Asian Journal of Applied Economics Vol. 32 No. 1 (January-June 2025): 125-152

ISSN: 2985-1610 (Online)

Received: 25 September 2024
Received in revised form: 20 February 2025
Accepted: 3 March 2025

Economic Implications of Digital Transformation on Pollution Reduction: A BART Analysis of Firm-Level Data

Xiaohui Xu*
Zhejiang Sci-tech University, People's Republic of China
Xiaoshi Chen
Zhejiang Sci-tech University, People's Republic of China

Abstract

Prior research has primarily examined the average economic effects of digital transformation, with limited attention to its heterogeneous treatment effects. This study applies the Bayesian Additive Regression Tree (BART) approach to analyze how digital transformation influences firms' pollution emissions and its broader economic implications. Using a dataset of 32,340 firm-year observations from Chinese A-share listed companies (2007–2022), we find that the impact of digital transformation on pollution emissions varies significantly across firms. The key economic mechanisms driving this effect include increased green innovation, more efficient factor allocation, and enhanced firm positioning within social networks. These findings offer new insights into the role of digital transformation in corporate environmental strategies, firm productivity, and economic sustainability, highlighting its differential effects across firms.

Keywords: digital transformation; enterprise emissions; green innovation; BART
JEL Codes: O33; Q53; C14

***Corresponding Author**, Address: Xuelin St, Jianggan District, Hangzhou, Zhejiang, China, 314423. Email: xuxiaohui@zstu.edu.cn

1. Introduction

China's initial phase of rapid economic expansion led to severe environmental pollution, making pollution control a crucial aspect of green development. In response, the Chinese government has set ambitious targets, including reaching carbon peak by 2030 and achieving carbon neutrality by 2060, demonstrating a strong commitment to combating climate change. As these goals are pursued, innovations in artificial intelligence, blockchain, cloud computing, and data analytics (Hallioui et al., 2022) are driving a digital transformation. This transformation is increasingly integrated into enterprise decision-making, leading to a substantial reorganization of traditional production resources and technological strategies (Miao & Chen, 2022).

Can digital transformation reduce pollution emissions? If so, what mechanisms drive this effect? Investigating these questions is crucial for maximizing the environmental benefits of digital transformation within firms and for providing valuable insights to countries pursuing sustainability and enhanced environmental stewardship.

Digital transformation involves adopting digital technologies and leveraging data elements. With characteristics such as scalability, editability, openness, and connectivity, digital transformation significantly reduces costs and mitigates information asymmetries in resource search, factor allocation, and green technology innovation (Bharadwaj et al., 2013; Vial, 2019). As a result, it enhances firms' ability to acquire, integrate, reorganize, and utilize resources for pollution control (Mikalef et al., 2019). Beyond improving the efficiency of internal resource allocation, digital transformation also strengthens external relationship networks, which are critical for accessing external resources and technologies, thereby reducing innovation risks (Borgatti & Foster, 2003). Through the strategic use of social networks, firms can access a broader spectrum of knowledge and capabilities that drive innovation (Tsai & Ghoshal, 1998). Additionally, these networks facilitate information exchange and collaboration among various stakeholders, which is essential for integrating diverse resources (Ahuja, 2000). Therefore, understanding the mechanisms by which digital transformation impacts pollution emissions is of paramount importance.

This study engages with two key strands of literature: one examining the factors influencing pollution reduction and the other analyzing the effects of digital transformation. In the context of pollution reduction, scholars have investigated macro- and micro-level factors, laying the groundwork for understanding how digital transformation might influence them. At the macro level, the literature has explored the effects of economic growth (Grossman & Krueger, 1995), environmental policies (Shahzad, 2020), technological innovation (Hao et al., 2020), and domestic and international trade (Copeland & Taylor, 2001) on environmental pollution and quality. At the micro level, research has focused on how factors such as financial performance (Trumpp & Guenther, 2017) and financing constraints (Goetz, 2018) influence corporate environmental performance.

Studies on corporate digital transformation have examined its implications for various areas, including economic development (Glinskiy et al., 2020), labor structure (Dou et al., 2023), industrial organization (Imran et al., 2021), innovation (Li et al., 2023), and environmental sustainability (Feroz & Chiravuri, 2021). Research on the relationship between digital transformation and environmental quality has primarily focused on its effects on green total factor productivity (Wang et al., 2023), green innovation (Feng et al., 2022), the green development of manufacturing firms (Chen et al., 2023), and carbon emissions (Shang et al., 2023). While much of the literature highlights the positive impact of digital transformation on green economic transformation and environmental quality, several gaps remain.

First, most studies focus on regional or industrial-level analyses, with limited attention to micro-level assessments of enterprise digitalization and pollution emissions. Second, existing research relies heavily on panel data and mean effect tests, lacking a detailed exploration of individual treatment effects of corporate digital transformation on pollution reduction. Third, while internal mechanisms, such as productivity improvements, are well-studied, external mechanisms, particularly the role of social networks, remain overlooked.

To bridge these gaps, this study leverages data from Chinese A-share listed manufacturing firms from 2007 to 2022. It develops firm-level indices for digital transformation, pollution emissions, and social network positions, utilizing the Bayesian Additive Regression Trees (BART) method to assess the individual treatment effects and internal mechanisms of corporate digital transformation on pollution reduction. The BART method is particularly advantageous because it can model complex, nonlinear relationships and interactions among variables, thereby reducing the risk of model misspecification. Furthermore, its ability to estimate individual treatment effects allows this study to identify variations in how digital transformation affects pollution emissions across firms. By prioritizing the relative importance of predictors, BART enables a deeper investigation into the underlying mechanisms driving these effects, making it a powerful tool for analyzing the multifaceted influence of digital transformation. The findings reveal that the impact of digital transformation on pollution emissions is not uniform across firms, with green innovation, optimized factor allocation, and enhanced social network positions serving as the key channels driving these effects.

This study makes several contributions to the literature. It provides a detailed micro-level analysis of both internal and external mechanisms associated with digital transformation, offering deeper insights into its impact on pollution reduction. Additionally, it introduces novel approaches, such as textual data analysis, to measure enterprise digitalization more precisely. By integrating these methodologies within a unified analytical framework, the study clarifies the key mechanisms through which digital transformation contributes to environmental improvements. Furthermore, by applying the Bayesian Additive Regression Trees (BART) method, this research advances existing methodologies and uncovers individual treatment effects that have previously been overlooked.

The paper is structured as follows. Section 2 outlines the theoretical framework and research hypotheses, providing the foundation for the study. Section 3 details the research design, including data sources, variable definitions, and methodology. Section 4 presents the empirical results, highlighting the key findings. Section 5 explores the underlying mechanisms, explaining how digital transformation influences pollution reduction through green innovation, factor allocation, and social networks. Finally, the last section summarizes the study, discusses policy implications, and suggests avenues for future research.

2. Theoretical Analysis and Hypotheses Development

2.1 Green Innovation Enhancement Effect

Green innovation refers to the creation and application of environmentally sustainable technologies, products, and processes that reduce pollution while maintaining economic and social benefits (Liu & Xiao, 2022). It encompasses both technological advancements and organizational practices aimed at minimizing environmental impacts. In the context of digital transformation, green innovation plays a particularly crucial role, as digital technologies enhance resource efficiency, improve collaboration, and accelerate innovation cycles.

Digital transformation facilitates the integration of digital technologies and data elements into production and business processes (Zhang, H. et al., 2024), thereby strengthening firms' ability to control pollution through enhanced innovation. First, digital transformation can increase R&D investment in digital technologies (Zhao et al., 2022), reducing reliance on traditional production factors and fostering innovation in clean and efficient production methods. The adoption of digital technologies and data elements supports the upgrading of corporate production models, leading to shorter R&D cycles for clean technologies (Tang et al., 2023).

Second, enterprise digitalization enhances resource openness and knowledge sharing, enabling the efficient dissemination of data elements and technology (Feng et al., 2022). This process encourages greater interdepartmental collaboration within enterprises during technology development while amplifying the spillover effects of technology and knowledge. As a result, enterprises improve their capacity for co-innovation and integration of green innovation resources (Hu et al., 2024).

Finally, digital transformation enables enterprises to rapidly access market competition information and anticipate technological trends, thereby intensifying R&D and innovation efforts in green and clean products (Ren et al., 2023). Based on this, we propose the following hypothesis:

H1: *Digital transformation reduces pollution emissions by enhancing green technology innovation.*

2.2 Factor Allocation Optimization Effect

The optimization of factor allocation in firm production is primarily reflected in enhanced input efficiency and improved utilization of resources and production factors. Digital transformation strengthens the integration of data elements and technology in resource allocation, improving overall resource efficiency and reducing pollution emissions.

First, the data-driven nature of digital transformation allows for the adoption of technologies that are low in pollution, energy consumption, and marginal costs (Chen & Hao, 2022). The increased incorporation of data, knowledge, and information resources into production processes helps firms reduce dependence on high-energy-consuming, inefficient, and polluting production factors, leading to a more sustainable input structure (Shang et al., 2023).

Second, digital transformation, powered by data and technology, fosters the creation of new forms of labor, new technologies, and innovative production elements (Bresciani et al., 2021). These elements interact dynamically with traditional production factors, encouraging firms to increase investments in digital technologies and data-driven decision-making. This iterative process optimizes resource allocation structures, increasing the proportion of knowledge-intensive and technology-driven inputs, which ultimately supports pollution reduction efforts (Lv & Wu, 2024).

Based on this, we propose the following hypothesis:

H2: *Digital transformation reduces pollution emissions by optimizing factor allocation.*

2.3 Social Network Position Improvement Effect

Network position arises from the relationships established among actors and is a pivotal concept in social network analysis (Qian et al., 2010). A firm's network position significantly influences its learning effectiveness, making it a crucial form of social capital, often referred to as location capital (Huang & Wang, 2008). An improved social network position enables firms to gain better access to innovation-related information while expanding their information channels and collaboration opportunities. Firms with strong network positions are more likely to acquire complementary skills from other enterprises and engage in strategic partnerships with high-performing firms.

Digital transformation plays a key role in enhancing firms' communication with neighboring businesses, fostering a sense of belonging and social recognition that strengthens trust-based relational networks. Within these networks, shared behavioral norms and social standards increase social proximity among firms, leading to greater informal cooperation (Capello, 2014; Granovetter, 1985). These cooperative relationships, supported by mutual trust, help reduce environmental governance costs and ensure that firms adhere to established norms and market regulations in pollution control practices (Ostas, 2003).

Furthermore, digitalized relational networks expand firms' access to diverse resources, including key production factors, technological knowledge, and information exchange platforms. This facilitates the discovery of new innovation opportunities and enhances firms' ability to identify and utilize valuable resources for environmental innovation and pollution control (Uzzi, 1996). By stabilizing inter-firm relational networks, digital transformation maximizes efficiency gains and pollution reduction benefits associated with network externalities and social proximity (Borgatti & Foster, 2003).

Based on this, we propose the following hypothesis:

H3: *Digital transformation reduces pollution emissions by improving firms' social network positions.*

2.4 Heterogeneous Effect of Digital Transformation on Pollution Emissions

Existing research suggests that enterprise digital transformation has the potential to reduce pollution emissions through multiple mechanisms, including green technology innovation, optimized factor allocation, and improved social network positions. However, the extent of these effects varies across firms, indicating heterogeneous impacts. These differences in outcomes may stem from varying levels of digital transformation adoption and its effectiveness in enhancing green innovation, optimizing resource allocation, and strengthening network positions.

First, the impact of digital transformation on green technology innovation differs depending on a firm's capacity to adopt and integrate digital technologies into its R&D processes. Firms with high digital maturity are more likely to experience greater advancements in green technology innovation, as they can effectively utilize digital tools to enhance sustainable production practices. Conversely, firms with low digital adoption may struggle to fully leverage digital transformation, leading to less impactful green innovations (Feng et al., 2022; Zhao et al., 2022).

Second, the ability of digital transformation to optimize factor allocation also varies across firms. Companies with better access to digital infrastructure, financial resources, and skilled labor are more likely to reallocate resources efficiently toward low-pollution, high-efficiency production processes. In contrast, firms with limited resources or inadequate digital infrastructure may face constraints in optimizing factor allocation, thereby limiting their ability to reduce emissions (Lv & Wu, 2024; Shang et al., 2023).

Third, digital transformation's impact on firms' social network positions is highly dependent on their level of integration into digital networks. Firms that are well-connected can leverage digital platforms to enhance collaboration, knowledge sharing, and access to cutting-edge technologies, all of which contribute to pollution reduction. On the other hand, firms that are less integrated into digital networks may not fully benefit from external knowledge spillovers and innovation-driven sustainability efforts, resulting in weaker pollution reduction outcomes (Capello, 2014; Granovetter, 1985).

Based on these varying effects, we propose the following hypothesis:

H4: *The impact of digital transformation on pollution emissions varies across firms, depending on differences in green innovation, factor allocation optimization, and social network position improvement.*

3. Research Design

3.1 Data

This study focuses on Chinese A-share manufacturing firms. Financial data were obtained from the China Stock Market & Accounting Research (CSMAR) database, while pollution emission data were sourced from annual reports, corporate social responsibility reports, and company websites. To ensure data reliability and representativeness, firms with Special Treatment (ST) or Particular Transfer (PT) statuses, along with observations containing significant missing data, were excluded from the sample.

In this study, "ST" (Special Treatment) refers to a status assigned by the stock exchange to companies that have suffered losses for two consecutive years or have abnormal financial conditions. These companies are subject to special regulatory requirements and trading restrictions. "PT" (Particular Transfer) refers to a transfer mechanism for stocks of companies that have encountered serious financial difficulties and are at risk of delisting. Such companies also fall under special regulatory oversight.

To manage potential data distortions, winsorization was applied at the 1% and 99% levels. This method addresses extreme values that could disproportionately impact statistical analyses. Specifically, values below the 1st percentile were replaced with the 1st percentile value, and values above the 99th percentile were replaced with the 99th percentile value. Winsorization reduces the influence of extreme outliers, thereby improving the accuracy of statistical inferences and regression analyses. Additionally, observations with missing values in key variables were excluded.

The final dataset comprises 3,645 firms, resulting in 32,340 firm-year observations spanning the period 2007 to 2022.

3.2 Variables

3.2.1 Firm Pollution Emissions

The pollution emission data primarily encompass five types of pollutants: chemical oxygen demand (COD) and ammonia nitrogen for water pollution, and sulfur dioxide (SO_2), nitrogen oxides (NO_x), and particulate matter for air pollution. To calculate a firm-level composite pollution emission index, we follow these steps:

First, the raw data is standardized using the extremum method, resulting in the standardized emission of the *k*th pollutant. The formula is as follows:

$$rpol_{kit} = \frac{pol_{kit} - min\, pol_{kit}}{max\, pol_{kit} - min\, pol_{kit}}, \tag{1}$$

Where subscriptions k, i, t represents the kth pollutant, firm and year, respectively.pol indicates the amount of the pollutant emission. Max and min respectively represent the maximum and minimum values of the kth pollutant emissions from all firms in a year.

Next, to calculate the coefficient ϕ for the *k*th pollutant for firm *i*, which is calculated as the ratio of the standardized emission of the *k*th pollutant to its mean standardized emission across firms. The formula is:

$$\phi_{nit} = \frac{rpol_{nit}}{\overline{rpol_{nit}}}, \tag{2}$$

Finally, the composite pollution emission index (*P*) for firm *i* is determined by summing these coefficients for all pollutants, see

$$P_{it} = \frac{1}{5} \sum_{n} (rpol_{nit} \times \phi_{nit}) \tag{3}$$

3.2.2 Enterprise digital transformation

Following existing studies (e.g., Gao et al., 2022; Huang et al., 2021), we employ a text-based methodology to develop an index of firm-level digital transformation. The construction of this index involves several steps. First, we collected the content of annual reports from 2007 to 2022. Next, we refined keywords in five key areas - Artificial Intelligence (AI), Blockchain, Cloud Computing, Big Data, and Digital Technology Applications - by reviewing relevant academic literature and policy documents to create a digital transformation dictionary. Using this dictionary, we then searched for, matched, and counted the frequency of these keywords in each annual report. Finally, we measured the level of digital transformation (lnSzh) by taking the logarithm of one plus the frequency of the identified keywords.[1]

[1] The detailed steps are presented in appendix.

3.2.3 Other variables

Following Peng & Tao (2022), we incorporate several control variables, including the number of employees, capital, return on assets (ROA), and the cash flow to assets ratio. Detailed descriptions of these variables are provided in Table 1.

The descriptive statistics for key variables are summarized in Table 2. The pollution emission index (P) has a mean of 1.821 and a standard deviation of 0.309, with values ranging from 0.858 to 2.393. This indicates considerable variation in pollution emissions across firms, suggesting that while most firms exhibit moderate pollution levels, some have significantly lower or higher emissions.

For labor (lnl) measured as the natural logarithm of labor amount, the mean value is 7.650, with a standard deviation of 1.387. The minimum value is 2.079, while the maximum reaches 13.223, reflecting significant variability in the labor size of firms in the sample.

The capital intensity (lnk), which reflects the amount of capital used by firms, has a mean of 22.123 and a standard deviation of 1.521, with values ranging from 14.108 to 30.967. This suggests that firms in the sample differ substantially in their capital requirements, with some being highly capital-intensive.

Return on Equity (ROE), a key measure of firm profitability, has a mean of 0.040 and a standard deviation of 0.033. The ROE values range from a negative minimum of -0.007 to a maximum of 0.100, reflecting a broad range of profitability among the firms. Some firms are experiencing losses, while others maintain positive profitability, though the majority fall within a narrow range around the mean.

The cash flow to assets ratio (Cr), an indicator of liquidity, has a mean of 0.047 and a standard deviation of 0.054, with values ranging from -0.039 to 0.134. This suggests that while many firms have low liquidity, some may face liquidity challenges, as indicated by the negative minimum value.

Finally, the digital transformation index (lnSzh), which captures the extent of digitalization within firms, has a mean of 0.982 and a standard deviation of 1.253. The values range from 0 to 5.690, indicating that many firms are still in the early stages of digital transformation. The wide variation in this index highlights significant differences in the level of digitalization across firms in the sample.

Table 1: Definition of control variables

Variables	Sign	Measurement
number of employees	lnl	The natural logarithm of the number of employees
capital	lnk	The natural logarithm of capital stock is represented by the total assets of the enterprise minus the net value of intangible assets and the net value of goodwill.
return to assets	roa	Total asset growth rate (%)
cash flow to assets ratio	Cr	The ratio of net cash flow from operating activities to total assets

Source: Peng and Tao (2022)

Table 2: Descriptive statistics

Var	Obs	Mean	SD	Min	Median	Max
P	32,340	1.821	0.309	0.858	1.883	2.393
lnl	32,340	7.650	1.387	2.079	7.609	13.223
lnk	32,340	22.123	1.521	14.108	21.906	30.967
roe	32,340	0.040	0.033	-0.007	0.034	0.100
Cr	32,340	0.047	0.054	-0.039	0.046	0.134
lnSzh	32,340	0.982	1.253	0.000	0.000	5.690

Source: Authors' calculation

3.3 Methodology

The heterogeneity of treatment effect (HTE) refers to the variation in individual responses to the same treatment. While most studies focus on estimating the average treatment effect (ATE) to summarize the overall impact across an entire sample, traditional methods for analyzing HTE often examine individual characteristics in isolation. These methods attempt to identify differences in treatment effects based on specific variables but often fail to account for the potential synergistic effects that may arise among multiple correlated characteristics. As a result, this approach risks overlooking nuanced relationships between variables and may lead to biased or incomplete conclusions.

To address these limitations, we employ the Bayesian Additive Regression Trees (BART) method, which offers a robust solution for capturing complex heterogeneity in treatment effects. Unlike conventional approaches, BART can identify nonlinear relationships and interactions within the data, hierarchically prioritizing them based on their relative importance. This method minimizes researcher bias and reduces the likelihood of model misspecification, a common issue in traditional interaction testing methods.

In this study, we utilize BART to assess the heterogeneous treatment effects of digital transformation on firms' pollution emissions. By leveraging the flexibility and adaptability of BART, we can uncover subtle variations in how digital transformation influences pollution outcomes across firms, providing a more precise and comprehensive understanding of its environmental impact.

BART is designed to estimate a model for the outcome *Y*, specified as $Y = f(z, x) + \varepsilon$, where *z* represents the treatment, *x* denotes the confounding covariates, and ε represents independent and identically distributed (*iid*) errors. Tree-based models explain the variation in an outcome variable by repeatedly partitioning the sample into increasingly homogeneous subgroups. The construction of a tree begins with the root node, which includes the entire sample. Nodes in the tree structure can be terminal (leaf nodes) or non-terminal (intermediate nodes), with non-terminal nodes always splitting into two daughter nodes. These splits are determined by Boolean questions about a single predictor. For example, a split is made based on a condition such as: $X_i \leq \theta_j$ where X_i is the value of a predictor variable for observation i, and c is a threshold value. Depending on whether the answer to the Boolean condition is "yes" or "no," each observation in the node is assigned to one of the two daughter nodes. Unlike single-decision tree models, BART employs an ensemble of m trees, where m typically ranges in the hundreds. This ensemble approach allows BART to capture complex relationships between variables while reducing the risk of overfitting. BART can be applied to both binary and continuous outcomes. For continuous outcomes, BART models the outcome Y as an unknown function f of a p-dimensional predictor vector x plus an *iid* error term:

$$Y = f(x) + \varepsilon \tag{4}$$

We begin by introducing the notation for a single tree. Let T represent a tree consisting of a set of interior nodes, terminal nodes, and the decision rules that connect these nodes. In other words, T encapsulates all the information necessary to define a decision tree model. Let M denote a set of parameters associated with the b terminal nodes of T. In a single tree model, these b parameters represent the fitted values for the terminal nodes. Given T and M, the output of the function $g(x;T,M)$ is obtained by first dropping an observation with characteristics x down the tree until it reaches a terminal node, and then reporting the value M associated with that terminal node:

$$g(x;T,M)=M(x) \tag{5}$$

Instead of fitting a single tree, BART fits an ensemble of mmm trees, typically in the hundreds. This sum-of-trees model is expressed as:

$$Y = \sum_{j=1}^{m} g_j\,(x; T_j, M_j) + \varepsilon \tag{6}$$

For each T_j and its associated set of terminal node parameters M_j , the output of $g_j(x;T_j,M_j)$ is the value obtained by dropping an observation with characteristics x down the tree until it reaches a terminal node, then reporting the appropriate terminal node parameter $M_j(x)$. Under this model, Y equals the sum of all terminal node parameters assigned to an observation with characteristics x by g_j . Each tree T_j may represent a main effect when $g_j(x)$ depends on only one component of x, or an interaction effect when $g_j(x)$ depends on more than one component. In this way, BART naturally incorporates both main and interaction effects, with some trees representing main effects and others capturing interactions. Since the trees in the ensemble can vary in size, the interactions can be of varying orders, allowing BART to provide extremely flexible fits, with each tree specializing in fitting a particular aspect of the data.

4.Empirical results

4.1 Baseline results

Figure 1 illustrates the individual treatment effect of digital transformation on pollution emissions for each observation. The results are categorized into three groups, represented by green, orange, and blue scatter points. The green scatter points, which are statistically significant at the 5% level, indicate that digital transformation reduces pollution emissions. This finding aligns with the theoretical perspective that digital transformation enhances production efficiency and resource utilization, ultimately leading to lower pollution levels. Firms that adopt advanced digital technologies can improve their production processes, optimize energy consumption, and thereby reduce pollutant emissions. This result is consistent with prior studies, which have demonstrated that digitally transformed firms are better equipped to manage resources efficiently and implement cleaner production methods (Sun et al., 2024; Zhang, C. et al., 2024).

The orange scatter points, which represent insignificant effects, suggest that for some firms, digital transformation may not have an immediate or direct impact on pollution emissions. Several factors could explain this result, including industry characteristics, the stage of digital adoption, and the presence of other confounding variables. In traditional manufacturing industries, for example, the adoption of digital technologies may be more challenging, and the expected pollution reduction benefits may take longer to materialize due to technological adaptation constraints and capital investment requirements.

The blue scatter points, which are statistically significant at the 5% level, indicate a positive relationship between digital transformation and pollution emissions. This seemingly counterintuitive finding may be attributed to short-term disruptions caused by the implementation of digital technologies. During the initial phase of digital transformation, firms may experience operational inefficiencies or increased energy consumption, as they adapt to new systems and integrate digital infrastructure (Li et al., 2021). For instance, the adoption of automated production processes and smart manufacturing technologies often requires significant power consumption, software adaptation, and workforce reskilling, which could lead to a temporary increase in emissions before the long-term benefits of efficiency and sustainability improvements are realized.

Overall, Figure 1 clearly illustrates the heterogeneous impact of digital transformation on pollution emissions across firms, emphasizing the complexity of this relationship. These findings highlight the necessity for further investigation into the underlying mechanisms driving these varying effects, particularly the role of industry characteristics, firm capabilities, and the temporal dynamics of digital transformation.

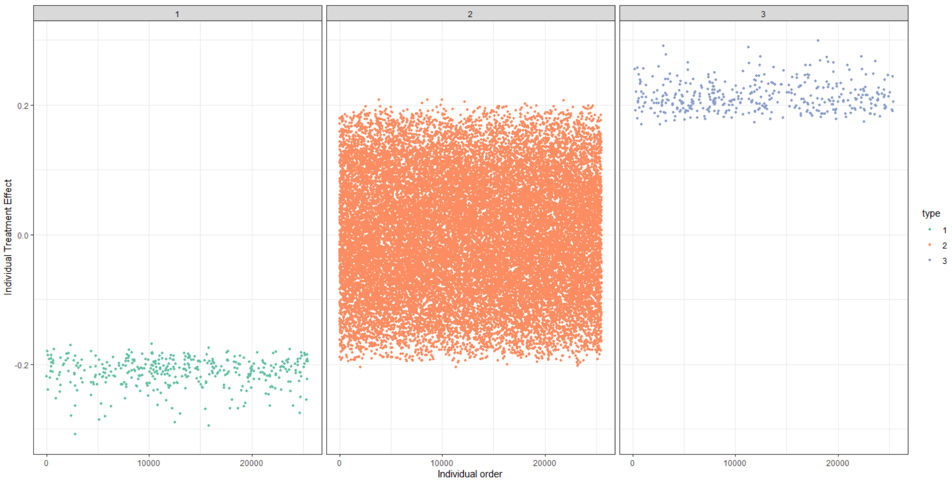

Figure 1: Individual treatment effect of digital transformation on pollution emissions

Source: Authors' calculation

Notes: "1" indicates the coefficients of digital transformation are significant negative at 5% level; "2" indicates the coefficients of digital transformation are insignificant; "3" indicates the coefficients of digital transformation are significant positive at 5% level.

4.2 Robustness checks

4.2.1 Replacing the measurement of pollution emission

The absolute value of pollution emissions may be influenced by firm size, making it less effective in assessing pollution control efficiency. In contrast, pollution emission intensity (PEI) provides a more accurate measure of a firm's pollution control efficiency relative to its production process. Therefore, to ensure the robustness of our findings, we construct a composite indicator based on pollution emission intensity and use it as an alternative dependent variable for robustness testing. The firm's composite pollution emission intensity (PEI) is calculated as follows:

First, we measure the emission intensity for the *k*th pollutant (*poli*) by using the ratio of its pollutant's emission volume (*pol*) to the firm's main business revenue (*revenue*).

$$poli_{kit} = \frac{pol_{kit}}{revenue_{it}} \tag{7}$$

Next, the *poli* are standardized, resulting in the standardized emission of the *k*th pollutant. The formula is as follows:

$$rpoli_{nit} = \frac{poli_{nit} - min\, poli_{nit}}{max\, poli_{nit} - min\, poli_{nit}}, \tag{8}$$

where max and min respectively represent the maximum and minimum values of the kth pollutant emissions intensity from all firms in a year.

Next, calculate the coefficient *w* for the *k*th pollutant for firm *i*, which is calculated as the ratio of the standardized emission intensity of the *k*th pollutant to its mean standardized emission intensity across firms. The formula is:

$$w_{nit} = \frac{rpoli_{nit}}{\overline{rpoli_{nit}}}, \tag{9}$$

Finally, *PEI* for firm *i* is determined by summing these coefficients for all pollutants:

$$PEI_{it} = \frac{1}{5} \sum_n (rpoli_{nit} \times w_{nit}) \tag{10}$$

Figure 2 demonstrates that the impact of digital transformation on pollution emissions varies across firms, which highlights the heterogeneous effects of digital adoption in the context of environmental performance. The varying impact of digital transformation on pollution emissions suggests that not all firms benefit equally from digital technologies. Some firms, particularly those in high-polluting industries or with older infrastructures, may face challenges in implementing digital solutions that significantly reduce emissions. Conversely, firms with greater technological maturity or those in less polluting sectors might experience more pronounced benefits from digitalization.

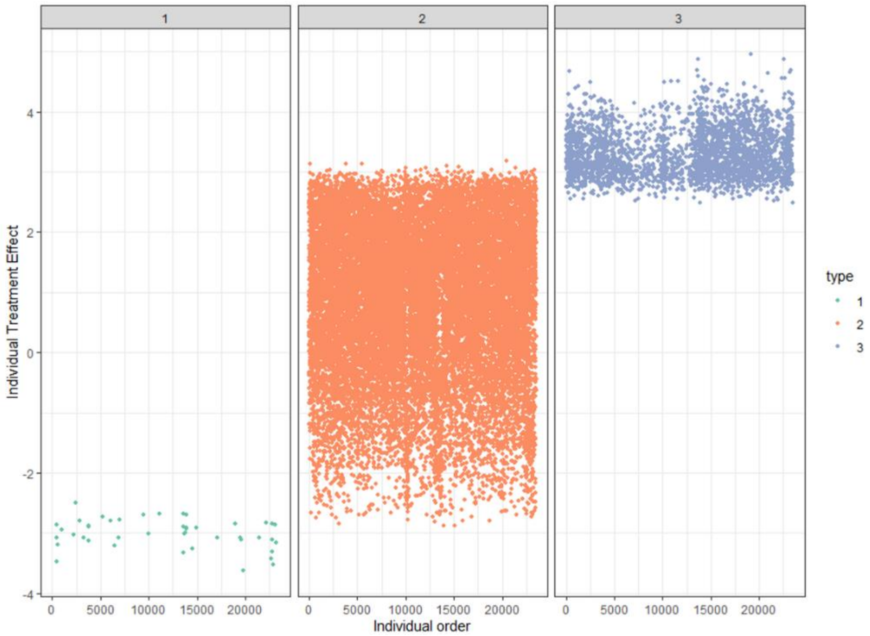

Figure 2: Robustness checks: replacing the measurement of dependent variable

Source: Authors' calculation

4.2.2 Replacing the measurement of digital transformation

Intangible assets, particularly those related to digital technologies—such as proprietary software, digital tools, and patents—are increasingly central to a firm's digital transformation. Incorporating this measure allows us to capture the depth of a firm's commitment to digitalization. Firms with higher proportions of intangible assets related to digital technology are likely to be further along in their digital transformation journey. In this section, we use the ratio of digital technology intangible assets to total assets at the end of the year for listed manufacturing firms as an indicator of digital transformation extent. Specifically, when the detailed items of intangible assets contain keywords related to digital economy technologies, such as "software," "network," "client," "management system," "intelligent platform," or related patents, these items are classified as digital technology intangible assets. The total value of digital technology intangible assets for a given firm in a given year is then aggregated and expressed as a proportion of the firm's total intangible assets for that year. This measure serves as a proxy variable for the degree of digitalization within the enterprise. The descriptive statistics, presented in Appendix Table 1, align with existing studies (Qi et al., 2020).

As illustrated in Figure 3, the impact of digital transformation on pollution emissions remains heterogeneous, with no observations indicating a negative impact.

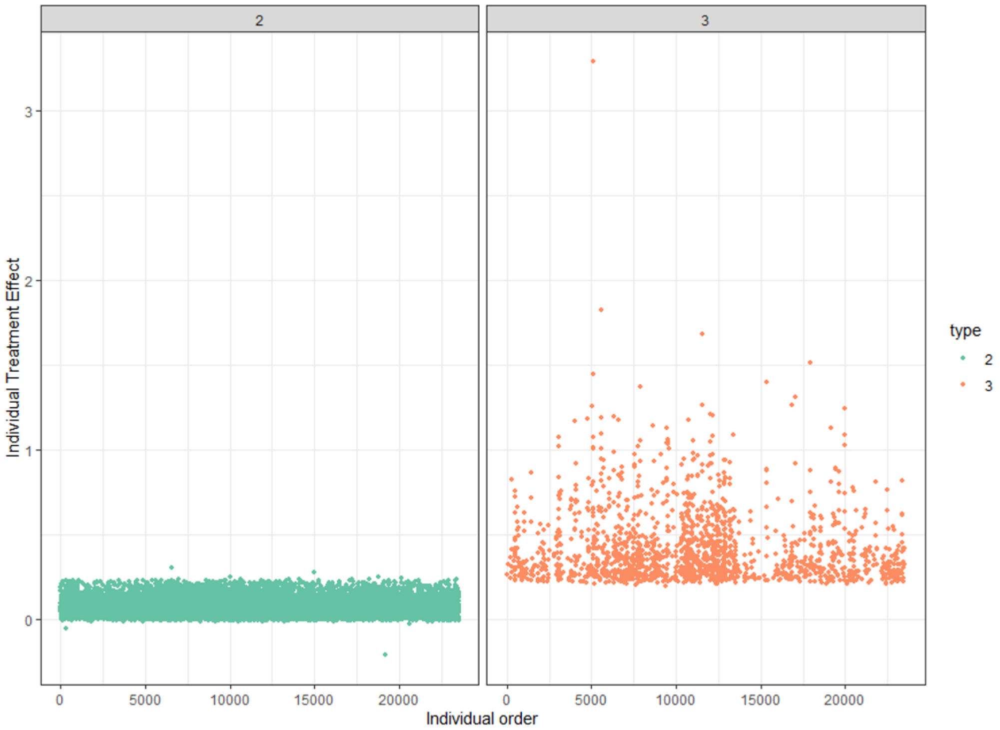

Figure 3: Robustness checks: replacing the measurement of independent variable

Source: Authors' calculation

4.2.3 Alternative samples

In the baseline regression analysis, the sample was trimmed at the 1% level. To further eliminate the influence of extreme values, this section applies a 1% level two-sided truncation to the firm pollution emission indicators. This approach serves as a robustness check to validate the baseline regression results. Figure 4 presents the results, demonstrating that even after applying the two-sided truncation, the baseline regression findings remain consistent and valid.

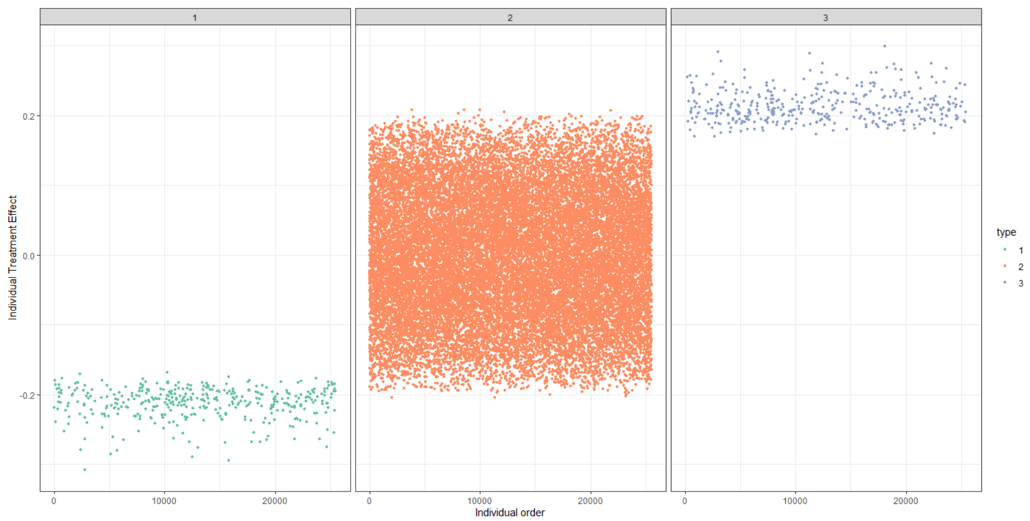

Figure 4: Robustness checks: 1% level two-sided truncation

Source: Authors' calculation

4.2.4 Alternative control variables

The proportion of corporate environmental protection expenditure to operating revenue reflects a firm's investment in environmental protection. If this variable is not controlled, it may lead to endogeneity issues. For example, firms with strong environmental awareness may simultaneously increase their investment in digital transformation and environmental protection expenditure. This would confuse the causal relationship between digital transformation and pollution emissions. The level of environmental regulation directly affects a firm's pollution emission behavior. A higher intensity of environmental regulation may prompt companies to increase their environmental protection investment and reduce pollution emissions, which is unrelated to digital transformation. If the level of environmental regulation is not controlled, it may overestimate or underestimate the impact of digital transformation on pollution emissions. Therefore, we also control the proportion of firm environmental protection expenditure to operating revenue and the level of environmental regulation. Figure 5 shows that the baseline regression results remain valid.

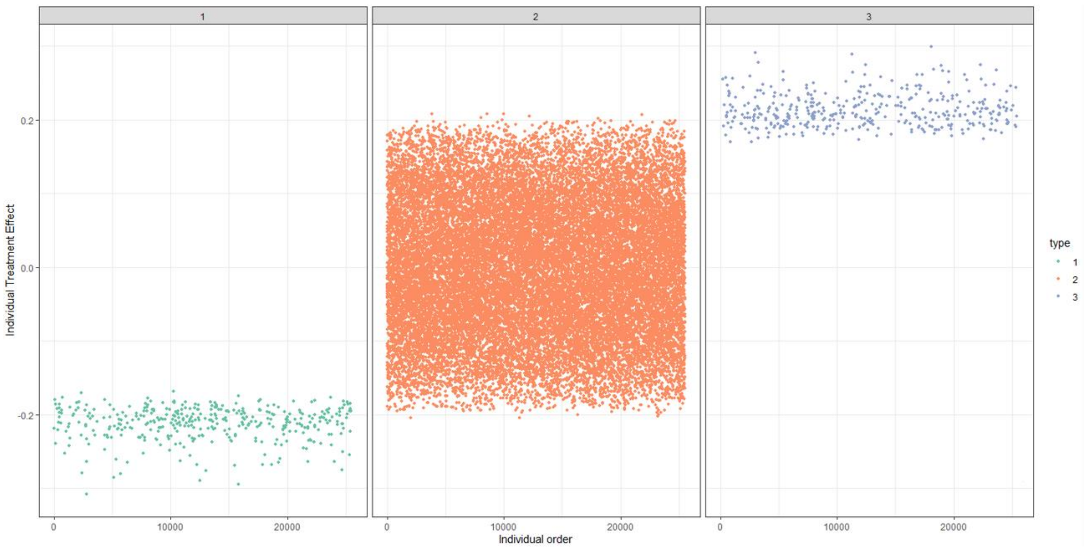

Figure 5: Robustness checks: new control variables

Source: Authors' calculation

5. Mechanism tests

5.1 Green innovation increasement effect

Based on Dangelico & Pujari (2010), Porter & Van der Linde (1995), this study reflects the green innovation increasement effects of enterprise digital transformation from three aspects: innovation investment, green innovation output, and green innovation spillover.

In terms of innovative investment, the measurement is conducted using the ratio of R&D expenditure, that is, the proportion of enterprise R&D expenditure to the total business revenue. Figure 6 shows that certain firms experience a significantly negative effect of digital transformation on pollution emissions suggest that digital tools and technologies are enabling these firms to optimize their operations in ways that substantially reduce their environmental footprint. Digital transformation likely helps these firms enhance operational efficiency, adopt cleaner production processes, and integrate sustainable technologies, which collectively contribute to pollution reduction.

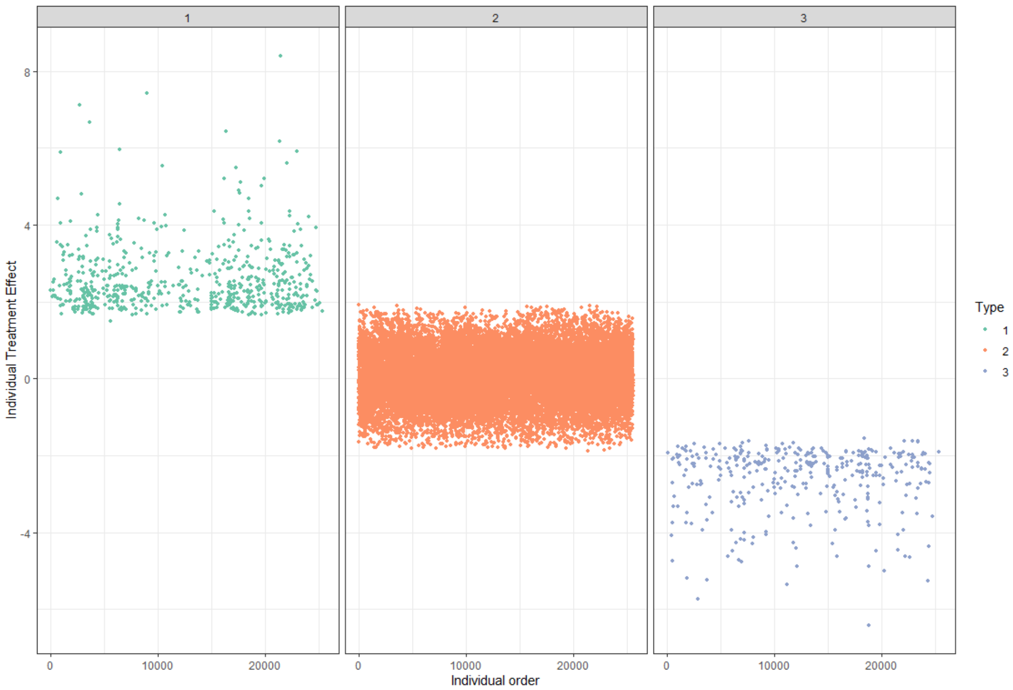

Figure 6: Green innovation increasement effect: technological innovation investment

Source: Authors' calculation

Green innovation output is represented by the number of green patents granted to the firm. The results in Figure 7 provide compelling evidence that green innovation output—as reflected in green patents—is a key driver of pollution emission reduction. Firms that engage in green innovation are more likely to introduce sustainable technologies and processes that directly reduce environmental harm. This supports the notion that innovation in sustainability is not only beneficial for the environment but also contributes to the firm's long-term success in a world increasingly focused on sustainability.

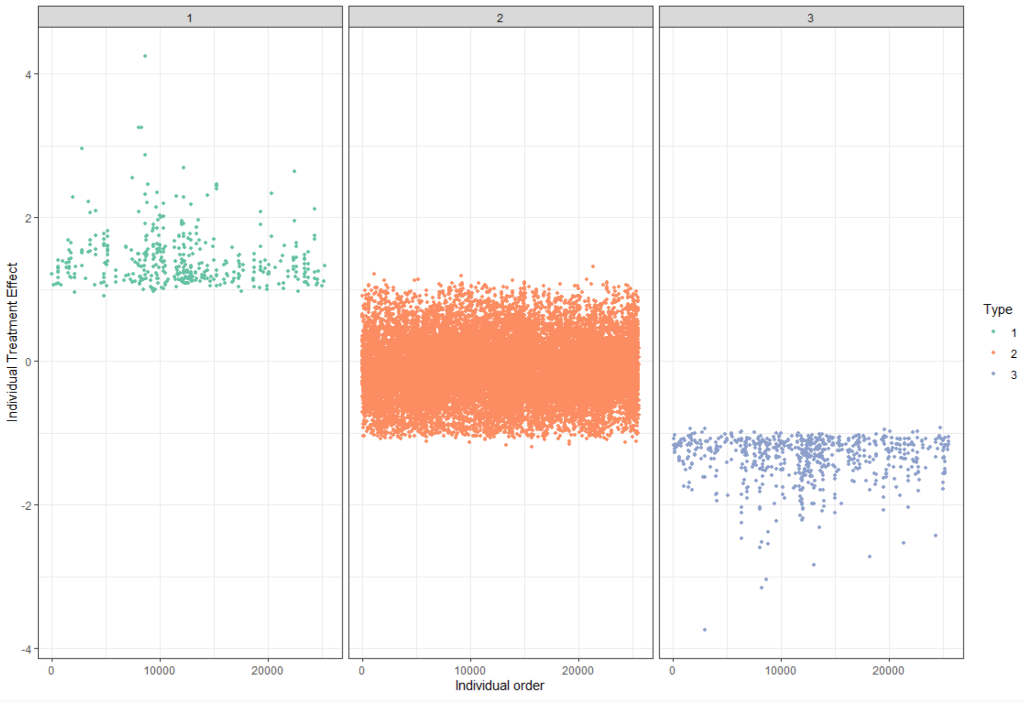

Figure 7: Green innovation increasement effect: green innovation output

Source: Authors' calculation

Green innovation spillover refers to the unintended dissemination or transfer of environmentally friendly innovations across firms, industries, or even regions. These innovations, once developed by a firm, may benefit other firms through knowledge sharing, collaboration, or competition, thus amplifying the environmental benefits across a broader set of players. Referring to Han et al. (2024), the green innovation spillover effect is measured using micro-geographic information data of listed companies, green patent data, and potential model. This index is then used to examine the mechanism by which digitalization reduces pollution emissions through green innovation spillover effects. The green innovation spillover index (*gs*) can be expressed as $gs_i = \sum(j=1, j \neq i)(R_j/d_{ij})$, where R_j represents the R&D investment of firm j in the green industry while d_{ij} represents the geographical distance between firm *i* and firm *j*. So, the green innovation spillover index measures the spillover effect of green innovation from other firms ($j \neq i$) on firm *i*, weighted by the inverse of the distance between them(d_{ij}). The assumption is that the spillover effect is stronger for firms that are geographically closer or more similar in terms of technology or industry. Figure 8 likely shows that firms experiencing positive green innovation spillovers—whether through partnerships, industry collaboration, or competitive pressures—tend to exhibit stronger reductions in pollution emissions. This suggests that when firms innovate in green technologies, these innovations do not remain isolated. Instead, they spread across other firms or industries, thereby multiplying the environmental benefits.

5.2 Factor allocation optimization effect

Building on the framework of Solow (1956), this study examines the factor allocation optimization effect by analyzing the impact of digital resource investment and its potential influence on factor misallocation.

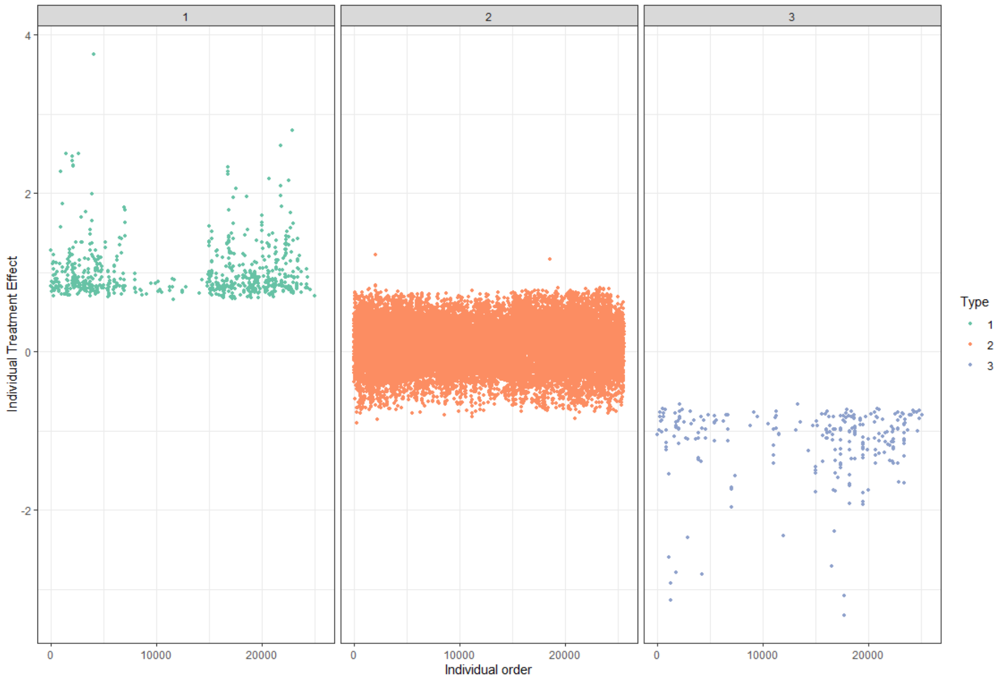

Figure 8: Green innovation increasement effect: green innovation spillover

Source: Authors' calculation

(1) Firm-level digital resource investment structure

Due to the lack of firm-level data on digital resource investment, we follow the approach of Han et al. (2024) and estimate this investment using provincial input-output tables. Specifically, we calculate the proportion of digital industry investment relative to the total intermediate input for each industry within a given province. This proportion is then allocated to the firm level, based on the firm's share of main business revenue within its industry, which serves as an approximation of its digital resource investment structure (DS). The estimation is formalized as follows:

$$DS_{ihpt} = \frac{mb_{ihpt}}{\overline{mb}_{hpt}} \times \frac{\sum_d \delta_{hdp}}{\sum_{h'} \delta_{hh'p}} \quad (11)$$

where *i, h, p,* and *t* represent firm, industry, province, and year respectively; δ_{hdp} is the complete consumption coefficient of industry *h* to the digital industry *d* in province *p*; $\sum_{h'} \delta_{hh'p}$ represents the complete consumption coefficient of industry *h* to all other industries h' in province *p*. $\frac{\sum_d \delta_{hdp}}{\sum_{h'} \delta_{hh'p}}$ represents the proportion of total intermediate input allocated to the digital industry within a specific industry in a province. It can be interpreted as a digital intensity of the industry in the province. A higher value of this proportion implies that the industry in that province is more reliant on digital inputs, signaling a more digitalized industry. A lower value would indicate that digital resources make up a smaller share of the industry's total inputs.The complete consumption coefficients for the years 2003 and 2004 are derived using the 2002 input-output table; for 2005-2009, the 2007 input-output table is used; for 2010-2013, the 2012 input-output table is used; for 2014-2016, the 2015 input-output table is applied; and for 2017-2019, the 2017 input-output table is used. mb_{ihpt} denotes the main business revenue of the enterprise, and $\overline{mb}_{hpt}$ is the average main business revenue for the industry. The coefficient ($\frac{mb_{ihpt}}{\overline{mb}_{hpt}}$) represents the firm's share of revenue within its industry, which is used to allocate the provincial-level digital resource investment proportion to the firm level. A higher share of business revenue suggests that the firm has a larger role within its industry, and thus a larger proportion of the industry-level digital investment is allocated to that firm. This ensures that the digital resource investment is distributed in proportion to the firm's relative size and importance in its sector. Figure 9 supports the conclusion that improvements in the digital resource investment structure contribute to the reduction of pollution emissions.

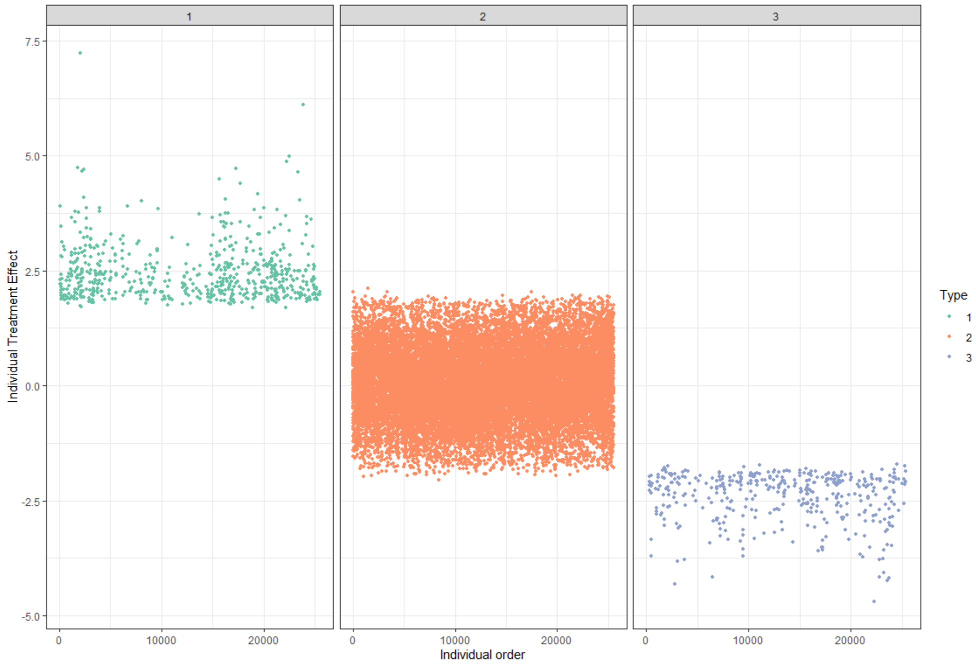

Figure 9: Factor allocation optimization effect: digital resource investment structure

Source: Authors' calculation

(2) Factor allocation status

In this study, the efficiency of factor allocation in enterprises is assessed by examining labor misallocation (*dist_l*) and capital misallocation (*dist_k*). Assuming that enterprises operate under a Cobb-Douglas (C-D) production function with constant returns to scale, labor allocation distortion is expressed as $dist\_l_{it} = (\frac{\hat{\alpha}Q_{it}}{\omega_{it}l_{it}}) - 1$, while capital allocation distortion is represented as $dist\_k_{it} =$

$[\frac{(1-\widehat{\alpha})Q_{it}}{r_{it}k_{it}}]-1$, where $\widehat{\alpha}$ represents the output elasticity of labor; Q_{it} is the industrial value added at the enterprise level, measured by the total operating revenue of listed companies; w_{it} denotes the employee wage, calculated as the ratio of the total staff remuneration payable to the number of employees; l_{it} is the number of employees; k_{it} represents the total fixed assets of listed companies; and r_{it} denotes the capital price or interest rate faced by the companies, measured by the annual average of the benchmark interest rate for enterprise loans with a term of 6 months to 1 year.

In this study, labor *(dist_l)* and capital (*dist_k)* misallocation indices are derived using their absolute values, which quantify the extent to which labor and capital are either under-allocated or over-allocated in listed companies. The absolute value approach is chosen to capture the overall misallocation magnitude, regardless of whether the misallocation results from under- or over-allocation. A higher value for *dist_l* or *dist_k* indicates greater misallocation, signifying that labor or capital resources are not being efficiently allocated across production activities. This inefficiency results in suboptimal resource utilization, leading to lower overall productivity and negatively affecting environmental performance, including higher pollution emissions.

Figure 10-a illustrates the relationship between labor misallocation and pollution emissions. As the labor misallocation index increases (i.e., higher values of *dist_l*), pollution emissions tend to rise, suggesting that inefficient labor allocation contributes to a greater environmental footprint. This indicates that when firms fail to optimize labor resources, their capacity to engage in sustainable practices is hindered, ultimately leading to higher emissions.

Similarly, Figure 10-b likely depicts the relationship between capital misallocation and pollution emissions. As capital misallocation increases (i.e., higher values of *dist_k*), pollution emissions also rise, demonstrating that inefficient capital allocation limits firms' ability to invest in cleaner technologies and energy-efficient processes, both of which are essential for reducing pollution.

The results presented in Figures 10-a and 10-b highlight the critical role of optimizing factor allocation—both labor and capital—in reducing pollution emissions. When firms efficiently allocate labor and capital to the appropriate activities, including investments in green technologies and energy-efficient production processes, they can significantly enhance environmental performance and mitigate their pollution footprint.

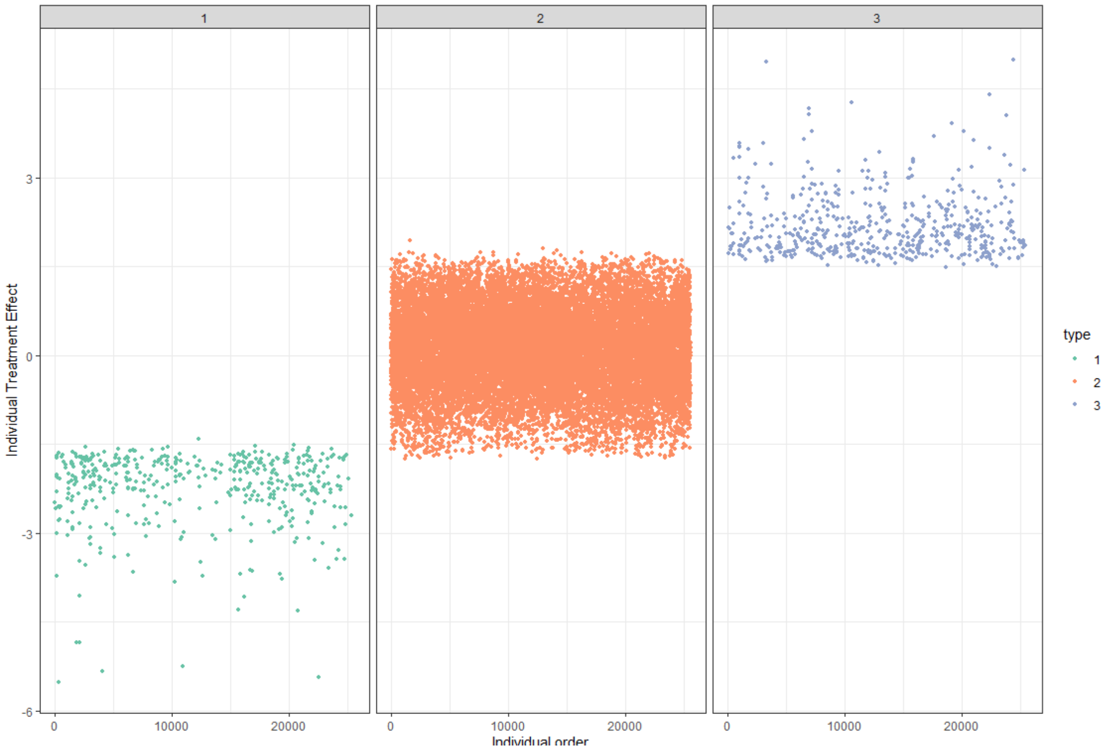

Figure 10-a: Factor allocation optimization effect: labor misallocation

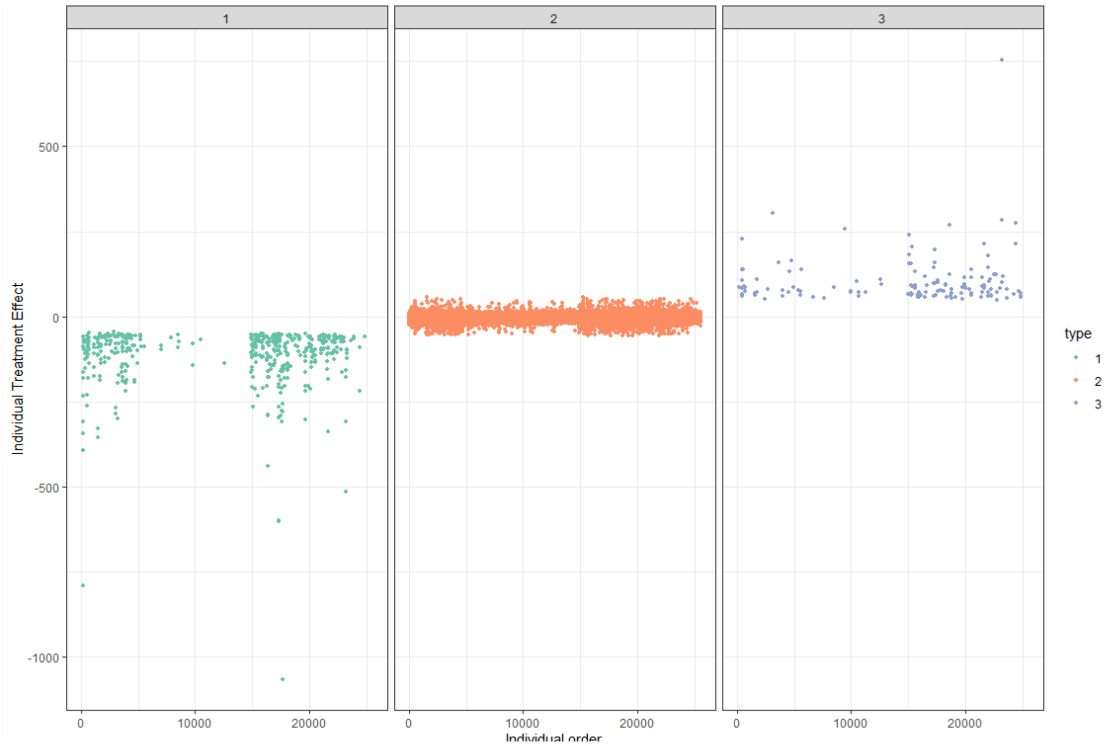

Figure 10-b: Factor allocation optimization effect: capital misallocation

Source: Authors' calculation

5.3 Social network position improvement effect

Following Han et al. (2024), the firm social network position (Net) is defined through the firm's social relationship network (SNet). This study measures the level of social network position based on the following steps:

First, information from the CSMAR database on whether directors of listed firms also serve on the boards of other companies is collected.

Second, a firm-firm network position relationship matrix is constructed annually. If listed firms *i* and *j* share at least one common director each year, the matrix element is assigned a value of 1; otherwise, it is assigned a value of 0.

Third, calculating network centrality G_i and structural hole S_i index to measure the position of a firm within the social network Network centrality reflects the extent of direct connections a firm has with other firms in the network, which is calculated as follows:

$$G_i = \frac{\sum_{j \neq i} \chi_{ij}}{N-1} \tag{12}$$

Where χ_{ij} indicates whether there is a network relationship between firm *i* and firm *j* ($\chi_{ij} = 1$ if yes, $\chi_{ij} = 0$ if no); $\sum_{j \neq i} \chi_{ij}$ is the sum of firm *i*'s direct network connections with other firms; and *N* is the total number of firms.

A structural hole refers to a gap in the information flow that is formed when two non-directly connected firms are both linked to the same firm. Firms located in structural holes act as “bridges” or “intermediaries” by connecting firms that have no direct contact with each other. Compared with firms not in structural holes, these firms can not only access more innovative resources and information from the entire network but also facilitate the flow of information and knowledge. Moreover, they can leverage the control of information flow to serve their own technological innovation.

Therefore, the network position of a firm (Net) is determined by a combination of network centrality and structural hole index, which can be expressed as:

$$Net_{it} = G_i \times S_i \tag{13}$$

The structural hole index (S_i) can be reflected by the level of constraint a firm faces within its relational network. The lower the level of constraint a firm faces, the greater the likelihood that it can act as a “bridge” or “intermediary,” and the higher its structural hole index (*Si*) will be. If we denote the level of constraint faced by a firm in its relational network as R_i, then the structural hole index can be expressed as $S_i=1/R_i$.

The constraint level R_i faced by firm i due to firm j in the network can be expressed as:

$$R_i = \sum_{i \neq j} (r_{ij} + \sum_{i \neq v,j} r_{iv} r_{vj})^2 \tag{14}$$

where r_{ij} is the direct connection strength between firm i and firm j, and $\sum_{i \neq v,j} r_{iv} r_{vj}$ is the total strength of all indirect connections through firm v between firm i and firm j. The overall constraint level $(r_{ij} + \sum_{i \neq v,j} r_{iv} r_{vj})^2$ faced by firm *i* in the network is the sum of the constraints R_i imposed by all other firms.

Using the network centrality and structural hole metrics, the social network position relationship matrix constructed from interlocking directorates is imported into social network analysis software (Ucinet). This enables the calculation of each firm's network centrality and structural hole index within both the social relationship network. The calculated values are then used to determine the firm's position within these networks. The findings underscore the importance of social network position in driving environmental outcomes. Firms that are more centrally located or occupy structural holes within their networks can leverage these positions to access and diffuse green technologies, which ultimately leads to pollution emission reductions. Figure 11 visually supports this conclusion by illustrating how firms with better social network positions are more successful in reducing emissions. This highlights the broader importance of network strategies for firms seeking to enhance their environmental sustainability and drive industry-wide changes.

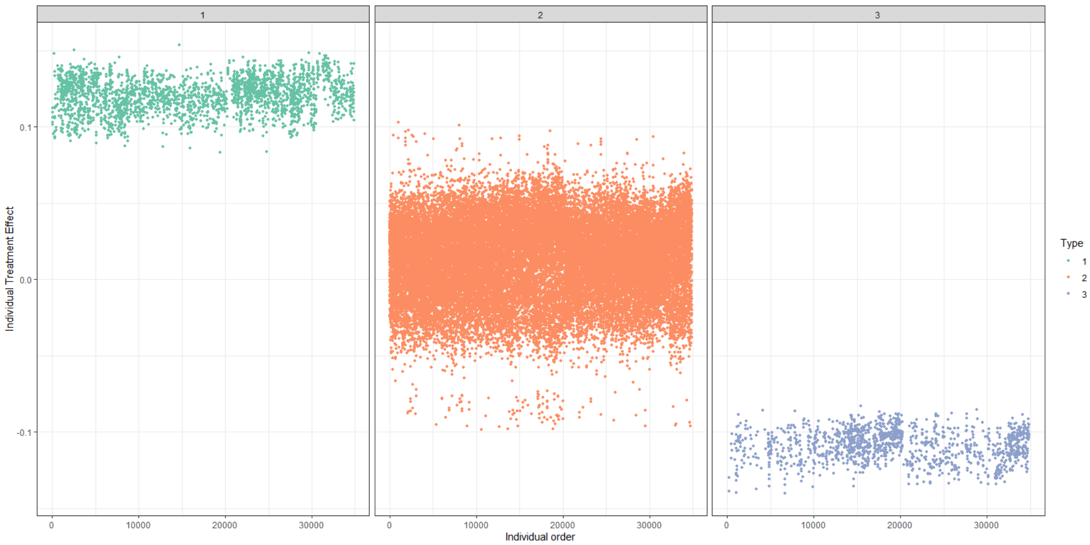

Figure 11: Social network position improvement

Source: Authors' calculation

6. Conclusion

This study examines the impact of enterprise digital transformation on pollution emissions, utilizing micro-level data from Chinese A-share listed manufacturing firms between 2007 and 2022. The findings reveal that the effects of digital transformation on pollution emissions are not uniform, highlighting significant heterogeneity across firms. Specifically, the analysis identifies three key mechanisms through which digital transformation contributes to pollution reduction:

1. Enhancement of green innovation,
2. Optimization of factor allocation, and
3. Improvements in social network positions.

These findings have important policy implications. First, to maximize the benefits of digital transformation, policymakers should accelerate the introduction and development of digital technologies, with a particular emphasis on enhancing connectivity and digitalizing production equipment. This approach will facilitate the comprehensive integration of digital technologies across all aspects of enterprise production, management, and organizational design, fostering new competitive advantages in the digital economy while promoting sustainable green growth.

Second, strengthening the utilization of data elements will further enhance resource allocation efficiency within enterprises, leading to greater pollution reduction. By leveraging data-driven decision-making, firms can optimize production processes, minimize waste, and improve environmental performance.

Finally, reinforcing social connections between enterprises will enable firms to capitalize on network externalities and social proximity, amplifying the positive impact of digital transformation on pollution reduction. Encouraging inter-firm collaboration, knowledge sharing, and digital ecosystem development will be essential in realizing the full environmental benefits of digital transformation.

References

Ahuja, G. (2000). Collaboration networks, structural holes, and innovation: A longitudinal study. *Administrative Science Quarterly*, *45*(3), 425–455.

Bharadwaj, A., El Sawy, O. A., Pavlou, P. A., & Venkatraman, N. V. (2013). Digital business strategy: Toward a next generation of insights. *MIS Quarterly*, *37*(2) 471–482.

Borgatti, S. P., & Foster, P. C. (2003). The network paradigm in organizational research: A review and typology. *Journal of Management*, *29*(6), 991–1013.

Bresciani, S., Huarng, K. H., Malhotra, A., & Ferraris, A. (2021). Digital transformation as a springboard for product, process and business model innovation. *Journal of Business Research*, *128,* 204–210.

Capello, R. (2014). Proximity and regional innovation processes: Is there space for new reflections? In A. Torre & F. Wallet (Eds.), *Regional development and proximity relations* (pp. 163-194). Cheltenham, UK: Edward Elgar Publishing.

Chen, P., & Hao, Y. (2022). Digital transformation and corporate environmental performance: The moderating role of board characteristics. *Corporate Social Responsibility and Environmental Management*, *29*(5), 1757–1767.

Chen, S., Yang, Y., & Wu, T. (2023). Research on the impact of digital transformation on green development of manufacturing enterprises. *Frontiers in Energy Research*, *10*, 1045328.

Copeland, B. R., & Taylor, M. S. (2001). *International trade and the environment: A framework for analysis* (NBER Working Paper No. 8540). Cambridge, MA: National Bureau of Economic Research.

Dangelico, R. M., & Pujari, D. (2010). Mainstreaming green product innovation: Why and how companies integrate environmental sustainability. *Journal of Business Ethics*, *95*(3), 471-486.

Dou, B., Guo, S., Chang, X., & Wang, Y. (2023). Corporate digital transformation and labor structure upgrading. *International Review of Financial Analysis, 90*, 102904.

Feng, H., Wang, F., Song, G., & Liu, L. (2022). Digital transformation on enterprise green innovation: Effect and transmission mechanism. *International journal of environmental research and public* *health*, *19*(17), 10614.

Feroz, A. K., Zo, H., & Chiravuri, A. (2021). Digital transformation and environmental sustainability: A review and research agenda. *Sustainability, 13*(3), 1530.

Gao, F., Lin, C., & Zhai, H. (2022). Digital transformation, corporate innovation, and international strategy: Empirical evidence from listed companies in China. *Sustainability, 14*(13), 8137.

Glinskiy, V., Serga, L., Tsurikov, S., & Plotnikov, V. (2020). Analysis of the digital transformation influence on the economic growth of the region. *Economic and Social Development: Book of Proceedings*, 260-269.

Goetz, M. R. (2018). *Financial constraints and corporate environmental responsibility* (SAFE Working Paper No. 241). Retrieved from https://ssrn.com/abstract=3230344

Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology, 91*(3), 481–510.

Grossman, G. M., & Krueger, A. B. (1995). Economic growth and the environment. *The Quarterly Journal of Economics, 110*(2), 353–377.

Hallioui, A., Herrou, B., Santos, R. S., Katina, P. F., & Egbue, O. (2022). Systems-based approach to contemporary business management: An enabler of business sustainability in a context of industry 4.0, circular economy, competitiveness and diverse stakeholders. *Journal of Cleaner Production, 373*, 133819.

Han, F., Huang, M., & Jiang Z. (2024). Enterprise digitalization, network status, and pollution reduction. *The World Economy, 47*(2), 204–232. (In Chinese)

Hao, Y., Wu, Y., Wu, H., & Ren, S. (2020). How do FDI and technical innovation affect environmental quality? Evidence from China. *Environmental Science and Pollution Research, 27*(8), 7835–7850.

Hu, X., Weng, X., & He, Z. (2024). Does enterprise digital transformation promote technological innovation? Empirical evidence from China. *Chinese Management Studies, 18*(4), 1164–1193.

Huang, D., Xie, H., Meng, X., & Zhang, Q. (2021). Digital transformation and enterprise value: Empirical evidence based on text analysis methods. *Economist, 33*(12), 41–51.

Huang, Z., & Wang, Y. (2008). Location embedding, social capital, and network learning of overseas subsidiaries in host country: An empirical study based on 123 subsidiaries of MNCs in China. *China Industrial Economics*, 12, 144–154.

Imran, F., Shahzad, K., Butt, A., & Kantola, J. (2021). Digital transformation of industrial organizations: Toward an integrated framework. *Journal of Change Management, 21*(4), 451–479.

Li, S., Gao, L., Han, C., Gupta, B., Alhalabi, W., & Almakdi, S. (2023). Exploring the effect of digital transformation on Firms' innovation performance. *Journal of Innovation & Knowledge, 8*(1), 100317.

Li, X., Liu, J., & Ni, P. (2021). The impact of the digital economy on CO2 emissions: A theoretical and empirical analysis. *Sustainability, 13*(13), 7267.

Liu, J., & Xiao, Y. (2022). China's environmental protection tax and green innovation: Leverage effect or crowding out effect? *Economic Research Journal, 57*(1), 72–88.

Lv, X., & Wu, Z. (2024). The green effect of digital transformation: The impact of digital transformation in fiscal and taxation on regional green development. *Economic Analysis and Policy, 81*, 787–800.

Miao, Z., & Chen, X. (2022). Combining parametric and non-parametric approach, variable & source-specific productivity changes and rebound effect of energy & environment. *Technological Forecasting and Social Change*, *175*, 121368.

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics capabilities and innovation: The mediating role of dynamic capabilities and moderating effect of the environment. *British Journal of Management, 30*(2), 272–298.

Ostas, D. T. (2003). Learning and innovation in organizations and economies. *Journal of Economic Issues, 37*(4), 1191–1192.

Peng, Y., & Tao, C. (2022). Can digital transformation promote enterprise performance?—From the perspective of public policy and innovation. *Journal of Innovation & Knowledge, 7*(3), 100198.

Porter, M., & Van der Linde, C. (1995). Green and competitive: Ending the stalemate. In E. F. M. Wubben (Ed.), *The Dynamics of the Eco-Efficient Economy* (pp. 33-56). Cheltenham, UK: Edward Elgar Publishing.

Qi, H. J., Cao, X. Q., & Liu Y. X. (2020). The impact of digital economy on corporate governance: A perspective of information asymmetry and managerial irrational behavior. *Reform, 4,* 50–64. (In Chinese)

Qian, X., Yang, Y., & Xu, W. (2010). Enterprise network location, absorption capacity, and innovation performance: An interactive effect model. *Journal of Management World, 15*(5), 118–129. (In Chinese)

Ren, X., Zeng, G., & Sun, X. (2023). The peer effect of digital transformation and corporate environmental performance: Empirical evidence from listed companies in China. *Economic Modelling, 128*, 106515.

Shahzad, U. (2020). Environmental taxes, energy consumption, and environmental quality: Theoretical survey with policy implications. *Environmental Science and Pollution Research, 27*(20), 24848–24862.

Shang, Y., Raza, S. A., Huo, Z., Shahzad, U., & Zhao, X. (2023). Does enterprise digital transformation contribute to the carbon emission reduction? Micro-level evidence from China. *International Review of Economics & Finance, 86*, 1–13.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, *70*(1), 65–94.

Sun, G., Fang, J., Li, J., & Wang, X. (2024). Research on the impact of the integration of digital economy and real economy on enterprise green innovation. *Technological Forecasting and Social Change, 200*, 123097.

Tang, M., Liu, Y., Hu, F., & Wu, B. (2023). Effect of digital transformation on enterprises' green innovation: Empirical evidence from listed companies in China. *Energy Economics, 128*, 107135.

Trumpp, C., & Guenther, T. (2017). Too little or too much? Exploring U-shaped relationships between corporate environmental performance and corporate financial performance. *Business Strategy and the Environment*, *26*(1), 49–68.

Tsai, W., & Ghoshal, S. (1998). Social capital and value creation: The role of intrafirm networks. *Academy of Management Journal*, *41*(4), 464–476.

Uzzi, B. (1996). The sources and consequences of embeddedness for the economic performance of organizations: The network effect. *American Sociological Review*, *61*(4), 674–698.

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *The Journal of Strategic Information Systems*, 22(8), 118–144.

Wang, J., Liu, Y., Wang, W., & Wu, H. (2023). How does digital transformation drive green total factor productivity? Evidence from Chinese listed enterprises. *Journal of Cleaner Production, 406*, 136954.

Zhang, C., Fang, J., Ge, S., & Sun, G. (2024). Research on the impact of enterprise digital transformation on carbon emissions in the manufacturing industry. *International Review of Economics & Finance, 92*, 211–227.

Zhang, H., Wu, J., Mei, Y., & Hong, X. (2024). Exploring the relationship between digital transformation and green innovation: The mediating role of financing modes. *Journal of Environmental Management, 356*, 120558.

Zhao, X., Sun, X., Zhao, L., & Xing, Y. (2022). Can the digital transformation of manufacturing enterprises promote enterprise innovation? *Business Process Management Journal, 28*(4), 960–982.

Appendix

Appendix A: Calculation of Digital Transformation Index

For the digital transformation index, the construction process is as follows:

Firstly, the annual reports of all A-share listed enterprises in the Shanghai Stock Exchange and Shenzhen Stock Exchange were summarized using a Python crawler function. Secondly, the text contents were extracted using the Java PDFbox library, serving as a data pool for the subsequent identification of key terms. To identify the key elements of enterprise digital transformation, this study examines a range of seminal literature on the subject and extracts specific keywords associated with digital transformation. Additionally, this research utilizes significant policy documents and research reports, such as the SMEs Digital Empowerment Special Action Program and On Promoting the Action of “Going to the Cloud, Using Digital Empowerment, and Fostering the Development of the New Economy”, as the foundation for the subsequent data screening process.

Based on the extracted keywords, we categorize them into distinct groups that reflect different aspects of digital transformation, specifically focusing on “underlying technology application” and “practical application of technology.” To ensure relevance, expressions containing negative words such as “no,” “none,” or “not” preceding the keywords are disregarded. Additionally, expressions unrelated to the company, such as references to shareholders, customers, and suppliers, are excluded.

Finally, utilizing Python to extract the text from annual reports of publicly traded companies, we conduct a search, match, and count of word frequencies based on the identified key technology terms. The categorized word frequencies are then aggregated to form the final cumulative word frequencies. This process enables us to construct an index system for measuring the digital transformation of enterprises.

Given the inherent “right-biased” nature of this type of data, the present study applies logarithmic transformation to obtain comprehensive indices that accurately depict the level of digital transformation among enterprises.

Appendix B

Appendix Table 1: Descriptive Analysis

VarName	Mean	SD	Min	Median	Max
PEI	7.019	2.055	2.356	7.210	11.978
Szh	0.005	0.014	0	0.001	0.604
Ecr	0.222	0.170	-0.206	0.190	0.971
dist_k	12.446	36.850	1.000	3.540	1344.776
SNet	0.116	0.047	0.000	0.130	0.187
lnGZ	-4.322	1.990	-14.526	-4.033	3.558
lnRDsr	18.875	1.717	9.718	18.747	26.772
lnGPA	21.003	1.845	12.570	20.828	31.217
lngs	21.306	1.392	12.223	21.184	28.888
DS	-1.202	1.252	-8.112	-1.022	3.911
dist_l	3.747	1.045	-5.728	3.639	12.949

Notes:

PEI represents the firm's composite pollution emission intensity.

Szh measures firm-level digital transformation, estimated as the ratio of digital technology intangible assets to total assets.

Ecr and lnGZ indicate corporate environmental protection expenditure and the level of environmental regulation in the city where the firm is located, respectively.

lnRDsr, lnGPA, and lngs respectively represent technological innovation input, green technological innovation output, and green technology spillover.

DS denotes the structure of digital resource input.

dist_l and dist_k measure the degree of labor and capital misallocation within firms.

SNet represents the firm's social network position.