forest is a basic learner, using a random subset of training points and features. Therefore, RF combines ordinary bagging with the random subspace method to reduce the variance of the model. The method constructs multiple DTs for classification forecasting, and each tree eventually votes for the predicted class. This study used the depth and number of trees, and number of covariates as the three hyper-parameters for RF and tuned them to build a prediction model.

XGBoost and LightGBM, two representative ensemble methods, are optimized distributed gradient-boosting algorithms [83]. created the former, a gradient-boosted algorithm capable of learning from parallel and distributed computing, used in various analyses, such as ranking, classification, and regression. Owing to its efficient and flexible characteristics, many researchers have widely adopted it in various fields [40,42]. XGBoost applies the second derivative, which lends more accuracy to the loss function. Regular terms in this model can help avoid the problem of overfitting. Compared to the level-wise growth in XGBoost (based on the contribution to the loss of a particular branch), the leaf-wise growth in LightGBM (splits nodes based on the contribution to global loss) is compatible with large and complex datasets while decreasing the execution time in data training [84]. Considering their efficiency, accuracy, and interpretability characteristics, these two algorithms can effectively help predict the CSR performance of family firms.

4. Results

4.1. Descriptive statistics and correlation analysis

In Table 2, we report the descriptive statistics and correlations. As

Table 2
Descriptive statistics and correlations.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
(1) CSR Rating	1.000														
(2) Firm Size	0.145*	1.000													
(3) Firm Age	−0.075*	0.183*	1.000												
(4) Market Value	−0.012	−0.148*	0.011	1.000											
(5) Leverage	−0.032*	0.037*	0.078*	0.238*	1.000										
(6) Profitability	0.073*	0.037*	−0.025*	−0.098*	−0.795*	1.000									
(7) Board Size	0.079*	0.155*	0.002	−0.041*	0.006	0.025*	1.000								
(8) Board Ownership	0.001	−0.222*	−0.270*	−0.032*	−0.114*	0.030*	−0.094*	1.000							
(9) CEO Gender	−0.006	−0.008	0.016	−0.006	−0.010	0.002	−0.059*	0.019*	1.000						
(10) CEO Age	−0.035*	0.055*	0.094*	−0.001	−0.027*	0.024*	0.011	−0.018*	−0.027*	1.000					
(11) CEO Duality	−0.038*	−0.099*	−0.093*	−0.008	−0.038*	0.012	−0.114*	0.148*	−0.069*	0.228*	1.000				
(12) CEO Tenure	0.061*	0.090*	0.011	−0.003	−0.039*	0.034*	−0.004	0.022*	−0.010	0.235*	0.173*	1.000			
(13) CEO Education	0.141*	−0.130*	−0.245*	−0.015	−0.029*	0.016	0.065*	0.045*	0.017*	−0.104*	0.018*	−0.001	1.000		
(14) CEO Family Membership	0.001	−0.106*	−0.135*	−0.026*	−0.074*	0.022*	−0.076*	0.299*	0.011	0.055*	0.597*	0.241*	0.039*	1.000	
(15) CEO Pay	0.089*	0.436*	0.172*	−0.021*	0.024*	0.025*	0.078*	−0.117*	0.005	0.099*	−0.023*	0.069*	−0.069*	−0.076*	1.000
Mean	2.065	19.730	15.740	2.636	0.398	0.030	8.218	0.207	0.078	3.873	0.378	1.257	0.128	0.486	0.098
Standard Deviation	0.485	1.133	5.941	16.690	0.712	0.557	1.481	0.214	0.270	0.150	0.485	0.768	0.334	0.500	0.094

Note: *$p < 0.05$.

the table shows, the standard deviation of the CSR rating is small (0.485). This result suggests little variation from the mean performance level (2.065) in the sample's CSR performance of Chinese family firms. Furthermore, the results show that 37.8 % of CEOs are also chairman of the board, 12.8 % hold a postgraduate degree, and 48.6 % are family members. Regarding the correlations, CSR rating has a relatively ideal and significant correlation with most variables.

4.2. Comparison of models' performance

The mean squared error (MSE), a widely used parameter that calculates the sum of squared errors between fitted data and original data, was applied to evaluate the performance of the ML models [74]. A smaller MSE represents better performance, and the MSE is defined as follows:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

where y_i is the actual data, $\hat{y}_i$ is the predicted data, and n is the total number of observations in the test set.

Table 3 presents the results of the model comparison of the 12 ML algorithms (Lasso, Ridge, ElasticNet, LR, MLP, KNN, BernoulliNB, SVM, DT, RF, XGBoost, and LightGBM). The coding of all the algorithms is in Python 3.0. Based on the scikit-learning (sklearn) package, first, we utilized the train_test_split function from the sklearn.model_selection to randomly split our dataset into two sets: one for training the model and the other for testing the predictive output. Then, we used each ML algorithm to train and test these models on the same data. Note that one training may fail to generate the most accurate algorithms, so we constructed 20 splitting schemes to identify the most suitable model for predicting CSR performance. Lastly, we used sklearn metrics to export the prediction performance (Table 3). The first column presents the splitting configurations; for example, 80:20 indicates 80 % of the data for training and 20 % as a test set. The average MSE for each model is in the last row. The results indicate that RF is the best method for predicting the CSR performance of family firms, with the smallest MSE of 0.2002 and the smallest average MSE of 0.2113, followed by Ridge (average MSE = 0.2130), LightGBM (average MSE = 0.2146), and XGBoost (average MSE = 0.2189). Therefore, RF fits well with the present data and has the best performance in predicting CSR performance.

4.3. Statistical test of ML models' performance

As shown in Table 3, RF outperformed the other 11 ML models. We applied the Friedman test to further compare the prediction performance among the ML models in the present data. The Friedman test is a method for testing whether multiple population distributions are significantly different. The null hypothesis assumed no significant difference among the 12 ML methods, while the alternative hypothesis postulated a significant difference in their comparison. The chi-squared value in the Friedman test was 187.2589, and the degree of freedom was 11. This result was significant at the p = 0.0000 level, rejecting the null hypothesis and accepting the alternative hypothesis, indicating that RF performs significantly better because of its excellent prediction ability.

4.4. Importance of factors influencing CSR performance

After confirming that RF is the most appropriate model for predicting CSR ratings, we employed it as the base algorithm to explore the factors that drive family firms to engage in CSR activities. To further interpret the current sample, we adopted the SHapley Additive exPlanations (SHAP) approach derived from the coalitional game theory [85]. introduced the approach to address the problems associated with interpreting results. The SHAP value represents a factor's contribution to the model's predictive output [40,42].

The results of the importance of the characteristics and their effects on CSR performance are shown in Figs. 2 and 3, respectively. Fig. 2 ranks all the SHAP values from high to low, where the x-axis shows the superposition of the effects of features on classes, and the y-axis shows all features. The results indicate that profitability is the most important driver of CSR performance, followed by firm size, CEO education, leverage, board ownership, CEO pay, market value, CEO tenure, CEO age, firm age, board size, CEO family membership, CEO duality, and CEO gender. This ranking determines each feature's importance in CSR performance. The results of the positive/negative effects of the features on family firms' CSR performance are in Fig. 3. Features that positively impact CSR are in red, and those that negatively impact CSR are in blue. For example, profitability (in red) indicates that increased profitability leads to a higher CSR performance in the family business. In other words, a significant and positive association exists between profitability and CSR performance. As shown in Fig. 3, profitability, board ownership, CEO age, firm age, CEO family membership, and CEO duality are positively associated with CSR performance. The effects of firm size, CEO education, leverage, market value, CEO pay, CEO tenure, board

Table 3
Performance comparison of machine learning models.

Splitting	Lasso	Ridge	ElasticNet	LR	MLP	KNN	BernoulliNB	SVM	DT	RF	XGBoost	LightGBM
80:20	0.2186	0.2098	0.2186	0.2224	0.2205	0.2169	0.2231	0.2235	0.3590	0.2071	0.2227	0.2053
78:22	0.2168	0.2078	0.2168	0.2207	0.2152	0.2161	0.2213	0.2217	0.3548	0.2042	0.2151	0.2141
76:24	0.2162	0.2087	0.2162	0.2199	0.2161	0.2162	0.2205	0.2208	0.3801	0.2002	0.2126	0.2114
74:26	0.2192	0.2123	0.2192	0.2234	0.6674	0.2230	0.2237	0.2242	0.3721	0.2136	0.2203	0.2139
72:28	0.2194	0.2114	0.2194	0.2235	0.2153	0.2264	0.2238	0.2243	0.3741	0.2121	0.2170	0.2129
70:30	0.2232	0.2149	0.2232	0.2273	0.2180	0.2309	0.2275	0.2280	0.3688	0.2137	0.2215	0.2115
68:32	0.2203	0.2119	0.2203	0.2240	0.2144	0.2274	0.2242	0.2247	0.3925	0.2085	0.2160	0.2165
66:34	0.2214	0.2126	0.2214	0.2247	0.2180	0.2304	0.2251	0.2257	0.3773	0.2103	0.2185	0.2063
64:36	0.2203	0.2111	0.2203	0.2237	0.2203	0.2302	0.2239	0.2245	0.3495	0.2067	0.2164	0.2096
62:38	0.2195	0.2100	0.2195	0.2226	0.2137	0.2284	0.2232	0.2236	0.3688	0.2072	0.2158	0.2119
60:40	0.2217	0.2118	0.2217	0.2251	0.2164	0.2319	0.2257	0.2261	0.3876	0.2092	0.2153	0.2113
58:42	0.2225	0.2122	0.2225	0.2259	0.2153	0.2310	0.2265	0.2269	0.3576	0.2085	0.2227	0.2155
56:44	0.2232	0.2128	0.2232	0.2267	0.4167	0.2306	0.2275	0.2278	0.4010	0.2098	0.2154	0.2095
54:46	0.2265	0.2161	0.2265	0.2301	0.2198	0.2339	0.2309	0.2312	0.3911	0.2143	0.2277	0.2170
52:48	0.2248	0.2147	0.2248	0.2284	0.2196	0.2320	0.2291	0.2294	0.3724	0.2120	0.2205	0.2187
50:50	0.2272	0.2164	0.2272	0.2309	0.2230	0.2357	0.2316	0.2319	0.3812	0.2197	0.2261	0.2208
48:52	0.2288	0.2171	0.2288	0.2325	0.2240	0.2367	0.2336	0.2333	0.4172	0.2203	0.2218	0.2169
46:54	0.2269	0.2150	0.2269	0.2304	0.2217	0.2353	0.2315	0.2313	0.3797	0.2170	0.2155	0.2221
44:56	0.2288	0.2167	0.2288	0.2321	0.4417	0.2362	0.2334	0.2333	0.3919	0.2173	0.2159	0.2205
42:58	0.2300	0.2167	0.2300	0.2334	0.2399	0.2386	0.2345	0.2346	0.3950	0.2141	0.2207	0.2262
Average	0.2228	0.2130	0.2228	0.2264	0.2629	0.2294	0.2270	0.2273	0.3786	0.2113	0.2189	0.2146

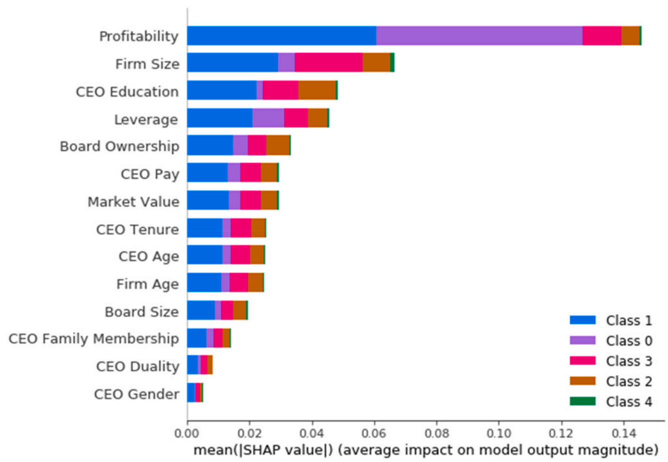

Fig. 2. Summary of SHAP values.

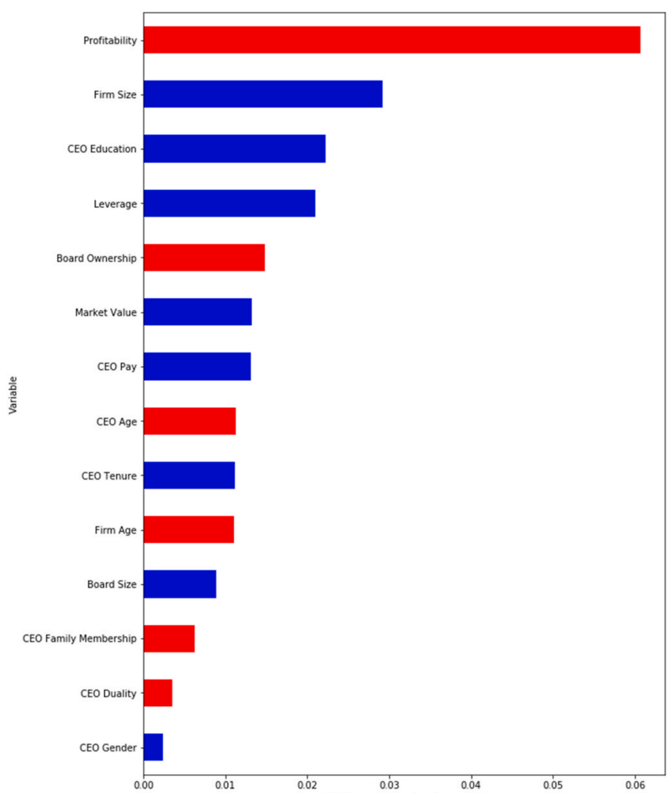

Fig. 3. SHAP values (red = positive effect, blue = negative effect).

size, and CEO gender on CSR performance are negative.

Furthermore, considering specific SHAP values of each variable, Table 4 indicates the importance and impact of family firm and CEO characteristics on CSR performance. The SHAP values of firm characteristics range from 0.0606 to 0.0089, while the SHAP values of CEO characteristics range from 0.0222 to 0.0023. Among family firm characteristics, the findings in this section suggest that a family firm's profitability positively impacts CSR performance, which is consistent with [30]. However, unlike [28], who pointed out that firm size has a positive relationship with CSR performance, our results indicate that increasing firm size decreases CSR performance. Regarding CEO characteristics, CEO age is positively related to CSR performance, supporting [32]. Our results show that the effects of CEO education and CEO tenure are negative. We discuss these results in the next section.

Table 4
The importance and impact of CSR determinants.

Category	Variable	Absolute SHAP value	Effects	Sign
Family firm characteristics	Profitability	0.0606	0.1684	Positive
	Firm Size	0.0292	−0.7057	Negative
	Leverage	0.0210	−0.1928	Negative
	Board Ownership	0.0148	0.2231	Positive
	Market Value	0.0132	−0.0289	Negative
	Firm Age	0.0110	0.5072	Positive
	Board Size	0.0089	−0.3921	Negative
CEO characteristics	CEO Education	0.0222	−0.8227	Negative
	CEO Pay	0.0131	−0.1401	Negative
	CEO Age	0.0112	0.3454	Positive
	CEO Tenure	0.0112	−0.3991	Negative
	CEO Family Membership	0.0063	0.5008	Positive
	CEO Duality	0.0035	0.3841	Positive
	CEO Gender	0.0023	−0.6830	Negative

5. Discussion

Drawing on RBV and BAT, this study developed a comprehensive prediction framework to investigate whether and how firm and CEO characteristics affect CSR performance by employing an ML approach on a large-scale sample of Chinese family firms from 2010 to 2020. Prior studies have mainly focused on the impact of board and firm characteristics on the CSR performance of family businesses [6,17,27]; [66]; [25], while this study conducted a relatively comprehensive study using ML methods on CEO-level characteristics. Concerning the widespread presence of family businesses and their key role in the world economy, as well as the fact that CSR actions contribute to the competitive advantage of any business, the present study results have several implications for the theoretical and practical application of CSR in family businesses.

5.1. Theoretical implications

This study contributes to the literature on the mechanisms linking family business characteristics and CSR in several ways. First, although prior literature has illustrated the difference between family firms and non-family firms in terms of social responsibility [22,60], little is known about whether and how various features of a family firm may impact its CSR performance. While scholars have proposed several antecedents, such as profitability and firm size [86,87], whether these findings apply to a family business research context remains unknown. The present study, from the perspective of RBV and BAT, contributes to filling this gap by proposing a detailed framework that examines various firm and CEO characteristics and analyzes their complexity concerning CSR performance. Moreover, the present study highlights the heterogeneous permutation importance and driving power of firms and CEOs to demonstrate how they promote or decrease family firms' CSR performance. Overall, the present findings identify that factors such as profitability, firm size, CEO education, leverage, board ownership, and CEO pay play an important role in family firms' CSR performance, while other factors such as CEO gender and CEO duality account for relatively less dominance over CSR performance.

Specifically, our study contributes to RBV by verifying how various firm characteristics play a role in a family firm's CSR performance. While previous studies, such as [88,89]; have used this theory to explain the determinants of CSR performance, their research factors did not address the factors discussed in this study, and their research context did not involve family firms, leaving a noticeable research gap. Therefore, we filled this gap and contributed knowledge to the RBV and CSR literature by employing an ML approach to investigate the permutation importance and heterogeneous impacts of firm characteristics on family firms' CSR performance. The present study found that profitability strongly impacts family firms' CSR performance, followed by firm age and board ownership. In contrast to prior literature that mostly investigated the impact of CSR performance on a firm's financial performance [90], the positive predictive effect of profitability and CSR performance is in line with previous studies such as that by Ref. [60]; which revealed a positive controlling effect of a firm's profit-adding ability on its CSR performance. However, the present findings partially contradict [60] research on the role of firm age. In their study, the impact of firm age varies for internal and external CSR. In contrast, the results of the present study show an overall positive relationship between firm age and CSR performance. This study also demonstrates that a higher proportion of the firm's total outstanding equity held by board members (board ownership) leads to better CSR performance in family firms. Furthermore, in descending order, firm size, leverage market value, and board size negatively impact family firms' CSR performance.

Moreover, we contribute to the BAT literature by demonstrating the permutation importance and heterogeneous impacts of CEO characteristics on family firms' CSR performance. Indeed, research on the influence of CEO characteristics on firms' visions, strategies, operations, and performance has long been a focus in the literature [69,91,92]. However, the literature has been silent on how CEO characteristics based on BAT can predict family firms' CSR performance. Hence, this study conducted an in-depth investigation of CEO characteristics in family firms from multiple dimensions. The present findings suggest that, in descending order, CEO education, CEO pay, CEO age, CEO tenure, CEO family membership, CEO duality, and CEO gender are important predictors of family firms' CSR performance. Among them, the CEO's age and duality positively affect family firms' CSR performance. In contrast, the CEO's education, pay, tenure, and gender negatively affect CSR performance.

The promoting effect of CEO age contradicts [28,32]; who argued that as the CEO's age increases, they tend to be more conservative in CSR initiatives and reduce CSR-related investments. Nevertheless, the present finding aligns with [33] argument, emphasizing older CEOs' interests in social ties and reputation. In addition, the positive impact of CEO duality on CSR performance indicates that family firms with the same CEO and chairman tend to have better CSR performance, supporting [60] argument that CEO duality leads to better social responsibility, such as environmental protection activities. While the present findings on CEO tenure and CEO pay align with recent advances [43,93,94], surprisingly, we noted a negative relationship between CEO education and CSR performance. In other words, well-educated CEOs discouraged family firms' CSR performance. This finding is inconsistent with conventional wisdom, which argues that educated CEOs have lower risk aversion and are more open-minded when making decisions [95]. The reason may be that, although CEOs care about CSR, due to the financial constraints of their firms, they tend to invest more in other areas, such as research and development (R&D) and technology innovation, which may result in a relatively small proportion of funds for developing and promoting CSR. Overall, using an ML approach, this study contributes to the literature by constructing a precise prediction model that integrates RBV and BAT to forecast the determinants of family firms' CSR performance.

5.2. Practical implications

This study has several important practical implications. As many scholars have emphasized, CSR performance can benefit companies in many ways, making it necessary for family businesses to achieve better social responsibility. First, the present findings highlight the importance of firm and CEO characteristics in predicting CSR performance, which provides recommendations for family firms' operations. For example, it has been highlighted in this study that profitability, firm size, CEO education, and leverage are the most significant factors in predicting CSR performance. This finding can enable firms to acknowledge that CSR-related investments require a certain percentage of a firm's financial capital and, by extension, can be reflected by profitability, firm size, and leverage. In addition, CEOs require a significant knowledge base regarding socially responsible activities.

Moreover, the present findings provide valuable information for investors and other potential shareholders [8]. findings indicate that poor CSR performance can hurt investors' assessments of firm value. Our findings can help investors establish a more comprehensive CSR performance evaluation system to assess corporate value. For example, managers in family firms should consider resource profitability, board ownership, and firm age as unique "resources" that may confer advantages for their CSR performance while being aware of the potential drawbacks of firm size, leverage, market value, and board size. While these expansionary activities might bring more resources to the firm, they may increase risks that could jeopardize the operations of CSR. More importantly, as RBV suggests, competitive advantage comes from resources that are unique and difficult to imitate, and firms should focus on developing unique resources rather than simply increasing financing.

In addition to firm characteristics, family firms should be attentive to CEO characteristics because of their decisive role in making major corporate decisions and strategic directions. We suggest firms seek older CEOs due to their stronger social ties and greater interest in reputation than younger ones. In addition, as our findings suggest, family members who are also CEOs and CEOs who are chairmen of boards of directors can do better in terms of CSR, so we recommend that family firms pay close attention to these characteristics. Moreover, as [96] noted, a better understanding of the impact of socially responsible investment on CSR can provide crucial feedback to non-governmental organizations (NGOs). Regarded as an important stakeholder of family businesses and key partners in implementing CSR strategy, NGOs can uniquely promote family firms to fulfill their social responsibilities. As the present findings can lead to a more comprehensive analysis of CSR investment, the results can offer useful information to NGOs and alleviate information asymmetry to help them make better decisions. In addition, this study identifies the determinants of CSR in family businesses according to importance, which is instrumental for NGOs in supervising family firms. For example, NGOs may invest more capital in family firms with high profitability and low leverage and be cautious about firms with less CEO duality. Finally, from the policymaker's perspective, governments must optimize the institutional environment and policy system to help family firms better promote CSR. For instance, governments can offer subsidies or tax privileges to firms that actively engage in socially responsible activities but lack capital.

5.3. Limitations and recommendations

Although the present study yields important findings, it has several limitations. First, although we tested many factors to predict the CSR performance of family firms, other antecedents or firm characteristics can drive CSR performance. In addition, while we integrated RBV and BAT to examine the determinants of family firms' CSR performance, we encourage future studies to employ other theories to develop their conceptual models. For instance, scholars might use the organizational information processing theory to explore how firms can achieve a fit between information processing needs and capabilities when pursuing CSR performance outcomes. Finally, scholars can use qualitative research methods such as interviews and case studies to explore the influencing characteristics of CSR in family businesses.

Funding

This work was supported by the Humanities and Social Sciences Foundation of the Ministry of Education of China [Grant number 21YJC630076].

CRediT authorship contribution statement

Feng Liu: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Validation, Visualization, Writing - original draft, Writing - review & editing. **Wanying Huang:** Conceptualization, Data curation, Formal analysis, Methodology, Writing - original draft, Writing - review & editing. **Jing Zhang:** Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Writing - original draft, Writing - review & editing. **Mingjie Fang:** Conceptualization, Methodology, Validation, Writing - original draft, Writing - review & editing, Investigation.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgements

Acknowledgement Our team gratefully acknowledges the insightful suggestions from the editors and the anonymous reviewers, which substantively improved this article. We would also like to thank the members of Star-lights Machine Learning Research Team for their comments on earlier versions of the manuscript.

References

[1] A. Awaysheh, R.A. Heron, T. Perry, J.I. Wilson, On the relation between corporate social responsibility and financial performance, Strat. Manag. J. 41 (6) (2020) 965–987.
[2] S.L. Gillan, A. Koch, L.T. Starks, Firms and social responsibility: a review of ESG and CSR research in corporate finance, J. Corp. Finance 66 (2021) 101889.
[3] B. Yuan, X. Cao, Do corporate social responsibility practices contribute to green innovation? The mediating role of green dynamic capability, Technol. Soc. 68 (2022) 101858.
[4] I. Ioannou, G. Serafeim, The impact of corporate social responsibility on investment recommendations: analysts' perceptions and shifting institutional logics, Strat. Manag. J. 36 (7) (2014) 1053–1081.
[5] M. Škare, T. Golja, The impact of government CSR supporting policies on economic growth, J. Pol. Model. 36 (3) (2014) 562–577.
[6] G. Campopiano, A. De Massis, Corporate social responsibility reporting: a content analysis in family and non-family firms, J. Bus. Ethics 129 (3) (2015) 511–534.
[7] L.S. Niehm, J. Swinney, N.J. Miller, Community social responsibility and its consequences for family business performance, J. Small Bus. Manag. 46 (3) (2008) 331–350.
[8] A. Hsu, K. Koh, S. Liu, Y.H. Tong, Corporate social responsibility and corporate disclosures: an investigation of investors' and analysts' perceptions, J. Bus. Ethics 158 (2019) 507–534.
[9] K. Koh, H. Li, Y.H. Tong, Corporate social responsibility (CSR) performance and stakeholder engagement: evidence from the quantity and quality of CSR disclosures, Corp. Soc. Responsib. Environ. Manag. 30 (2) (2023) 504–517.
[10] M. Chen, J.Z. Xiao, Y. Zhao, Confucianism, successor choice, and firm performance in family firms: evidence from China, J. Corp. Finance 69 (2021) 102023.
[11] Z. Qin, Q. Wang, Father-daughter succession in China: the conceptual framework and a case study, The Business Review, Cambridge 20 (1) (2012) 68–75.
[12] J. Tan, Institutional structure and firm social performance in transitional economies: evidence of multinational corporations in China, J. Bus. Ethics 86 (2009) 171–189.
[13] C. Provis, Business ethics, Confucianism and the different faces of ritual, J. Bus. Ethics 165 (2) (2020) 191–204.
[14] B.I. Park, P.N. Ghauri, Determinants influencing CSR practices in small and medium sized MNE subsidiaries: a stakeholder perspective, J. World Bus. 50 (1) (2015) 192–204.
[15] L. Perez-Batres, J.P. Doh, V.V. Miller, M.J. Pisani, Stakeholder pressures as determinants of CSR strategic choice: why do firms choose symbolic versus substantive self-regulatory codes of conduct? J. Bus. Ethics 110 (2012) 157–172.
[16] Y. Kim, S.Y. Kim, The influence of cultural values on perceptions of corporate social responsibility: application of Hofstede's dimensions to Korean public relations practitioners, J. Bus. Ethics 91 (2010) 485–500.
[17] P. Berrone, C. Cruz, L.R. Gomez-Mejia, M. Larraza-Kintana, Socioemotional wealth and corporate responses to institutional pressures: do family-controlled firms pollute less? Adm. Sci. Q. 55 (1) (2010) 82–113.
[18] S. Noor, A. Saeed, M.S. Baloch, M. Awais, CSR permanency, family ownership, and firm value: evidence from emerging economies, Corp. Soc. Responsib. Environ. Manag. 27 (5) (2020) 2135–2149.
[19] I.M. García-Sánchez, J. Martín-Moreno, S.A. Khan, N. Hussain, Socio-emotional wealth and corporate responses to environmental hostility: are family firms more stakeholder oriented? Bus. Strat. Environ. 30 (2) (2021) 1003–1018.
[20] L.R. Gomez-Mejia, C. Cruz, P. Berrone, J. De Castro, The bind that ties: socioemotional wealth preservation in family firms, Acad. Manag. Ann. 5 (1) (2011) 653–707.
[21] A. Venturelli, S. Principale, L. Ligorio, S. Cosma, Walking the talk in family firms. An empirical investigation of CSR communication and practices, Corp. Soc. Responsib. Environ. Manag. 28 (1) (2021) 497–510.
[22] A.P. Abeysekera, C.S. Fernando, Corporate social responsibility versus corporate shareholder responsibility: a family firm perspective, J. Corp. Finance 61 (2020) 101370.
[23] A. Yu, H.B. Ding, H.M. Chung, Corporate social responsibility performance in family and non-family firms: the perspective of socio-emotional wealth, Asian Bus. Manag. 14 (5) (2015) 383–412.
[24] P. Vazquez, Family business ethics: at the crossroads of business ethics and family business, J. Bus. Ethics 150 (3) (2018) 691–709.
[25] S. Veltri, R. Mazzotta, F.E. Rubino, Board diversity and corporate social performance: does the family firm status matter? Corp. Soc. Responsib. Environ. Manag. 28 (6) (2021) 1664–1679.
[26] T.C.T. Hou, The relationship between corporate social responsibility and sustainable financial performance: firm-level evidence from Taiwan, Corp. Soc. Responsib. Environ. Manag. 26 (1) (2019) 19–28.
[27] M. Farooq, A. Noor, M. Naeem, Does family ownership moderate the relationship between board characteristics and corporate social responsibility? Evidence from an emerging market, Asian Journal of Business Ethics 12 (2023) 71–99.
[28] W.Y. Oh, Y.K. Chang, R. Jung, Board characteristics and corporate social responsibility: does family involvement in management matter? J. Bus. Res. 103 (2019) 23–33.
[29] J.H. Block, M. Wagner, The effect of family ownership on different dimensions of corporate social responsibility: evidence from large US firms, Bus. Strat. Environ. 23 (7) (2014) 475–492.
[30] J.J. Cordeiro, M. Tewari, Firm characteristics, industry context, and investor reactions to environmental CSR: a stakeholder theory approach, J. Bus. Ethics 130 (4) (2015) 833–849.
[31] L. Scholes, M. Hughes, M. Wright, A. De Massis, J. Kotlar, Family management and family guardianship: governance effects on family firm innovation strategy, Journal of Family Business Strategy 12 (4) (2021) 100389.
[32] O. Meier, G. Schier, CSR and family CEO: the moderating role of CEO's age, J. Bus. Ethics 174 (3) (2021) 595–612.
[33] M. Fabrizi, C. Mallin, G. Michelon, The role of CEO's personal incentives in driving corporate social responsibility, J. Bus. Ethics 124 (2) (2014) 311–326.
[34] F. Malik, F. Wang, M.A. Naseem, A. Ikram, S. Ali, Determinants of corporate social responsibility related to CEO attributes: an empirical study, Sage Open 10 (1) (2020) 2158244019899093.
[35] J. Barney, Firm resources and sustained competitive advantage, J. Manag. 17 (1) (1991) 99–120.
[36] L. Madden, A. McMillan, O. Harris, Drivers of selectivity in family firms: understanding the impact of age and ownership on CSR, Journal of Family Business Strategy 11 (2) (2020) 100335.
[37] E.N. Lim, M.H. Lubatkin, R.M. Wiseman, A family firm variant of the behavioral agency theory, Strateg. Entrep. J. 4 (3) (2010) 197–211.
[38] T.M. Zellweger, F.W. Kellermanns, J.J. Chrisman, J.H. Chua, Family control and family firm valuation by family CEOs: the importance of intentions for transgenerational control, Organ. Sci. 23 (3) (2012) 851–868.
[39] Y.R. Shrestha, V.F. He, P. Puranam, G. von Krogh, Algorithm supported induction for building theory: how can we use prediction models to theorize? Organ. Sci. 32 (3) (2021) 856–880.
[40] F. Liu, R. Wang, M. Fang, Mapping green innovation with machine learning: evidence from China, Technol. Forecast. Soc. Change 200 (2024) 123107.
[41] M. Wang, Y. Yu, F. Liu, Does Digital Transformation Curb the Formation of Zombie Firms? A Machine Learning Approach, Technology Analysis & Strategic Management, 2023, https://doi.org/10.1080/09537325.2023.2296007. Ahead of Print.
[42] J. Zhang, M. Zhu, F. Liu, Find who is doing social good: using machine learning to predict corporate social responsibility performance, Operations Management Research (2023) 1–14.
[43] V. Cui, S. Ding, M. Liu, Z. Wu, Revisiting the effect of family involvement on corporate social responsibility: a behavioral agency perspective, J. Bus. Ethics 152 (1) (2018) 291–309.
[44] E. López-González, J. Martínez-Ferrero, E. García-Meca, Does corporate social responsibility affect tax avoidance: evidence from family firms, Corp. Soc. Responsib. Environ. Manag. 26 (4) (2019) 819–831.
[45] F. Shahzad, I.U. Rehman, F. Nawaz, N. Nawab, Does family control explain why corporate social responsibility affects investment efficiency? Corp. Soc. Responsib. Environ. Manag. 25 (5) (2018) 880–888.
[46] M.M. Mariani, K. Al-Sultan, A. De Massis, Corporate social responsibility in family firms: a systematic literature review, J. Small Bus. Manag. (2021) 1–55.
[47] M.D.L.C.D. Déniz, M.K.C. Suárez, Corporate social responsibility and family business in Spain, J. Bus. Ethics 56 (1) (2005) 27–41.

[48] L. Rodríguez-Ariza, B. Cuadrado-Ballesteros, J. Martínez-Ferrero, I.M. García-Sánchez, The role of female directors in promoting CSR practices: an international comparison between family and non-family businesses, Bus. Ethics Eur. Rev. 26 (2) (2017) 162–174.
[49] L.M. Uhlaner, H.A. van Goor-Balk, E. Masurel, Family business and corporate social responsibility in a sample of Dutch firms, J. Small Bus. Enterprise Dev. 11 (2) (2004) 186–194.
[50] R. Nason, A. Mazzelli, M. Carney, The ties that unbind: socialization and business-owning family reference point shift, Acad. Manag. Rev. 44 (4) (2019) 846–870.
[51] H. Gao, K. Li, A comparison of CEO pay–performance sensitivity in privately-held and public firms, J. Corp. Finance 35 (2015) 370–388.
[52] P.B. McGuinness, J.P. Vieito, M. Wang, The role of board gender and foreign ownership in the CSR performance of Chinese listed firms, J. Corp. Finance 42 (2017) 75–99.
[53] E.T. Penrose, The Theory of the Growth of the Firm, Oxford University Press, 1959.
[54] B. Wernerfelt, A resource-based view of the firm, Strat. Manag. J. 5 (2) (1984) 171–180.
[55] J. Yang, K. Basile, Communicating corporate social responsibility: external stakeholder involvement, productivity and firm performance, J. Bus. Ethics (2021) 1–17.
[56] T.G. Habbershon, M.L. Williams, A resource-based framework for assessing the strategic advantages of family firms, Fam. Bus. Rev. 12 (1) (1999) 1–25.
[57] A. Pepper, J. Gore, Behavioral agency theory: new foundations for theorizing about executive compensation, J. Manag. 41 (4) (2015) 1045–1068.
[58] J. Martínez-Ferrero, L. Rodríguez-Ariza, I.M. García-Sánchez, B. Cuadrado-Ballesteros, Corporate social responsibility disclosure and information asymmetry: the role of family ownership, Review of Managerial Science 12 (4) (2018) 885–916.
[59] R. Wang, P.P. Heugens, F. Wijen, Green by affiliation? Ownership identity and environmental management system adoption in Chinese business groups, J. Manag. (2023), https://doi.org/10.1177/01492063221143721. Ahead of Print.
[60] Y. Ye, K. Li, Impact of family involvement on internal and external corporate social responsibilities: evidence from Chinese publicly listed firms, Corp. Soc. Responsib. Environ. Manag. 28 (1) (2021) 352–365.
[61] S. Ding, B. Qu, Z. Zhuang, Accounting properties of Chinese family firms, J. Account. Audit Finance 26 (4) (2011) 623–640.
[62] H. Huang, E. Lee, C. Lyu, Y. Zhao, Bequest motive, information transparency, and family firm value: a natural experiment, J. Corp. Finance 65 (2020) 101751.
[63] R. Labelle, T. Hafsi, C. Francoeur, W. Ben Amar, Family firms' corporate social performance: a calculated quest for socioemotional wealth, J. Bus. Ethics 148 (2018) 511–525.
[64] H. Liang, L. Renneboog, On the foundations of corporate social responsibility, J. Finance 72 (2) (2017) 853–910.
[65] C. Lin, D. Su, Industrial diversification, partial privatization and firm valuation: evidence from publicly listed firms in China, J. Corp. Finance 14 (4) (2008) 405–417.
[66] J.K. Kang, J. Kim, Do family firms invest more than non-family firms in employee-friendly policies? Manag. Sci. 66 (3) (2020) 1300–1324.
[67] S.P. Ferris, M. Jagannathan, A.C. Pritchard, Too busy to mind the business? Monitoring by directors with multiple board appointments, J. Finance 58 (3) (2003) 1087–1111.
[68] Y. Yuan, G. Tian, L.Y. Lu, Y. Yu, CEO ability and corporate social responsibility, J. Bus. Ethics 157 (2) (2019) 391–411.
[69] P. Xu, X. Xu, G. Bai, Corporate environmental responsibility, CEO's tenure and innovation legitimacy: evidence from Chinese listed companies, Technol. Soc. 70 (2022) 102056.
[70] H. Farag, C. Mallin, The influence of CEO demographic characteristics on corporate risk-taking: evidence from Chinese IPOs, Eur. J. Finance 24 (16) (2018) 1528–1551.
[71] S. Zhang, J. Luo, S. Wang, F. Liu, Oil price forecasting: A hybrid GRU neural network based on decomposition–reconstruction methods, Expert Systems with Applications 218 (2023) 119617.
[72] C. Liu, Y. Chen, S. Huang, X. Chen, F. Liu, Assessing the determinants of corporate risk-taking using machine learning algorithms, Systems 11 (5) (2023) 263.
[73] C. Liu, Y. Li, M. Fang, F. Liu, Using machine learning to explore the determinants of service satisfaction with online healthcare platforms during the COVID-19 pandemic, Service Business 17 (2023) 449–476.
[74] F. Liu, X. Long, L. Dong, M. Fang, What makes you entrepreneurial? Using machine learning to investigate the determinants of entrepreneurship in China, China Econ. Rev. 81 (2023) 102029.
[75] M. Kuhn, K. Johnson, Applied Predictive Modeling, Springer, New York, 2013.
[76] J. Friedman, T. Hastie, R. Tibshirani, Regularization paths for generalized linear models via coordinate descent, J. Stat. Software 33 (1) (2010) 1–22.
[77] H. Zou, T. Hastie, Regularization and variable selection via the elastic net, J. Roy. Stat. Soc. B Stat. Methodol. 67 (2) (2005) 301–320.
[78] S. Haykin, Neural Networks and Learning Machines, Pearson Education, Upper Saddle River, 2009.
[79] X. Wu, V. Kumar, J. Ross Quinlan, J. Ghosh, Q. Yang, H. Motoda, D. Steinberg, Top 10 algorithms in data mining, Knowl. Inf. Syst. 14 (2008) 1–37.
[80] G. Mistikoglu, I.H. Gerek, E. Erdis, P.M. Usmen, H. Cakan, E.E. Kazan, Decision tree analysis of construction fall accidents involving roofers, Expert Syst. Appl. 42 (4) (2015) 2256–2263.
[81] S. Nijs, N. Dries, V. Van Vlasselaer, L. Sels, Reframing talent identification as a status-organising process: examining talent hierarchies through data mining, Hum. Resour. Manag. J. 32 (1) (2022) 169–193.
[82] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32.
[83] T. Chen, C. Guestrin, Xgboost: a scalable tree boosting system, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 785–794.
[84] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, T.Y. Liu, LightGBM: a highly efficient gradient boosting decision tree, Adv. Neural Inf. Process. Syst. (2017) 30.
[85] S.M. Lundberg, S.I. Lee, A unified approach to interpreting model predictions, Adv. Neural Inf. Process. Syst. 30 (2017) 4768–4777.
[86] C. Reverte, Determinants of corporate social responsibility disclosure ratings by Spanish listed firms, J. Bus. Ethics 88 (2) (2009) 351–366.
[87] J. Surroca, J.A. Tribó, S. Waddock, Corporate responsibility and financial performance: the role of intangible resources, Strat. Manag. J. 31 (5) (2010) 463–490.
[88] E. Battisti, N. Nirino, E. Leonidou, A. Thrassou, Corporate venture capital and CSR performance: an extended resource based view's perspective, J. Bus. Res. 139 (2022) 1058–1066.
[89] A. Shaukat, Y. Qiu, G. Trojanowski, Board attributes, corporate social responsibility strategy, and corporate environmental and social performance, J. Bus. Ethics 135 (2016) 569–585.
[90] X. Partalidou, E. Zafeiriou, G. Giannarakis, N. Sariannidis, The effect of corporate social responsibility performance on financial performance: the case of food industry, Benchmark Int. J. 27 (10) (2020) 2701–2720.
[91] M.H. Manner, The impact of CEO characteristics on corporate social performance, J. Bus. Ethics 93 (2010) 53–72.
[92] T. Nawaz, What's in an education? Implications of CEO education for financial inclusion, Int. J. Finance Econ. 27 (3) (2022) 3741–3753.
[93] R. Borghesi, J.F. Houston, A. Naranjo, Corporate socially responsible investments: CEO altruism, reputation, and shareholder interests, J. Corp. Finance 26 (2014) 164–181.
[94] W. Chen, G. Zhou, X. Zhu, CEO tenure and corporate social responsibility performance, J. Bus. Res. 95 (2019) 292–302.
[95] R. Orens, A. Reheul, Do CEO demographics explain cash holdings in SMEs? Eur. Manag. J. 31 (6) (2013) 549–563.
[96] T. Guay, J.P. Doh, G. Sinclair, Non-governmental organizations, shareholder activism, and socially responsible investments: ethical, strategic, and governance implications, J. Bus. Ethics 52 (1) (2004) 125–139.
[97] G.B. Derchi, L. Zoni, A. Dossi, Corporate social responsibility performance, incentives, and learning effects, J. Bus. Ethics 173 (2021) 617–641.