



Article

Machine learning-driven pay gap analysis: predicting corporate innovation performance using XGBoost and SHAP interpretability

Qunyu Zhao*, Dayao Zhou

School of Economics and Management, Guizhou Qiannan College of Science and Technology, China

ARTICLE INFO

Article history:

Received 20 November 2025

Received in revised form

10 February 2026

Accepted 16 March 2026

Keywords:

Pay gap, Innovation performance, XGBoost, SHAP, Tournament theory, Social comparison theory

*Corresponding author

Email address:

leodaley2333@163.com

DOI: 10.55670/fpll.futech.5.2.24

ABSTRACT

The relationship between pay gap and corporate innovation has been the focus of significant theoretical discussion, with tournament theory and social comparison theory generating contrasting predictions. Traditional linear methods are ill-suited to capturing the nonlinear nature of the relationship. This study proposes an XGBoost-SHAP approach to predict innovation performance using a sample of 26,815 firm-year observations from Chinese A-share-listed firms from 2010 to 2023. The results show that the XGBoost model achieves an R2 of 0.382, which is 65.4% higher than OLS (R2=0.231). The SHAP value analysis indicates that the vertical pay gap ranks as the third most important factor, following firm size and firm age. The SHAP dependence plot shows an inverted U-shaped relationship between the vertical pay gap and innovation performance, with a turning point at approximately 8.7 times. The heterogeneity analysis indicates that state-owned enterprises attain their turning point earlier (7.2 times) than non-state-owned enterprises (10.1 times), suggesting that employees are more responsive to pay inequality. These findings provide practical insights that may guide managers in designing their firms' compensation schemes. Firms that fall below the threshold may consider expanding their pay gaps, while those that fall above may consider compressing their pay gaps. This XGBoost-SHAP approach translates statistical evidence into practical diagnostic tools that managers may use to assess the optimality of their firms' compensation schemes in supporting innovation.

1. Introduction

On average, the compensation package for the CEOs of the S&P 500 companies in 2024 is \$18.9 million, 285 times that of the ordinary employees, a figure that rose from a mere 20 times in 1965 [1]. The compensation package for the new Starbucks CEO is \$95.8 million, or 6,666 times the median employee salary [2]. An ordinary employee would need to work from 1740 to the present to earn what he makes in a single year. Is the huge disparity between the compensation packages of CEOs and ordinary employees the driving force behind organizational innovation or the poison that kills organizational cohesion? The answer is unclear. According to tournament theory, the large disparity between CEOs' and ordinary employees' compensation packages drives employees to innovate to reach the top [3]. In contrast, social comparison theory suggests that individuals are not purely rational economic actors; they naturally compare their outcomes to those of others [4]. Perceptions of distributive injustice can lead to passive resistance or even sabotage as a means to restore equity [5]. These two mechanisms operate simultaneously within every organization, and their net effect

depends on their relative strength. Empirical research based on 12,052 firm-year observations reveals an inverted U-shaped relationship between the CEO-employee pay gap and innovation performance, indicating a "sweet spot" that balances motivational and detrimental effects [6]. Data from the pharmaceutical industry support this inverted U-shaped relationship between the CEO-employee pay gap and innovation [7], and fuzzy-set qualitative comparative analysis yields comparable conclusions [8]. The discovery of the inverted U-shaped relationship itself reveals the limitations of traditional methods. The linear regression model can estimate the average effect but cannot determine the positions of the turning points. Similarly, while they can test the statistical significance, they are ill-equipped to forecast the optimal pay gap for a specific firm. In the presence of complex threshold effects and interactions among variables, models based on linear assumptions are like trying to measure curves with a straight ruler. The XGBoost algorithm uses the iterative process of constructing decision trees to approximate arbitrary complex functions without requiring a pre-specified mathematical form for variable relationships

[9]. It has demonstrated a high level of accuracy, much better than traditional models, in the tasks of financial risk early warning [10] and ESG rating prediction [11]. However, high predictive accuracy does not guarantee interpretability of the underlying variables. The "black box" problem in machine learning means that decision-makers still do not know the underlying reasons for a prediction. The SHAP method uses Shapley values from the theory of cooperative games to address the black-box problem of the XGBoost model [12]. While these techniques have been applied in other corporate finance domains, the pay gap-innovation relationship presents a unique challenge: managers do not need to know whether pay gaps "on average" promote innovation, but rather whether their specific firm's pay gap is optimally calibrated. The traditional regression approach provides population-level estimates, which are of little use to firms operating in different competitive environments, with varying ownership structures and cultural settings. The combination of XGBoost's ability to account for firm-specific conditional effects and SHAP's ability to provide predictions based on feature contributions makes this approach highly suitable for providing firm-level diagnostics from aggregated statistical models.

This study aims to develop a new analytical framework based on the XGBoost-SHAP model to address the following research inquiries: (1) Can the pay gap function as an effective predictor for corporate innovation performance? (2) What is the extent of the pay gap's contribution as a predictor for corporate innovation performance compared to other factors? (3) Is there a turning point in the marginal effect of the pay gap for corporate innovation performance, and what is its location?

2. Problem analysis and related work

2.1 The double-edged sword effect of pay incentives

The discourse on the pay gap is often characterized by a contest between two presumptions about human behavior. On the one hand, tournament theory posits that individuals behave as rational actors who aim to maximize utility, while, on the other hand, social comparison theory presumes that individuals are social beings who respond to the idea of fairness. An essential challenge in the discourse has been that the two theories are not mutually exclusive, as they are presumed to operate simultaneously. The relative importance of these two mechanisms varies depending on the size of the pay disparity. When the pay gap is moderate, tournament incentives dominate, and employees perceive promotion rewards as attainable, thus increasing effort to compete for promotion. As pay gaps widen beyond a certain threshold, social comparison effects become dominant, and employees shift their focus from the desire for rewards to the perception of unfairness, triggering psychological withdrawal. Consequently, the net effect on innovation depends on which mechanism dominates at a given level of pay disparity.

Empirical studies prove the existence of the tournament effect. The pay gap between executives and other employees may encourage innovation by influencing team stability. Executives who know they will be highly rewarded for promotion are more likely to stay in their current position and enhance their qualifications [13]. The tournament effect also encourages executives to prove their abilities to outside stakeholders, thus enhancing the media image of the company [14]. However, this also has a negative effect on the company: although it increases the executive's probability of promotion, the promotion will come at the cost of the firm's future success, as the executive will overinvest in projects

that yield quick returns but are of questionable long-term value [15]. This short-termism can be understood through the lens of temporal discounting: when the rewards for promotion are immediate and substantial while the costs of failed long-term projects are distant and diffuse, executives rationally discount future consequences in favor of near-term gains. Innovation is driven by the ability to think long-term, and if incentive structures encourage executives to think only in the short term, breakthrough innovation is hindered.

The data from listed companies in Indonesia show a peculiar relationship whereby the vertical pay gap is positively correlated with patent quantity but negatively correlated with patent citations [16]. A larger pay gap is associated with a greater willingness among R&D staff to produce more patents to prove their competence, but, in turn, patent quality decreases. This is because the employees have adopted a "quantity for quality" approach to compensate for the perceived unfairness of the wage gap. The horizontal wage gap has an overall negative effect on innovation by affecting the basic trust required for interdepartmental collaboration. The presence of similar trends in the Chinese context, with its significantly different environmental conditions and cultural perspectives on inequality, is an empirical issue to be addressed here.

Several other contextual factors also moderate the relationship between pay dispersion and innovation. Government regulations on executive compensation in SOEs, although intended to reduce income inequality, may actually stifle innovation. However, the impact of the CEO's career horizon and shareholding structure may offset this potential drawback [17]. The increase in the remuneration gap among high-ESG-rated firms is mainly attributable to the rise in the pay-performance sensitivity of top executives. This is particularly true for non-state-owned and larger firms [18]. Other research has found that external remuneration gaps have an innovative effect if the executive compensation is higher than the industry average, while the opposite is true if the gap is wider and the compensation is lower than the industry average [19]. This change in the direction of the effect can only lead to an unclear average in traditional linear approaches, which do not show the critical points.

2.2 The introduction of machine learning methods

An analysis of panel data on S&P 1500 constituents over 25 years shows the limitations of traditional approaches: the relationship between the CEO-employee pay gap and productivity is nonlinear and cubic, varying by firm type. High-salary large firms can still achieve positive returns even with substantial pay gaps, while low-salary, labor-intensive small firms are highly sensitive [20]. While polynomial regression can capture such nonlinearity in principle, pre-specifying uniform quadratic or cubic terms across all firms fails to accommodate the full heterogeneity in threshold locations and curve shapes. The field of pay-performance research needs a methodological paradigm shift, as traditional studies have been overly fixated on "average treatment effects" while neglecting the heterogeneous distribution of effects, and overly concerned with statistical significance while neglecting predictive accuracy [21].

Ensemble learning methods based on decision trees offer a way forward. Such methods do not require pre-specification of functional forms among variables but let the data "speak for themselves." When gradient boosting is applied to predicting corporate green innovation performance, it not only significantly outperforms linear models in accuracy but also reveals an inverted U-shaped effect of environmental tax

on green innovation [10]. If only linear models were used for estimation, this non-monotonic effect would be drowned in an insignificant average effect. XGBoost stands out among algorithms because it simultaneously utilizes first- and second-order derivative information of the loss function for Taylor expansion, achieving a more precise search for optimal split points, while regularization terms effectively control model complexity [9]. The computational efficiency of LightGBM is improved by the Gradient-based One-Side Sampling and Exclusive Feature Bundling techniques [22].

These techniques have been used extensively within corporate finance. Gradient Boosting Machines achieve an accuracy of 96% in recognizing 72 out of 75 firms that are specifically treated while analyzing the financial condition of companies listed on the ChiNext market [23]. Random Forest and CatBoost show significant performance over linear regression in predicting corporate innovation capacity [24]. XGBoost shows better performance in classifying firms according to performance categories (high, medium, and low) [25]. A Random Forest model for salary prediction in the IT industry shows an accuracy of 99.6%, indicating that experience, education, and skill are the main drivers, while the effect of gender disappears almost completely after controlling for other variables [26].

2.3 Pathways to interpretability

The ability of machine learning models to make accurate predictions often comes at the cost of interpretability. In the context of financial decision-making, interpretability has shifted from "nice to have" to "must have," where the decisions taken by the algorithm should be interpretable, and the stakeholders should be provided with actionable information rather than the mysterious numbers of the prediction model [27, 28].

The SHAP method is theoretically elegant because it is rooted in the theory of cooperative games. In cooperative games, the Shapley value is the profit allocation that is unique for the entire class of games satisfying the following four axioms: efficiency, symmetry, dummy, and additivity. In the context of machine learning interpretation, each feature is a "player," and the SHAP values represent the average marginal contribution of each feature across all possible coalitions of the features [12]. Feature importance plots, dependence plots, and force plots can intuitively present analytical results [29]. The "feature packing" method proposed in the medical field is also highly relevant for providing more intuitive results for the SHAP values, which is equally important for the context of high-dimensional data in the corporate world [30].

The combined use of XGBoost and SHAP analysis has been extensively and successfully applied to predict various corporate behaviors. In the context of prediction for financialization of corporations, ESG greenwashing, and credit risk prediction, this method not only achieves excellent results in terms of prediction accuracy, but more importantly, the SHAP dependence plot is capable of detecting nonlinear threshold effects that are difficult for other methods to detect [31-33]. This analytical capability represents a leap from "average effects" to "conditional effects," from "statistical significance" to "individual heterogeneity." This progression aligns with the growing focus on conditional average treatment effects (CATE) in the causal inference literature [34]. For compensation system design, what managers truly need is not "pay gap on average promotes innovation," but rather "what level should the pay gap be set at for my firm?" Based on the theoretical analysis above, this study proposes the following hypotheses:

H1: There is an inverted U-shaped relationship between the vertical pay gap and corporate innovation performance.

H2: The turning point of the inverted U-shaped relationship is lower for state-owned enterprises than for non-state-owned enterprises.

3. Methodology

3.1 Data sources and sample selection

The study is based on Chinese A-share listed companies from 2010 to 2023, and the data is derived from the China Stock Market & Accounting Research (CSMAR) database. Patent data are also obtained from CSMAR, which compiles patent information originally disclosed by the China National Intellectual Property Administration (CNIPA). The reason for choosing 2010 as the initial year is twofold. First, although the new enterprise accounting standards implemented in 2007 required listed companies to disclose their R&D expenses, disclosure quality was inconsistent during the first two years after implementation, with quality improving significantly after 2010. Second, the global financial crisis of 2008 had a severe impact on the compensation system and innovation investment. Therefore, the study starts from 2010 to exclude the influence of the abnormal period. The end year is 2023, representing the most recent complete annual report available in the database.

The criteria for the sample screening are as follows: (1) excluding financial and insurance companies, since their asset structure and compensation system are quite different from general enterprises; (2) excluding ST, *ST, and delisted companies, since these companies are in abnormal financial status and their compensation system may be influenced by irregular factors; (3) excluding the sample with missing values for key variables such as executive compensation, employees, and patents; and (4) excluding the sample with fewer than 50 employees to avoid extreme biases for the calculation of the pay gap for micro-enterprises. This threshold is used to address distortions that can occur in pay structures at firms with smaller employee counts, where the 'average employee pay' can be skewed by a few individuals, affecting the accuracy of the vertical pay gap calculation. In addition, the sample consists of 26,815 firm-year observations from 3,126 listed companies.

Table 1 shows the distribution of the sample by industry. Manufacturing companies account for the largest share at 63.24%, consistent with the industry distribution of the entire Chinese A-share market. Information technology companies account for 9.22%, a figure in line with the increasing pace of listings in this industry in recent years. Traditional sectors such as wholesale and retail trade, construction, and mining also constitute notable proportions. From a time-series perspective, the sample size has grown progressively from 1,287 in 2010 to 2,416 in 2023, reflecting an average annual growth rate of 4.9%. This growth rate is consistent with the overall expansion trend of the Chinese A-share market during the same period.

3.2 Variable definitions and measurement

Corporate innovation performance is typically measured from both input and output dimensions. This study adopts patent application counts as the primary indicator for the following reasons: compared to R&D investment, patent output better reflects the actual effectiveness of innovation activities; compared to patent grants, patent applications are closer to the timing of innovation decisions with less lag [6]. In the calculation, the natural logarithm is taken after adding

1 to the total patent applications to mitigate the right-skewed distribution:

$$Innovation = \ln(Patent_{apply} + 1) \tag{1}$$

where $Patent_{apply}$ represents the firm's total patent applications in the current year, including invention patents, utility models, and design patents. In robustness tests, invention patent applications and R&D intensity (R&D expenditure/operating revenue) will be used as alternative indicators.

Table 1. Sample industry distribution

Industry Classification	Sample Size	Proportion (%)
Manufacturing	16,958	63.24
Information Transmission, Software and IT Services	2,473	9.22
Wholesale and Retail Trade	1,502	5.60
Construction	1,139	4.25
Mining	897	3.35
Electricity, Heat, Gas and Water Production and Supply	863	3.22
Real Estate	791	2.95
Transportation, Warehousing and Postal Services	668	2.49
Other Industries	1,524	5.68
Total	26,815	100.00

Note: Industry classification is based on the "Guidelines for Industry Classification of Listed Companies" revised by the CSRC in 2012.

Drawing on existing research practices, this study characterizes internal pay gaps from three dimensions [8, 16]. Vertical pay gap measures the pay multiple between executives and ordinary employees:

$$Gap_{vertical} = \frac{AvgPay_{exec}}{AvgPay_{emp}} \tag{2}$$

where $AvgPay_{exec}$ is the average compensation of directors, supervisors, and senior managers, and $AvgPay_{emp}$ is the average compensation of ordinary employees, calculated as (total employee compensation payable – total executive compensation) / (total employees – number of executives). The top management team internal gap measures the degree of pay differentiation within the executive team, using the ratio of the mean compensation of the top three executives to the average compensation of all executives:

$$Gap_{TMT} = \frac{Top3Pay/3}{AvgPay_{exec}} \tag{3}$$

Pay gap coefficient of variation characterizes the unevenness of pay distribution from a dispersion perspective, calculated as the coefficient of variation of executive compensation (standard deviation/mean). Referring to mainstream literature in the compensation and innovation field, this study incorporates the following control variables: firm size (Size) measured by the natural logarithm of total assets; leverage (Lev) reflecting financial leverage level; profitability (ROA) calculated as net profit/total assets; growth (Growth) using operating revenue growth rate;

ownership concentration (Top1) as the shareholding proportion of the largest shareholder; CEO duality (Dual) as a dummy variable equal to 1 when the chairman and general manager are the same person; firm age (Age) expressed as the natural logarithm of listing years; ownership type (SOE) as a dummy variable equal to 1 for state-owned enterprises. In addition, industry and year dummy variables are included to control for fixed effects. Table 2 describes the definitions of all variables and their computation methods.

Table 2. Variable definitions

Variable Type	Variable Name	Symbol	Calculation Method
Dependent Variable	Innovation Performance	Innovation	$\ln(\text{Patent applications} + 1)$
Core Independent Variables	Vertical Pay Gap	Gap_vertical	Average executive pay / Average employee pay
	TMT Internal Gap	Gap_TMT	(Top 3 executive pay / 3) / Average executive pay
	Pay Gap Coefficient of Variation	Gap_CV	Executive pay std. dev. / Executive pay mean
Control Variables	Firm Size	Size	$\ln(\text{Total assets})$
	Leverage	Lev	Total liabilities / Total assets
	Profitability	ROA	Net profit / Total assets
	Growth	Growth	(Current revenue – Prior revenue) / Prior revenue
	Ownership Concentration	Top1	Largest shareholder's shareholding proportion
	CEO Duality	Dual	Chairman also serves as GM = 1, otherwise = 0
	Firm Age	Age	$\ln(\text{Listing years})$
	Ownership Type	SOE	State-owned = 1, otherwise = 0

To control for the influence of extreme values, all continuous variables are winsorized at the 1st and 99th percentiles. For missing value treatment, variables with missing proportions below 5% are filled with median values, while observations with missing proportions exceeding 5% are directly excluded. Categorical variables (industry, year) are converted to dummy variables through one-hot encoding. All continuous features are standardized before entering machine learning models, with mean of 0 and standard deviation of 1, to eliminate the impact of scale differences on model training.

3.3 Prediction model construction

XGBoost (eXtreme Gradient Boosting) is an efficient implementation of the gradient boosting framework, with the core idea of iteratively adding decision trees to progressively approximate the objective function [9]. Unlike traditional gradient boosting methods that only utilize first-order derivative information, XGBoost simultaneously uses both first-order derivatives (gradients) and second-order derivatives (Hessian matrix) of the loss function for Taylor expansion, thereby achieving more precise search for optimal split points. Assuming the dataset contains n samples and m

features, XGBoost aims to learn an additive model of K decision trees:

$$\hat{y}_i = \sum_k 1^k f_k(x_i), \quad f_k \in \mathcal{F} \tag{4}$$

where \mathcal{F} represents the space of all possible regression trees, and f_k represents the prediction function of the k -th tree. The model's objective function consists of two parts: the loss function and the regularization term:

$$\mathcal{L} = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_k 1^k \Omega(f_k) \tag{5}$$

where $l(y_i, \hat{y}_i)$ is the loss function measuring the deviation between predicted and actual values; $\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \sum_j 1^T w_j^2$ is the regularization term, with T being the number of leaf nodes, w_j the weight of the j -th leaf node, and γ and λ controlling the penalty strength on leaf count and weights, respectively. The introduction of regularization effectively suppresses the model's tendency to overfit. In the t -th iteration, the model finds the optimal tree structure by performing a second-order Taylor expansion on the loss function:

$$\mathcal{L}^{(t)} \approx \sum_{i=1}^n \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) \tag{6}$$

where $g_i = \partial_{\hat{y}_i(t-1)} l(y_i, \hat{y}_i^{(t-1)})$ is the first-order derivative and $h_i = \partial_{\hat{y}_i(t-1)}^2 l(y_i, \hat{y}_i^{(t-1)})$ is the second-order derivative. This second-order approximation enables XGBoost to optimize complex non-convex objective functions more efficiently.

To assess the predictive capability of the XGBoost model, this study proposes three models for comparison. Ordinary Least Squares Regression (OLS) is used to represent conventional models, and the purpose of this model is to determine to what extent machine learning models are superior to conventional economic models. Random Forest uses a voting or averaging process, whereas XGBoost uses a sequential learning process, also known as boosting. LightGBM is another efficient gradient boosting library that uses Gradient-based One-Side Sampling and Exclusive Feature Bundling to improve computational efficiency [22]. The above discussion provides a framework for understanding differences in prediction results across various gradient boosting algorithms. The performance of XGBoost prediction is significantly influenced by hyperparameters. In this study, a grid search with five-fold cross-validation is utilized for hyperparameter tuning, and the specific search space is shown in Table 3. The optimal combination of hyperparameters is then used to fit the model on the entire dataset.

3.4 Model evaluation and interpretability analysis

To ensure the reliability of the evaluation results, the current study uses stratified sampling to split the dataset into training and test sets while maintaining the industry and year proportions. During training, five-fold cross-validation is used to evaluate the model's ability to generalize. In five-fold cross-validation, the training set is randomly divided into five equal-sized subsets, four of which are used to train the model and the remaining one is used to validate it, repeating the process five times. This can effectively eliminate bias in the evaluation results caused by random data splitting.

To evaluate the innovation performance prediction model as a continuous variable, three metrics are used. Root Mean Square Error (RMSE) measures the average magnitude of deviation between predicted and actual values:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{7}$$

Mean Absolute Error (MAE) is found to be less sensitive to extreme values than RMSE. It is an alternative perspective for evaluating errors:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{8}$$

The coefficient of determination (R^2) is the fraction of variance in the dependent variable explained by the model, and values close to 1 indicate better model fit:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n 1^n (y_i - \bar{y})^2} \tag{9}$$

Table 3. XGBoost hyperparameter search space and optimal values

Hyperparameter	Meaning	Search Range	Optimal Value
n_estimators	Number of trees	[100, 200, 300, 500]	300
max_depth	Maximum tree depth	[3, 5, 7, 9]	5
learning_rate	Learning rate	[0.01, 0.05, 0.1, 0.2]	0.05
min_child_weight	Minimum sum of instance weight in leaf	[1, 3, 5]	3
subsample	Sample sampling ratio	[0.7, 0.8, 0.9, 1.0]	0.8
colsample_bytree	Feature sampling ratio	[0.7, 0.8, 0.9, 1.0]	0.8
gamma	Minimum loss reduction for split	[0, 0.1, 0.2]	0.1
reg_lambda	L2 regularization coefficient	[1, 1.5, 2]	1.5

SHAP (SHapley Additive exPlanations) is inspired by the idea of the Shapley value in cooperative game theory to provide explanations for predictions of machine learning models in a consistent manner [12]. The idea behind it is to decompose the predictions of a machine learning model into a sum of contributions of individual features. For a given sample x , its predicted value can be expressed as:

$$f(x) = \phi_0 + \sum_{j=1}^M \phi_j \tag{10}$$

where ϕ_0 is the baseline prediction value (typically the mean of the dependent variable in the training set), and ϕ_j is the SHAP value of the j -th feature, representing that feature's marginal contribution to the prediction result. The evaluation of SHAP value computation involves measuring the change in predictions as feature j is included in all possible feature subsets, weighted by the size of each subset. For tree-based models, the TreeSHAP method leverages the structure of decision trees to reduce the computational complexity of SHAP value computation from exponential to polynomial time, thus making it practical for large datasets.

The research adopts SHAP analysis on three levels. The global feature importance plot sorts features by the average absolute SHAP value in descending order, indicating which factors are most important for predicting innovation performance. The SHAP dependence plot shows how SHAP values for a feature vary with that feature's values. To facilitate interpretation, the horizontal axis of all SHAP dependence plots shows the original (unstandardized) feature values, whereas SHAP values are computed from the model's standardized inputs. This plot can intuitively

illustrate changes in the marginal effects of the pay gap variables. Moreover, it may be used to show the possible locations of turning points, as indicated by an inverted U-shaped pattern, to verify the hypothesis of an inverted U-shaped relationship. The local attribution analysis uses a force plot to show the decomposition of feature contributions in predicting a specific firm’s innovation performance. This plot may be used to intuitively show why a firm has a high or low innovation performance.

4. Empirical results

4.1 Descriptive statistics and correlation analysis

Table 4 lists the descriptive statistics of the key variables. The average value of innovation performance, defined as the natural logarithm of patent applications plus one, is 3.127, and the standard deviation is 1.536. The minimum value is 0, implying that the firm has no patent applications that year. The maximum value is 8.217. The average vertical pay gap is 6.782, suggesting that the compensation paid to the executives of the sample firms is approximately 6.8 times that paid to ordinary employees. The standard deviation is as high as 4.693. The variable ranges from 1.217 to 31.856, providing substantial variation to explore potential nonlinear relationships. The average TMT internal gap is 1.843, suggesting that the average compensation paid to the top three managers is approximately 1.8 times that paid to all the managers. The average firm size is 22.137, while the median is 21.964. This suggests that the distribution of firm scale is relatively balanced. The average leverage ratio is 0.418. State-owned enterprises constitute 32.6% of the sample, and the proportion of CEO duality is 31.3%.

Table 5 presents the Pearson correlation coefficients for the key variables. The correlation coefficient for innovation performance and vertical pay gap is 0.147, significantly positive at 1%, preliminarily indicating that firms with larger pay gaps have higher innovation output. However, the absolute value of the correlation coefficient is relatively low, suggesting a nonlinear relationship between the two variables that linear correlation may not fully capture.

The correlation coefficient for innovation performance is highest with firm size (0.412), consistent with the notion that larger firms have more financial resources available for R&D investment. Firm age is also found to have a positive correlation with innovation performance (0.236), possibly because firms with longer listing ages have more experience in innovation. The correlation coefficient between profitability and innovation performance is 0.183, possibly because firms with higher profitability have greater financial capacity to support innovation investment. The correlation coefficients among the three pay gap variables are relatively strong and positive, but none exceeds 0.6.

Table 4. Descriptive statistics (N=26,815)

Variable	Mean	Std. Dev.	Min	Median	Max
Innovation	3.127	1.536	0.000	3.091	8.217
Gap_vertical	6.782	4.693	1.217	5.436	31.856
Gap_TMT	1.843	0.547	1.008	1.719	4.582
Gap_CV	0.568	0.279	0.083	0.517	1.793
Size	22.137	1.271	19.862	21.964	26.283
Lev	0.418	0.196	0.047	0.409	0.887
ROA	0.039	0.064	-0.276	0.037	0.208
Growth	0.156	0.382	-0.563	0.102	2.791
Top1	0.341	0.146	0.084	0.319	0.752
Age	2.183	0.678	0.693	2.197	3.401
SOE	0.326	0.469	0.000	0.000	1.000
Dual	0.313	0.464	0.000	0.000	1.000

Table 5. Correlation matrix

	Innov	Gap_v	Gap_T	Gap_C	Size	Lev	ROA	Growth	Top1	Age
Innov	1.000									
Gap_v	0.147***	1.000								
Gap_T	0.089***	0.463***	1.000							
Gap_C	0.052***	0.387***	0.524***	1.000						
Size	0.412***	0.218***	0.103***	0.067***	1.000					
Lev	0.093***	0.072***	0.031***	0.019**	0.487***	1.000				
ROA	0.183***	0.156***	0.087***	0.042***	-0.017***	-0.378***	1.000			
Growth	0.107***	0.063***	0.048***	0.029***	0.021***	0.038***	0.162***	1.000		
Top1	0.068***	0.094***	0.057***	0.023***	0.186***	-0.023***	0.102***	0.014**	1.000	
Age	0.236***	0.078***	0.041***	0.018**	0.341***	0.203***	-0.086***	-0.127***	-0.052***	1.000

Note: ***, **, * denote significance at the 1%, 5%, and 10% levels respectively; variable names are abbreviated due to space constraints.

4.2 Model prediction performance comparison

The prediction performance of the four models on the test set is presented in Table 6. XGBoost achieves the best predictive performance across all evaluation metrics. The RMSE and MAE for XGBoost are 1.208 and 0.912, respectively, and the R² is 0.382. On the other hand, the R² value for the OLS model, which represents traditional linear models, is only 0.231. Moreover, the RMSE value for the OLS model is as high as 1.347. The improvement in the R² value for the XGBoost model over the OLS model is 65.4%, further supporting the effectiveness of machine learning models over traditional models in capturing nonlinear relationships. Both the gradient boosting models outperform the Random Forest model. The R² value for the LightGBM model is 0.359, which is close to the R² value for the XGBoost model but with faster training speed. The R² value for the Random Forest model is only 0.331, which falls between the linear and gradient boosting models. This shows that although the random forest approach improves the models' prediction accuracy, it is still less effective than the boosting approach at capturing interactions among variables.

Table 6. Model prediction performance comparison

Model	RMSE	MAE	R ²	R ² Improvement over OLS
OLS	1.347	1.051	0.231	—
Random Forest	1.256	0.967	0.331	43.3%
LightGBM	1.230	0.935	0.359	55.4%
XGBoost	1.208	0.912	0.382	65.4%

Note: Evaluation metrics are calculated based on the test set (n=8,044); R² improvement = (Model R² - OLS R²) / OLS R².

Figure 1 shows a scatter plot of XGBoost-predicted values versus actual values for the test data. The data points lie along the diagonal line, implying that predictions are accurate. In the low-to-medium innovation performance range (0 to 4), predicted values closely match actual values. In the high innovation performance range (>6), there is some underestimation, which may be attributed to the influence of unobservable factors on firms with very high levels of innovation output. Table 7 shows the results of a fivefold cross-validation, a method for evaluating model stability. The R² values of XGBoost across the five folds range from 0.371 to 0.394, with a standard deviation of 0.009, indicating robust performance and low overfitting. The minor changes in RMSE and MAE across the five folds confirm the model's strong generalization capacity.

Table 7. XGBoost five-fold cross-validation results

Fold	RMSE	MAE	R ²
Fold 1	1.212	0.917	0.378
Fold 2	1.195	0.902	0.394
Fold 3	1.218	0.923	0.371
Fold 4	1.204	0.909	0.385
Fold 5	1.210	0.914	0.379
Mean	1.208	0.913	0.381
Std. Dev.	0.009	0.008	0.009

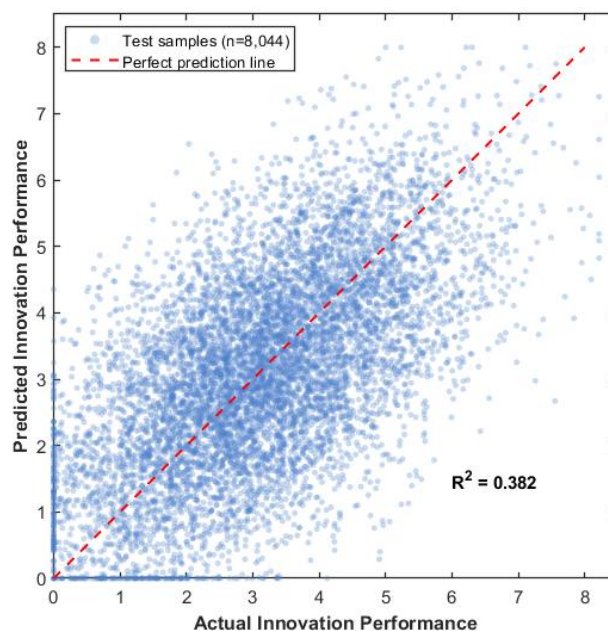


Figure 1. XGBoost predicted vs actual values scatter plot

4.3 SHAP interpretability analysis

Figure 2 shows the global feature importance ranking based on SHAP values. Firm size has the largest mean absolute SHAP value, 0.487, making it the most important variable for predicting innovation performance. This is consistent with economic logic because larger firms possess more substantial resources, stronger R&D infrastructure, and stronger patent application capabilities. Firm age is the second most important factor, with a mean absolute SHAP value of 0.312. This implies that firms with a longer listing history tend to benefit more from accumulated innovation.

The vertical pay gap ranks as the third-highest variable, with a mean absolute SHAP value of 0.276, the highest among the compensation-related variables. This suggests that the pay gap between executives and employees has strong predictive power for evaluating innovation performance, even more than profitability (ranked fourth, with a mean absolute SHAP value of 0.251) and leverage (ranked fifth, with a mean absolute SHAP value of 0.198). The internal gap within the TMT ranks seventh, while the coefficient of variation for the pay gap ranks ninth, implying that the vertical pay gap has better predictive ability for innovation performance than the internal gap within the TMT. The SHAP values indicate predictive significance rather than causal relevance. A high SHAP value indicates a high degree of predictability.

Figure 3 presents the SHAP summary plot, which shows the importance and effect of each feature. Each point in the plot represents a sample, and its position along the horizontal axis indicates the sample's SHAP value for this feature. The color of the point represents whether the value of this feature is high or low. For firm size, the red points indicating high values are concentrated in the area of positive SHAP values on the right. Therefore, the larger the firm size, the higher the predicted innovation performance. For the vertical pay gap, the median value has the highest SHAP value, while the extremely high value shows a declining SHAP value, indicating a nonlinear relationship.

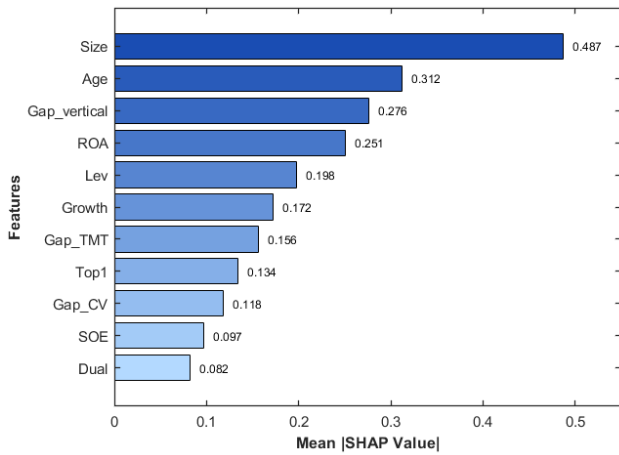


Figure 2. SHAP feature importance ranking

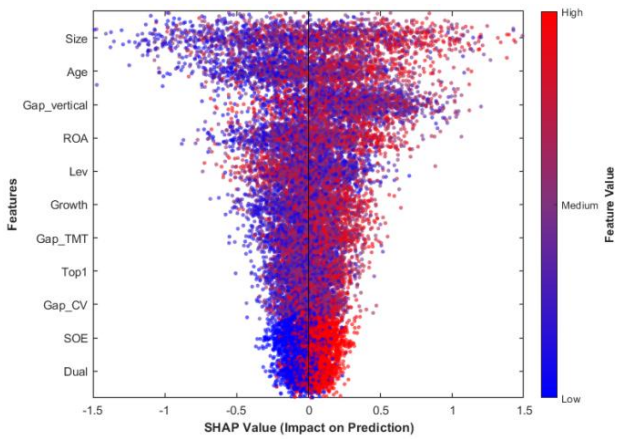


Figure 3. SHAP summary plot

To further analyze the nonlinear impact of the pay gap on innovation performance, Figure 4 shows the SHAP dependence plot of the vertical pay gap. The horizontal axis represents the actual values of the vertical pay gap, while the vertical axis represents the SHAP values. It is clear that the SHAP dependence plot for the vertical pay gap shows an inverted U-shape. As the pay gap increases from 1 to 9 times, the SHAP values continue to increase, indicating that a widening pay gap positively impacts innovation performance. However, once the pay gap exceeds 9 times, the SHAP values decline, indicating that a widening pay gap negatively impacts innovation performance. The inflection point is at 8.7 times, where the SHAP value reaches its maximum of 0.31. Bootstrap resampling with 1,000 iterations yields a 95% confidence interval of [7.9, 9.5] for this turning point, indicating reasonable precision in the threshold estimate.

This phenomenon is consistent with the dual effects of the tournament theory and social comparison theory. As the pay gap is relatively small, the tournament-theory effect is more pronounced. In this case, the increasing pay gap can motivate employees' innovation performance in the hope of promotion. Nevertheless, if the gap exceeds the critical level, the negative consequences of the social comparison theory are found. Employees feel unfairly treated, their willingness to cooperate decreases, and their innovation performance suffers. These findings support H1, confirming an inverted U-

shaped relationship between vertical pay gap and innovation performance. To provide a more intuitive sense of the economic magnitude, we calculate the predicted change in innovation performance across different pay-gap levels. Holding other variables at their mean values, increasing the vertical pay gap from 5 times to 8 times is associated with an increase of approximately 0.18 in predicted innovation performance (equivalent to about 20% more patent applications). However, further increasing the pay gap from 9 times to 15 times is associated with a decrease of approximately 0.23 in predicted innovation performance. These estimates illustrate that the marginal effect of the pay gap on innovation is not only nonlinear but also economically meaningful.

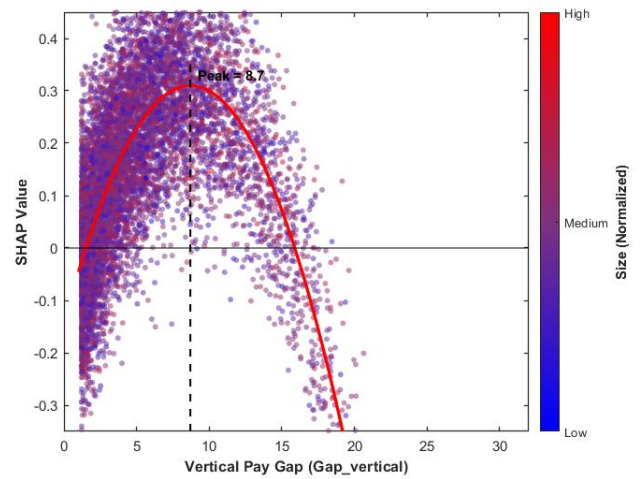


Figure 4. SHAP dependence plot for vertical pay gap

Figure 5 further compares the effects of pay gaps across different ownership types. It divides the samples into state-owned enterprises and non-state-owned enterprises with sample sizes of 8,742 and 18,073, respectively. It creates a SHAP dependence plot for each group. Both groups show an inverted U-shaped curve but differ in the locations of the turning points and the slopes of the curves. The turning point for SOEs is approximately 7.2 times, while for non-SOEs it is approximately 10.1 times. A permutation test with 1,000 iterations confirms that this 2.9-fold difference is statistically significant ($p < 0.01$). This means that SOEs have lower "tolerance" for pay gaps and will experience negative effects from social comparison with smaller pay gaps, whereas non-SOEs can tolerate larger pay gaps without harming innovation performance. These results support H2, indicating that the turning point is significantly lower for state-owned enterprises than for non-state-owned enterprises.

Figure 6 presents the local attribution analysis for two typical firms. Sample A is a large-scale state-owned manufacturer with an actual innovation performance of 5.83 and a predicted value of 5.67. The force plot indicates that firm size (+0.89) and firm age (+0.52) are the primary positive drivers. The vertical pay gap (6.2 times) contributes +0.28, placing it within the positive incentive range identified earlier. The combined contribution of all other features is +0.79. Starting from the baseline prediction of 3.127, the sum of these SHAP contributions yields the predicted value of 5.67. Sample B is a medium-scale private information technology firm with an actual innovation performance of 4.21 and a predicted value of 4.08.

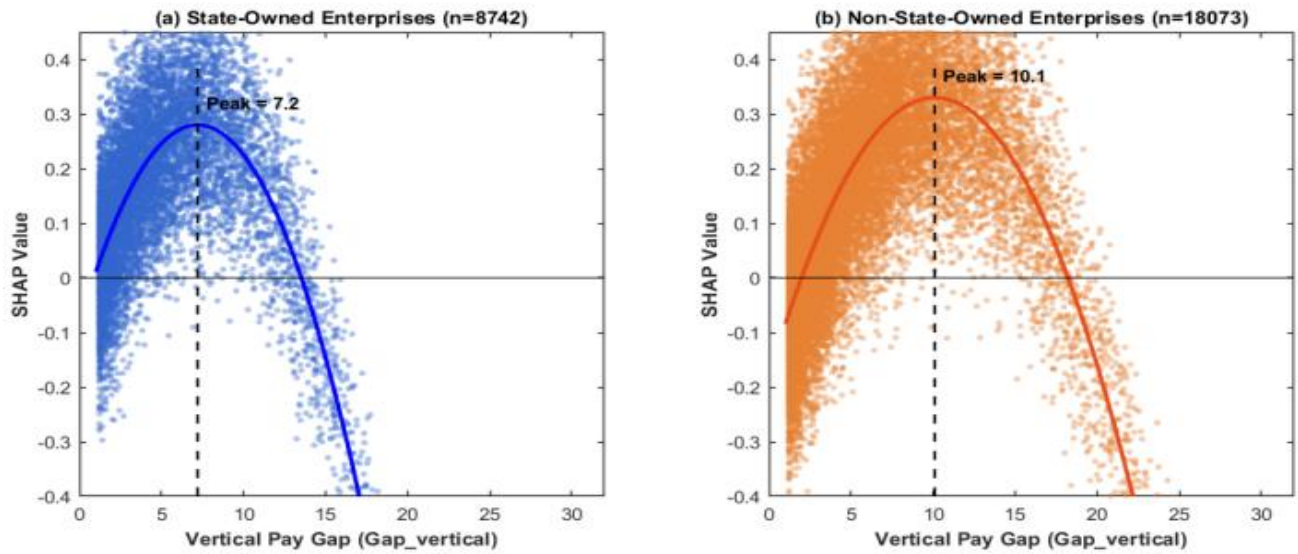


Figure 5. SHAP dependence plot by ownership type

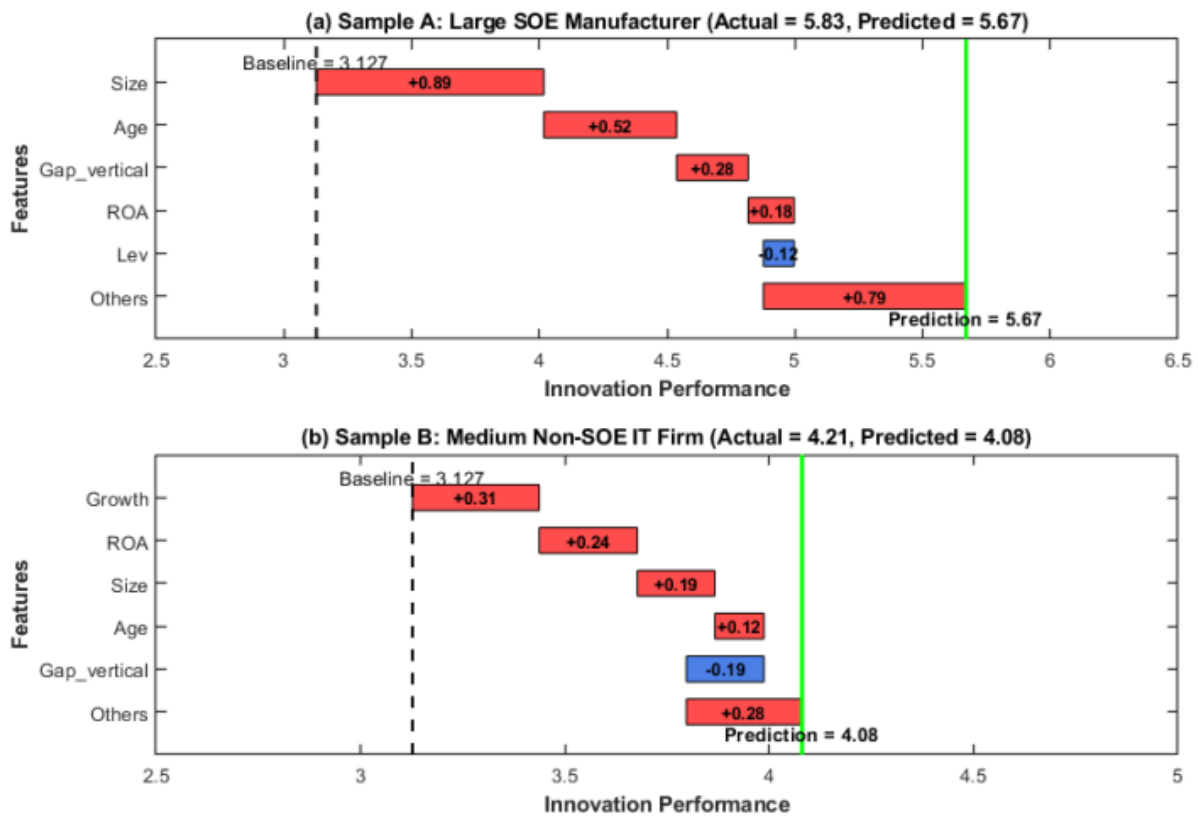


Figure 6. SHAP force plots for typical firms

Its vertical pay gap is 18.3 times, and the SHAP value is -0.19, showing a negative impact on its innovation performance due to the excessive pay gap; high growth contributes +0.31, and profitability contributes +0.24; other features contribute +0.59; the sum of all contributions is 0.95, and the baseline prediction is 3.127, so the predicted value is 4.08.

4.4 Robustness Checks

The main findings are verified through several robustness tests (Table 8). Using invention patent applications as an alternative measure of innovation performance, the XGBoost model maintains strong predictive performance ($R^2 = 0.361$), and the SHAP dependence plot still displays an inverted U-shaped pattern. The turning point shifts slightly to 8.3 times, which falls within the 95% confidence interval of the main analysis, suggesting that the core findings are not driven by the inclusion of utility models and design patents. Relaxing the employee threshold from 50 to 30 increases the sample size to 29,847 firm-year observations. The inverted U-shaped relationship remains robust, with a turning point of 8.5 times. The stability of results across different thresholds alleviates concerns that the findings are artifacts of arbitrary sample restrictions.

Table 8. Robustness check results

Test	Sample Size	R ²	Turning Point	95% CI
Main analysis (total patents)	26,815	0.382	8.7	[7.9, 9.5]
Invention patents only	26,815	0.361	8.3	[7.4, 9.2]
Employee threshold ≥ 30	29,847	0.378	8.5	[7.7, 9.3]

Note: 95% confidence intervals are based on bootstrap resampling with 1,000 iterations.

5. Discussion

The fundamental finding of this study is an inverted U-shaped relationship between the vertical pay gap and innovation performance, with a turning point at 8.7. Such a finding not only confirms the theoretical prediction of the combined effect of tournament theory and social comparison theory but also reveals the limitations of the traditional linear research framework. The turning point of the inverted U-shape relationship also needs to be discussed. The threshold of 8.7 implies that when the pay gap between executives and ordinary employees reaches 8.7 times, the tournament effect and the social comparison effect are balanced, and when the pay gap exceeds this value, the marginal effect of the pay gap on innovation performance becomes negative. Such a value is particularly important because the average value of the vertical pay gap of Chinese A-share listed companies is 6.782, and the median value is 5.436, indicating that more than half of the companies are within the positive incentive zone, but the remaining companies have already surpassed the turning point. The implication is that for firms beyond this threshold, further widening of the pay gap is associated with lower, rather than higher, predicted innovation performance.

The difference in the number of turning points between state-owned and non-state-owned enterprises (7.2 and 10.1, respectively) illustrates the role that organizational context can play in the design of incentive systems. For workers in SOEs, the importance of pay fairness appears to be higher. This may be explained by several factors: workers in SOEs have more social responsibility, hence the desire for more equitable distribution of prosperity; the promotion system is

less flexible in SOEs, thereby the pay gap is more difficult to rationalize as reflecting differences in ability; the government regulations on the pay gap between executives and workers in SOEs reinforce the idea that there are typical and reasonable gaps between abilities. For workers in non-SOEs, there is greater market pressure, and they may be more inclined to accept the idea that “you get more pay for more ability,” hence a higher tolerance for the pay gap. This is consistent with the findings from the digital transformation study, where local tournament incentives are more important in digitalization in non-SOEs, hence the importance of ownership type [35].

The 65.4% improvement in R^2 for XGBoost over OLS carries significant methodological implications. In the context of the pay-innovation relationship, the traditional econometric approaches commonly used require a priori specification of the relationship's form (e.g., linear, quadratic, or cubic). This “a priori assumption” approach has obvious drawbacks: if the true relationship is non-monotonic, forcing a linear specification can lead to severely biased inferences. The advantage of machine learning methods lies precisely in their “data-driven” nature, with algorithms automatically identifying complex relationships among variables without human intervention. Improved prediction accuracy, however, does not automatically equate to causal identification; XGBoost captures conditional correlations rather than causal effects. For the practical question of “what level might pay gap be set at,” accurate conditional prediction itself holds significant value.

The SHAP results show that the vertical pay gap ranks as the third most important predictor, after firm size and firm age. This ranking aligns with intuition, as innovation requires more resources, and larger, older firms have greater initial innovation capacity. Even after controlling for such strong factors, economic intuition suggests that compensation incentives, as an indicator of soft power, contribute to predictive power. What is more surprising is that the vertical pay gap is more important than profitability, contrary to some research on financing constraints. Some research argues that patent stocks can aid smaller firms in breaking financing constraints for R&D investment [36]. This study's findings suggest that, in the Chinese context, the design of internal incentive mechanisms may better predict innovation output than external financing conditions do, perhaps due to the prevalent “soft financing constraints” faced by Chinese firms.

The two typical cases identified in the local attribution analysis have important implications. For Sample A, a large SOE, the vertical pay gap is 6.2, which falls within the range associated with positive predicted innovation performance. For Sample B, a medium-sized private firm, the pay gap is as high as 18.3, far beyond its turning point of 10.1. Its pay structure is associated with lower predicted innovation performance. It is difficult to conduct such an analysis on an individual firm level using traditional regression analysis. For managers, rather than wondering whether the average pay gap is associated with innovation, they want to know whether their firm's pay gap is appropriately calibrated. SHAP force plots transform abstract statistical results into concrete diagnostic tools. This research has some limitations. Patent counts may not be the best way to measure the quality of innovation. Other research on Indonesian firms found that pay gaps may lead to a “quantity for quality” approach [16]. Future research could explore the robustness of the results with patent citations. This research only examines internal pay gaps and does not account for external gaps, namely the difference between executive compensation and industry

averages. Other research has found that when industry averages are low, external gaps may negatively affect innovation [19]. Future research could explore the relationship between internal and external gaps. This research is about prediction and does not seek to prove causality. SHAP is used to identify patterns of correlation but not causation. There may be reverse causality or omitted variable bias in the relationship between pay gaps and innovation. Future research seeking causal identification could exploit exogenous policy shocks as instrumental variables or employ regression discontinuity designs. This research examines only Chinese A-share-listed companies. The findings may be relevant for other emerging markets with similar institutional features, such as significant state ownership and cultural emphasis on fairness, but threshold values are likely to differ depending on local labor market conditions and regulations.

6. Conclusion

The current study develops and applies an XGBoost-SHAP analytical framework to forecast the innovation performance of corporations using data on Chinese A-share-listed companies from 2010 to 2023. The results showed that the XGBoost model achieved prediction accuracy far surpassing that of the traditional linear regression method, with a difference of 65.4%. Moreover, the results showed that the vertical pay gap was the third most important factor in predicting innovative performance, after firm size and firm age. The dependence plots of SHAP values indicated that the relationship between the pay gap and innovation performance was inverted U-shaped, with a turning point at 8.7 times, supporting H1. Moreover, the turning point of state-owned enterprises was lower than that of non-state-owned enterprises. That is, the turning point for state-owned enterprises was 7.2 times, while for non-state-owned enterprises it was 10.1 times, supporting H2. These findings provide actionable guidance for compensation system design. For firms whose pay gaps have not yet reached the turning point, moderately widening gaps is associated with higher predicted innovation performance; for firms that have already crossed the turning point, narrowing gaps may be beneficial. State-owned enterprises may need to be more cautious when designing compensation structures, as their employees appear to be more sensitive to fairness. The strength of the XGBoost-SHAP framework, therefore, is not just that it provides a way to improve the accuracy of predictions, but more importantly, that it provides a way to translate abstract statistical findings into a tool for diagnosis at the individual level, and hence allows the manager to address the critical question of "whether my firm's pay gap is appropriately calibrated."

Ethical issue

The authors are aware of and comply with best practices in publication ethics, specifically regarding authorship (avoidance of guest authorship), dual submission, manipulation of figures, competing interests, and compliance with research ethics policies. The authors adhere to publication requirements that the submitted work is original and has not been published elsewhere.

Data availability statement

The manuscript contains all the data. However, more data will be available upon request from the authors.

Conflict of interest

The authors declare no potential conflict of interest.

References

- [1] AFL-CIO, "Executive Paywatch 2025," AFL-CIO, Washington, DC, 2025. [Online]. Available: <https://aflcio.org/paywatch>.
- [2] S. Anderson, "Executive Excess 2025: CEO Pay at America's 100 Largest Low-Wage Employers," Institute for Policy Studies, Washington, DC, 2025. [Online]. Available: <https://ips-dc.org/release-executive-excess-2025/>.
- [3] E. P. Lazear and S. Rosen, "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, vol. 89, no. 5, pp. 841-864, 1981. DOI: <https://doi.org/10.1086/261010>.
- [4] L. Festinger, "A theory of social comparison processes," *Human Relations*, vol. 7, no. 2, pp. 117-140, 1954. DOI: <https://doi.org/10.1177/001872675400700202>.
- [5] J. S. Adams, "Inequity In Social Exchange," in *Advances in Experimental Social Psychology*, vol. 2, L. Berkowitz Ed.: Academic Press, 1965, pp. 267-299.
- [6] Z. Tan, X. Wu and R. Chu, "Impact of Pay Gap on Innovation Performance: The Moderating Role of Top Management Team Diversity," *Sustainability*, vol. 16, no. 17, p. 7459, 2024. DOI: <https://doi.org/10.3390/su16177459>.
- [7] L. Fu, S. Zhang and F. Wu, "The Impact of Compensation Gap on Corporate Innovation: Evidence from China's Pharmaceutical Industry," *International Journal of Environmental Research and Public Health*, vol. 19, no. 3, p. 1756, 2022. DOI: <https://doi.org/10.3390/ijerph19031756>.
- [8] S. Wang and Z. Lin, "Pay structure and firm technological innovation: comparative research based on three pay gaps," *Humanities and Social Sciences Communications*, vol. 11, no. 1, p. 373, 2024. DOI: <https://doi.org/10.1057/s41599-024-02820-0>.
- [9] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, California, USA, 2016. DOI: <https://doi.org/10.1145/2939672.2939785>.
- [10] J. Zhang and K. Yin, "Application of gradient boosting model to forecast corporate green innovation performance," *Frontiers in Environmental Science*, vol. 11, 2023. DOI: <https://doi.org/10.3389/fenvs.2023.1252271>.
- [11] J. Zhang and Z. Zhao, "Corporate ESG rating prediction based on XGBoost-SHAP interpretable machine learning model," *Expert Systems with Applications*, vol. 295, p. 128809, 2026. DOI: <https://doi.org/10.1016/j.eswa.2025.128809>.
- [12] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in Neural Information Processing Systems*, vol. 30, 2017. DOI: <https://doi.org/10.48550/arXiv.1705.07874>.
- [13] W. Zheng, Y. Li and X. Guan, "How the executive pay gap affects corporate innovation," *Technology Analysis & Strategic Management*, vol. 36, no. 11, pp. 3973-3986, 2024. DOI: <https://doi.org/10.1080/09537325.2023.2280536>.

- [14] J. Zheng, H. Chowdhury, M. S. Hossain and K. Gupta, "Tournament-based incentives and media sentiment," *Journal of Contemporary Accounting & Economics*, vol. 19, no. 2, p. 100353, 2023. DOI: <https://doi.org/10.1016/j.jcae.2023.100353>.
- [15] H. Choi, K. Karim and Y. Liu, "Tournament incentives, corporate overinvestment, and economic consequences," *International Review of Financial Analysis*, vol. 102, p. 104090, 2025. DOI: <https://doi.org/10.1016/j.irfa.2025.104090>.
- [16] B. Budiandriani and M. Fahlevi, "How Vertical and Horizontal Pay Gaps in Research and Development Affect Corporate Innovation in Indonesian Public Firms," *Economics - Innovative and Economics Research Journal*, vol. 13, no. 2, pp. 367-387, 2025. DOI: <https://doi.org/10.2478/eoik-2025-0048>.
- [17] U. S. Maqsood, Q. Li, S. Wang and R. M. A. Zahid, "Government CEO pay regulation and corporate innovation performance: The role of CEO's career horizon and shareholding," *BRQ Business Research Quarterly*, vol. 28, no. 2, pp. 543-568, 2025. DOI: <https://doi.org/10.1177/23409444251314891>.
- [18] X. Li, X. Wang, Z. Zhao and Q. Zhao, "ESG ratings, executive pay-for-performance sensitivity and within-firm pay gap," *Humanities and Social Sciences Communications*, vol. 12, no. 1, p. 599, 2025. DOI: <https://doi.org/10.1057/s41599-025-04908-7>.
- [19] Y. Jingyi, F. Sijia and J. Hui, "The executive external pay gap and firm innovation performance: An empirical study from the perspective of the market competition environment," *Academic Journal of Business & Management*, vol. 7, no. 1, pp. 37-44, 2025. DOI: <https://doi.org/10.25236/AJBM.2025.070105>.
- [20] W. Przychodzen and F. Gómez-Bezares, "CEO–Employee Pay Gap, Productivity and Value Creation," *Journal of Risk and Financial Management*, vol. 14, no. 5, p. 196, 2021. DOI: <https://doi.org/10.3390/jrfm14050196>.
- [21] I. S. Fulmer, B. Gerhart and J. H. Kim, "Compensation and performance: A review and recommendations for the future," *Personnel Psychology*, vol. 76, no. 2, pp. 687-718, 2023. DOI: <https://doi.org/10.1111/peps.12583>.
- [22] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye and T.-Y. Liu, "LightGBM: A highly efficient gradient boosting decision tree," *Advances in Neural Information Processing Systems*, vol. 30, 2017. DOI: <https://doi.org/10.5555/3294996.3295074>.
- [23] W. Zhu, M. Li, C. Wu and S. Liu, "Comprehensive financial health assessment using advanced machine learning techniques: Evidence based on private companies listed on ChiNext," *PLOS ONE*, vol. 19, no. 12, p. e0314966, 2024. DOI: <https://doi.org/10.1371/journal.pone.0314966>.
- [24] X. Wang and T. Kong, "Exploring the Drivers of Innovation Capability in Listed Companies: A Machine Learning Approach to Identifying Synergistic Factors," *Computational Economics*, 2026. DOI: <https://doi.org/10.1007/s10614-025-11193-8>.
- [25] M. A. Omar, I. I. Goma, S. H. Sabry and H. Moubarak, "Artificial Intelligence's Role in Predicting Corporate Financial Performance: Evidence from the MENA Region," *Journal of Risk and Financial Management*, vol. 19, no. 1, p. 51, 2026. DOI: <https://doi.org/10.3390/jrfm19010051>.
- [26] R. Jaiswal, S. Gupta and A. K. Tiwari, "Dissecting the compensation conundrum: a machine learning-based prognostication of key determinants in a complex labor market," *Management Decision*, vol. 61, no. 8, pp. 2322-2353, 2023. DOI: <https://doi.org/10.1108/MD-07-2022-0976>.
- [27] F. S. Khan, S. S. Mazhar, K. Mazhar, D. A. AlSaleh and A. Mazhar, "Model-agnostic explainable artificial intelligence methods in finance: a systematic review, recent developments, limitations, challenges and future directions," *Artificial Intelligence Review*, vol. 58, no. 8, p. 232, 2025. DOI: <https://doi.org/10.1007/s10462-025-11215-9>.
- [28] P.-D. Arsenault, S. Wang and J.-M. Patenaude, "A Survey of Explainable Artificial Intelligence (XAI) in Financial Time Series Forecasting," *ACM Computing Surveys*, vol. 57, no. 10, p. 265, 2025. DOI: <https://doi.org/10.1145/3729531>.
- [29] S. Ergün, "Explaining XGBoost predictions with SHAP value: A comprehensive guide to interpreting decision tree-based models," *New Trends in Computer Sciences*, vol. 1, no. 1, pp. 19-31, 2023. DOI: <https://doi.org/10.3846/ntcs.2023.17901>.
- [30] Y. Nohara, K. Matsumoto, H. Soejima and N. Nakashima, "Explanation of machine learning models using shapley additive explanation and application for real data in hospital," *Computer Methods and Programs in Biomedicine*, vol. 214, p. 106584, 2022. DOI: <https://doi.org/10.1016/j.cmpb.2021.106584>.
- [31] Y. Wang, W. Wei, Z. Liu, J. Liu, Y. Lv and X. Li, "Interpretable Machine Learning Framework for Corporate Financialization Prediction: A SHAP-Based Analysis of High-Dimensional Data," *Mathematics*, vol. 13, no. 15, p. 2526, 2025. DOI: <https://doi.org/10.3390/math13152526>.
- [32] F. Zeng, J. Wang and C. Zeng, "An optimized machine learning framework for predicting and interpreting corporate ESG greenwashing behavior," *PLOS ONE*, vol. 20, no. 3, p. e0316287, 2025. DOI: <https://doi.org/10.1371/journal.pone.0316287>.
- [33] J. A. Darwish, "Optimization and prediction of corporate credit rating through advanced feature selection based on AI and deep learning," *Alexandria Engineering Journal*, vol. 127, pp. 586-594, 2025. DOI: <https://doi.org/10.1016/j.aej.2025.05.043>.
- [34] S. Athey and G. Imbens, "Recursive partitioning for heterogeneous causal effects," *Proceedings of the National Academy of Sciences*, vol. 113, no. 27, pp. 7353-7360, 2016. DOI: <https://doi.org/10.1073/pnas.1510489113>.
- [35] Z. Hua and Y. Yu, "Digital transformation and the impact of local tournament incentives: Evidence from publicly listed companies in China," *Finance Research Letters*, vol. 57, p. 104204, 2023. DOI: <https://doi.org/10.1016/j.frl.2023.104204>.

- [36] S. Milani and R. Neumann, "R&D, patents, and financing constraints of the top global innovative firms," *Journal of Economic Behavior & Organization*, vol. 196, pp. 546-567, 2022. DOI: <https://doi.org/10.1016/j.jebo.2022.02.016>.



This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).