Information Systems and Economics (2023)
Clausius Scientific Press, Canada

DOI: 10.23977/infse.2023.040918
ISSN 2523-6407 Vol. 4 Num. 9

# *Research on the influence of science and technology innovation policy on innovation performance based on text mining*

**Li Zhenzhen***

*School of Public Administration, Dalian University of Technology, Dalian, Liaoning, China*
**Corresponding author: lzz17709829136@163.com*

***Keywords:*** Technological innovation policies, text sentiment analysis technology, innovation performance nationwide

***Abstract:*** In order to promote better development in the field of technological innovation, the government needs to timely establish and enhance an adaptive policy framework to maximize the effectiveness of policies. Against this backdrop, studying the impact of technological innovation policies on corporate innovation performance holds both theoretical and practical significance. This study utilizes text mining techniques to conduct thematic analysis and keyword extraction on national technological innovation policies from 2009 to 2018. It further compares and analyzes the similarities and differences in technological innovation policies among selected provinces over the past decade. Employing text sentiment analysis technology and considering the inherent characteristics of technological innovation policies, a policy polarity analysis is conducted to obtain incentive scores, responsibility scores, and policy tones from policy texts, exploring the emotional orientation of policies. Taking manufacturing-related policies in various provinces as an example, the study combines text mining techniques with the quantification of policy texts to empirically investigate the role of technological innovation policy polarity and tone in innovation performance nationwide.

## 1. Introduction

Technological innovation has to a certain extent propelled economic and social development, fostering a transformation in the quality and outcomes of economic growth. The focal point of reforms in technological innovation lies in the construction of a more comprehensive policy framework. In order to promote technological innovation within enterprises and enhance its effectiveness, governments at all levels have increased their investment in technological innovation and implemented a substantial number of policies[1]. The government, in various official reports, has repeatedly emphasized the need to "accelerate the industrialization of technological achievements." However, concerns persist regarding the outcomes of technological innovation and its industrialization. The actual number of technological achievements transformed into production is relatively low, and the conversion rate of technological achievements is insufficient. Therefore, studying the impact of technological innovation policies on corporate innovation performance holds significant practical significance[2].

Text Mining (TM) is the process of extracting implicit, previously unknown, and potentially useful information and knowledge from large-scale text databases. Fieldsman first proposed the term text mining, Justicia and Kobayashi gave an overview of text mining and its applications, and Kayser summarized the advantages and contributions of text mining. The early research of text mining focused on text mining models and text feature extraction, and has now extended to the application of text mining technology in many fields, including scientific and technological innovation. The research on text mining started relatively late in China and was introduced into academia around 2000. With the development of science and technology and the expansion of theory, the research and application of Chinese text mining technology have also come into the attention of scholars[3]. Yang Huiji analyzes the situation of climate policy in China and other countries by using semantic topic mining and word frequency analysis. In terms of entrepreneurship policies, Liu Xiaotong uses the entrepreneurship policies issued by national institutions in the past decade, uses the LDA model to analyze the evolution process of entrepreneurship policy themes, and presents its evolution process. In the field of innovation policy, Huang CAI conducts co-word analysis and cluster analysis of China's science and technology policy based on text mining technology[4]. Zhang Yongan used R language to carry out topic mining on the technological innovation policy of Zhongguancun, and made a comprehensive analysis on the word frequency and high frequency words, popular topics and theme change trends, so as to put forward improvement suggestions for the formulation of regional technological innovation policy and regulation system and provide data support for policy formulation and introduction[5]. Huang Lucheng et al. used basic keywords and semantically similar words to construct a lexeme representing innovation policy ideas. On this basis, they evaluated the balance state of innovation policy ideas during 2014-2016, providing a new perspective and method for evaluating innovation policy ideas, and further improving the quality of innovation policies. Based on text mining technology, Zhang Baojian analyzed the text data of China's national science and technology innovation policy in recent ten years according to the content and nature by using keywords and clustering methods[6]. Text mining has become one of the effective means to solve practical problems in various fields, but compared with foreign countries, domestic research is still not deeply applied in policy, research data sources and methods are more limited, and there is still a large space for development. Therefore, it is of good theoretical and practical value to strengthen the research and exploration in this aspect[7].

In order to better promote the development of the technological innovation sector, the government needs to promptly establish and improve an adaptive policy framework to maximize policy effectiveness. Against this backdrop, studying the impact of technological innovation policies on the innovation performance of enterprises holds significant theoretical and practical significance. This research employs text mining techniques to conduct thematic analysis and keyword extraction on national technological innovation policies from 2009 to 2018. Furthermore, it conducts a comparative analysis of the similarities and differences in technological innovation policies among various provinces and cities over the past decade. By utilizing text sentiment analysis techniques and combining them with the inherent features of science and technology innovation policies, the study performs a policy polarity analysis to extract incentive scores, responsibility scores, and policy tones from policy texts, aiming to explore the emotional orientation of the policies. Using manufacturing-related policies from various provinces as examples, this research integrates text mining techniques with the quantification of policy texts to empirically study the impact of the nationwide technological innovation policy polarity and tone on innovation performance.

## 2. Analysis of policy polarity of science and technology innovation based on text mining

### 2.1. An algorithm for polarity values

In addition, this study uses the semantic rule-based emotion analysis method to quantify text emotion by combining the processing rules of degree adverbs and negative words. In the process of studying the text of science and technology innovation policy, when there are negative words in the text sentence, the meaning of the sentence will be opposite. Example 1: The deep-seated problems affecting the development of the equipment manufacturing industry in our province have not been completely solved, the ability to participate in international competition is weak, there are unreasonable industrial and product structure, independent innovation ability is not strong, and the proportion of high-end industries is small. "Strong" and "reasonable" are originally positive words, but the existence of the negative word "no" changes the emotion of the sentence. The influence of negative words on sentences should be considered in the analysis of affective tendency.

In Chinese grammar, situations involving multiple negations can occur. Therefore, when the number of negation words is odd, the emotional polarity of the sentence is reversed; when the number of negation words is even, the emotional polarity of the sentence remains unchanged. Formula (1) is used to calculate the weight of negation words, where t represents the number of negation words.

$$W_t = \begin{cases} 1, t = 2n \\ -1, t = 2n+1 \end{cases} \tag{1}$$

When emotional words are modified by adverbs of degree, the degree of the emotional tendency of the words will correspondingly change. Taken from a sentence in a technology innovation policy text:

Example 2: By 2015, the overall strength of the equipment manufacturing industry in the province had significantly improved. The main goals were: a noticeable enhancement in independent innovation capabilities, further optimization of organizational structures, and a significant strengthening of basic supporting capabilities.

In the sentence, the degree words before "upgrade" are "large" and "obvious", "further" before "optimization" and "significant" before "enhancement", and the emotional degree of the modified sentence is more intense.

When calculating the intensity of emotional tendencies, Formula (2) can be used to compute the weight of adverbs of degree:

$$W_k = \begin{cases} 4, k \in C1 \\ 3, k \in C2 \\ 2, k \in C3 \\ 1, k \in C4 \end{cases} \tag{2}$$

Based on this, sentiment analysis scores are determined through the following semantic rules. By counting the number of negation words before emotional words, when the count of negation words preceding emotional words is odd, it indicates an opposite sentiment, i.e., multiply by -1. When the count of negation words before emotional words is even, it indicates that the sentiment remains unchanged. Additionally, when scanning for emotional words, if adverbs of degree appear before them, the sentiment is multiplied by corresponding weights based on different degree levels. The specific calculation method for sentiment according to these semantic rules is shown in Table 1.

Table 1: Emotional computation of semantic rules

| Serial number | Combination | Pblicity |
|---|---|---|
| 1 | Emotional words | S=word(j) |
| 2 | Negative word + emotion word | S=W(t)*word(j) |
| 3 | Degree words + emotion words | S=W(k)*word(j) |
| 4 | Degree word + negative word + emotion word | S=W(k)*W(t)*word(j) |
| 5 | Negative word + degree word + emotion word | S=W(k)*W(t)*word(j)*0.5 |

The expression "word(j)" represents the weight of emotional words, all set to 1. The emotional values are then summed up to obtain the overall incentive tendency value (E(pos)) and responsibility tendency value (E(neg)) for the entire policy text, as calculated in formulas (3) and (4).

$$E(pos) = \sum_{i \in Pos}^{n} S(i) \tag{3}$$

$$E(neg) = \sum_{i \in Neg}^{n} S(i) \tag{4}$$

## 2.2. Analysis of policy polarity of science and technology innovation

This article first conducts a polarity analysis on the technology innovation policies for each year, calculating incentive scores and responsibility scores for each policy text. After visual processing, a grouped comparative analysis is performed on the polarity analysis scores of technology innovation policies over the past decade. In the graph, blue represents incentive scores, red represents responsibility scores, the vertical axis represents the scores, and the horizontal axis represents the number of policy texts[8]. The graph illustrates the incentive and responsibility scores for each policy text in each year. Figure 1 presents the "Incentive-Responsibility" polarity analysis for technology innovation policies in 2010, 2012, 2017, and 2018. From the graph, it can be visually observed that in each policy text of these four years, the blue area is significantly larger than the red area. This indicates that the incentive scores are higher than the responsibility scores for the technology innovation policy texts in 2010, 2012, 2017, and 2018. It can be inferred that the policies in these four years lean towards incentive-oriented technology innovation policies. In other words, the technology innovation policies during these four years emphasize incentives for innovation, aiming to enhance innovation levels and speed up economic development.

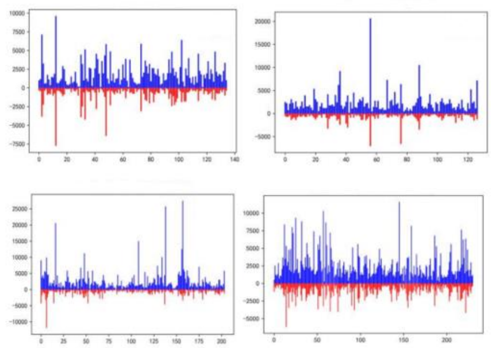

Figure 1: Analysis of "incentive-responsibility" polarity of science and technology innovation policies in 2010, 2012, 2017 and 2018

Figure 2 presents the "Incentive-Responsibility" polarity analysis for technology innovation policies in 2009, 2013, 2014, and 2015. From the graph, it can be visually observed that among all policies in these four years, there are a few where the blue area (incentive) is greater than the red area (responsibility), and a few where the red area is smaller than the blue area. However, for the majority of policies, the blue area is approximately equal to the red area. In other words, for the technology innovation policy texts in 2009, 2013, 2014, and 2015, there are a few policies with incentive scores greater than responsibility scores, indicating incentive-oriented policies. Similarly, there are a few policies with responsibility scores greater than incentive scores, indicating responsibility-oriented policies. For the majority of policies, incentive scores are roughly equal to responsibility scores, suggesting a balanced policy approach. This indicates that, overall, the technology innovation policies in 2009, 2013, 2014, and 2015 tend to be balanced, with a mix of responsibility-oriented, incentive-oriented, and balanced policies. It can be observed that during the initial years following the release of the national "Eleventh Five-Year" and "Twelfth Five-Year" development plans, various provinces formulated robust technology innovation policies to promote industrial development in accordance with national planning. These policies emphasized incentives to boost development speed. After achieving certain results, local governments began to stress the quality of development, moving away from solely focusing on speed. By implementing responsibility-oriented policies to regulate innovation behavior, they aimed to improve development quality and enhance corporate social responsibility.

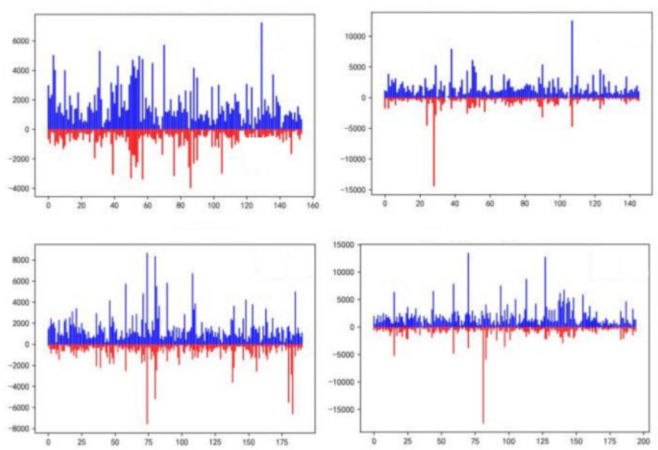

Figure 2: Analysis of "incentive-responsibility" polarity of science and technology innovation policy in 2009, 2013, 2014 and 2015

Figure 3 illustrates the "Incentive-Responsibility" polarity analysis for technology innovation policies in 2011 and 2016. From the graph, it can be readily observed that, in each technology innovation policy in these two years, the blue area (incentive) is consistently larger than the red area (responsibility). This indicates that the incentive scores for technology innovation policies in these two years are greater than the responsibility scores, classifying them as incentive-oriented policies. Additionally, it is notable that the incentive scores for technology innovation policies in these two years are generally high. When compared to other years, the maximum incentive scores are significantly higher, with the peak reaching close to 50,000 in 2011 and around 40,000 in 2016. This reflects a strong emphasis on innovation incentives in various regions during 2011 and 2016, aimed at encouraging the development of different industries and promoting economic growth. This emphasis is closely tied to the national "Twelfth Five-Year" and "Thirteenth Five-Year" plans. During the "Twelfth Five-Year" and "Thirteenth Five-Year" periods, rapid development occurred in various fields across China. The years 2011 and 2016 mark the beginning of these respective plans, during which the country placed significant emphasis on development. Regions formulated technology innovation policies in line with national development plans, with a greater focus on innovation incentives to accelerate both development speed and innovation levels. Consequently, higher incentive scores are observed in 2011 and 2016. The elevated incentive and responsibility scores in these years are also attributed to the enactment of numerous technology innovation policies. By comparing the "Incentive-Responsibility" polarity analysis graphs for 2011 and 2016, it is evident that, overall, the blue area in 2016 is larger than that in 2011. Moreover, the red area in 2016 has increased compared to 2011. This indicates that, overall, the incentive scores for technology innovation policies in 2016 are higher than those in 2011, and there is a noticeable improvement in responsibility scores in 2016 compared to 2011. This reflects the gradual shift of various provinces towards recognizing the role of responsibility in innovation activities while still emphasizing the incentivizing aspect.

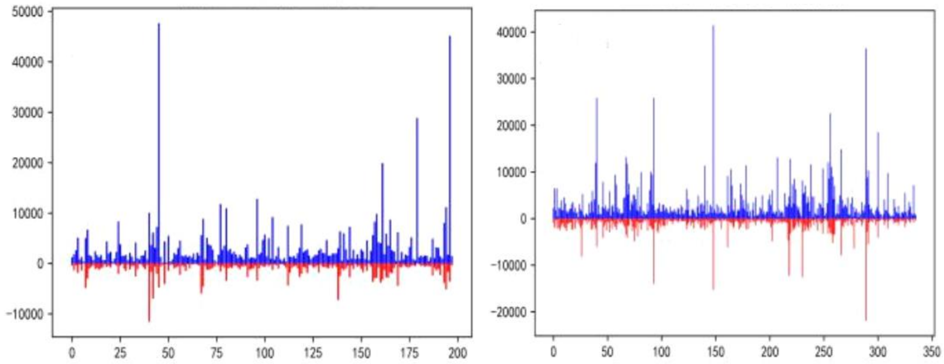


Figure 3: Analysis of "incentive-responsibility" polarity of science and technology innovation policy in 2011 and 2016

Through a comparative analysis of visualized graphs resulting from the polarity analysis of technology innovation policies from 2009 to 2018, it is evident that, in the majority of cases, the blue area (representing incentive scores) is larger than the red area (representing responsibility scores). There are instances where the blue area is equal to the red area, indicating a balance between incentive and responsibility. Only a small number of technology innovation policies exhibit a red area greater than the blue area. This pattern reflects that, over the past decade, most technology innovation policies tend to be incentive-oriented, with a notable presence of balanced policies and a relatively smaller number of responsibility-oriented policies.

## 3. Analysis of the influence of science and technology innovation policy on innovation performance

The government, by formulating policies, directly guides and supports innovative activities to promote economic development. Existing research indicates that technology innovation policies have a dual impact on innovation activities. Through quantitative and polarity analysis of technology innovation policy texts in the preceding chapters, a visual understanding of the development status and characteristics of technology innovation policies over the past decade can be gained[9]. This chapter will employ empirical methods to analyze the impact of technology innovation policies on innovation performance, drawing insights from relevant literature and theoretical perspectives put forth by scholars. The goal is to provide recommendations and strategies for enhancing policy effectiveness and fostering innovation performance in enterprises.

### 3.1. Sample selection

This study utilizes data on the number of patents from state-owned enterprises, with the samples sourced from the China Intellectual Property Office website, while other observational data primarily come from the Guotai An Financial Database (CSMAR) and the WIND Database. Relevant factors that may impact the data have been excluded in the sample selection process. Therefore, the following considerations were taken into account in sample selection: (1) Exclusion of ST, ST*, and PT companies, as well as companies with extreme values each year. (2) Selection of companies that only issue A shares, avoiding differences between B shares and H shares. The final research sample consists of 390 state-owned publicly listed companies, resulting in 3900 observations.

### 3.2. Variable design

#### 3.2.1. Explained variable

Enterprise innovation performance (ZL) is often measured using patents, which is a common indicator for assessing regional innovation performance. Metrics such as patent applications, patent output, sales revenue from new products, and the output value of new products have been commonly employed in previous literature to gauge innovation performance. Generally, sales revenue and output from new products are more accurately reflective after undergoing commercialization and generating benefits. Given the difficulty in obtaining company-level data, and the good availability and generality of patent data, applying for invention patents is considered an appropriate metric for measuring enterprise innovation performance. Additionally, when policies are enacted, companies take time to innovate and generate innovation performance in accordance with the policies, and a consideration for time lag effects should be taken into account. For example, when policies include aspects of innovation incentives, companies need time from the initiation of innovation to the actual innovation output. Therefore, this study selects companies in the sample for analysis based on the innovation patent quantity without lag, with a one-year lag, respectively, to consider time lag effects.

#### 3.2.2. Explanatory variable

Incentive score (POS), responsibility score (POS), and policy tone (Tone). The incentive score is derived from the total incentive score for manufacturing policies in each province across the nation from 2009 to 2018, as analyzed in Chapter 4. The responsibility score is similarly obtained from the total responsibility score for manufacturing policies in each province across the nation from 2009 to 2018, as analyzed in Chapter 4. This study draws inspiration from the relevant research on managerial tone by Fan Libo and Shang Duo (2020). Managerial tone refers to the tone employed by managers in the process of information disclosure. "Management Discussion and Analysis," as a crucial component of annual reports for publicly listed companies, conveys the management's judgment and optimism regarding the company's operational status through the tone of its textual content. In this study, the policy tone is defined as the degree of incentive or responsibility conveyed in the textual content of policy documents, reflecting the emotions of encouragement or responsibility expressed by policymakers in formulating policies. Through polarity analysis of technology innovation policies from 2009 to 2018, each policy is preliminarily categorized as incentive-oriented, responsibility-oriented, or balanced, with corresponding scores for incentive and responsibility aspects obtained. By aggregating the total incentive and responsibility scores for manufacturing policies in each province across the nation from 2009 to 2018, the policy tone for each year is quantified using Formula (5).

$$Tone = (POS - NEG) / (POS + NEG) \quad (5)$$

Where POS represents the total incentive score for manufacturing policies in each year from 2009 to 2018, NEG represents the total responsibility score for manufacturing policies in each year from 2009 to 2018, and policy tone is a positive indicator with values ranging from -1 to 1. A higher value of policy tone indicates a policy more inclined towards incentive-oriented policies, reflecting stronger incentivization. Conversely, a lower value of policy tone indicates a policy more inclined towards responsibility-oriented policies, indicating a stronger sense of responsibility.

#### 3.2.3. Control variable

Drawing on the research experiences of other scholars, this study controls for the following variables in the model: firm size (Size), research and development expenditure (R&D), profitability (Roa), proportion of independent directors (Ind), board size (Board), proportion of the largest shareholder (Top1), leverage ratio (Lev), and cash holdings (Cash).

### 3.3. Empirical model construction

#### 3.3.1. Research hypothesis

The government, through policy formulation, directly guides and supports innovation activities to promote economic development. Government intervention in innovation activities is primarily achieved through the form of science and technology policies, which intervene, guide, and regulate innovation activities from the perspective of direct resource acquisition. For businesses, innovation activities require substantial financial support, and government policies of encouragement and support can directly provide financial assistance, reduce research and development costs, offset certain market uncertainties, and thereby enhance the enthusiasm and motivation for innovation. This, in turn, directly improves a company's innovation performance. Furthermore, from the perspective of signal transmission, when a company receives support and funding from government policies, it indicates that the company is in a favorable market situation or has significant development potential. This can serve as a positive signal to the market, encouraging investment and fostering innovation activity. This, in turn, stimulates increased investment in innovation activities by companies and promotes improvements in innovation performance. Based on the analysis above, the following research hypotheses are proposed.

H1: The higher the incentive score in the technology innovation policy text, the more conducive it is to promoting the improvement of the current innovation performance of enterprises.

H2: The higher the responsibility score in the technology innovation policy text, the more conducive it is to promoting the improvement of the current innovation performance of enterprises.

H3: The higher the policy tone in the technology innovation policy text, the more conducive it is to promoting the improvement of the current innovation performance of enterprises.

H4: The higher the policy tone in the technology innovation policy text, the more favorable it is for promoting the improvement of the subsequent-period innovation performance of enterprises.

After the government formulates policies and provides financial support, it has not established a robust system to ensure the conduct of innovation activities. There is a lack of supervision and management of enterprises' innovation activities, which, to a certain extent, hinders the progress of such activities. Based on the above analysis, the following research hypothesis is proposed.

H5: The higher the incentive score in the technology innovation policy text, the less favorable it is for promoting the improvement of subsequent-period innovation performance of enterprises.

H6: The higher the responsibility score in the technology innovation policy text, the less favorable it is for promoting the improvement of subsequent-period innovation performance of enterprises.

### 3.3.2. Model building

To test the research hypotheses, this paper employs the econometric method of panel regression and constructs models as shown in equations (6) and (7):

$$ZL_{it} = a_0 + a_1 POS_{it} + a_2 NEG_{it} + a_3 Tone_{it} + \sum a_i Controls_{it} + \varepsilon_{it} \tag{6}$$

$$ZL_{1it} = a_4 + a_5 POS_{it} + a_6 NEG_{it} + a_7 Tone_{it} + \sum a_i Controls_{it} + \varepsilon_{it} \tag{7}$$

Where model (1) is employed to examine the impact effects of incentive scores, responsibility scores, and policy tone on innovation performance, and model (2) is used to test the impact effects on subsequent-year innovation performance. The dependent variables are innovation performance (ZL) and subsequent-year innovation performance (ZL1), while the independent variables include incentive score (POS), responsibility score (NEG), and policy tone (Tone). Here, the subscript i represents the enterprise, t denotes the year, a0 and a4 are intercept terms, a1 and a5 are the coefficients for the impact of incentive scores on the number of patents, a2 and a6 are the coefficients for the impact of responsibility scores on the number of patents, a3 and a7 are the coefficients for the impact of policy tone on the number of patents, and ai represents the regression coefficients for various control variables. εit represents the disturbance term.

## 3.4. Empirical test

### 3.4.1. Descriptive statistics

Figure 4 depicts a scatter plot of incentive and responsibility scores for technology innovation policies related to the manufacturing industry in various provinces across China from 2009 to 2018. The statistical results indicate that for the majority of provinces, the incentive scores for technology innovation policies related to the manufacturing industry are higher than the responsibility scores. However, there are some regions where responsibility scores surpass incentive scores, suggesting that technology innovation policies related to the manufacturing industry across various provinces in China tend to be incentive-oriented, but also incorporate elements of responsibility-oriented policies. Notably, in 2011 and 2016, the incentive scores reached around 40,000 points, and the responsibility scores were close to 25,000 points. This is closely related to China's "Twelfth Five-Year" and "Thirteenth Five-Year" plans. During these periods, various sectors in China experienced rapid development, with 2011 and 2016 marking the beginning years of these plans. The country strongly emphasized development, and each region formulated technology innovation policies in accordance with national development plans. These policies placed greater emphasis on innovation incentives and responsibility, focusing on accelerating development speed and improving innovation levels and quality. Therefore, higher incentive and responsibility scores were observed in 2011 and 2016.

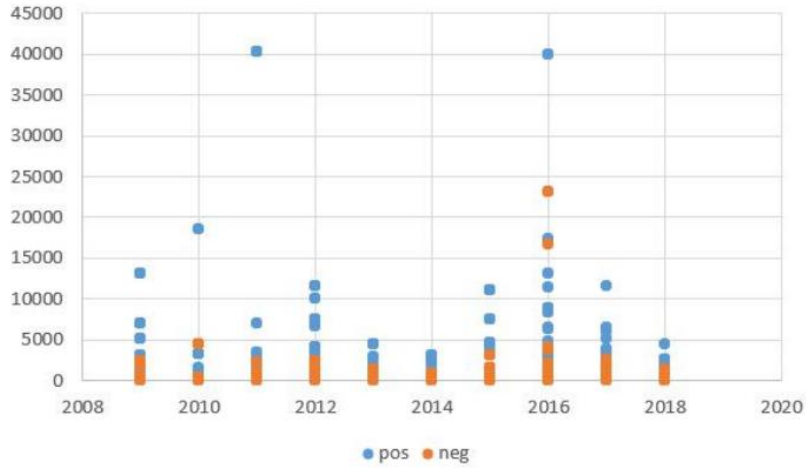


Figure 4: The scientific and technological innovation policy incentive and responsibility of the manufacturing industry in each province of the country are distributed

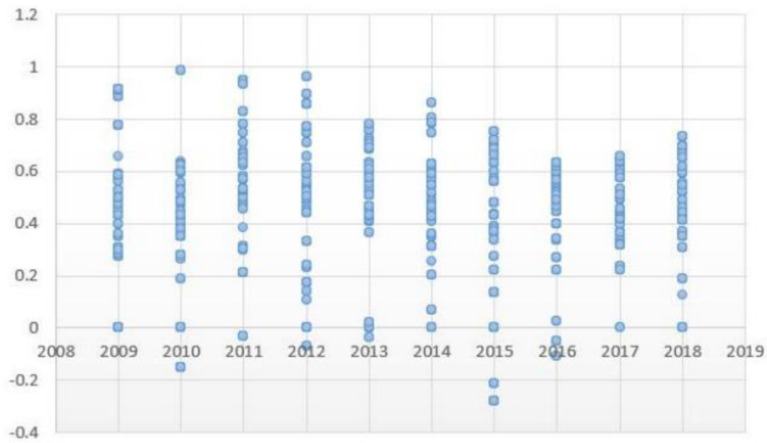


Figure 5: Scatterplot of the national science and technology innovation policy tone of manufacturing industry

Figure 5 illustrates a scatter plot of policy tones for technology innovation policies related to the manufacturing industry in China from 2009 to 2018. The statistical results indicate that the majority of policy tones for technology innovation policies related to the manufacturing industry across China are mostly greater than 0, with the majority falling between 0.4 and 0.8. This suggests that the policy tones of these policies predominantly emphasize innovation incentives. Each year, there are policies with a tone of 0, indicating the presence of balanced technology innovation policies. Additionally, in most years, there are policies with negative tones, signifying a stronger sense of responsibility, indicating the existence of responsibility-oriented technology innovation policies.

Furthermore, considering the possibility of multicollinearity among the relevant variables, this study conducted multicollinearity tests. The results indicate that the variance inflation factor (VIF) for each variable is less than 3. Therefore, this study concludes that there is no multicollinearity among the various relevant variables.

### 3.4.2. Regression analysis of the model

(1) The effects of incentive score, responsibility score and policy tone on current innovation performance

This study initially employed F-tests and the Hausman test to determine the appropriate model for sample selection. The results were generally significant, indicating that a fixed-effects model should be used. Building on this foundation, the study conducted panel regression with the current innovation patent quantity as the dependent variable. Table 2 displays the regression results of incentive score, responsibility score, and policy tone on current innovation performance, aiming to validate hypotheses 1, 2, and 3. The empirical results reveal a significant positive correlation between incentive score and the current quantity of patents ($\beta=0.0318$, $p<0.01$), a significant positive correlation between responsibility score and the current quantity of patents ($\beta=0.0334$, $p<0.01$), and a significant positive correlation between policy tone and the current quantity of patents ($\beta=0.160$, $p<0.05$). For each unit increase in incentive score, the current innovation patent quantity increases by 0.0318%, for each unit increase in responsibility score, the current innovation patent quantity increases by 0.0334%, and for each unit increase in policy tone, the current innovation patent quantity increases by 0.160%. Hypotheses H1, H2, and H3 are thus confirmed.

The incentive score has a significantly positive impact on the current innovation performance of enterprises, indicating that increasing the incentive score of technology innovation policy texts, using more incentive vocabulary such as "stimulate," "enhance," "call for," etc., can convey positive signals for innovation to enterprises, having an encouraging effect on their innovation activities. Emphasizing innovation incentives in technology innovation policies supports and encourages enterprise innovation activities, enhancing the enthusiasm of enterprises and increasing the number of innovation patents in the current year, thereby improving innovation performance. Additionally, the responsibility score has a significantly positive effect on the current innovation performance of enterprises. This suggests that increasing the responsibility score of technology innovation policy texts, using more responsibility-related vocabulary such as "integrate," "maintain," "implement," etc., sends signals of responsibility to enterprises, urging them to undertake responsible innovation. By establishing effective mechanisms to improve innovation quality and ensuring the effective implementation of innovation activities, the responsibility score contributes to an increase in the number of innovation patents in the current year, thus enhancing innovation performance. Both incentive score and responsibility score have a significantly positive impact on enterprise innovation performance, with relatively similar coefficient magnitudes. This implies that both the incentive and responsibility aspects of policy texts are crucial for influencing enterprise innovation performance.

The policy tone has a significantly positive impact on the current innovation performance of enterprises. This indicates that when the technology innovation policy text demonstrates a stronger inclination toward innovation incentives, i.e., when the policy's incentive score is greater than its responsibility score, it conveys a clear signal of innovation encouragement. This strong incentive orientation in technology innovation policy signals a substantial motivating effect on enterprise innovation activities. The policy's support and encouragement of innovation activities enhance the enthusiasm of enterprises for innovation, leading to an increase in the number of innovation patents in the current year and thereby boosting the current innovation performance of enterprises. Conversely, when the technology innovation policy text exhibits a stronger inclination toward innovation responsibility, i.e., when the responsibility score is greater than the incentive score, signaling a stronger policy responsibility orientation, it suggests that the government places emphasis on innovation responsibility. This imparts a signal to enterprises that responsibility is prioritized over incentives, potentially dampening the enthusiasm of enterprises for innovation activities. While this emphasis on responsibility may enhance the quality of innovation activities to some extent, it could slow down the pace and efficiency of innovation activities. This has a restraining effect on increasing the number of innovation patents for enterprises, thereby hindering the improvement of the current innovation performance of enterprises.

Table 2: The regression results of incentive score, responsibility score and policy tone on current innovation performance

| | ZL | ZL | ZL |
|---|---|---|---|
| POS | 0.0318* | | |
| | (3.33) | | |
| NEG | | 0.0334*** | |
| | | (3.14) | |
| Tone | | | 0.160** |
| | | | (1.98) |
| Size | 0.406*** | 0.303*** | 0.420*** |
| | (11.94) | (10.80) | (12.41) |
| Lev | 1.514* | 1.495* | 1.552* |
| | (1.54) | (1.42) | (3.23) |
| Rd | 1.216 | 1.115 | 1.317 |
| | (1.44) | (1.33) | (1.67) |
| Roa | -3.940*** | -2.947*** | -3.770*** |
| | (-3.76) | (-2.76) | (-3.59) |
| Ind | -0.0477 | -0.0488 | -0.124 |
| | (-0.25) | (-0.25) | (-0.24) |
| Board | -0.125 | 0.0638 | -0.0498* |
| | (-0.24) | (0.14) | (-0.26) |
| Top1 | 0.0299 | 0.0312 | 0.0135 |
| | (0.28) | (0.39) | (0.13) |
| _cons | -3.202** | -2.107** | -3.466** |
| | (-2.34) | (-1.27) | (-2.53) |
| A given year | Control | Control | Control |
| Model type | Fixed effect | Fixed effect | Fixed effect |
| *N* | 3900 | 3900 | 3900 |

(2) The effect of incentive score, responsibility score and policy tone on innovation performance with one-stage lag

Utilizing F-test and Hausman test for sample model selection, the results are generally significant, indicating the suitability of employing a fixed-effects model. Building upon this, the selection of lagged innovation patent quantity as the dependent variable is made for panel regression. Table 3 displays the regression results of incentive score, responsibility score, and policy tone on the lagged innovation performance, aiming to validate hypotheses 4, 5, and 6. The empirical results demonstrate that the incentive score exhibits a significantly negative correlation with the lagged quantity of patents ($\beta$=-0.0257, $p<0.01$), the responsibility score shows a significantly negative correlation with the lagged quantity of patents ($\beta$=-0.0329, $p<0.01$), and the policy tone presents a significantly positive correlation with the lagged quantity of patents ($\beta$=0.0416, $p<0.05$). For each additional unit of incentive intensity, the lagged quantity of enterprise innovation patents decreases by 0.0257%. For each additional unit of responsibility intensity, the lagged quantity of enterprise innovation patents decreases by 0.0329%. Conversely, for each additional unit of incentive intensity in policy tone, the lagged quantity of enterprise innovation patents increases by 0.0416%. Hypotheses H4, H5, and H6 are thus validated.

The incentive score exhibits a significantly negative impact on the lagged innovation performance, indicating that elevating the incentive score in the science and technology innovation policy text—utilizing more incentive terms such as "strong support," "firm," "forge," etc.—conveys a positive signal to enterprises in the current period, boosting their enthusiasm for innovation activities. As innovation activities are long-term endeavors that are challenging to complete in the short term, the government continues to enact corresponding science and technology innovation policies in the second year to promote innovation. Coupled with the incentive effects of the first-year policy on innovation activities, this may lead to an innovation crowding-out effect. This effect stimulates the factor demand for research and development (R&D) activities, thereby increasing factor prices and elevating R&D costs. Consequently, this limits enterprises' innovation activities, resulting in a reduction in the quantity of innovation patents and a decrease in overall innovation performance.

Table 3: The regression results of incentive score, responsibility score and policy tone on innovation performance with one-stage lag

| | ZL1 | ZL1 | ZL1 |
|---|---|---|---|
| POS | -0.0257* | | |
| | (-2.64) | | |
| NEG | | 0.0329*** | |
| | | (-3.03) | |
| Tone | | | 0.0416** |
| | | | (0.51) |
| Size | 0.337*** | 0.443*** | 0.326*** |
| | (9.73) | (9.84) | (9.47) |
| Lev | 2.120* | 2.144* | 4.303* |
| | (2.11) | (2.14) | (3.09) |
| Rd | -1.283 | -1.080 | -1.300 |
| | (-1.49) | (-1.08) | (-1.71) |
| Roa | 1.242 | 1.277 | 1.125 |
| | (1.16) | (1.29) | (1.05) |
| Ind | -0.376 | -0.370 | -0.584 |
| | (-0.70) | (-0.51) | (-0.32) |
| Board | -0.0599 | 0.1596 | -0.0753 |
| | (-0.30) | (0.40) | (-0.28) |
| Top1 | -0.123 | -0.228 | -0.077 |
| | (-1.13) | (-1.28) | (-0.98) |
| _cons | -5.165** | -5.289** | -5.053** |
| | (-3.70) | (-3.79) | (-3.62) |
| A given year | Control | Control | Control |
| Model type | Fixed effect | Fixed effect | Fixed effect |
| *N* | 3900 | 3900 | 3900 |

The responsibility score exhibits a significantly negative impact on the lagged innovation performance, indicating that increasing the responsibility score in the science and technology innovation policy text—using more responsibility-related terms such as "promulgate," "strictly," "should be changed," etc.—sends a signal of responsibility to enterprises in the current period, demanding them to engage in responsible innovation. In the second year, if the government emphasizes innovation incentives, increasing the incentive score in the science and technology innovation policy, the effect of the incentive score on the current enterprise innovation performance is smaller compared to the responsibility score. This conveys a signal to enterprises of "emphasizing responsibility while neglecting incentives," reducing their enthusiasm. While it may enhance the quality of innovation activities to some extent, it slows down the pace and efficiency of innovation activities, exerting a restraining effect on increasing the quantity of innovation patents for enterprises and hindering the improvement of innovation performance.

The policy tone has a significant positive impact on the lagged innovation performance, but the coefficient is noticeably smaller compared to its impact on the current enterprise innovation performance. This suggests that the positive effect of policy tone on enterprise innovation performance diminishes gradually over time. When the science and technology innovation policy text demonstrates a stronger inclination toward innovation incentives, i.e., when the incentive score is higher than the responsibility score, it sends a positive signal to enterprises in the current period, boosting their enthusiasm. However, as the increase in the incentive score leads to a reduction in the innovation performance of enterprises in the second year, the positive impact of policy tone on the lagged enterprise innovation performance weakens.

## 4. Conclusions

To foster the development of the technological innovation sector, the government needs to swiftly establish a flexible and adaptive policy framework to maximize policy effectiveness. Investigating the impact of technology innovation policies on corporate innovation performance holds significant theoretical and practical importance. This study employs text mining techniques to conduct thematic analysis and keyword extraction on national technology innovation policies from 2009 to 2018. It delves into a detailed comparison of the similarities and differences in technology innovation policies among various provinces and cities over the past decade. Combining text sentiment analysis techniques with the inherent features of science and technology innovation policies, a policy polarity analysis is performed to extract incentive scores, responsibility scores, and policy tones, exploring the emotional orientation of the policies. Using manufacturing-related policies in each province as an example, this study integrates text mining techniques with the quantification of policy texts to empirically investigate the impact of the polarity and tone of national technology innovation policies on innovation performance. Although this paper constructs "incentive" and "responsibility" lexicons through literature review, there are some limitations. With the advancement of machine learning technologies, especially the widespread application of techniques like Word2Vec, we plan to employ machine learning methods in future research to enhance the "incentive" and "responsibility" lexicons, making them more comprehensive for more robust research outcomes.

## References

*[1] Wang M, Li Z. Level, type and firm innovation of science and technology innovation policy: an empirical analysis based on survey data[J]. Science and Management of Science and Technology, 2019, 39(11): 20-30.*
*[2] Du W, Song Y, Li J, et al. An analysis of policy evolution and regional differences in the transformation of scientific and technological achievements: A case study of Beijing-Tianjin-Hebei and Yangtze River Delta[J]. Science and Management of Science and Technology, 2020, 38(02): 3-11.*
*[3] Amankwah-Amoah J. The evolution of science, technology and innovation policies: A review of the Ghanaian experience [J]. Technological Forecasting and Social Change, 2021, 30(110): 134- 142.*
*[4] Zhang Y, Geng J, Wang Y. Systematic classification of regional science and technology innovation policies in China: a study based on Zhongguancun data[J]. Journal of Systems Science, 2022, 24(02): 92-95.*
*[5] Justicia De La Torre C, Sanchez D, Blanco I, et al. Text: mining: Techniques, applications, and challenges[J]. International Journal of Uncertainty Fuzziness and Knowledge-Based Systems, 2021, 26(04): 553-582.*
*[6] Zhang B, Li Peng, Chen J, et al. The topic analysis and evolution process of national science and technology innovation policy: the perspective of basic text mining[J]. Science and Management of Science and Technology, 2019, 40(11): 15-31.*
*[7] Kayser V, Blind K. Extending the knowledge base of foresight: The contribution of text mining[J]. Technological Forecasting and Social Change, 2018, 116(18): 208-215.*
*[8] Yang Y, Zhou Q, Li Wei. The impact of pilot policies on innovation performance of innovation-oriented enterprises: Empirical evidence from micro-enterprises[J]. Economic Review, 2018, 20(01): 91-105.*
*[9] He K, Wang Y, Zhang L, et al. Tax incentives, innovation output and innovation efficiency: An empirical test based on the policy of additional deduction for R&D expenses[J]. East China Economic Management, 2020, 34(01): 37-48.*