Proceedings of ICEMGD 2025 Symposium: Digital Transformation in Global Human Resource Management
DOI: 10.54254/2754-1169/2025.LD25044

Digital Alienation: How Data Elements Exacerbate the Green Innovation Bubble—Based on the Dual Mechanisms of Information Manipulation and Technological Lock-in

Yiqing Yue

School of International Economics, China Foreign Affairs University, Beijing, China
elaineyuewin@163.com

Abstract. This study explores the paradox where elevated data element utilization level intensifies green innovation bubbles (GIB) in Chinese A-share listed firms from 2012 to 2023. Employing time-fixed effect regression on 36,229 firm-year panel observations, this paper finds a significant positive association between data element utilization level (DEUL) and GIB (β=0.161, p<0.01), indicating that data-driven strategies may unexpectedly widen innovation quantity-quality gaps. Moreover, mechanism analysis reveals DEUL worsens GIB through dual pathways: (1) worsening information asymmetry via selective environmental disclosure (EID coefficient: -0.382, p<0.01), and (2) reinforcing technological path dependence (IID coefficient: 0.090, p<0.05). Notably, superstar inventors positively moderate this effect (β=0.162, p<0.05), speeding up bubble formation on grounds of overreliance on models. These findings challenge "data-driven efficiency," demonstrating that digital tools may cause strategic rigidity and resource misallocation. In terms of this issue, three solutions are proposed: policy-driven algorithm transparency as well as diversified R&D investment, corporate long-term evaluation reforms, and financial bubble-warning systems.

Keywords: Data element utilization, green innovation bubble, Information asymmetry, Technological path dependence, Digital alienation

1. Introduction

Under the dual-carbon goal, green technology innovation is an indispensable part of enterprise transformation and high-quality development. Nevertheless, current literature has indicated that the "green innovation bubble" phenomenon has been prominent in China in recent years. The "green innovation bubble" restricts the green innovation quality, which is ultimately detrimental to the sustainable development. Hence, it’s urgent to solve the problem of the "green innovation bubble" through studying its contributing factors and formation mechanism.

Information asymmetry, financing constraints, and market pressure are important theoretical foundations for the formation of green innovation bubbles. Based on the above theories, the existing literature mainly explores the influencing factors of green innovation bubbles from three perspectives: policy, market, and microstructure.

Unfortunately, few scholars have so far examined the role of the utilization level of new production factors in generating green innovation bubbles in enterprises. Does the level of data element utilization have an impact on firms' green innovation activities? Using the relevant data of A-share listed

DOI: 10.54254/2754-1169/2025.LD25044

companies from 2012 to 2023, this study constructs the analytical framework of "data factor-intermediary mechanism-green innovation bubble” and systematically deconstructs the effects of data factor on green innovation bubble.

The contribution of this research is reflected in two aspects. Theoretically, this paper breaks through the unidimensional focus on "green innovation quality" and reveals for the first time the bubble-generating mechanism driven by data elements. With the above findings, the paper demonstrates the "double-edged sword" effect of the utilization level of data elements, challenges the intuitive cognition that "data elements inevitably enhance efficiency", and provides theoretical explanations for the unintended consequences of digital transformation. Practically, the information asymmetry and technology path-dependent mediation effects argued in this study provide insights and lessons for regulatory policy design, corporate management innovation, and market risk regulation.

2. Literature review

2.1. Main drivers of green innovation bubbles

A corporate green innovation bubble is an illusory prosperity in which innovation activities appear to be booming due to various factors but lack sustainable value creation. In recent years, academics have conducted in-depth research on its drivers and influences from various perspectives.

From the policy perspective, Ma Lina and Chen Yuhan found the role of environmental regulation in promoting the green innovation bubble [1]. While Xu Baochang et al. discussed the effect of green credit policy on inhibiting the quality of green innovation [2]. In addition, the QFII (Qualified Foreign Institutional Investor) system suppresses the creation of "green innovation bubbles" through the governance and financing paths [3]. Talent policy support can promote the "quantitative and qualitative leap" of green innovation [4]. From a market perspective, Li Qing and Chen Lin suggest that market ESG rating uncertainty has a positive effect on green innovation bubbles [5].

2.2. Level of data element utilization

The level of data element utilization refers to the enterprise's ability to collect, store, clean, analyze and utilize data elements. Regarding the level of data factor utilization, the research at home and abroad is still in the primary exploration stage.

In terms of influencing factors, Cao Ping and Chen Youbin explored the synergistic effect of certain factors on increasing the level of data factor utilization [6]. In terms of the value, Shi Qingchun et al. found from the micro perspective that the level of data factor utilization has a positive impact on enterprise investment efficiency [7]. And at the macro level, some studies have shown that the level of data factor utilization promotes the modernization of industrial structure [8].

In January 2024, the National Data Bureau issued the Three-Year Action Plan of “Data Factor × ”, emphasizing the construction of a digital economy as an inevitable requirement for promoting high-quality development. In this context, it is urgent to explore how to regulate the utilization level of data elements to energize green innovation, so as to realize the high-level combination of economic and social benefits.

3. Mechanism analysis and hypothesis

3.1. The direct impact of data utilization on the green innovation bubble

Excessive reliance on data elements may contribute to the expansion of green innovation bubbles through the rigidity of strategic planning, homogenization of resource allocation, short-sightedness in performance evaluation, and risk perception bias.

DOI: 10.54254/2754-1169/2025.LD25044

First, AI-driven decision-making systems may cause firms to ignore the nonlinear changes in technology and policy, spawning far advanced investment bubbles. Besides, the "data cocoon" effect may lead to the concentration of funds in a few areas with good data performance. At the level of performance evaluation, "greenwashing" strategies may be adopted to pursue instant optimization of data indicators. What's more, the inherent "black box" nature of big data models can give birth to new types of risks. Notably, the dynamic feedback nature of data elements may create a self-fulfilling prophecy. In summary, this paper proposes the following hypothesis:

Hypothesis 1: The level of data element utilization has a significant positive correlation to green innovation bubbles.

3.2. Mediating effects

3.2.1. Information asymmetry

Information asymmetry can lead to a variety of unfavorable outcomes: on the one hand, it can lead to the presence of speculators, pushing up the valuation bubble of green innovation projects. On the other hand, it will increase market uncertainty and financing costs, under which condition enterprises tend to pursue short-term benefits and ignore long-term development.

The high utilization level of data elements provides enterprises with strategic information screening tools, through which they can generate data that can pass regulatory scrutiny but has no substantive innovation. Faced with massive amounts of data, investors will rely on simple quantitative indicators rather than technical content, driving valuations away from fundamentals.

Hypothesis 2a: The level of data factor utilization contributes to green innovation bubbles by exacerbating information asymmetry.

3.2.2. Technology path dependence

Decisions made by artificial intelligence models are based on historical data, which presumably limits the innovation to the established technological trajectory. This kind of data-driven incremental innovation appears to be the prosperity of technology iteration, but it is, in reality, caught in the "diminishing efficiency trap". The resources invested can only realize micro-improvement but significantly crowd out the R&D space for breakthrough innovation. Accordingly, the following hypothesis is proposed:

Hypothesis 2b: The level of data factor utilization contributes to green innovation bubbles by reinforcing technological path dependence.

3.3. The moderating effect

The combination of "superstars + data elements" may send out overly optimistic signals, triggering a herd effect among investors. When capital is tilted towards star projects with short-term results, it creates a mismatch of resources, distorts the innovation ecosystem, and leads to irrational valuations. Moreover, the decision-making model based on historical R&D success rates of star inventors may exacerbate technology path lock-in. Accordingly, the following hypothesis is proposed:

Hypothesis 3: Superstar inventors play a positive moderating role in the process of promoting green innovation bubbles.

DOI: 10.54254/2754-1169/2025.LD25044

4. Research design

4.1. Sample and data

This paper takes A-share listed companies from 2012 to 2023 as a sample. The data mainly comes from the Cathay Pacific Database (CSMAR), China Research Data Service Platform (CNRDS), and annual reports of listed companies. To ensure the homogeneity and validity, this study excluded: (1) samples from the financial industry and real estate industry; (2) all ST, *ST, and PT samples; (3) samples with abnormal financial data (gearing greater than 1 or less than 0); and (4) samples with undisclosed relevant indicators. In addition, samples in the software and information technology service industry, whose business is closely related to big data, are excluded [7]. After screening, 36,229 annual observations are finally obtained, and all continuous variables in this paper are shrink-tailed at the 1% level.

4.2. Variable setting

4.2.1. The explained variable

For Green Innovation Bubbles (GIB), referring to Geng Yuan et al., this paper calculates the standardized difference between the quantity and quality of green patents [9]. The larger the value, the more serious the green innovation bubbles of enterprises.

4.2.2. The explanatory variable

Data Element Utilization Level (DEUL) is measured by referencing the approach of Wu Fei et al., which involves counting the disclosure frequencies of subdivided items under five categories: artificial intelligence technology, blockchain technology, cloud computing technology, big data technology, and big data application technology [10]. The total count is first multiplied by 0.01, added to 1, and then transformed by natural logarithm to construct the DEUL index.

4.2.3. Intermediary variables

a. Information Asymmetry

Referring to previous studies, this paper constructs an indicator of the quality of corporate environmental information disclosure (EID). The information is rated from four dimensions: completeness, specificity, comparability, and transparency. And the EID quality index is the natural logarithm of the values that are obtained by adding 1 to the tenth of the total score. The higher the index, the lower the degree of information asymmetry, and the more transparent the environmental governance.

b. Technology Path Dependence

Following previous work, this study selects the Incremental Innovation Degree (IID) to measure the degree of technology path dependence. The IID index is generated by quantifying technological continuity through the semantic similarity of patent texts. The higher the index, the more the enterprise innovation relies on the existing technology paradigm.

4.2.4. The moderating variable

Based on previous literature, this paper first calculates the top 10% of patent inventors in terms of the number of patent applications per year, which are defined as Superstar Inventors (Star). Then, a dummy variable is generated according to whether the enterprise has a superstar inventor in that year, which is 1 if it does, and 0 if it does not.

DOI: 10.54254/2754-1169/2025.LD25044

4.2.5. Control variables

Six control variables are selected in this study: the gearing ratio (Lev), return on total assets (Roa) and TobinQ value (TobinQ), the proportion of shares held by the first largest shareholder (Top1), the proportion of shares held by management (Mshare) and the percentage of independent directors (Indep).

4.3. Modeling

4.3.1. Benchmark model

This study introduces time fixed effects λ_t , ε_{it} as the random error term.

$$\mathrm{GIB_{it}} = \alpha_0 + \beta_1 \mathrm{DEUL_{it}} + \sum_{j=1}^{n} \gamma_j \mathrm{Controls_{jit}} + \lambda_t + \varepsilon_{it} \tag{1}$$

4.3.2. Mediation model of information asymmetry

This study adopts two-step approach to test the mediation effect.

$$\mathrm{GIB_{it}} = \alpha_0 + \beta_1 \mathrm{DEUL_{it}} + \sum_{j=1}^{n} \gamma_j \mathrm{Control_{jit}} + \lambda_t + \varepsilon_{it} \tag{2}$$

$$\mathrm{EID_{it}} = \alpha_1 + \beta_2 \mathrm{DEUL_{it}} + \sum_{j=1}^{n} \gamma_j \mathrm{Control_{jit}} + \lambda_t + \varepsilon_{it} \tag{3}$$

4.3.3. Mediation models of technology path dependence

$$GIB_{\mathrm{it}} = \alpha_0 + \beta_1 DEUL_{\mathrm{it}} + \sum_{j=1}^{n} \gamma_j \mathrm{Control_{jit}} + \lambda_t + \varepsilon_{it} \tag{4}$$

$$\mathrm{IID_{it}} = \alpha_1 + \beta_2 DEUL_{\mathrm{it}} + \sum_{j=1}^{n} \gamma_j \mathrm{Control_{jit}} + \lambda_t + \varepsilon_{it} \tag{5}$$

4.3.4. Moderated modeling

$$\mathrm{GIB_{it}} = \alpha_0 + \beta_1 \mathrm{DEUL\_c_{it}} + \beta_2 \mathrm{Star_{it}} + \sum_{j=1}^{n} \gamma_j \mathrm{Control_{jit}} + \lambda_t + \varepsilon_{it} \tag{6}$$

$$\mathrm{GIB_{it}} = \alpha_1 + \beta_1 \mathrm{DEUL\_c_{it}} + \beta_2 \mathrm{Star_{it}} + \beta_3 (\mathrm{DEUL\_c_{it}} \times \mathrm{Star_{it}}) + \sum_{j=1}^{n} \gamma_j \mathrm{Control_{jit}} + \lambda_t + \varepsilon_{it} \tag{7}$$

5. Empirical results and analysis

5.1. Descriptive statistics

Table 1 shows the results of descriptive statistics of the main variables. The overall level of data element utilization of the sample firms is low, and the differences between firms are small. Most enterprises in the sample have lower green innovation bubbles, but there is significant differentiation. The quality of EID of the sample enterprises is low and there are some differences among enterprises. The IID is low, with insignificant differences between samples. 22.5% of the sample firms have superstar inventors.

DOI: 10.54254/2754-1169/2025.LD25044

Table 1: Descriptive statistics

	Mean	SD	Min	Max
GIB	-.032	.304	-0.605	1.968
DEUL	.08	.138	0.000	.751
EID	.208	.63	0.000	1.792
IID	.105	.102	0.000	.535
Star	.225	.417	0.000	1
Lev	.399	.199	0.038	.934
ROA	.036	.07	-0.578	.22
TobinQ	1.998	1.302	0.000	17.676
Indep	.378	.054	0.286	.6
Mshare	.167	.206	0.000	.706
Top1	.332	.144	0.076	.758

5.2. Benchmark regression

In Table 2, the regression coefficient between DEUL and GIB is significantly positive ($\beta = 0.161,\ p < 0.01$) . This indicates that the higher the level of data factor utilization, the more serious the "bubble" phenomenon in its green innovation activities, which verifies Hypothesis 1. For control variables, Lev and Roa are significantly and positively related to GIB, while TobinQ is significantly and negatively related to GIB, which is consistent with the results of existing research [2].

DOI: 10.54254/2754-1169/2025.LD25044

Table 2: Benchmark regression

	(1)	(2)	(3)
	GIB	GIB	GIB
DEUL	0.149***	0.159***	0.161***
	(0.011)	(0.011)	(0.028)
Lev		0.121***	0.122***
		(0.008)	(0.015)
ROA		0.215***	0.215***
		(0.023)	(0.033)
TobinQ		-0.006***	-0.006***
		(0.001)	(0.001)
Indep		0.027	0.026
		(0.028)	(0.048)
Mshare		-0.005	-0.005
		(0.008)	(0.012)
Top1		0.035***	0.035
		(0.010)	(0.025)
_cons	-0.053***	-0.119***	-0.100***
	(0.002)	(0.012)	(0.022)
N	36241.000	36239.000	36239.000
r2	0.005	0.014	0.015
year	yes	yes	yes

5.3. Mediation effects test

According to columns (1)-(3) of Table 3, the coefficient between the DEUL and the quality of EID and the IID is respectively 0.407 ($p<0.01$) and 0.090 ($p<0.05$). This suggests that over-reliance on data elements increases the degree of information asymmetry and reinforces technological path dependence, which in turn creates a green innovation bubble. Hypothesis 2a and hypothesis 2b are tested.

5.4. Moderated effects test

Column (5) of Table 3 shows that both coefficients are significantly positive, suggesting that superstar inventors play a positive moderating role in the process of generating green innovation bubbles. Firms with superstar inventors (Star=1) are more likely to fall into the traps of information asymmetry and technological path lock-in by relying excessively on data elements, thus triggering innovation bubbles.

DOI: 10.54254/2754-1169/2025.LD25044

Table 3: Mediation and moderated effects

	(1)	(2)	(3)	(4)	(5)
	GIB	EID	IID	GIB	GIB
DEUL(_c)	0.161***	-0.382***	0.090**	0.142***	0.107***
	(0.028)	(0.040)	(0.035)	(0.026)	(0.024)
0.Star				0.000	0.000
				(.)	(.)
1.Star				0.141***	0.139***
				(0.012)	(0.012)
inter					0.162**
					(0.080)
controls	yes	yes	yes	yes	yes
_cons	-0.100***	0.477***	0.889***	-0.106***	-0.105***
	(0.022)	(0.043)	(0.043)	(0.021)	(0.021)
N	36239.000	30089.000	25153.000	36239.000	36239.000
r2	0.015	0.238	0.020	0.051	0.052
year	yes	yes	yes	yes	yes

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

5.5. Robustness test

Based on benchmark regression, this paper: (1) replaces the measurement index of the independent variables; (2) takes the lagged one-period values of the independent variables and all control variables; (3) adds three control variables including Equity Multiplier(EM), ROE and Institutional Investor Share(Mshare). As shown in Table 4, the regression coefficients are all significantly positive in the above cases, indicating a robust positive correlation between the DEUl and GIB.

DOI: 10.54254/2754-1169/2025.LD25044

Table 4: Robustness test

	(1)	(2)	(3)
	GIB	GIB	GIB
DEUL	0.762***	0.159***	0.157***
	(0.182)	(0.029)	(0.028)
controls	yes	yes	yes
EM			-0.010***
			(0.003)
Inst			0.076***
			(0.016)
_cons	-0.101***	-0.119***	-0.126***
	(0.022)	(0.022)	(0.024)
N	36238.000	30671.000	36239.000
r2	0.013	0.016	0.018
year	yes	yes	yes

6. Conclusions and outlook

6.1. Main conclusions

Using the relevant data of A-share listed companies from 2012 to 2023, this paper empirically examines the impact of data factor utilization level on green innovation bubbles. The main conclusions are as follows:

Briefly, the elevated utilization level of data element significantly spawns the bubble phenomenon in firms' green innovation activities. The mediation mechanisms are as follows: On one hand, the increase in the utilization level of data factor exacerbates the information asymmetry problem. On the other hand, the elevated level of data factor utilization reinforces the technological path lock-in. Initially, the number of green patents increases through selective environmental information disclosure and incremental innovation. Then the actual value of green innovation is restrained through the intrinsic defects of the model and resource mismatch. As a result, the bubble between quantity and quality inflates. Additionally, superstar inventors positively reinforce the negative effect of data elements.

6.2. Revelations

Based on the above findings, this paper obtains revealing suggestions from three aspects: policy regulation, management reform, and market risk control.

First, the government should: (1) enforce penetrating disclosure of technical details and environmental benefit verification data of green patents; (2) pay attention to algorithmic transparency to avoid black-box operation; (3) implement incentives for diversification of technology routes.

Secondly, companies should: (1)establish a long-term evaluation mechanism to curb short-term speculative behaviors; (2) strengthen the ethical responsibility binding mechanism, especially for superstar inventors; (3) ensure the balanced allocation of resources.

Thirdly, financial institutions should: (1) establish early bubble-warning indicators that integrate dimensions such as patent quality, environmental benefit realization rate, and technological maturity; (2) prompt institutional investors to pay attention to the match between the number of patents and the progress of industrialization.

DOI: 10.54254/2754-1169/2025.LD25044

References

[1] MA L N, CHEN Y H. Impact of environmental regulation on the quality and quantity of green innovation [J/OL]. Science Research Management, 1-19 [2025-05-27]. http: //kns.cnki.net/kcms/detail/11.1567.g3.20241205.1634.011.html.
[2] XU B C, LI J H, LI S H. Does China's green credit policy stimulate "innovation bubbles"? Evidence from the quality of corporate green innovation [J]. Journal of China University of Geosciences (Social Sciences Edition), 2023, 23(5): 44-60. doi: 10.16493/j.cnki.42-1627/c.20230818.002.
[3] DENG K. Can QFII inhibit corporate green innovation "bubbles"? [J]. Cooperation Economy and Science, 2024, (21): 124-125. doi: 10.13665/j.cnki.hzjjykj.2024.21.067.
[4] LUO S C, LIU J J, XIONG Z Q. Talent policy support and green innovation performance of polluting enterprises: Empirical evidence from high-level talent subsidies [J]. Industrial Economics Research, 2024, (1): 56-70. doi: 10.13269/j.cnki.ier.2024.01.010.
[5] LI Q, CHEN L. Impact of ESG rating uncertainty on corporate green innovation [J]. Chinese Journal of Management, 2024, 21(12): 1820-1829.
[6] CAO P, CHEN Y B. Research on influencing factors of enterprise data element utilization from a configurational perspective [J]. Science Research Management, 2025, 46(1): 1-11. doi: 10.19571/j.cnki.1000-2995.2025.01.001.
[7] SHI Q C, NIU Y, XU H. Mechanism of enterprise data element utilization affecting investment efficiency: Mediating role of activating redundant resources [J]. Journal of Central University of Finance & Economics, 2023, (11): 105-115. doi: 10.19681/j.cnki.jcufe.2023.11.009.
[8] ZHAO S M, XU H H, GAO W, et al. Enterprise data element utilization, digital innovation and industrial structure modernization [J]. China Soft Science, 2024, (S2): 398-407.
[9] GENG Y, CHEN J Y, LIU R. ESG rating disagreement and corporate green innovation bubbles: Evidence from Chinese A-share listed firms [J]. International Review of Financial Analysis, 2024, 95.
[10] WU F, HU H Z, LIN H Y, et al. Corporate digital transformation and capital market performance: Evidence from stock liquidity [J]. Management World, 2021, 37(7): 130-144+10. doi: 10.19744/j.cnki.11-1235/f.2021.0097.