

Can the Opening of Government Data Improve the Level of Enterprise Innovation?

Xinyu Long

School of Business, Hunan University of Science and Technology, Xiangtan 411100, China

Abstract: In the era of big data, the massive amount of government data, as an important national strategic resource, has a significant role in improving governance efficiency and promoting economic growth. However, whether the opening of government data can gather innovation factors, improve the efficiency of enterprise resource utilization, and thus improve the level of enterprise innovation needs further in-depth argumentation. Therefore, this paper establishes a dynamic difference-in-differences model and uses the enterprise sample data from 2011 to 2021 to study whether the opening of government data can improve the level of enterprise innovation. Based on the results of benchmark regression and various robustness tests, the empirical results show that the construction of government data open platforms can significantly and continuously improve the level of enterprise innovation, and the opening of government data can promote enterprise innovation in non-state-owned enterprises. This paper reveals the value of government data opening from the perspective of enterprise innovation value creation and provides an important decision-making reference for promoting the high-quality opening of government data and the innovation-driven development strategy.

Keywords: Government data opening; Enterprise innovation; Dynamic difference-in-differences model.

1. Introduction

With the continuous development of digital technology, data, as a new element, has its value repeatedly emphasized. In the digital age, the government's digital transformation to enhance governance capabilities is becoming increasingly important. The data resources generated by the government in performing public duties cover areas such as resource reserves, industrial and commercial administrative penalties, and the credit status of enterprises. These data are not only rich in type and huge in scale but also key, basic, and authoritative in the field, and are regarded as national basic strategic resources.

As an important part of digital government construction, the open application of massive government data in government decision-making, public services, social governance, etc. is a major step in promoting the modernization of the national governance system. In October 2019, the Fourth Plenary Session of the 19th Central Committee of the Communist Party of China proposed to "improve the mechanism for the market to evaluate the contribution of production factors such as labor, capital, land, knowledge, technology, management, and data and determine remuneration based on contribution", explicitly proposing for the first time that data can be used as a factor of production to participate in production and allocation. In the process of data element circulation, data coordinates with traditional elements to optimize production relations and promote the development of productivity. At this stage, data element circulation, as a prerequisite for the market-oriented allocation of data elements, mainly takes the form of data open activities.

China's first government data open platform, the "Shanghai Municipal Government Data Service Network" (now the "Shanghai Public Data Open Platform"), was launched in 2012. It wasn't until the State Council issued the "Action Outline for Promoting Big Data Development" in 2015 that related research and practice entered a rapid development

period. In recent years, Shandong, Hunan and other places have successively released systems for public data opening, which reflects the high degree of emphasis on data application, value mining and precise guidance at the institutional level. However, data can only exert its maximum effect when fully utilized, yet in reality, a large amount of government data can't be fully used as competent authorities are unwilling or unable to open it to the public.

To promote public use of government data, governments worldwide have actively promoted government data opening in recent years, that is, providing original government data to the public free of charge and without discrimination. Since the "Action Outline for Promoting Big Data Development" issued by the State Council in 2015 proposed "accelerating the open sharing of government data," it has been elevated to a national strategic plan as a major initiative to promote the digital economy strategy and digital government construction. For example, the "Opinions of the CPC Central Committee and the State Council on Building and Perfecting the Market-Oriented Allocation System and Mechanism of Factors" regards "promoting the open sharing of government data" as the primary measure to "accelerate the cultivation of the factor market." In addition, documents such as the "14th Five-Year Plan and the Outline of the Long-Range Objectives for 2035" and the "14th Five-Year Plan for Digital Economy Development" also emphasize the important role of government data opening in promoting the development of the digital economy and digital government construction.

Overall, government data opening provides enterprises with rich data resources and cooperation opportunities for innovation, enabling them to better adapt to the market, improve competitiveness, and make breakthroughs in technological and business model innovation. Practically, some studies suggest that government data opening can promote innovation by enhancing the entrepreneurial vitality and talent gathering degree of the urban information data industry (Huang Xianhai et al., 2023).

Public data opening led by the government started earlier,

with local governments being more active in building data opening platforms. According to the statistics in the "China Local Government Data Opening Report (2022)", as of the second half of 2022, 208 provincial and municipal local governments in China had launched government data opening platforms. Then, in the current context where the value of data elements is prominent, whether the construction of such platforms can help accelerate the implementation of the innovation-driven development strategy and improve enterprise innovation vitality is particularly important.

2. Policy Background and Literature Review

Existing literature on government data opening mainly focuses on public management, researching its theoretical basis, platform construction, and the realization of data's public value. Government data opening is an extension of government functions and citizens' right to know in the big data era. However, there are practical difficulties in transforming government information disclosure to government data disclosure in the big data era (Chen Shanglong et al., 2016). These include weak legal system protection, low effectiveness of State Council administrative regulations, difficulty in grasping data opening ownership and scope, ensuring data security, and inadequate supervision of data utilization (Shi Yanjun, 2023).

Research on government data disclosure platform construction has been carried out, covering platform composition, construction paths, policies and regulations, and privacy protection. Some scholars have also studied the evaluation of disclosed data quality and open platform construction. Additionally, some literature explores the co-creation of public value from perspectives like the ecosystem and life cycle (Zhao Xuyao et al., 2019) and embedded symbiosis (Men Lixiang and Wang Conghu, 2021). Meanwhile, some studies have pointed out that different subjects' expectations during government data opening in China are diverse, hindering value co-creation (Zheng Lei, 2015). Solutions have been proposed, such as a value co-creation model from the data supply, use sides, and cross-interaction field (Li Zhijie et al., 2023). The government should provide essential data resources to reduce investment redundancy, strengthen guidance while ensuring platform exposure, and promote data resource utilization (Zhao

Longwen, 2023).

Research on the value release of government data disclosure mainly includes: promoting economic growth (Zhou Quan et al., 2022; Zhang Chen et al., 2023; Du Zhenhua, 2023), improving economic development quality (Zhang Li and Lin Anran, 2023), and promoting regional coordinated development (Fang Jincheng et al., 2023). From a micro perspective, some scholars consider that government data opening can enhance enterprise total factor productivity (Yang Xiuyun and Han Qi, 2023), and from the perspectives of enterprise digital transformation (Cai Feiying and Huang Xiuru, 2022), the entrepreneurial vitality of the information and data industry, and human capital aggregation to explore its driving force on innovation (Huang Xianhai et al., 2023; Ruyu Yan et al., 2023). Some literature also studies how enterprises use open government data, which industries benefit more, and how to achieve service innovation (Magalhaes Gand Roseira C, 2020).

Based on existing research, this paper examines whether the government data opening policy can empower innovation-driven development. Starting from micro enterprises, it analyzes the impact of provincial government data opening on the innovation vitality of listed companies in the local area.

3. Research Design

Innovation activities include innovation of production factors, process innovation, product innovation, technological innovation etc. (Joseph Schumpeter, 1939). When data becomes available and accessible, existing knowledge can be combined with different types of data resources opened by the government to generate new knowledge, thereby promoting technological innovation. Therefore, the opening of government data is considered a favorable infrastructure and important resource for innovative development. Empirical research also shows that, the opening of government data is conducive to promoting the absorption, creation, and dissemination of knowledge, and thus promotes knowledge development of technological innovation. Therefore, the opening of government data. There is an inherent relationship between the implementation of and innovation. Based on previous research, we propose the hypothesis: H1 Government data opening actively enhances increases corporate innovation vitality.

The basic difference-in-differences model is as follows:

$$Innovation_{i,t-1} = \alpha_0 + \alpha_1 OpenData_{it} + \alpha_2 \sum X_{it} + year_t + \mu_i + \varepsilon_{it} \quad (1)$$

Due to the long cycle of patent application authorization, which lags behind government data opening, this paper will lag the explained variable, the number of patent applications, by one period. Control variables at the enterprise and provincial levels are included. At the enterprise level, they include: the proportion of enterprise R&D investment in total revenue (as R&D investment directly affects the number of patent applications, this variable needs to be controlled); the asset-liability ratio and revenue growth rate (to control the financial condition of enterprises). At the provincial level, they include: the industrial structure of the province and per capita GDP. This model focuses on the net impact of government data opening on corporate innovation vitality, which is significantly positive, indicating that government data opening can significantly improve corporate innovation vitality. Year fixed effects and firm-specific fixed effects are controlled respectively.

Measurement of key indicators in the model: The measurement of government data opening is a dummy variable. If the business location (province, autonomous region, municipality) opens after the launch of the government data open platform, it is assigned a value of 1; otherwise, it is assigned a value of 0. Among them, Shanghai and Beijing opened in 2012; Zhejiang Province in 2015; Guangdong and Guizhou Provinces in 2016; Henan, Jiangxi, Shandong, Shaanxi, and Ningxia Hui Autonomous Regions in 2018; Tianjin, Fujian, Jiangsu, Sichuan, Hainan, Qinghai, and Xinjiang Uygur Autonomous Regions in 2019; Chongqing, Hubei, Hunan, and Guangxi Zhuang Autonomous Regions in 2020; Anhui, Hebei, and Gansu Provinces in 2021. The government data platforms in Tibet Autonomous Region and Yunnan Province are under construction and thus not included in this study. The number of patents granted one year later is used as an indicator to measure the innovation vitality of

enterprises.

The data source selects A-share listed companies from 2011 to 2021 as the initial sample, and eliminates the samples from the financial and insurance industries and those with missing data for the variables involved in the basic regression, following the practices of Song Min (2021) et al. and Qian Xuesong et al. (2018). Enterprise-level data come from the database, and provincial-level indicator data are all from the statistical yearbook.

4. Empirical Results Analysis

4.1. Baseline Regression Results Analysis

Table 1 shows the regression results of government data opening and enterprise innovation vitality. Without controlling variables and fixed effects, the regression coefficients of government data opening are significantly positive at the 1% level, regardless of the number of enterprise patent grants or the number of professional applications. After adding control variables and fixing the year, enterprise, and industry effects, the regression coefficient of government data opening has decreased somewhat, but is still significant at the 5% level. The regression results indicate that government data opening has a promoting effect on improving enterprise innovation vitality.

Table 1. Baseline Regression Results

Variables	(1)	(2)	(3)	(4)	(5)	(6)
	license	application	license	application	license	application
OpenData	0.075*** (0.024)	0.056*** (0.023)	0.049** (0.012)	0.038** (0.015)	0.042*** (0.010)	0.023** (0.011)
Control variables	No	No	Yes	Yes	Yes	Yes
Year-to-year fixed	No	No	Yes	Yes	Yes	Yes
Enterprise fixed	No	No	Yes	Yes	Yes	Yes
Sample size	24754	24754	24754	24754	24754	24754
R ²	0.004	0.003	0.120	0.160	0.241	0.238

4.2. Parallel Trend Test

According to the principle of the parallel trend test, if government data is not open, the innovation levels of enterprises in non-open regions and open regions will have the same changing trend. Therefore, this paper conducts a counterfactual test and draws on Qian Xuesong's method to

Considering the large financial differences among enterprises, we control the asset liability ratio and revenue growth rate. We also consider different R&D investments for each enterprise by controlling the proportion of R&D investment to total revenue. Given the economic conditions, population size, and industrial structure of the region where an enterprise is located, which may affect the innovation enthusiasm of enterprises in that region, we control regional-level variables such as per capital GDP and industrial structure. However, many variables that should be considered are still not controlled and need improvement later.

According to the regression results, after considering the control variables, the significance level of the regression coefficients has decreased, but they are still significant at the 5% significance level. From a practical perspective, when the number of enterprise patent authorizations is the dependent variable, a one-percent increase in the degree of government data openness will increase the average number of enterprise patent authorizations by 4.9 percentage points. There are differences in the development status and trends of different industries. This paper further controls industry trends and finds that the regression coefficients of government data openness remain significant, whether for the number of patent authorizations or the number of patent applications, corroborating the hypothesis to some extent.

investigate the component changes of the dynamic difference-in-differences model. If there is no significant difference in the changing trends of enterprise innovation vitality between non-open regions and open regions before the opening of government data, it is considered that the parallel trend test is passed. The model is set as follows:

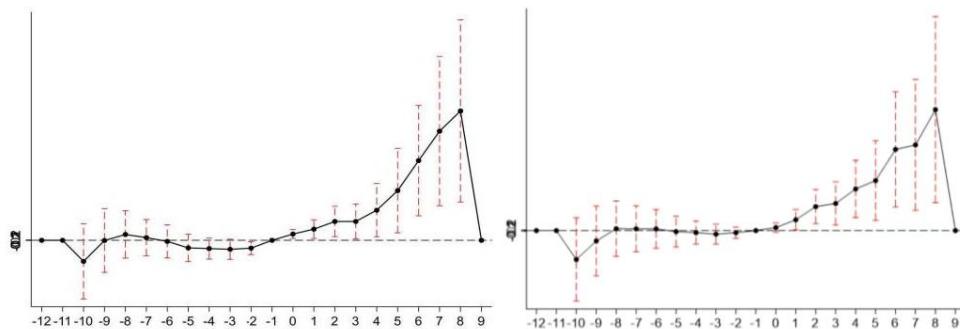


Figure 1. Parallel trend test chart

4.3. Placebo Test

Omitted time and regional variables will lead to biased regression results. Considering that it is difficult to completely control the variables, this paper sets the opening year and region to randomly select sample data for placebo

testing. 1000 sample data were extracted, and the regression results of the extracted samples were obtained to verify that the regression results were not caused by accidental factors of time and region.

Figure 2 shows the regression results of randomly sampled data. Most of the regression results are clustered near zero,

which is far less than the regression coefficient results in Table 1. It can be concluded that the sampled government open data cannot significantly affect enterprise innovation vitality. It can be seen that,The opening of government data to enterprises' innovation vitality.The improvement is not caused by the missing factors of region and time.

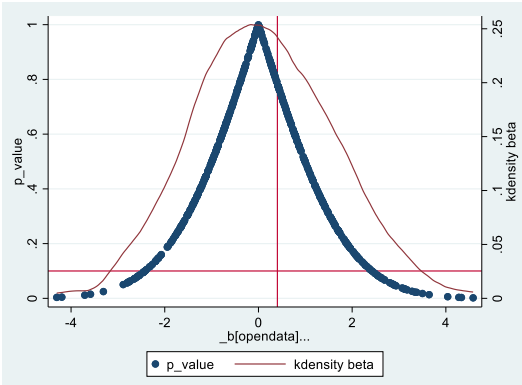


Figure 2. Placebo Test Result Chart

Since most of China's provincial regions have opened government data after 2015, this paper proposes data before 2015, and the regression results are still significantly positive, which further confirms the positive role of government data opening in improving enterprise innovation vitality.

4.4. Heterogeneity Test

The heterogeneity test considers the enterprise level.

Table 2. Grouping tests based on the nature of enterprise ownership

Variables	state-owned enterprise	non-state-owned enterprise	state-owned enterprise	non-state-owned enterprise
	(1)	(2)	(3)	(4)
	license	application	license	application
Open data	0.012 (0.021)	0.019** (0.013)	0.009 (0.022)	0.018** (0.015)
Control variables	Yes	Yes	Yes	Yes
Year-to-year fixed	No	No	Yes	Yes
Enterprise fixed	No	No	Yes	Yes
Sample size	13921	10833	13921	10833
R ²	0.209	0.257	0.213	0.195

5. Conclusion and Suggestions

This study explores and verifies the impact of government data opening measures on enterprise innovation vitality. Based on the data of 29 provinces (autonomous regions, municipalities directly under the central government) and 24,754 listed companies in China, a dynamic difference-in-differences model was employed to study the theoretical assumptions of government data opening in China under the implementation background. The results show that government data opening has a significant positive impact on urban innovation ability, increasing patent authorization (application), and significantly improves the innovation level of non-state-owned enterprises. The empirical results also indicate that the impact of government data opening on improving regional innovation vitality is sustainable, though specific mechanism analysis needs improvement. Based on the above research, the innovation level of government data opening can be further enhanced from the following aspects.

The government should clearly recognize the value of data elements. Whether in the development of the digital economy

Whether the ownership nature of different enterprises has different impacts on the improvement of enterprise innovation vitality in the opening of government data. The existing enterprises in the sample are divided into state-owned enterprises and non-state-owned enterprises. The two are in line. There are differences in many aspects such as government approval efficiency and enterprise performance. Private enterprises generally. They are more flexible in terms of organizational structure and experience management (Peng Yuanhuai, 2023), and can quickly make strategic adjustments to gain greater benefits from the opening of public data; SOEs play a leading role in the socialist market economy and have closer relations with the government, making it easier to obtain public information. According to Table 2results, the improvement of the innovation level of non-state-owned enterprises under the background of government data open platform construction is significant, while government data opening has no obvious effect on promoting enterprise innovation for state-owned enterprises. This shows that the opening of public data does not necessarily bring newer and more unique information to SOEs, but it can make private enterprises with a poorer information environment get information that was not easily available before data opening more quickly and conveniently, is conducive to the gathering of innovation elements by non-state-owned enterprises, improving the circulation efficiency of innovation elements, and thus improving the level of enterprise innovation.

or the construction of a digital government, data serves as its underlying logical support. In particular, as a basic strategic resource of the country, government data, characterized by being key, basic, and authoritative in the field, accounts for over 80% of total social data resources.

Enterprises should enhance data talents' data knowledge skills and competency potential, encouraging them to learn data development professional knowledge and innovation culture from advanced regions and enterprises. Enterprises should make full use of the empowering role of government data opening in improving enterprise innovation efficiency.

While improving the level of data openness, it is necessary to strengthen data supervision to prevent potential risks of data privacy leaks and illegal data applications. The opening of data elements is a "double-edged sword" that, while improving the total factor productivity of enterprises and empowering industrial structure upgrading, also contains unknown risks at present. The government should prepare for potential problems by doing a good job in supervision and legislation.

References

- [1] Shanglong Chen. On the Theoretical Basis of Government Data Opening [J]. Theory and Reform, 2016(06): 104-107.
- [2] Yanjun Shi. Dilemma of Open Utilization of Government Data and Its Legal Guarantee Path in Era of Big Data. [J]. Journal of Beijing University of Posts and Telecommunications (Social Sciences Edition), 2023, 25(06): 70-78.
- [3] Xiaotong Li, Jun Zhai, Guifu Zheng. Research on the Data Quality Evaluation of Open Government Data in Chinese Local Governments: A Case Study of Beijing, Guangzhou and Harbin [J]. Journal of Intelligence, 2018, 37(06): 141-145.
- [4] Xuyao Zhao, Xiaoli Hou, Tangjie Xu et al. The Ecosystem Chain of Government Open Data: Concept, Essence and Types [J]. Journal of Intelligence Science and Technology, 2019, 42(06): 22-28.
- [5] Lixiang MEN, Conghu WANG. An Exploration of the Influencing Factors of the Construction Achievements of Local Government Data Opening in China: Based on the Ecological System Theory Framework [J]. Modern Information, 2021, 41(02): 152-161.
- [6] Lei Zheng. The Value Creation Mechanism of Open Government Data: From the Perspective of Ecosystem [J]. E-Government, 2015(07): 2-7.
- [7] Lixiang Men, Zhimo Zhao, Yalan Li, et al. The Public Value Co-creation Logic, Current Situation and Optimization Path of Government Data Opening in China: Based on the Public Value Strategic Triangle Model [J/OL]. Journal of Intelligence Science and Technology: 1-9.
- [8] Zhijie Li, Zhimei Sun, Pingping Huang et al. Research on the Value Co-creation System and Model of Government Data Open Platform [J/OL]. Journal of Intelligence Science, 2023: 1-22.
- [9] Longwen Zhao, Yifei Hong, Xueqi Zhao. Research on the Evaluation of the Utilization and Innovation Application Efficiency of Government Open Data Resources [J]. Modern Information, 2023, 43(07): 133-144.
- [10] Quan Zhou, Erqi Ge, Xilu Chen, et al. Digital Government Promotes Economic Development: Theoretical Mechanism and Cross-country Evidence [J]. The Journal of Quantitative & Technical Economics, 2023, 40(12): 49-68.
- [11] Chen Zhang, Xiangyu Wan, Zhichao Jiang, et al. Research on the Economic Growth Effect of Open Government Data [J]. China Soft Science, 2023(02): 1-11.
- [12] Li Zhang, Anran Lin. Local Government Data Opening and the Quality of Economic Development in China [J]. Public Administration Review, 2023, 16(02): 120-137.
- [13] Magalhaes G, Roseira C. Open government data and the private sector: An empirical view on business models and value creation [J]. Government Information Quarterly, 2020, 37(3): 101-248.
- [14] Ruyu Yan, Chengju Xing, Xiaojun Chen, Yupan Zhao, Is it real or illusory? An empirical examination of the impact of open government data on innovation capability in the case of China, Technology in Society, Volume 75, 2023, 102396, ISSN 0160-791X, <https://doi.org/10.1016/j.techsoc.2023.102396>.
- [15] Jincheng Fang, Ying Liu, Haoyu Gao, et al. Can the Opening of Public Data Promote Regional Coordinated Development? — A Quasi-natural Experiment Based on the Launch of Government Data Platforms [J]. Management World, 2023, 39(09): 124-142.
- [16] Xianhai Huang, Liming Yu, Ling Dai. Government Data Opening and Innovation-driven: Connotation, Mechanism and Practical Path [J]. Southeast Academic Research, 2023, (02): 102-113.
- [17] Zhenhua Du. The Relationship between Government Data Opening and Innovation-driven Economic Growth [J]. Journal of Capital Normal University (Social Sciences Edition), 2020(02): 63-71.
- [18] Xiuyun Yang, Qi Han. Can the Opening of Public Data Improve the Total Factor Productivity of Enterprises? [J]. Securities Market Herald, 2023, (12): 18-30.
- [19] Feiying Cai, Xiuru Huang. Government Data Opening, Enterprise Digital Transformation and Enterprise Innovation [J]. Statistics and Decision-making, 2022, 38(23): 175-179.
- [20] Joseph Alois Schumpeter. The Theory of Economic Development [M]. China Huaxia Publishing House: 201501. 242.
- [21] Xuesong Qian, Jin Kang, Yinglun Tang, et al. Industrial Policy, Capital Allocation Efficiency and Enterprise Total Factor Productivity — An Empirical Study Based on the Quasi-natural Experiment of China's 2009 Ten Major Industrial Revitalization Plans [J]. China Industrial Economics, 2018(08): 42-59.
- [22] Min Song, Peng Zhou, Haitao Si. Fintech and Enterprise Total Factor Productivity — From the Perspectives of "Empowerment" and Credit Rationing [J]. China Industrial Economics, 2021, (04): 138-155.
- [23] Yuanhuai Peng. The Value Creation Role of Government Data Opening: From the Perspective of Enterprise Total Factor Productivity [J]. The Journal of Quantitative & Technical Economics, 2023, 40(09): 50-70.