The Structure and Substantive Innovation of Corporate Digital Transformation

Minghan Li *

School of Labor and Economics, Capital University of Economics and Business, Beijing 100070, China

* Corresponding Author Email: GWDQLBGYDPL@163.com

Abstract. Under the circumstances of today's digital economy, there have been many discrepancies regarding the success of digital transformation within companies. It is hard for firms to choose between slogans and real actions of digital transformation. The research in previous literature tends to use the digital index, which is more general, to assess the innovation effect, without considering the differences inside the structure within the four dimensions of strategic articulation, technological drive, organizational empowerment, and application implementation, which makes it impossible to explore the differential effects of digital transformation. Therefore, in this paper, I employ the sample of A-share listed companies from 2011 to 2023. Under the framework of the logic of digital implementation, the study decomposes digital transformation into four dimensions, namely strategic articulation, technological drive, organizational empowerment, and application implementation, and explores their heterogeneous effects on substantive and strategic innovation. The empirical results show that the degree of digital transformation has a significant positive relationship with substantive innovation but no significant effect on strategic innovation. By dimensions, technological push and application implementation are significantly positively related to substantive innovation, but organizational empowerment does not have a significant effect, and strategic declarations have a minimal marginal negative effect. More importantly, further investigation finds that technological push can contribute to both forms of innovations, while strategic declarations have a significant negative effect on both forms of innovations. On the whole, the benefits of innovation gained through digital transformation is mostly embodied in the improved quality of substantive innovation, while there is a huge gap between the two dimensions. This paper offers an empirical supplement to digital empowerment of innovation at the micro level from the perspective of structural deconstruction.

Keywords: Digital transformation; digital structure; substantive innovation; strategic innovation.

1. Introduction

In the current situation, innovation within Chinese enterprises encounters several practical problems, including low-quality patent explosion and inability to achieve major innovations in core technologies [11]. In addition, current research only studies the average effect of digital transformation on corporate innovation and does not account for why there is a difference in high-quality innovation performance [7]. Considering the multiple and heterogeneous dimensions of digital transformation, various models have a considerable difference in terms of their ability to facilitate innovation; there cannot be only one measure to properly understand the value of the connection between the two.

In view of the above-mentioned practical issues and the lack of relevant literature, this paper will mainly address three kinds of research questions. Firstly, does the total level of digital transformation facilitate substantive innovation? Secondly, within the multiple dimensions of digital transformation, which one can act as effective channels to facilitate substantive innovation? Finally, does the effect of digital transformation on substantive innovation (such as invention patents) differ significantly from its effect on strategic innovation (utility models, design patents).

The listed companies are used as the sample to conduct empirical analysis, accounting for industry and year fixed effects in the study. According to the findings, there exists a significant positive effect of the level of digital transformation on companies' substantive innovation, which is mainly explained by the variation in the levels of digital capabilities across firms. Specific dimension tests confirm the

positive influence of the technology-driven and application-driven digital transformation on substantive innovation, while the strategic-declaration-based digital transformation exhibits an insignificant negative impact and the organizational empowerment-based digital transformation does not affect the innovation performance significantly. Thus, only substantive transformations that involve developing the underlying technologies and business scenarios can promote innovation, but not slogan-driven transformations. The driving effects of digital transformation on both substantive and strategic innovation show significant heterogeneity.

Research contributions of this paper can be mainly attributed to four aspects. First, theory contribution. The majority of previous studies concentrate on the effect of digital transformation on the innovation of companies as a whole without taking into account the distinctions. This paper makes the distinction between substantive innovation and strategic innovation, analyzes the differences in innovation results of four different types of digital transformation, specifies the distinctions in innovation enabling among different digital approaches, and contributes to further development of theoretical and empirical research on digital innovation [23]. Second, framework contribution. In this paper, the multidimensional structural features of digital transformation and the diversity of innovation behavior are included in a unified research framework, which overcomes the limitation of previous studies using only a single holistic index. Third, practice contribution. Through this paper, it is clarified that there is a difference in innovation performance between formal and substantive digitalization, which solves the dilemma of enterprises in digitalization: form comes first. It gives empirical evidence for optimizing the transformation process of enterprises and formulating digital support policies for the government. Fourth, measurement contribution. This research paper departs from the previous approach which relies on the use of one comprehensive index for digitalization, breaks down the four sub-dimensions of digital transformation, and defines its internal characteristics, accordingly, thereby offering an additional contribution to the measurement of digital transformation [19].

2. Literature Review

The development of digital technologies like artificial intelligence and big data is spurring corporate transformations [14], and digital transformations have the potential to boost corporate governance as well as innovation capabilities through the reformation of technology and business models [4, 20, 22]. Patents, according to innovation motives, can be categorized into substantive innovation and strategic innovation: The former revolves around innovations in core technologies, while the latter is about the volume of patents and their benefits. Strategic innovation is not likely to lead to the iteration of technology [6]. In this study, utility model patents and design patents will be used as indicators of strategic innovation [11, 23]. The proxies are just approximations, but they cannot screen out improvement innovation.

There is an extensive number of studies analyzing the connection between digital transformation and corporate innovation and, in general, proving that digital transformation may help to boost corporate innovation performance and optimize management and product innovation [5, 7, 9, 17, 21]. Nevertheless, existing research gives few insights into the connection between discrepancies in digitalization structure and discrepancies in innovation quality. Researchers have classified innovation into two categories – substantive and strategic, and showed that digital transformation is much more useful when talking about promoting substantive core innovation and helping to cope with the problem of patents "prioritizing quantity over quality." This statement is also confirmed by the studies dedicated to the topic of citation-based patents and digital finance as digital transformation helps the creation of high-quality invention patents and does not create any incentives for low-end strategic innovation [11, 12, 13, 15, 23].

Research on the dimensions of digital transformation divides types of digitalization into groups on the basis of their technical features and shows that innovations in underpinning technologies such as big data and artificial intelligence have more enabling influence than scenario-based innovations, and

simply constructing information systems will not increase innovation efficiency [3, 18, 20, 21]. Nonetheless, current literature makes use of only a simplistic binary division and does not perform an in-depth deconstruction by taking into consideration the whole process of digital transformation implementation, and, therefore, does not allow one to distinguish between different innovation impacts of different types of digitalization. Therefore, in this paper, following the sequence of digital transformation from top-down design to business implementation, a four-dimensional framework consisting of strategic awareness, technological investment, organizational embedding, and business transformation is created for the purpose of performing an in-depth deconstruction of digital transformation [19, 24].

The existing literature tends to divide substantive innovation from strategic innovation on the basis of types of patents. Specifically, the number of invention patent applications measures substantive innovation, while the number of utility model and design patent applications represents strategic innovation. This paper uses the number of patent applications for empirical analysis since the number of applications can be used to gauge the current innovative intentions of a firm, while the number of patents issued is easy to be affected by the patent examination cycle [6, 11, 23]. In the case where a difference exists between substantive innovation and strategic innovation, digital transformation cannot be a unidimensional concept, and different modes of digitalization have different effects on these two innovations. Previous literature has identified strategic digitalization that may involve hype from substantive digitalization that involves hardware investment, which demonstrates that the two have different effects on the innovation process [2, 22].

However, many studies indicate that the incentivization effect of digital transformation towards substantive innovation is higher; implementation of underlying digital technologies can perfect patent structures and reduce patent bubbles, and the advantages brought by digital innovations are achieved exclusively by technology-oriented firms, firms that are well-developed from the standpoint of infrastructural issues and eastern-region firms [7, 10, 15, 17]. The external digital activities also support high-quality innovations; digital mergers and acquisitions (M&A) and digital consumption help R&D of core technologies, although cross-border digital M&A is limited due to institutional reasons and spatial distance [16, 23].

Furthermore, some studies have indicated that digital transformation does not just encourage high-quality innovations but is likely to motivate patent growth strategy. For SMEs which lack resources, digital transformation results in strategic patent growth; although digital transformation itself motivates strategic innovation, it enhances substantive innovation only in conjunction with R&D expenditures [1, 8, 15].

To conclude, current classification systems for digital transformation face some limitations, namely, one system utilizes a composite index approach, which is capable of showing only average effect and concealing heterogeneity within the phenomenon [5], whereas the second system involves the dichotomy of classification with respect to technology features, which means comparing the role of digital technologies only [3, 18]. In addition, both systems overemphasize the role of the technological factor, ignoring other dimensions like corporate strategy and restructuring of organizations, hence, not allowing innovative impacts of non-technological digitalization to be identified properly. The current paper deviates from the logic of technological classification by developing a framework involving four dimensions and reflecting the sequential process of implementing digitalization in an organization.

Although there is research to differentiate the motivations of substantive and strategic innovations, it seems to be lacking in comparative analysis of the intrinsic structure of digitalization and both of these innovations. Strategic innovation has not been used in research as the basis for identifying the role played by different digitalization paths. This makes it hard to determine whether different digitalization strategies motivate innovations to achieve breakthroughs or patent-hunting tendencies and hence identify critical digitalization paths towards innovation improvement.

3. Hypothesis Analysis

According to the analysis of the literature, this paper formulates the following research hypotheses, which include three levels: the impact of digital transformation on the whole, its heterogeneous influence on the two kinds of innovation behaviors, and the heterogeneity of the inner four-dimensional structure:

H1: The more advanced digital transformation is, the higher the level of substantive innovation of the company will be. Digital transformation makes it possible for companies to use data to improve their R&D processes and cut down costs of trial and error in R&D. Those companies with higher digitalization levels have the advantage in performing risky and long-term R&D of core technologies to contribute to substantive innovation results.

H2: Compared with strategic innovation, the relationship between digital transformation and substantive innovation is more positive. While substantive innovation aims at innovations in core technologies, strategic innovation targets gaining policy support and increasing the quantity of patents. The value of digital transformation mainly lies in the innovation of paradigms of R&D, and its contribution to high-quality innovation is much more significant than strategic innovation, which increases the number of patents.

H3: The effect of different dimensions of digital transformation on substantive innovation is not the same.

H3a: The higher the dimension of technology-driven digital transformation, the higher the level of substantive innovation in the enterprise. Underlying digital technologies may increase the capacity of R&D tools, integrate the data of R&D activities, give rise to innovations in core technology, and therefore make substantive innovation possible.

H3b: The higher the dimension of organizational-empowerment-driven digital transformation, the higher the level of substantive innovation in the enterprise. Organizational-level digital transformation will optimize the internal information process and management process; from theoretical point of view, this organizational dimension will facilitate R&D activities positively, and thus there will be positive correlation with substantive innovation.

H3c: The dimension of strategic declaration-driven digital transformation has less positive correlation with substantive innovation than other types of digital transformation, even the negative one. Strategic declarations are often only a digital slogan without any resource input, and excessive hype of digital concepts will cost R&D resources.

H3d: An increase in the degree of application-oriented digitalization leads to an increase in the degree of substantive innovation in the enterprise. The embedding of digital technologies within research and development and production business scenarios converts digital competencies into business competencies, facilitating technological iteration and positively correlating with substantive innovation.

4. Research Design

4.1. Sample and Data Sources

In this study, Chinese A-share listed firms during 2011 to 2023 have been chosen as the research sample. Excluding the firms from the financial industry, ST-listed firms, and those samples that had key variable missing, a total of 7,868 firm-year unbalanced panel observations have been collected. In order to control the effect of outliers on the study results, the data of patents and other continuous variables have been trimmed two ways at 1% and 99%, and the logarithm of the patents variables has been used to deal with right-skewed data. The financial, patents, governance, and digital transformation variables have been collected from CSMAR.

4.2. Variable Definitions

4.2.1. Dependent Variables

Substantive Innovation (InvPatent): Following the study by Li Wenjing and Zheng Manni [6], substantive innovation is measured by the number of annual invention patent applications filed by a firm, calculated as ln (1 + number of annual invention patent applications). Invention patents represent high-tech, high-quality innovation outputs and effectively characterize a firm's substantive innovation behavior in pursuing breakthroughs in core technologies.

Strategic Innovation (StrPatent): Drawing on the research by Li Wenjing and Zheng Manni [6], strategic innovation is measured by the sum of a firm's annual utility model and design patent applications, calculated as ln (1 + number of utility model patent applications + number of design patent applications). These types of patents have relatively lower technical barriers and shorter application cycles, and are more likely to reflect a company's strategic innovation behavior driven by a focus on patent volume. This indicator serves only as a proxy measure and cannot completely rule out the possibility that it includes genuine incremental innovations.

4.2.2. Core Explanatory Variables

Overall Level of Digital Transformation (Digital): Drawing on the research by Zhen Hongxian et al. [19], we adopt the comprehensive digital transformation index constructed in that study as the core metric. This indicator is derived from text mining of listed companies' annual reports. It is synthesized using the entropy-weighted method to weight the word frequency metrics across four sub-dimensions—strategic declaration, technology-driven, organizational empowerment, and application implementation. The field name is `Digitaltransindex` with a value range of 0–100; no additional logarithmic transformation is required. Higher values indicate a higher overall level of digital transformation for the enterprise.

The four sub-dimensions of digital transformation are also derived from this entropy-weighted composite index:

(1) Strategy-led Digital Transformation (Digital_strategy): The score for the strategic leadership dimension reflects the extent to which a company's annual report articulates its top-level digital strategy and planning;

(2) Technology-driven Digital Transformation (Digital_tech): The score for the technology-driven dimension reflects the level of a company's foundational digital technology capability development and investment in technical resources;

(3) Organizational Empowerment-Driven Digital Transformation (Digital_org): The score for the organizational empowerment dimension reflects the extent to which internal organizational management processes have been digitized;

(4) Application-Driven Digital Transformation (Digital_app): The score for the digital outcomes application dimension reflects the level of actual implementation and transformation of digital technologies in the enterprise's production, R&D, and business scenarios.

4.2.3. Control Variables

Drawing on existing literature on corporate innovation, this study selects the following firm-level control variables: firm size (Size), debt-to-equity ratio (Lev), return on assets (ROA), firm growth (Growth), cash flow level (Cashflow), proportion of fixed assets (PPE), largest shareholder's ownership ratio (Top1), board size (Board), and proportion of independent directors (Indep). The model controls for year-specific fixed effects and industry-specific fixed effects, and standard errors are clustered at the firm level. Due to sample data limitations, the baseline model does not include the ownership nature (SOE) variable. As shown in Table 1, the definitions of the variables are presented below.

Table 1. Table of Variable Definitions.

Variable Name	Variable Symbol	Specific Calculation Method
Dependent Variable		
Substantive Innovation	InvPatent	The natural logarithm of a company's annual number of invention patent applications is calculated as ln(1 + annual number of invention patent applications), representing high-tech, high-quality, substantive innovation output.
		The natural logarithm of the sum of a company's annual utility model and design patent applications is calculated as ln (1 + number of utility model patent applications + number of design patent applications), measuring the company's low-quality, superficial innovation output; This indicator serves only as a proxy measure and cannot completely rule out the inclusion of genuine, improved innovations.
Strategic Innovation	StrPatent	The natural logarithm of a company's annual number of invention patent applications is calculated as ln(1 + annual number of invention patent applications), representing high-tech, high-quality, substantive innovation output.
		The natural logarithm of the sum of a company's annual utility model and design patent applications is calculated as ln (1 + number of utility model patent applications + number of design patent applications), measuring the company's low-quality, superficial innovation output; This indicator serves only as a proxy measure and cannot completely rule out the inclusion of genuine, improved innovations.
Key Explanatory Variable		
Overall Level of Digital Transformation	Digital	Referring to the research by Zhen Hongxian et al. [19], the Digitaltransindex—a comprehensive index of corporate digital transformation developed in that study—is adopted as the core metric. Based on the text of listed companies' annual reports, this index evaluates multiple dimensions—including strategic declarations, technological drivers, organizational empowerment, application of digital outcomes, and external environmental support—through text preprocessing and entropy-weighted aggregation. It ranges from 0 to 100 and does not require a logarithmic transformation; a higher value indicates a higher overall level of a company's digital transformation.
Strategy-Driven Digital Transformation	Digital_strategy	The score for the "Strategic Leadership" dimension of the Digital Transformation Index reflects the extent to which a company's annual report articulates its overarching digital strategy and outlines its plans.
Technology-Driven Digital Transformation	Digital_tech	The score for the "Technology-Driven" dimension of the Digital Transformation Index reflects the level of a company's foundational digital technology capabilities and investment in technological resources.
Organizational Empowerment-Driven Digital Transformation	Digital_org	The score for the "Organizational Empowerment" dimension of the Digital Transformation Index reflects the extent to which a company's internal organizational management processes have been digitized.
Application-Driven Digital Transformation	Digital_app	The score for the "Application of Digital Outcomes" dimension of the Digital Transformation Index reflects the level of actual implementation and transformation of digital technologies in the company's production, R&D, and business operations.

Control Variables		
Company Size	Size	Natural logarithm of a company's total assets at year-end
Debt-to-Equity Ratio	Lev	Total liabilities at year-end / Total assets at year-end
Return on Total Assets	ROA	Annual net profit / Total assets at year-end
Company Growth Potential	Growth	Revenue growth rate
Operating Cash Flow Level	Cashflow	Annual net cash flow from operating activities / Total assets at year-end
Proportion of Fixed Assets	PPE	Net fixed assets at year-end / Total assets at year-end
Company Age	Age	Natural logarithm of the number of years the company has been publicly listed
Concentration of Shareholding	Top1	Number of shares held by the largest shareholder at year-end / Total number of shares outstanding
Board Size	Board	Natural logarithm of the total number of board members at year-end
Proportion of Independent Directors	Indep	Number of independent directors at year-end / Total number of board members

4.3. Model Specification

4.3.1. Baseline Regression Model

To examine the impact of digital transformation on firms' substantive innovation, this study constructs the following baseline regression model:

$$InvPatent_{i,t} = \beta_0 + \beta_1 Digital_{i,t} + \gamma Controls_{i,t} + IndustryFE + YearFE + \varepsilon_{i,t} \quad (1)$$

Where $InvPatent_{i,t}$ represents firm i's level of substantive innovation in year t; $Digital_{i,t}$ represents firm i's overall level of digital transformation in year t; $Controls_{i,t}$ is the set of all control variables mentioned above; IndustryFE is the industry fixed effect, used to account for industry-level characteristics that do not change over time; YearFE is the year fixed effect, used to control for common factors that vary over time, such as macroeconomic shocks and policy changes; $\varepsilon_{i,t}$ is the random error term. During the regression process, standard errors are clustered at the firm level to mitigate heteroscedasticity and serial correlation issues.

4.3.2. Decomposition Model of Digital Transformation Sub-Dimensions

To further deconstruct the internal mechanisms of digital transformation and clarify the heterogeneous impacts of different digital implementation paths on substantive innovation in enterprises, this paper simultaneously incorporates four sub-dimensions into the regression model for empirical testing:

$$\begin{aligned} InvPatent_{i,t} = {} & \beta_0 + \beta_1 Digital_strategy_{i,t} + \beta_2 Digital_tech_{i,t} \\ & + \beta_3 Digital_org_{i,t} + \beta_4 Digital_app_{i,t} + \gamma Controls_{i,t} \\ & + IndustryFE + YearFE + \varepsilon_{i,t} \end{aligned} \quad (2)$$

Where $InvPatent_{i,t}$ represents the level of substantive innovation of firm i in year t; $Digital_strategy_{i,t}$ represents the strategic declaration dimension of firm i's digital transformation in year t; $Digital_tech_{i,t}$ represents the technology-driven dimension of firm i's digital transformation in year t; $Digital_org_{i,t}$ represents the organizational empowerment dimension of firm i's digital transformation in year t; $Digital_app_{i,t}$ represents the application and implementation dimension of firm i's digital transformation in year t.

This model incorporates four sub-indicators—strategic declaration, technological drive, organizational empowerment, and application implementation—to compare the differing effects of the four types of digital transformation on high-quality substantive innovation.

4.3.3. Testing Model for the Impact of Digital Transformation on Strategic Innovation

To distinguish the effects of digital transformation on different types of innovation, this study further constructs a regression model with strategic innovation as the dependent variable. First, the model for the overall digital transformation index:

$$StrPatent_{i,t} = \beta_0 + \beta_1 Digital_{i,t} + \gamma Controls_{i,t} + IndustryFE + YearFE + \varepsilon_{i,t} \quad (3)$$

Where $StrPatent_{i,t}$ represents the level of strategic innovation for firm i in year t; $Digital_{i,t}$ represents the overall level of digital transformation for firm i in year t.

Building on this, we simultaneously incorporate the four digital sub-dimensions into the model to construct a decomposition model of digital transformation sub-dimensions, thereby testing the differentiated effects of various digital structures on strategic innovation:

$$\begin{aligned} StrPatent_{i,t} = {} & \beta_0 + \beta_1 Digital_strategy_{i,t} + \beta_2 Digital_tech_{i,t} \\ & + \beta_3 Digital_org_{i,t} + \beta_4 Digital_app_{i,t} + \gamma Controls_{i,t} \\ & + IndustryFE + YearFE + \varepsilon_{i,t} \end{aligned} \quad (4)$$

Where $StrPatent_{i,t}$ represents firm i's level of strategic innovation in year t; $Digital_strategy_{i,t}$ represents the strategic declaration dimension of firm i's digital transformation in year t; $Digital_tech_{i,t}$ represents the technology-driven dimension of firm i's digital transformation in year t; $Digital_org_{i,t}$ represents the organizational empowerment dimension of firm i's digital transformation in year t; $Digital_app_{i,t}$ represents the application and implementation dimension of firm i's digital transformation in year t.

This model incorporates four sub-indicators—strategic declaration, technological drive, organizational empowerment, and application implementation—to compare the differing effects of the four types of digital transformation on strategic innovation. The regression results for strategic innovation serve as a benchmark, illustrating that digital transformation does not simply stimulate firms to expand all types of patents but is more closely associated with high-quality, substantive innovation.

5. Empirical Results

5.1. Descriptive Statistics

Table 2. Descriptive Statistics.

	N	mean	sd	min	max
Digital	7868	37.652	10.528	23.711	66.381
InvPatent	7868	2.032	1.676	0.000	6.227
Size	7868	22.108	1.270	20.035	26.186
Age	7868	2.042	0.799	0.693	3.367
Lev	7868	0.372	0.199	0.050	0.861
ROA	7868	0.043	0.060	-0.208	0.211
Growth	7868	0.161	0.336	-0.490	1.923
Cashflow	7868	0.047	0.069	-0.156	0.249
PPE	7868	0.201	0.138	0.005	0.615
Top1	7868	33.436	14.554	8.860	74.090
Board	7868	2.230	0.166	1.792	2.639
Indep	7868	0.375	0.052	0.333	0.571

Outliers were addressed by using Winsorization of the 1st and 99th percentiles in all continuous variables, descriptive statistics of which can be found in Table 2. The central explanatory variable of digital transformation, Digital (Digitaltransindex), was based on the composite index used by Zhen Hongxian et al. (2023). This composite index is compiled using the text of listed companies' annual reports through text preprocessing and entropy-weighted averaging; it comprises four dimensions, including strategic commitment, technology-based innovation, organizational empowerment, and

digital results application. It is bound between 0 and 100 and does not need any logarithmic transformations; thus, the higher its score, the higher the digital transformation level in the firm.

The dependent variable, substantive innovation (InvPatent), has an average of 2.032, a standard deviation of 1.676, a minimum of 0.000, and a maximum of 6.227. Some enterprises have zero invention patents in the given period, and the average is far lower than the maximum value. Even after taking the logarithmic transformation, large differences remain among the substantive innovation outputs of the firms; the logarithmic transformation can solve the problem of the right-skewed data distribution of the patent data. The variable of digital transformation (Digital) has an average of 37.652, a standard deviation of 10.528, a minimum of 23.711, and a maximum of 66.381. On the whole, the degree of digitalization of the sampled enterprises is at the medium-low level, with most of the enterprises still being at the beginning stage of digitalization; meanwhile, the disparity of the degree of digitalization among the enterprises is significant, which provides rich data for this paper to analyze the innovation effect of digitalization.

5.2. Benchmark Regression

Table 3. Benchmark Regression and Structural Effects.

Variable	(1) InvPatent	(2) InvPatent	(3) StrPatent	(4) StrPatent
Digital	0.016***		-0.003	
	(0.004)		(0.004)	
Digital_app		0.004**		0.002
		(0.002)		(0.002)
Digital_tech		0.012***		0.007***
		(0.002)		(0.002)
Digital_org		0.008		0.004
		(0.005)		(0.004)
Digital_strategy		-0.003*		-0.007***
		(0.002)		(0.002)
Size	0.312***	0.315***	0.015	0.010
	(0.042)	(0.042)	(0.032)	(0.032)
Age	-0.298***	-0.246***	-0.367***	-0.330***
	(0.044)	(0.043)	(0.042)	(0.042)
Lev	-0.122	-0.126	0.311*	0.305*
	(0.185)	(0.182)	(0.170)	(0.168)
ROA	0.123	0.210	0.223	0.296
	(0.512)	(0.508)	(0.494)	(0.492)
Growth	0.101*	0.094*	0.145***	0.142***
	(0.054)	(0.054)	(0.051)	(0.051)
Cashflow	-0.171	-0.179	0.249	0.233
	(0.360)	(0.357)	(0.344)	(0.342)
PPE	-0.977***	-0.954***	-0.004	0.050
	(0.251)	(0.247)	(0.231)	(0.230)
Top1	0.001	0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)
Board	0.557**	0.530**	-0.315*	-0.336*
	(0.227)	(0.227)	(0.188)	(0.187)
Indep	1.226*	1.272*	0.102	0.128
	(0.715)	(0.716)	(0.594)	(0.588)
_cons	-6.406***	-6.576***	2.261***	2.101***
	(0.987)	(0.984)	(0.763)	(0.755)
N	7868	7868	7868	7868
adj. R^2	0.153	0.160	0.129	0.135

Note: Figures in parentheses represent robust standard errors clustered at the firm level; ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. The same applies below.

Column (1) in Table 3 shows the outcomes of the baseline regression analysis, where the industry fixed effect and year fixed effect are considered together. Clustering standard errors is based on the firm level, and the impact of digital transformation on substantive innovation is measured through the cross-sectional variance in firms in the same industry. In regression analysis, the coefficient of digital transformation (Digital) is 0.016, and it is positively significant at the 1% significance level. It implies that the higher the level of digital transformation in the firm, the greater is its innovation output (InvPatent). The adjusted goodness-of-fit of the regression model is 0.153, and it shows that the industry-level and year-level factors, along with different variables, are the important explanatory factors in measuring firm innovation output.

The baseline regression results confirm the validity of H1: The overall level of digital transformation is significantly positively associated with the substantive innovation in the firm. Furthermore, regression results for the substantive innovation and strategic innovation were compared to prove H2: The overall level of digital transformation has no significant impact on the strategic innovation in the firm, showing that H2 is supported.

5.3. Structural Effects

In order to deconstruct the internal process of digital transformation, we decompose digital transformation into four dimensions, namely Digital transformation from technology dimension (Digital_tech), Digital transformation from application dimension (Digital_app), Digital transformation from organizational dimension (Digital_org), and Digital transformation from strategy dimension (Digital_strategy). The model includes both industry fixed effects and year fixed effects with standard error clustering by firms. The impacts of the four different dimensions on substantive innovation (InvPatent) and strategic innovation (StrPatent) are tested respectively; the regression results are reported in columns (1), (2), (3), and (4) of Table 3.

First of all, we analyze the regression results for the overall digital transformation index. The regression coefficients for the overall digital transformation index (Digital) reported in Column (1) of Table 3 are 0.016, which is positive and significant at 1% significance level, suggesting that the digital transformation index as a whole may effectively boost the innovation output of an organization. In Column (3) of Table 3, the regression coefficients of the overall digital transformation index (Digital) on strategic innovation (StrPatent) are -0.003, which do not pass the significance test. By comparing the above baseline regression results, we conclude that the incentives provided by digital transformation only affect high-quality substantive innovation, but not strategic innovation.

As to the sub-dimension regression results for substantive innovation, as presented in Column (2) of Table 3: the coefficient of technology-driven innovation (Digital_tech) is 0.012, significantly positive at the 1% level; the coefficient of application implementation (Digital_app) is 0.004, significantly positive at the 5% level. These regression results show that the key factors driving the improvement of companies' substantive innovation capabilities come from the construction of digital technology foundation and the application of digital technologies in real business situations like R&D and production processes. The coefficient of strategic declarations (Digital_strategy) is -0.003, slightly negative at the 10% level. This suggests that the digital strategy declarations that remain just annual report statements rather than concrete actions are not only ineffective in promoting the generation of high-quality invention patents but also slightly inhibiting the substantive innovation activities. The coefficient of Organizational Empowerment (Digital_org) is insignificant, which means there is no significant relationship between the organizational empowerment dimension and substantive innovation; that is, the digital transformation of internal organizational management processes does not lead to substantive innovation achievements.

The details of the regression analysis of the different dimensions of strategic innovation are shown in column 3 and 4 of Table 3. The regression coefficient for the technology dimension of innovation (Digital_tech) is 0.007, which is statistically significant at the 1% level; the regression coefficient for the strategic articulation dimension of innovation (Digital_strategy) is -0.007, which is statistically significant at the 1% level; while the regression coefficients for the implementation dimension

(Digital_app) and the organizational empowerment dimension (Digital_org) of innovation are statistically insignificant. From the comparison of the results of the regression analysis between the two kinds of innovations, there are distinctions between the effects of different dimensions of digital transformation on innovation. The technology dimension of digital transformation facilitates substantive innovation and strategic innovation but positively affects substantive innovation more significantly, while implementation significantly boosts substantive innovation but not strategic innovation. Strategic declarations significantly inhibit both kinds of innovation, and the inhibiting effect on strategic innovation is stronger, and organizational empowerment does not significantly affect either of the two kinds of innovation.

From the perspective of the research hypotheses presented in this paper, the first hypothesis, H3a, is found to be validated because investment in digital technology infrastructure would help organizations to address technological issues faced by them and would have a positive effect on innovation. The second hypothesis, H3b, is rejected because digitization of organization's management activities is unlikely to result in good innovation performance. Thirdly, the strategic declaration variable shows a weak negative effect at the significance level of 10 percent and does not result in any expected innovation benefits but rather has an inhibiting effect. Lastly, H3d is validated, meaning that the use of digital technologies in business environments can significantly enhance innovation levels.

Taking into account the above findings from all the dimensions, it becomes clear that what will actually lead to the improvement in the quality of innovation as well as innovation growth is something significant in terms of digitalization – like development of digital technologies and use of such technologies in practical situations within the business setting. Empty digital strategies represented only by texts will not be able to improve the quality of innovation and might even result in reducing the number of patents of a company. In other words, it turns out that the incentive effect of digital transformation on innovation is not equal everywhere.

5.4. Robustness Results

5.4.1. Basic Robustness

Table 4. Basic Robustness Test.

Variables	(1) Number of invention patents	(2) InvPatent
Digital	3.105***	
	(1.037)	
Percentage of Technology-Driven Initiatives		3.150***
		(0.836)
Percentage of Digital Applications		2.136***
		(0.774)
Percentage of Strategic Announcements		0.979
		(0.713)
Size	79.892***	0.339***
	(26.265)	(0.042)
Age	-32.978**	-0.234***
	(14.282)	(0.043)
Lev	-41.533*	-0.118
	(22.490)	(0.182)
ROA	-9.637	0.154
	(96.969)	(0.510)
Growth	-8.487	0.086
	(10.658)	(0.054)
Cashflow	110.123	-0.150
	(83.639)	(0.360)
PPE	-71.090*	-1.082***
	(39.445)	(0.249)

Top1	-0.802 (0.764)	0.001 (0.002)
Board	-57.533 (69.902)	0.525** (0.228)
Indep	60.883 (141.661)	1.296* (0.715)
N	7868	7868
adj. R^2	0.093	0.153

For the robustness test of the benchmark regression results, this paper adopts the robustness test in terms of the variable measurement methods and structural dimensions of the core explanatory variables. Specifically, first, the measurement method of the dependent variable is modified by directly adopting the absolute number of invention patent applications in the regression. Second, the total index of digital transformation is disaggregated into four indices of proportions: technology-driven, digital application, strategic commitment, and organizational empowerment. Given that the four proportions' indices take the same value of 1, there is a problem of structural perfect multicollinearity. In order to overcome the above issue and analyze the structural configuration effects of digital transformation, this paper uses the proportion of organizational empowerment as the benchmark reference group and adopts only the other three proportions in the regression analysis.

As shown in Table 4, results of robustness regression analysis indicate that after the alteration of the measure of the dependent variable, the coefficient of the overall degree of digital transformation becomes significantly positive at the 1% level, in line with the baseline results. The results of the regression analysis on the structure dimension demonstrate that, in comparison to the digital transformation of the type of organization empowerment, the coefficients of the ratios of the technology implementation and digital applications become significantly positive, whereas the coefficient of the ratio of the strategic declaration becomes insignificant. This means that technology implementation oriented digital transformation becomes more efficient for the realization of substantial innovation, while digital transformation strategies which are realized only at the strategic level cannot produce any significant innovations. Therefore, the conclusions of empirical verification of research hypotheses put forward in this paper are stable, confirming the robustness of the research findings concerning the innovation effects of digital transformation.

5.4.2. Propensity Score Matching (PSM) Test

Table 5: Overall Matching Statistics

Sample	Pseudo R^2	LR Chi-Square	p-value	Mean Standardized Deviation	Median Standardized Deviation
Before matching	0.053	584.740	0.000	14.700	15.500
After matching	0.002	26.420	0.003	3.900	3.700

In order to address the potential self-selection bias arising from the observable difference in firm characteristics, this research uses propensity score matching (PSM) to conduct robustness tests. More specifically, based on the annual median of the level of digital transformation for a firm, we create the dummy variable for the treatment group HighDigital where if the level of digital transformation of a firm in a year is higher than the annual median, then its value is equal to 1, meaning that it is a treatment group; otherwise, its value is 0, and it is a control group. We choose Size, Lev, ROA, Growth, Cashflow, PPE, Age, Top1, Board, and Indep as matching covariates. One-to-one nearest neighbor matching is conducted, and the threshold is set to be 0.05.

The findings from the balance test are presented in Table 5. Prior to the matching procedure, the treatment and control groups showed a certain degree of significant inter-group differences in terms of many covariates; after the matching, the sample mean standardized difference reduced from 14.7%

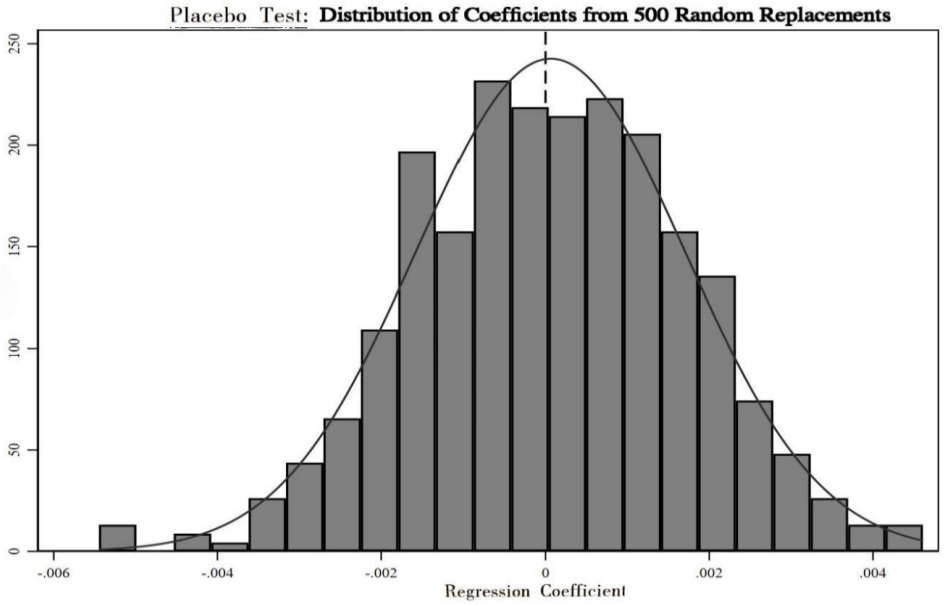
prior to matching to 3.9%, the median standardized difference reduced from 15.5% to 3.7%, pseudo R^2 reduced from 0.053 to 0.002, and even the value of the LR chi-square statistic fell significantly. For the majority of covariates, the significance levels of the between-group t-tests post-matching fell significantly, and the standardized differences for each of the variables became close to zero, suggesting that the observable characteristics of firms in the two groups became more similar after the matching procedure than before.

The baseline regression was estimated using the post-matching sample that complied with the common support assumption (support == 1). Both industry and year fixed effects were controlled for in the regression, with the estimation clustered at the firm level. The regression output revealed that the regression coefficient of digital transformation (Digital) is 0.01537, and the coefficient is significantly positive at the 1% significance level. Therefore, the result of the baseline regression still holds true in the matched sample under the common support assumption. It is necessary to state that the propensity score matching (PSM) method can only address selection bias due to observable characteristics but not endogeneity due to unobservable omitted variables.

5.4.3. Placebo Test

In order to validate if the baseline regression results occurred due to any spurious correlations owing to randomness, this study conducted the random replacement placebo test 500 times. This study kept the same specification of the model in terms of the dependent variable, all control variables, fixed effects of industry and year, and firm-level clustering while randomizing the core independent variable, "Digital," (digital transformation). The baseline regression was run once again with each randomization, thus obtaining 500 instances of spurious regression coefficients. The placebo test mainly helps reject the possibility of spurious correlations owing to randomness but not owing to omitted variable bias.

Table 6: Placebo Test.



As shown in Table 6, the findings of the random permutation test reveal that the simulated pseudo-regression coefficients have values which cluster around 0 and are approximately normally distributed, as well, where the mean of the simulated 500 coefficients is merely 0.0000648 with a standard deviation of 0.0016442. The true coefficient of 0.0155714 in this paper's benchmark regression analysis is obviously situated outside the right tail of the distribution of random coefficients and can therefore be easily distinguished from the spurious coefficients created via random permutation; the p-value obtained through this random permutation test was less than 0.002.

The above findings suggest that the positive relationship between digital transformation and substantive innovation in firms does not originate from any form of random match or random shock. This has effectively ruled out the possible interference of any random elements in the benchmark findings and has further confirmed the validity of the key findings of this paper. Along with the propensity score matching test findings presented earlier, this paper has validated the positive relationship between digital transformation and substantive innovation via two forms of auxiliary robustness tests.

6. Conclusions and Implications

6.1. Main Research Conclusions

This study selects A-share listed firms in China during 2011-2023 as the research sample. Based on the logical structure of digital transformation—from design to practice, it decomposes digital transformation into four dimensions: strategic declaration, technological impetus, organizational empowerment, and application implementation. Empirical investigation is made of the differential effects of both the degree of digital transformation as a whole and each of its sub-dimensions on both substantive innovation and strategic innovation, with a focus on exploring the different innovation effects of "paper-based" digital transformation and "implementation-based" digital transformation. The findings are mainly as follows:

First, a positive relationship exists between the digital transformation of an enterprise as a whole and the innovation development of high quality. Results of benchmark regressions show that the degree of digital transformation has a significant positive correlation with the substantive innovation outcome of a firm; that is, enterprises with higher degrees of digital transformation tend to have higher substantive innovation outcomes. This implies that the primary function of digital transformation is to improve the quality of corporate innovation to a certain extent, but not the extensive expansion of patents.

Second, a dimension-by-dimension analysis shows that technology-focused and application-focused substantive digitalization increases the level of substantive innovation, while strategic declaration type of digitalization does not have any positive impact on innovation and has a slight negative influence on it. Technology-focused digital transformation deals with the development of basic technologies like big data and artificial intelligence; therefore, this type of digitalization leads to an increase in both types of innovations, while application-focused digital transformation is responsible for the implementation of digital technologies in R&D and production and increases only the level of substantive innovation. Strategic declaration type of digital transformation, which is limited only by the declaration, has a negative impact on both types of innovations, while organizational empowerment type of digitalization has no effect on them.

Third, the positive link between digital transformation and innovation mainly exists due to cross-firm differences, where the digital gap across firms leads to different levels of innovation. This paper, at least at this point in time, does not consider the impacts of innovation caused by dynamic changes in the level of digitalization of the firm. After running robustness and endogeneity tests such as variable substitution, structural share regression, PSM, and 500 placebo tests, the results remain unchanged.

6.2. Points of Innovation

From the angle of the structural heterogeneity of digital transformation, this article tries to analyze the different innovative values brought by the formal and substantial parts of digital development. The innovations in this article can be summarized as follows:

Firstly, this article decomposes the digital transformation of companies from the internal structure perspective, refines its internal dimensions, and reviews the intrinsic logic about the impacts of digital transformation on company innovation. In most previous studies, indicators of comprehensive digital transformation are used to analyze its impact on innovation; however, the roles played by different

types of digital developments have not been considered. Also, the innovation effect brought by different types of digital transformation has not been distinguished from the traditional binary classification between technology and application. Therefore, there is definitely a gap in current research that this article could fill.

Secondly, this article systematically decomposes digital transformation into four dimensions: strategic declaration, technological activities, organizational empowerment, and application activities, and uses substantive innovation and strategic innovation as the comparison criteria.

Thirdly, from the perspective of the empirical findings, it is clear that the innovative impacts of diverse digital dimensions are different from each other. Substantive digital efforts—technologically-oriented and business-related implementations—enable high-quality innovations within firms. Digital strategy announcements alone not only do not facilitate innovation but inhibit innovation outcomes. This article demonstrates the differences in value creation between diverse digital development paths, refutes the popular notion of "average impact" in current research, provides empirical findings on the structural level, and expands the research on the micro-level mechanisms through which digitalization enables corporate innovation.

6.3. Practical Implications

According to the empirical evidence obtained through the deconstruction of the concept of digital transformation into four dimensions and the existing problems related to the homogenization of corporate digital transformations and favoring form over execution in reality, this paper provides insights for enterprises to help them promote their digital transformation accurately and create innovations through digital technologies.

The first insight for enterprises is not to have superficial digital activities. Digital activities which are only at the strategic level cannot create innovation and may cause inefficiency. Enterprises need to cut superficial digital activities and use resources for investing in digital technologies and business scenario implementation, thereby making digital activities an endogenous source of innovation.

Secondly, enterprises should improve the allocation of digital resources and pay attention to innovative ways. Enterprises should make more efforts in investing in technology-driven programs and business digitalization, advocate digital technologies in production and R&D, carry out digital transformation cautiously, cut off inefficient strategic investments and capitalize on the innovation-enabling power of digitalization.

Thirdly, enterprises should take advantage of substantive digitalization to avoid patent bubbles and increase innovation quality. Through implementation-oriented digital models, enterprises can make fewer investments in strategic patent expansion and concentrate on producing invention patents of high quality; meanwhile, enterprises can improve their evaluation system to focus on technology implementation, and innovation results in evaluation, thus avoiding digital bandwagoning and hype.

Lastly, the government and regulators should improve the digital innovation ecosystem. They should regulate the disclosure of digital information, ban concept speculation, change supporting modes to implementation-oriented empowerment, and support basic technologies and scenario implementation projects, guiding the real economy to high-quality digital transformation.

7. Limitations of the Study and Future Prospects

Despite the above findings regarding the structure of digital transformation and its differentiated innovation effects, there are some issues that should be mentioned here: First, this paper does not consider firm-level fixed effects; Second, inferences have been made based on cross-sections, which means there is limited causal inference; And third, this paper statically analyzes the effects of the independent variables, without considering the dynamic evolution process of those dimensions, the synergies of those dimensions or heterogeneities. Meanwhile, the measures of digital dimensions through text-mining of annual reports may be influenced by companies' strategic disclosures and fail to reflect the real digitalization of those firms. Moreover, using patents as an indicator of innovation

just reflects the quantity of innovations, but cannot reflect the quality or breakthrough innovations of those companies. In future researches, researchers may continue exploring the dynamic logic of digital transformation, mechanisms of multidimensional synergies of digital transformation and boundary conditions for the innovation effects of digital transformation and improving digital measures with the help of objective data, including investment in hardware and digital talents. Moreover, the quality of innovation could be better estimated by using patent shares and patent citations as indicators.

Reference

[1] Guo, Y. (2018). Signal transmission mechanism of government innovation subsidies and firm innovation. China *Industrial Economics*, (9), 98–116.

[2] Hu, H. F., Dou, B., Zhao, J. L., et al. (2025). Corporate digital catering and business performance. *The* Journal *of World Economy*, 48(12), 209–236.

[3] Huang, B., Li, H. T., Liu, J. Q., et al. (2023). Digital technology innovation and high-quality development of Chinese enterprises: Evidence from corporate digital patents. *Economic Research Journal*, 58(3), 97–115.

[4] Huang, X. H., Sun, Y. M., & Chen, M. T. (2024). Corporate digital transformation and disruptive technological innovation: Micro evidence from patent networks and SBERT model. *China Industrial* Economics, (10), 137–154.

[5] Ji, Y. Y., Zhou, X., & Zhang, Q. (2023). Digital transformation and firm innovation: An analysis from the perspective of R&D investment and R&D efficiency. *Journal of Financial Research*, (4), 111–129.

[6] Li, W. J., & Zheng, M. N. (2016). Substantive innovation or strategic innovation? The impact of macro industrial policy on micro firm innovation. *Economic Research Journal*, 51(4), 60–73.

[7] Pang, R. Z., Liu, L., & Zhang, S. (2023). How does digitalization affect firm innovation? From the perspective of human capital and transaction cost transmission mechanisms. *Nankai Economic Studies*, (2), 102–120.

[8] Shang, H., Yang, Y., & Tian, X. R. (2026). Supply chain linkage effects of substantive innovation in key digital technologies. Science Research Management, 1–16. Advance online publication.

[9] Shen, G. B., & Yuan, Z. Y. (2020). Internetization, innovation protection, and the improvement of export product quality of Chinese firms. *The Journal of World Economy*, 43(11), 127–151.

[10] Song, H. S., Xue, X. R., & Huang, J. (2025). Research on the impact of corporate digital transformation structure on innovation strategies. *Science Research Management*, 46(3), 16–27.

[11] Tan, X. F., & Qian, J. Q. (2020). Capital market pressure and corporate strategic patent behavior: The perspective of short-selling mechanism. *China Industrial Economics*, (5), 156–173.

[12] Tang, S., Wu, X. C., & Zhu, J. (2020). Digital finance and enterprise technological innovation: Structural characteristics, mechanism identification, and effect differences under financial regulation. Management *World*, 36(5), 52–66+9.

[13] Wang, D., & Zhang, Y. (2024). Corporate digital transformation, strategic green innovation, and corporate environmental performance. *Economic Research Journal*, 59(10), 113–131.

[14] Wu, F., Hu, H. Z., Lin, H. Y., et al. (2021). Corporate digital transformation and capital market performance: Empirical evidence from stock liquidity. *Management World*, 37(7), 130–144+10.

[15] Xiao, X., Wang, J. M., & Dong, X. S. (2023). How does digital transformation affect substantive innovation in manufacturing? From the perspectives of "digital empowerment" and "digital divide". *Journal of Zhejiang University (Humanities and Social Sciences)*, 53(10), 28–50.

[16] Xie, H. J., & Qi, X. (2026). Global cross-border digital M&A: Does distance still matter? *Journal of Finance and Economic*s, 52(7), 124–138.

[17] Xie, J. P., Zheng, Y. S., & Han, Z. S. (2025). Business environment empowering digital transformation of manufacturing enterprises and improvement of independent innovation quality. *R&D* Management, 37(1), 14–30.

[18] Zhang, Z. N., Qian, X. Y., & Cao, X. W. (2023). Research on the green innovation effect of corporate digital transformation: Substantive innovation or strategic innovation? *Industrial Economics Research*, (1), 86–100.

[19] Zhen, H. X., Wang, X., & Fang, H. X. (2023). Administrative protection of intellectual property rights and corporate digital transformation. *Economic Research Journal*, 58(11), 62–79.

[20] Fan, Y. Z. (2025). The innovation effect of digital business model transformation orientation: Evidence from corporate cross-boundary innovation. *International Review of Economics & Finance*, 103.

[21] Ge, C. W., Lv, W. D., & Wang, J. L. (2023). The impact of digital technology innovation network embedding on firms' innovation performance: The role of knowledge acquisition and digital transformation. Sustainability, 15(8).

[22] Khanagha, S., Ansari, S., Paroutis, S., & Oviedo, L. (2022). Mutualism and the dynamics of new platform creation: A study of Cisco and Fog computing. *Strategic Management Journal*, 43(3), 476–506.

[23] Pi, A. B., Deng, S. Y., Zhao, W. F., & Wang, J. H. (2025). How does digital consumption affect corporate innovation activity? Evidence from China's information consumption pilot policy. *Sustainability*, 17(21).

[24] Putra, F. H. R., Pandza, K., & Khanagha, S. (2024). Strategic leadership in liminal space: Framing exploration of digital opportunities at hierarchical interfaces. *Strategic Entrepreneurship Journal*, 18(1), 165–199.