



Research article

Modeling and optimization of data trading supply chain regulatory strategies based on system dynamics

Jian Yang¹, Wenyu Jia¹, Qianfei Guo¹, Pengfei Pu¹ and Jie Zhang^{2,3,4,*}

¹ School of Information, Shanxi University of Finance and Economics, Taiyuan 030006, Shanxi, China

² Graduate School, Sehan University, Yeongam-gu 58447, Jeollanam-do, Republic of Korea

³ School of Management, Tianjin Normal University, Tianjin, 300387, China

⁴ Technology Innovation Center, Shanxi Information Industry Technology Research Institute Co., Ltd., Taiyuan 030012, Shanxi, China

*** Correspondence:** Email: zhangjie19920717@gmail.com.

Abstract: With the rapid development of the data market and the emergence of new business models, data trading has become a paradigm for allocating production elements. However, most existing studies focused on traditional production elements or single-sided regulation, which did not consider the unique risks of data trading, such as usage tracking difficulties and potential data leakage, and rarely applied dynamic system theory to the analysis of multi-party quality regulation in the data market. In this paper, we first proposed a four-party evolutionary game model of data providers, data trading platforms, data demanders, and regulatory departments. Second, we focused on the evolutionary stability strategies and evolutionary trends of different participants, as well as the influential factors, such as regulatory strategies, regulatory costs, punishment intensity, and the prudent analysis strategy of the data demand side, to analyze the influence of different behavioral strategies on the regulation of data product quality. Third, we verified the stability of the equilibrium point of the four-party game system through system dynamics simulation experiments. Finally, we summarized our work and provided related recommendations for governing agencies and market stakeholders to facilitate the healthy development of the data market and maximize social welfare.

Keywords: data quality; data trading; evolutionary game; system dynamics

Mathematics Subject Classification: Primary: 91A80

1. Introduction

With the rapid development of emerging technologies [1] such as the Internet of Things and artificial intelligence, data has become an indispensable asset and core resource across various industries. It is not only an important tool for driving production and industrial upgrading [2], but also can invigorate the big data industry, helping us realize the true value of data. As people gradually recognize the commercial value of data, data trading has evolved into a new emerging business model. High-quality data products are the cornerstone of ensuring the healthy development of the data trading market [3]. Currently, the demand for data quality in the market has shifted from “self-use” to “external use” and “regulation” [4]. Therefore, the quality of data products is increasingly becoming a focus of attention for data consumers and regulatory agencies.

Compared to traditional production factors, data elements are characterized by low replication costs, fast growth rates, and difficulty in estimating value [5]. However, challenges exist in ensuring data quality and tracking data usage during transactions [6], which may even lead to data breaches and other security issues [7]. To address such risks, [8] utilized evolutionary game analysis to explore stakeholder privacy management in the context of Artificial Intelligence Generated Content (AIGC). These challenges pose significant obstacles to the regulation of data product quality. Centralized trading and decentralized trading are currently the two basic methods in the data trading market [9]. The data market as an impartial third party, enhances trust between buyers and sellers by providing new trading technologies and ensuring the security of transaction evidence, thereby reducing disputes and maximizing the fair value of data, facilitating traceability and regulation, which are advantages that cannot be matched by decentralized trading [10]. Therefore, the establishment of data markets and trading centers has been regarded by governments as an important task in building a data market. By guiding data products into centralized trading, the legality and orderliness of data transactions are realized, and the introduction and improvement of relevant laws and regulations are accelerated. Data quality standards and norms are established, data certification systems are promoted, and regulatory technologies are innovated. Currently, academic research on data product quality regulation is mainly focused on theoretical analysis, involving compliance review [11], data governance frameworks [12], technologies and tools [13], and data quality assessment standards [14]. However, from an economic perspective, data trading is a continuous, bounded rationality multi-agent game. Previous studies have not yet explored the evolutionary mechanism of data product quality regulation [15], especially the analysis of the behavior and interaction strategies of various stakeholders in data trading [16], for example, how the behaviors of data providers, data markets, regulatory agencies, and data consumers influence the evolution of data product quality regulation.

Therefore, this paper establishes a four-party evolutionary game model involving data providers, data markets, data consumers, and regulatory agencies to analyze the decision evolution and optimization strategies of each participant in the process of data product quality regulation. In practical applications, combined with system dynamics for simulation, key factors affecting data product quality regulation are identified. Based on this, feasible policy recommendations are proposed to assist government regulatory agencies in improving legislation and enforcement.

The rest of this paper is organized as follows. In Section 2, we summarize previous research on data sharing and trading, and the application of system dynamics and evolutionary game theory to data quality regulation. We analyze various scenarios of strategic stability among data providers, data

trading markets, regulators, and data consumers in Section 3. In Section 4, we construct a system dynamics model and explore the impact of the cost of strict government regulation, fines for data providers and data markets, and the cost of prudent analysis for data consumers on the evolution of data quality regulation through numerical simulation methods. In Section 5, we summarize the main conclusions of this paper, put forward relevant recommendations, and list the future research agenda.

2. Literature review

2.1. Data sharing and trading

Data trading as a key aspect of the market-oriented allocation of data elements, has been the focus of research by scholars from various perspectives, including the types of trading platforms [17], trading processes [18], pricing of data elements [19], and the development paths of the data element market [20]. Data has become another important production factor following labor, capital, and land, and its quality issues have gained wide attention in academia. Related research primarily focuses on evaluation metrics and models for data quality, as well as governance mechanisms. Regarding data quality assessment, [21] used factor analysis on user perception surveys to construct a data resource quality evaluation system that includes five dimensions: connotation, service, expectation, reputation, and potential value, and 18 secondary indicators. [22], from a panoramic big data quality assessment perspective composed of “human logic event logic-data logic-mechanism logic,” proposed a data quality assessment framework with 56 indicators, such as usefulness and data modeling quality, based on a review of literature related to data quality. [23] indicated that for the evaluation of the asset value of data elements in China, the random forest model is more suitable for data asset value assessment. In terms of optimizing data quality, [24] suggested improving the data quality governance structure, implementing scientific data quality audits, and establishing a collaborative governance system for scientific data quality among stakeholders. [25] proposed that the management of open data quality by the Chinese government should emphasize metadata standards and build automated quality assessment systems to monitor data quality in real-time. [26] reviewed technological methods that can be used to improve data quality, such as using data visualization techniques to integrate fragmented data and transform it into data products that meet the requirements of demand-side users. [27] constructed a data element system architecture that includes technical, data, and policy layers, and explored four technical paths for supporting the cultivation of the national integrated big data center in the data market.

However, [28] pointed out through a systematic review of data trading literature that existing research is mostly macro-level judgments or empirical analyses, lacking observation of data trading practices and the construction and testing of models for influencing factors. Therefore, studying the operational mechanisms of data trading quality regulation from the perspective of micro-entities has become a new research direction.

2.2. System dynamics and evolutionary games in data quality governance

System dynamics is an effective tool for exploring complex dynamic management [29] and can complement evolutionary game models by testing the validity of evolutionary game models through simulation [30]. It allows for the dynamic analysis of decision-making behaviors of various entities

involved in data trading quality regulation from a micro perspective. [31] modeled the dynamic changes in data quality using system dynamics, combined with evolutionary game analysis, to examine participants' strategic choices under different data quality conditions, revealing the positive effects of data quality improvement on market stability. [32] constructed a data trading model that integrates system dynamics (SDs) and evolutionary game theory (EGT), simulating the strategic evolution and market dynamics of data providers and consumers under different incentive mechanisms. Similarly, [33] employed evolutionary game theory to analyze the strategic interactions among three stakeholders in the big data trading market.

However, there are few studies that apply system dynamics and evolutionary game theory to data quality supervision at the same time [34]. In addition, the above studies usually only involve two or three subjects among data providers: trading platforms, consumers, and government departments. In reality, data trading is a complex process involving the interaction of multiple subjects. Therefore, the evolutionary trend and micro-mechanism of data product quality supervision deserve in-depth exploration.

Based on the above analysis, we constructed a four-party evolutionary game model of data providers, data markets, regulatory authorities, and data demanders under bounded rationality conditions, and used the system dynamics method to examine the causal relationship and feedback loop between relevant variables and subject strategies, identify key variables affecting data quality supervision, and conduct further simulation analysis to clarify the strategic focus of government regulatory authorities.

3. Model construction

With the continuous growth in data demand, the economic value of data is becoming increasingly high, and market competition is becoming more intense. Some data providers focus only on immediate interests without paying attention to data quality. Moreover, professional and independent data markets should provide strict quality inspection services and act as "referees" of data product quality together with regulatory agencies, maintaining market order. However, rent-seeking behavior between data markets and data providers and weak government regulation bring uncertainty to market order and consumer protection. To address these issues, a multi-agent quality supervision game model is proposed in this paper, as shown in Figure 1. The data suppliers produce, collect, and process data, and they are supervised and subsidized by government regulatory authorities, which also conduct quality inspections. The government regulatory authorities are supervised and incentivized by higher-level regulatory bodies and receive financial support. Data suppliers provide data products to data exchanges, which undergo regulatory testing. Afterward, data consumers obtain the data products from the exchange, analyze and use them, and provide feedback or complaints. There is a potential for rent-seeking behavior from data consumers, which the government regulatory authorities address by improving data governance. The process concludes when data consumers purchase and use the data products, and the cycle continues.

It is assumed that the data trading process mainly involves four participants, namely data providers, data markets, regulatory authorities, and data demanders. There is a probability of x that high-quality data is provided by data providers, and a probability of $1 - x$ that low-quality data is provided. Quality checks are performed by the data market before the circulation of data products, with a probability of y

that strict checks are conducted and a probability of $1 - y$ that lenient checks are conducted. To resolve market failures and in the interest of the public, regulatory authorities adopt strong regulation with a probability of z and conventional regulation with a probability of $1 - z$. Data demanders carefully analyze the quality information released by the data market with a probability of q , and directly trust this information with a probability of $1 - q$. Under bounded rationality, strategies are continuously adjusted by the participants through trial and error, learning, and imitation, ultimately reaching a stable equilibrium state.

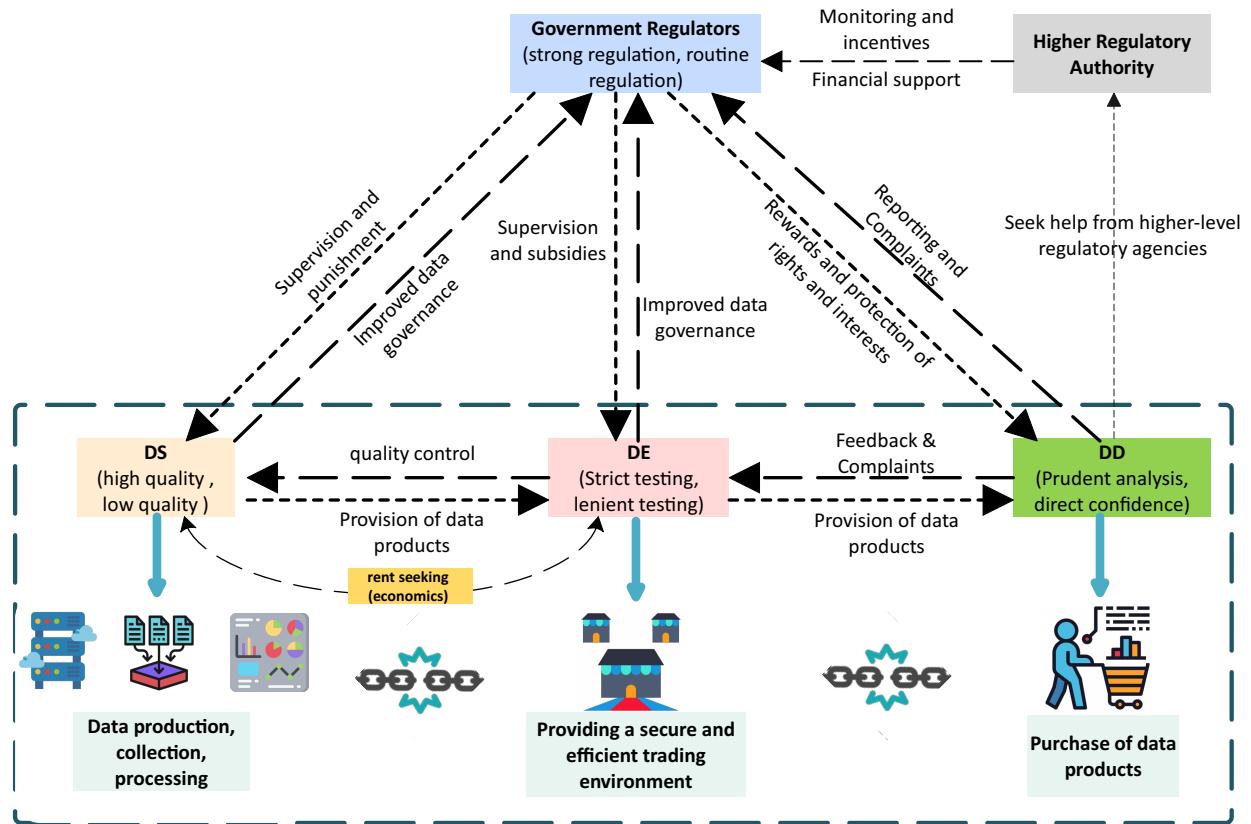


Figure 1. Logical relationship among the four game players.

When data providers focus on quality ethics and provide high-quality data products, the cost is Cp_1 . If low-quality data products are provided, the cost is Cp_2 . High-quality data can pass data market inspections without incurring extra costs; however, low-quality data will incur rent-seeking costs Cp_3 , with the condition $Cp_2 + Cp_3 < Cp_1$. Data product revenue is Rp . The cost for a loose inspection by the data market is Ce_1 , while the cost for a strict inspection is Ce_2 . If data quality is not up to standard, data providers will be fined $\lambda * Fp$ by regulatory authorities. A data market that refuses rent-seeking can receive a reward Re . If the data provider offers high-quality data and the data market opts for a loose inspection, even if no consumer rights are harmed, this behavior can undermine quality regulation and result in a fine of $\eta * Fe$, where λ and η are the tolerance coefficients of the regulatory authority for quality regulation failures by data providers and data markets, respectively ($0 < \lambda < 1, 0 < \eta < 1$). If the data market and data provider collude to seek rents and conduct loose inspections, they will be fined Fe by the regulatory authority.

Table 1. Parameters and symbols.

Parameter	Meaning
C_{p_1}	The cost of providing high-quality products
C_{p_2}	The cost of providing low-quality products
C_{p_3}	Rent-seeking cost paid by data providers to data markets
R_p	Revenue generated by data products
F_p	Fine imposed on data providers for failing quality checks
L_p	Reputational loss for data providers
Ce_1	The cost of conducting lenient quality checks
Ce_2	The cost of conducting strict quality checks
R_e	Government reward for rejecting rent-seeking
F_e	Fine imposed on data markets by the government
L_e	Reputational loss for data markets
Cg_1	The cost of implementing strong regulatory strategies
Cg_2	The cost of routine regulation
F_g	Administrative penalties and loss of credibility for regulatory departments
L_g	Loss of trust in government regulatory departments
R_g	Benefit from increased government trust
γ	Degree of ethical deficiency identified in data providers after government quality review
η	Degree of ethical deficiency identified in data markets after government quality review
C_d	The cost of conducting careful analysis
L_d	Loss incurred due to low-quality data products
L_s	Social loss caused by the entry of low-quality products into the market
R_d	Revenue generated from high-quality products
R_s	Social benefit from the entry of high-quality products into the market
α	Utility coefficient of low-quality products
β	Probability of requesting higher-level government oversight of regulatory departments

The cost of conducting cautious analysis by data demanders is C_d , and the identification of false quality information will bring reputational losses of L_p to data providers and L_e to the data market. The cost of implementing strict regulation by regulatory agencies is Cg_1 , and the cost of daily regulation is Cg_2 . Due to the abstract nature of rules and significant discretionary power, the intensity of regulation by agencies may be unstable, necessitating supervision by higher-level government to ensure the proper exercise of power and selection of regulatory strategies. Higher-level government is introduced as a supervisor to ensure the effective implementation of quality regulation. When data demanders conduct cautious analysis, there is a probability β ($0 < \beta < 1$, positively correlated with q) that higher-level government will be asked to supervise the exercise of power by regulatory agencies. If it is found that the agency has failed to effectively regulate quality ethics, it will face administrative penalties and reputational losses, with a total loss of βF_g .

Low-quality data products reduce overall value and utility, with a utility coefficient of α ($0 < \alpha < 1$). Once in the market, data demanders incur a loss of $\alpha * L_d$, while market disruption causes a social loss of L_s and undermines government trust, resulting in a loss of L_g . Conversely, high-quality data products can maximize their value and utility, generating a revenue of R_d for data demanders. The quality signal also brings a societal benefit of R_s and enhances government trust, with an associated benefit of R_g .

For the convenience of understanding, the key symbols used in this paper are shown in Table 1. Based on the above assumptions, we construct the payoff matrix of the evolutionary game among data providers, data markets, regulatory authorities, and data demanders, as shown in Table 2.

Table 2. Payoff matrix.

Provider	Market	Regulator	Demander	Payoff Vector (U_p, U_m, U_g, U_d)
High-Quality Data (x)	Strict Detection (y)	Strong (z)	Careful (q)	$(R_p - C_{p1}, -C_{e2}, -C_{g1} + R_s + R_g, R_d - C_d)$
		Conv. ($1 - z$)	Believe ($1 - q$)	$(R_p - C_{p1}, -C_{e2}, -C_{g1} + R_s + R_g, R_d)$
	Loose Detection ($1 - y$)	Strong (z)	Careful (q)	$(R_p - C_{p1}, -C_{e2}, -C_{g2} + \beta F_g + R_s + R_g, R_d - C_d)$
		Conv. ($1 - z$)	Believe ($1 - q$)	$(R_p - C_{p1}, -C_{e2}, -C_{g2} + R_s + R_g, R_d)$
	Strict Detection (y)	Strong (z)	Careful (q)	$(R_p - C_{p1}, -C_{e1} - \eta F_e, -C_{g1} + R_s + R_g + \eta F_e, R_d - C_d)$
		Conv. ($1 - z$)	Believe ($1 - q$)	$(R_p - C_{p1}, -C_{e1} - \eta F_e, -C_{g1} + R_s + R_g + \eta F_e, R_d)$
Low-Quality Data ($1 - x$)	Strict Detection (y)	Strong (z)	Careful (q)	$(-C_{p2} - C_{p3} - \lambda F_p, R_e - C_{e2}, -R_e - C_{g1} + \lambda F_p, 0)$
		Conv. ($1 - z$)	Believe ($1 - q$)	$(-C_{p2} - C_{p3} - \lambda F_p, R_e - C_{e2}, -R_e - C_{g1} + \lambda F_p, 0)$
	Loose Detection ($1 - y$)	Strong (z)	Careful (q)	$(-C_{p2} - C_{p3}, -C_{e2}, -C_{g2}, 0)$
		Conv. ($1 - z$)	Believe ($1 - q$)	$(-C_{p2} - C_{p3}, -C_{e2}, -C_{g2}, 0)$
	Strict Detection (y)	Strong (z)	Careful (q)	$(R_p - C_{p2} - C_{p3} - L_p - \lambda F_p, C_{p3} - C_{e1} - F_e - L_e, \Delta G_1, -C_d)$
		Conv. ($1 - z$)	Believe ($1 - q$)	$(R_p - C_{p2} - C_{p3} - L_p - \lambda F_p, C_{p3} - C_{e1} - F_e, \Delta G_2, -\alpha L_d)$
	Loose Detection ($1 - y$)	Strong (z)	Careful (q)	$(R_p - C_{p2} - C_{p3} - L_p, C_{p3} - C_{e1} - L_e, -C_{g2} - \beta F_g - L_g - L_s, -C_d)$
		Conv. ($1 - z$)	Believe ($1 - q$)	$(R_p - C_{p2} - C_{p3} - C_{e1}, C_{p3} - C_{e1}, -C_{g2} - L_g - L_s, -\alpha L_d)$

Note: U_p, U_m, U_g, U_d represent the utility functions of the data provider, data market, regulatory authority, and data demander, respectively.

For brevity in the table, let $\Delta G_1 = -C_{g1} + \lambda F_p + F_e - L_g - L_s$ and $\Delta G_2 = -C_{g1} + \lambda F_p + F_e - L_g - L_s$.

4. Analysis of the evolutionary game model

4.1. Stability analysis

4.1.1. Stability analysis of data provider strategies

Based on Table 2, the expected payoff for data providers providing high-quality data products E_{11} , the expected payoff for providing low-quality products E_{12} , and the average expected payoff E_1 are as follows:

$$E_{11} = R_p - Cp_1 \quad (4.1)$$

$$E_{12} = -Cp_2 - Cp_3 - z\lambda F_p + (1 - y)(R_p - qL_p) \quad (4.2)$$

$$E_1 = xE_{11} + (1 - x)E_{12} \quad (4.3)$$

Therefore, the replicator dynamic equation for the data providers and the first derivative with respect to x are as follows:

$$\begin{aligned} F(x) &= \frac{dx}{dt} = x \cdot (1 - x) \cdot (E_{11} - E_{12}) \\ &= x \cdot (1 - x) \cdot (y \cdot R_p - Cp_1 + Cp_2 + Cp_3 + z \cdot \lambda F_p + (1 - y) \cdot q \cdot L_p) \end{aligned} \quad (4.4)$$

$$\frac{dF(x)}{dx} = (1 - 2x) \cdot (y \cdot R_p - Cp_1 + Cp_2 + Cp_3 + z \cdot \lambda F_p + (1 - y) \cdot q \cdot L_p) \quad (4.5)$$

According to the stability theorem of differential equations, the probability that the data provider will choose to provide high-quality data is in a stable state if and only if: $F(x) = 0$ and $\frac{dF(x)}{dx} < 0$.

Proposition 1. *When $y = y^*$, the stable strategy of data providers cannot be determined. When $0 < y < y^*$, the stable strategy for data providers is to provide low-quality data products. When $y^* < y < 1$, the stable strategy for data providers is to provide high-quality data products. Here,*

$$y^* = Cp_1 - Cp_2 - Cp_3 - z \cdot \lambda F_p - y \cdot L_p \cdot R_p - q \cdot L_p \quad (4.6)$$

Proof. Let $G(y) = y^*R_p - Cp_1 + Cp_2 + Cp_3 + z \cdot \lambda F_p + (1 - y)q \cdot L_p$. Since $\frac{dG(y)}{dy} > 0$, $G(y)$ is an increasing function with respect to y . When $y = y^*$, we have $F(x) = 0$ and $\frac{dF(x)}{dx} = 0$. In this situation, all $x \in [0, 1]$ are in a stable state, so a stable strategy cannot be determined. When $0 < y < y^*$, we find that $F(x)|_{x=0} = 0$ and $\frac{dF(x)}{dx}|_{x=0} < 0$. This indicates that $x = 0$ is stable. When $y^* < y < 1$, we have $F(x)|_{x=1} = 0$ and $\frac{dF(x)}{dx}|_{x=1} < 0$. This indicates that $x = 1$ is stable. \square

Proposition 1 indicates that if the data market can place greater emphasis on data quality and conduct quality assessments with a rigorous and responsible attitude, rather than lowering standards to cater to the interests of data providers, then when data providers choose to provide low-quality data products, they will not only be unable to achieve the expected returns but may also face penalties from regulatory authorities and damage to their reputation. Based on this, a replicative dynamic phase diagram for data providers can be derived, as shown in Figure 2. Specifically, the volume of section V_{x0} represents the probability that data providers choose to provide low-quality data products, while the volume of section V_{x1} represents the probability of choosing to provide high-quality products.

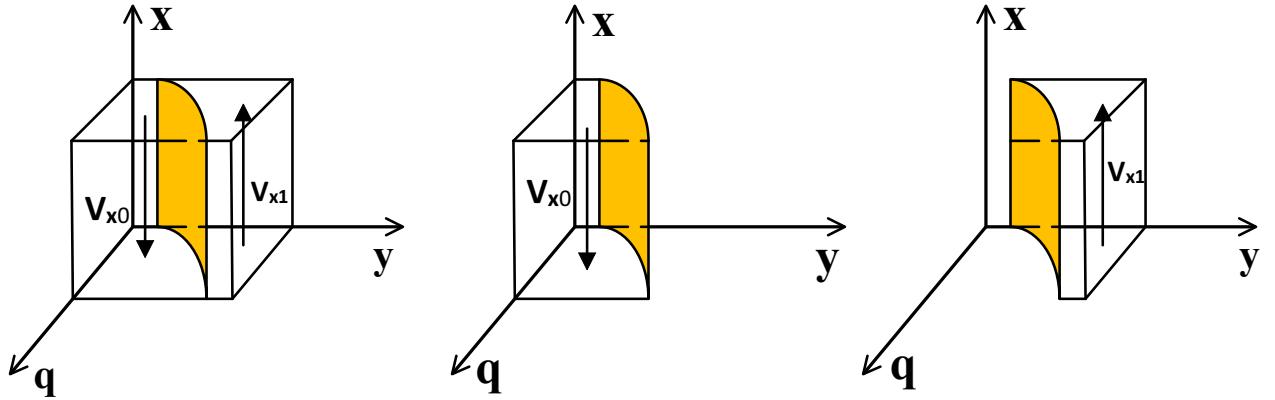


Figure 2. Replicator dynamics phase diagram for data providers.

4.1.2. Stability analysis of the data market strategy

The expected returns for a data market choosing strict verification, lenient verification, and the average expected return (E_{21}, E_{22}, E_2) are, respectively:

$$E_{21} = z \cdot (1 - x) \cdot R_e - C e_2 \quad (4.7)$$

$$E_{22} = (1 - x) \cdot C p_3 - C e_1 - (1 - x) \cdot q \cdot L_e - F_e \cdot (x \cdot \eta + (1 - x)) \quad (4.8)$$

$$E_2 = y \cdot E_{21} + (1 - y) \cdot E_{22} \quad (4.9)$$

Hence, the replication dynamic equation of the data market and the first derivative with respect to y are:

$$\begin{aligned} F(y) &= y \cdot (E_{21} - E_{22}) \\ &= y(1 - y)((1 - x)(z \cdot R_e - C p_3 + q \cdot L_e) - C e_2 + C e_1 + z \cdot F_e(x \cdot \eta + (1 - x))) \end{aligned} \quad (4.10)$$

$$\frac{dF(y)}{dy} = (1 - 2y)((1 - x)(z \cdot R_e - C p_3 + q \cdot L_e) - C e_2 + C e_1 + z \cdot F_e(x \cdot \eta + (1 - x))) \quad (4.11)$$

Proposition 2. *When $z = z^*$, the stable strategy of the data market cannot be determined. When $0 < z < z^*$, the stable strategy is lenient verification. When $z^* < z < 1$, the stable strategy is strict verification. The threshold is given by: $z^* = \frac{(1-x)(C p_3 - q \cdot L_e) + C e_2 - C e_1}{(1-x)R_e + F_e[x\eta + (1-x)]}$.*

Proof. We define $G(z)$ as:

$$G(z) = (1 - x)(z \cdot R_e - C p_3 + q \cdot L_e) - C e_2 + C e_1 + z F_e[x\eta + (1 - x)] \quad (4.12)$$

Since $\frac{dG(z)}{dz} > 0$, $G(z)$ is an increasing function of z . When $z = z^*$, we have $F(y) = 0$ and $\frac{dF(y)}{dy} = 0$. In this case, $y \in [0, 1]$ is in a stable state, so the stable strategy cannot be determined. When $0 < z < z^*$, $F(y)$ at $y = 0$ is $F(y)|_{y=0} = 0$ and $\frac{dF(y)}{dy}|_{y=0} < 0$. Here, $y = 0$ is stable. When $z^* < z < 1$, $F(y)$ at $y = 1$ is $F(y)|_{y=1} = 0$ and $\frac{dF(y)}{dy}|_{y=1} < 0$. In this scenario, $y = 1$ is stable. \square

Proposition 2 states that if government regulatory agencies can adapt flexibly to new regulatory environments, innovate regulatory methods, and prevent potential violations instead of sticking to

routine regulation, then when data markets neglect quality ethics, relax inspections, and engage in rent-seeking with data providers, they will face government penalties and resistance from data demanders, leading to a decline in reputation. Therefore, as the likelihood of increased government regulation rises, data markets should adopt a stable strategy of strict inspections and reject rent-seeking.

Therefore, the replicator dynamic phase diagram of the data market can be derived, as shown in Figure 3. The volume of V_{y0} represents the probability of the data market choosing lenient detection, and the volume of the V_{y1} represents the probability of choosing strict detection.

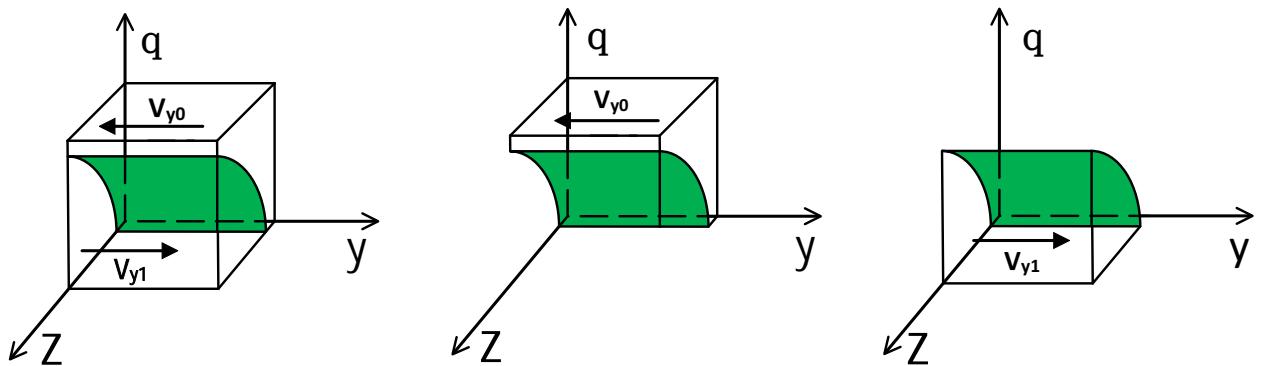


Figure 3. Dynamic phase diagram of the replication of the data market.

4.1.3. Stability analysis of regulatory agencies' strategies

The expected returns and average expected returns for a government regulatory agency choosing either strict or regular regulation (E_{31}, E_{32}, E_3) are:

$$E_{31} = -Cg_1 + x(R_s + R_g) + (1 - x)(\lambda F_p - yR_e) - (1 - x)(1 - y)(L_g + L_s) + (1 - y)F_e(x\eta + (1 - x)) \quad (4.13)$$

$$E_{32} = -Cg_2 + x(R_s + R_g) + (y - xy - 1)q\beta F_g - (1 - x)(1 - y)(L_g + L_s) \quad (4.14)$$

$$E_3 = zE_{31} + (1 - z)E_{32} \quad (4.15)$$

The replicator dynamic equation for the government regulatory agency and the first derivative with respect to z are obtained as follows:

$$\begin{aligned} F(z) &= z(E_{31} - E_3) \\ &= z(1 - z)(-Cg_1 + Cg_2 + (1 - x)(\lambda F_p - yR_e) \\ &\quad + (1 - y)F_e(x\eta + (1 - x)) + (xy + 1 - y)q\beta F_g) \end{aligned} \quad (4.16)$$

$$\begin{aligned} \frac{dF(z)}{dz} &= (1 - 2z)(-Cg_1 + Cg_2 + (1 - x)(\lambda F_p - yR_e) \\ &\quad + (1 - y)F_e(x\eta + (1 - x)) + (xy + 1 - y)q\beta F_g) \end{aligned} \quad (4.17)$$

Proposition 3. When $q = q^*$, the stable strategy for the data market cannot be determined. When $q = q^*$, the stable strategy for the government regulatory agency is regular regulation. When $q^* < q < 1$,

the stable strategy is strict regulation. The threshold is given by:

$$q^* = \frac{C_{g1} - C_{g2} - (1-x)\lambda F_p + (1-x)yR_e - (1-y)[x\eta + (1-x)]F_e}{(xy + 1 - y)q\beta F_g} \quad (4.18)$$

Proof. We start with the function $G(q)$:

$$G(q) = -Cg_1 + Cg_2 + (1-x)(\lambda F_p - yR_e) + (1-y)F_e(x\eta + (1-x)) + (xy + 1 - y)q\beta F_g \quad (4.19)$$

Since $\frac{dG(q)}{dq} > 0$, we conclude that $G(q)$ is an increasing function of q . When $z = z^*$, we have $F(z) = 0$ and $\frac{dF(z)}{dz} = 0$, which indicates that all states $z \in [0, 1]$ are stable, and the stable strategy cannot be determined. When $0 < z < z^*$, we observe that $F(z)|_{z=0} = 0$ and $\frac{dF(z)}{dz}|_{z=0} < 0$, which indicates that $z = 0$ is stable. When $z^* < z < 1$, we have $F(z)|_{z=1} = 0$ and $\frac{dF(z)}{dz}|_{z=1} < 0$, which indicates that $z = 1$ is stable. \square

Proposition 3 indicates that if data consumers become more quality-conscious and carefully analyze the quality inspection information of data products, rather than blindly trusting the data market and not monitoring the exercise of regulatory power by government agencies, the government regulatory agencies will lose credibility. Therefore, as data consumers increasingly analyze carefully and seek oversight from higher government authorities on the exercise of power by regulatory agencies, the stable strategy for government regulatory agencies becomes strict regulation. Based on the analysis of the strategy evolution of government regulatory agencies, a replicator dynamic phase diagram can be derived, as shown in Figure 4.

According to Figure 4, the volume of section V_{z0} represents the probability of the government regulatory agency choosing regular regulation, while the volume of section V_{z1} represents the probability of choosing strict regulation.

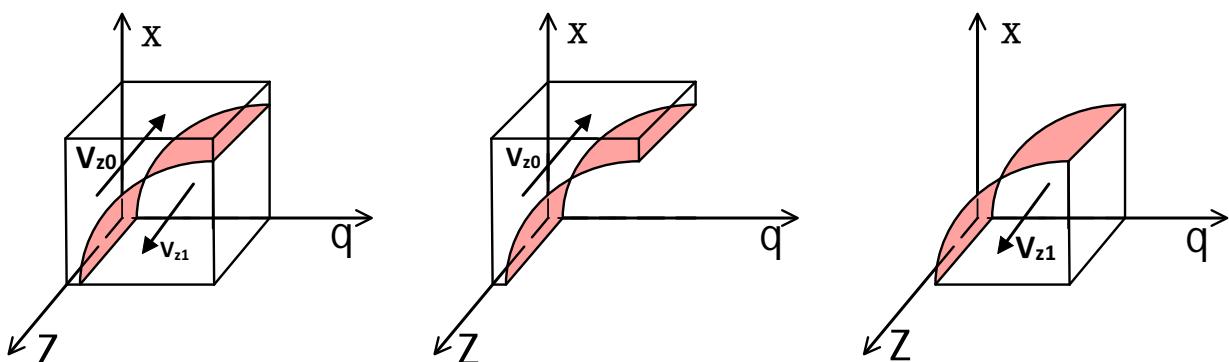


Figure 4. Dynamic phase diagram of replication by government regulators.

4.1.4. Stability analysis of data demander strategies

The expected returns and average expected returns for data demanders choosing cautious analysis or direct trust (E_{41}, E_{42}, E_4) are as follows: From this, the replicator dynamic equation for data demanders and its first derivative with respect to q can be obtained as follows:

$$\begin{aligned} F(q) &= q(E_{41} - E_4) \\ &= q(1-q)(1-x)(1-y)\alpha L_d - (x + (1-x)(1-e))C_d \end{aligned} \quad (4.20)$$

$$\frac{dF(q)}{dq} = (1-2q)((1-x)(1-y)\alpha L_d - (x + (1-x)(1-e))C_d) \quad (4.21)$$

Proposition 4. *When $y = y^*$, the stable strategy for data demanders cannot be determined. When $0 < y < y^*$, the stable strategy is cautious analysis. When $y^* < y < 1$, the stable strategy is direct trust. The threshold is given by:*

$$y^* = \frac{C_d + (x-1)\alpha L_d}{(x-1)(-C_d + \alpha L_d)} \quad (4.22)$$

Proof. Given the function:

$$G(y) = (1-x)(1-y)\alpha L_d - [x + (1-x)(1-e)]C_d, \quad (4.23)$$

we have $\frac{dG(y)}{dy} < 0$, which shows that $G(y)$ is a decreasing function of y . When $y = y^*$, we find that $F(q) = 0$ and $\frac{dF(q)}{dq} = 0$. In this case, all $q \in [0, 1]$ are stable, and the stable strategy cannot be determined. When $0 < y < y^*$, we have $F(q)|_{q=1} = 0$ and $\frac{dF(q)}{dq}|_{q=1} < 0$, which indicates that $q = 1$ is stable. When $y^* < y < 1$, we find $F(q)|_{q=0} = 0$ and $\frac{dF(q)}{dq}|_{q=0} < 0$, which indicates that $q = 0$ is stable. \square

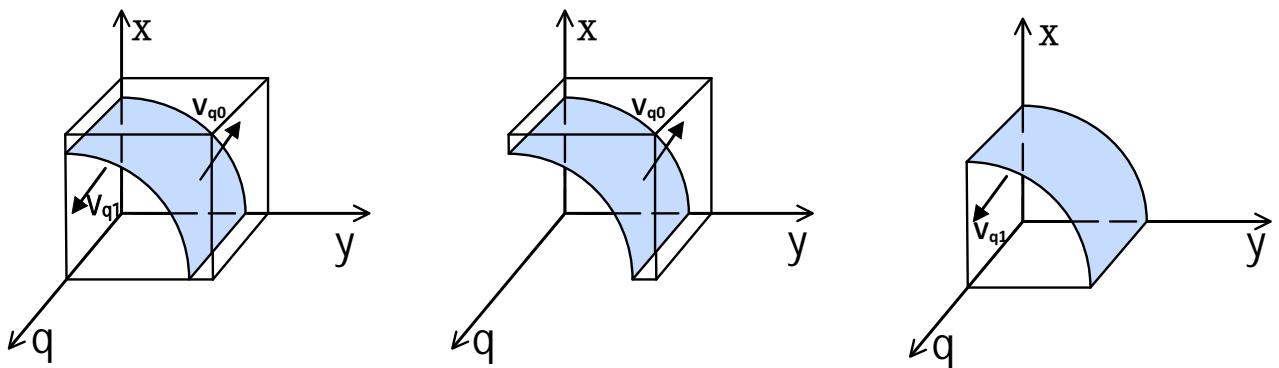


Figure 5. Dynamic phase diagram of replication on the data demand side.

Proposition 4 indicates that if a data market emphasizes data quality management and conducts rigorous inspections, data demanders will be more inclined to trust directly. This is because strict inspections can filter out low-quality data products, increasing the expected returns for data demanders,

leading them to choose direct trust. Conversely, if a data market neglects quality management, conducts lax inspections, or colludes with data providers to let low-quality data pass, the credibility of the data market will decrease, and data demanders will tend to opt for cautious analysis. This leads to the replication dynamic phase diagram for data demanders, as shown in Figure 5.

The volume of the Vq_0 region represents the probability that data demanders choose to trust directly. The volume of the Vq_1 region represents the probability that they choose cautious analysis.

4.2. Stability analysis of the equilibrium points

In asymmetric dynamic games, mixed strategy equilibria are not evolutionarily stable, so we only analyze the pure strategy equilibrium points of the evolutionary game system. From the equation $F(x) = F(y) = F(z) = F(q) = 0$, we find that the game system has 16 pure strategy equilibrium points. Although there are 16 potential equilibrium points in the proposed four-party game, not all possess practical significance or stability properties. To analyze the stability of equilibrium points, we use the Lyapunov indirect method for determination.

When all the eigenvalues of the Jacobian matrix are negative, the equilibrium point is an evolutionarily stable strategy for the system. If at least one eigenvalue is positive, the equilibrium point is unstable. When the eigenvalues are both zero and negative, the equilibrium point is in a critical state, and its stability cannot be determined solely by the sign of the eigenvalues. By substituting each equilibrium point into the Jacobian matrix, we obtain the corresponding eigenvalues, as shown in Table 3.

Table 3. Eigenvalues of equilibrium points.

	Γ_1	Γ_2	Γ_3	Γ_4
(0,0,0,0)	$\alpha L_d - C_d$	$Cp_2 - Cp_1 + Cp_3$	$Ce_2 - Ce_1 + Cp_3$	$Cg_2 - Cg_1 + F_e + \lambda F_p$
(0,1,0,0)	0	$Ce_2 - Ce_1 + Cp_3$	$Cg_2 - Cg_1 - R_e + \lambda F_p$	$Cp_2 - Cp_1 + Cp_3 + R_p$
(0,0,1,0)	$\alpha L_d - C_d$	$Cp_2 - Cp_1 + Cp_3 + F_p \lambda$	$Cg_1 - Cg_2 - F_e - F_p \lambda$	$Ce_1 - Ce_2 - Cp_3 + F_e + R_e$
(0,0,0,1)	$C_d - L_d \alpha$	$Ce_1 - Ce_2 - Cp_3 + L_e$	$Cp_2 - Cp_1 + Cp_3 + L_p$	$Cg_2 - Cg_1 + F_e + F_g \beta + F_p \lambda$
(0,1,1,0)	0	$Cg_1 - Cg_2 + R_e - F_p \lambda$	$Ce_2 - Ce_1 + Cp_3 - F_e - R_e$	$Cp_2 - Cp_1 + Cp_3 + R_p + F_p \lambda$
(0,1,0,1)	0	$Cg_2 - Cg_1 - R_e + F_p \lambda$	$Ce_2 - Ce_1 + Cp_3 - L_e$	$Cp_2 - Cp_1 + Cp_3 + R_p$
(0,0,1,1)	$C_d - L_d \alpha$	$Cg_1 - Cg_2 - F_e - F_g \beta - F_p \lambda$	$Cp_2 - Cp_1 + Cp_3 + L_p + F_p \lambda$	$Ce_1 - Ce_2 - Cp_3 + F_e + L_e + R_e$
(1,0,0,0)	$Ce_1 - Ce_2$	$-C_d$	$Cg_2 - Cg_1 + F_e \eta$	$Cp_1 - Cp_2 - Cp_3$
(1,1,0,0)	$Ce_2 - Ce_1$	$Cg_2 - Cg_1$	$-C_d$	$Cp_1 - Cp_2 - Cp_3 - R_p$
(1,0,1,0)	$-C_d$	$Ce_1 - Ce_2 + F_e \eta$	$Cg_1 - Cg_2 - F_e \eta$	$Cp_1 - Cp_2 - Cp_3 - F_p \lambda$
(1,0,0,1)	C_d	$Ce_1 - Ce_2$	$Cg_2 - Cg_1 + F_g \beta + F_e \eta$	$Cp_1 - Cp_2 - Cp_3 - L_p$
(1,1,1,0)	$Cg_1 - Cg_2$	$-C_d$	$Ce_2 - Ce_1 - F_g \beta$	$Cp_1 - Cp_2 - Cp_3 - R_p$
(1,1,0,1)	C_d	$Ce_2 - Ce_1$	$Cg_2 - Cg_1 + F_g \beta$	$Cp_1 - Cp_2 - Cp_3 - R_p$
(1,0,1,1)	C_d	$Ce_1 - Ce_2 + F_e \eta$	$Cg_1 - Cg_2 + F_g \beta + F_e \eta$	$Cp_1 - Cp_2 - Cp_3 - L_p - F_p \lambda$
(1,1,1,1)	C_d	$Cg_1 - Cg_2 - F_g \beta$	$Ce_2 - Ce_1 + F_e \beta$	$Cp_1 - Cp_2 - Cp_3 - R_p - F_p \lambda$

$$J = \begin{bmatrix} \frac{\partial F(x)}{\partial x} & \frac{\partial F(x)}{\partial y} & \frac{\partial F(x)}{\partial z} & \frac{\partial F(x)}{\partial q} \\ \frac{\partial x}{\partial F(y)} & \frac{\partial y}{\partial F(y)} & \frac{\partial z}{\partial F(y)} & \frac{\partial q}{\partial F(y)} \\ \frac{\partial x}{\partial F(z)} & \frac{\partial y}{\partial F(z)} & \frac{\partial z}{\partial F(z)} & \frac{\partial q}{\partial F(z)} \\ \frac{\partial x}{\partial F(q)} & \frac{\partial y}{\partial F(q)} & \frac{\partial z}{\partial F(q)} & \frac{\partial q}{\partial F(q)} \end{bmatrix} \quad (4.24)$$

Based on Lyapunov stability theory and economic rationality, we explicitly exclude points that are mathematically unstable or economically contradictory. Consequently, this study focuses on analyzing the most representative asymptotic stable states: the “Market Failure” scenario and the “Ideal Equilibrium” scenario. This selection allows for a focused discussion on the transformation mechanisms between suboptimal and optimal regulatory outcomes. By analyzing the stability of the 16 equilibrium points in Table 3, we can derive the following five types of stability conditions, as shown in Table 4.

Table 4. Possible stable points and their stable conditions.

	Equilibrium Point	Stability	Stability Condition
Situation 1	(0,0,1,0)	ESS	$\alpha L_d - C_d < 0$
			$Cp_2 - Cp_1 + Cp_3 + F_p\lambda < 0$
			$Cg_1 - Cg_2 - F_e - F_p\lambda < 0$
			$Ce_1 - Ce_2 - Cp_3 + F_e + R_e < 0$
			$Ce_1 - Ce_2 + F_e\eta < 0$
Situation 2	(1,0,1,0)	ESS	$Cg_1 - Cg_2 - F_e\eta < 0$
			$Cp_1 - Cp_2 - Cp_3 - F_p\lambda < 0$
			$C_d - L_d\alpha < 0$
Situation 3	(0,0,1,1)	ESS	$Cg_1 - Cg_2 - F_e - F_g\beta - F_p\lambda < 0$
			$Cp_2 - Cp_1 + Cp_3 + L_p + F_p\lambda < 0$
			$Ce_1 - Ce_2 - Cp_3 + F_e + L_e + R_e < 0$
			$\alpha L_d - C_d < 0$
			$Cp_2 - Cp_1 + Cp_3 < 0$
Situation 4	(0,0,0,0)	ESS	$Ce_1 - Ce_2 - Cp_3 < 0$
			$Cg_2 - Cg_1 + F_e + \lambda F_p < 0$
			$C_d - L_d\alpha < 0$
Situation 5	(0,0,0,1)	ESS	$Ce_1 - Ce_2 - Cp_3 + L_e < 0$
			$Cp_2 - Cp_1 + Cp_3 + L_p < 0$
			$Cg_2 - Cg_1 + F_e + F_g\beta + F_p\lambda < 0$

From Tables 3 and 4, it is evident that when the government regulatory agency chooses a strong regulation strategy, there are three possible stable strategy combinations: (0,0,1,0), (1,0,1,0), and (0,0,1,1). Among these, the strategy combination (1,0,1,0), where the data provider provides

high-quality data products, becomes the evolutionarily stable strategy combination for the system. When the government opts for a conventional regulation strategy, there are two possible stable strategy combinations: (0,0,0,0) and (0,0,0,1). In this case, there is no stable strategy combination where the data provider provides high-quality products.

In the context of a rapidly developing but not yet mature data trading market, effective government intervention and strong regulation are crucial factors in ensuring the healthy development of the data market. Therefore, this paper focuses only on analyzing situations 1, 2, and 3.

Situation 1. When $\alpha L_d - C_d < 0$, $Ce_1 - Ce_2 + Cp_3 + F_e + R_e < 0$, and $Cg_1 - Cg_2 - F_e - \lambda F_p < 0$, $Cp_2 - Cp_1 + Cp_3 + \lambda F_p < 0$, the system gradually stabilizes at the equilibrium point (0, 0, 1, 0). The equilibrium strategy for the four-party evolutionary game is as follows (provide low-quality data products, lenient inspection, strong regulation, direct trust):

- $\alpha L_d - C_d < 0$: The cost for data demanders to carefully analyze data is greater than the losses caused by low-quality data products. As a result, the data demanders choose the “direct trust” strategy.
- $Ce_1 - Ce_2 - Cp_3 + F_e + R_e < 0$: The cost of strict inspection by the data market, minus the reward for rejecting rent-seeking behavior and receiving government incentives, is still greater than the sum of the costs of lenient inspection and fines, minus the benefits from rent-seeking.
- $Cg_1 - Cg_2 - F_e - \lambda F_p < 0$: The cost of strong regulation by the government regulatory agency is less than the cost of conventional regulation plus the total fines collected from data markets and data providers. Hence, the government chooses the “strong regulation” strategy.
- $Cp_2 - Cp_1 + Cp_3 + \lambda F_p < 0$: When the government chooses the “strong regulation” strategy, the combined cost for the data provider to provide low-quality products and engage in rent-seeking, along with the fines for failing the quality inspection, is less than the cost of providing high-quality products. Therefore, the data provider chooses to “provide low-quality data products.”

Situation 2. When $Ce_1 - Ce_2 + \eta F_e < 0$, $Cg_1 - Cg_2 - \eta F_e < 0$, and $Cp_1 - Cp_2 - Cp_3 - \lambda F_p < 0$, the system gradually stabilizes at the equilibrium point (1, 0, 1, 0). The equilibrium strategy for the four-party evolutionary game is as follows (provide high-quality data products, lenient inspection, strong regulation, direct trust):

- $Cp_1 - Cp_2 - Cp_3 - \lambda F_p < 0$: The cost for the data provider to provide high-quality data products is less than the cost of providing low-quality products, the rent-seeking cost, and the fines for failing the quality inspection.
- $Ce_1 - Ce_2 + \eta F_e < 0$: The cost of lenient inspection by the data market, along with the government fines, is still less than the cost of strict inspection. Therefore, the data market chooses lenient inspection.
- $Cg_1 - Cg_2 - \eta F_e < 0$: The routine regulatory costs for government regulatory agencies are greater than the difference between the costs of strict regulation and the fines collected, so regulatory agencies tend to choose a stringent regulatory strategy.

Situation 3. When $C_d - \alpha L_d < 0$, $Cg_1 - Cg_2 - F_e - \beta F_g - \lambda F_p < 0$, and $Cp_2 - Cp_1 + Cp_3 + L_p + \lambda F_p < 0$, $Ce_1 - Ce_2 - Cp_3 + F_e + L_e + R_e < 0$, the system gradually stabilizes at the equilibrium point $(0, 0, 1, 1)$, and the equilibrium strategy of the four-party evolutionary game is (providing low-quality data products, lenient testing, stringent regulation, and prudent analysis):

- $Cp_2 - Cp_1 + Cp_3 + L_p + \lambda F_p < 0$: The combined cost of providing low-quality data products, rent-seeking costs, reputation loss, and fines is less than the cost of providing high-quality data products.
- $Ce_1 - Ce_2 - Cp_3 + F_e + L_e + R_e < 0$: The sum of the loose detection cost, fines, and reputation loss, minus the rent-seeking gains, is smaller than the difference between the strict detection cost and government rewards, so the data market platform chooses loose detection.
- $Cg_1 - Cg_2 - F_e - \beta F_g - \lambda F_p < 0$: The cost of strong regulation is smaller than the sum of the regular regulation costs, the fines collected, and the administrative penalties from higher government authorities.
- $C_d - \alpha L_d < 0$: The cost of cautious analysis by the data demander is smaller than the potential losses incurred from low-quality data products, so the data demander adopts the “cautious analysis” strategy.

In summary, in situations 1 and 3, the stable strategy combination does not involve the data provider providing high-quality data products. The regulatory agency should increase penalties for data providers that provide low-quality data products. Data demanders should enhance public opinion supervision, and data markets should strengthen their monitoring intensity. Increasing penalties for data providers that provide low-quality products is a necessary measure to ensure the quality and safety of products in the data trading market. Strengthening public opinion supervision and ensuring that data markets strictly enforce quality inspections are also key to effectively preventing non-Pareto optimal evolutionary strategy combinations from becoming the stable strategy in the system’s game.

5. Simulation analysis

5.1. SD model

This study on the regulation of data product quality involves four participating entities: data providers, data markets, government regulatory agencies, and data demanders. The transaction process is influenced by various factors. To better validate the theoretical analysis, a simulation analysis using system dynamics will provide a more systematic and intuitive examination of the evolutionary dynamics of the game participants. As shown in Figure 6, this section will integrate evolutionary game theory with system dynamics to establish an SD model for the four-party evolutionary game of data product quality regulation. The specific definitions of the state variables, rate variables, and intermediate variables involved in the system dynamics model are comprehensively detailed in Table 5. Based on the theoretical analysis results mentioned above, we will first validate situations 1, 2, and 3, and then conduct a sensitivity analysis on key parameters.

As shown in Figure 6 and Table 5, the SD model of the evolutionary game for product quality regulation in the data trading market mainly consists of 4 stocks, 4 rate variables, 8 intermediate

variables, and 22 exogenous variables. The four rate variables in the system refer to the rate of change for each strategy. The eight intermediate variables in the system refer to the expected payoffs under different scenarios.

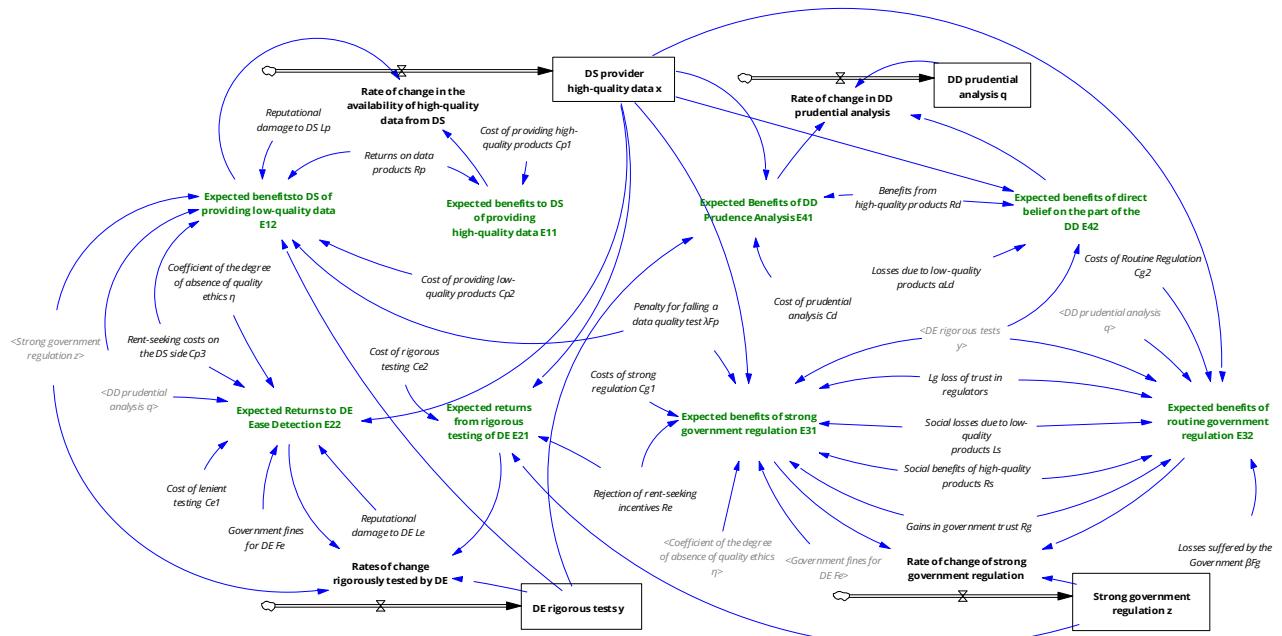


Figure 6. System dynamics model.

Furthermore, in the system dynamics model, the functional relationships between the stock variables, rate variables, intermediate variables, and external variables are derived from the replicator dynamic equations obtained from the aforementioned evolutionary game analysis. Therefore, by using Vensim software to build the SD evolutionary game model as shown in Figure 6, we can visually represent the factors influencing the probability of strategy selection by the four game participants. As detailed in Table 5, the system consists of four primary variable categories: (1) state variables, representing the strategy selection probabilities (x, y, z, q) of the four participants; (2) rate variables, which are governed by the replicator dynamic equations derived in Section 4; (3) intermediate variables, corresponding to the expected payoffs (E_{11} – E_{42}); and (4) exogenous variables, which align with the 22 parameters defined in the payoff matrix (Table 1). This comprehensive structure allows for a systematic simulation of how strategy choices evolve under different regulatory scenarios.

5.2. Scenario simulation

Based on the system stability analysis, initial values are assigned to the strategy probabilities and exogenous variables in the SD evolutionary game model for data transaction quality management. It is important to note that some parameters are quite abstract and difficult to obtain; therefore, this study draws on related cases and expert consultations for these values. Specifically, the utility

coefficient, representing the residual utility of low-quality data, can be operationalized using standard data quality indicators like the accuracy, completeness, and consistency ratios. Similarly, parameters representing the degree of violation can be measured through proxies such as corporate social credit scores or historical non-compliance rates. The model primarily reflects the relationships between variables, which does not affect the research conclusions. The initial values of key parameters in the simulation are unified as shown in Table 6. These values are determined by referring to industry practices and expert advice to ensure rationality. Since Vensim can only perform two-dimensional simulation experiments, Python is used to better display the simulation results. The simulation results for Scenarios 1, 2, and 3 are shown in Figure 7.

Table 5. Detailed classifications and definitions of variables in the SD Model.

Category	Symbol	Description	Logic/Basis
State Variables (Stocks)	x	Prob. that providers choose “provide high-quality data”	Integral of the rate of change
	y	Prob. that markets choose “strict monitoring”	
	z	Prob. that government chooses “strong regulation”	
	q	Prob. that demanders choose “cautious analysis”	
Rate Variables (Flows)	dx/dt	Rate of change for providers’ strategy (x)	Replicator Dynamic Equations $F(x)$, $F(y)$, $F(z)$, $F(q)$
	dy/dt	Rate of change for markets’ strategy (y)	
	dz/dt	Rate of change for government’s strategy (z)	
	dq/dt	Rate of change for demanders’ strategy (q)	
Intermediate Variables (Exp. Payoffs)	E_{11}	Payoff for providers (High-quality strategy)	Derived from the Payoff Matrix
	E_{12}	Payoff for providers (Low-quality strategy)	
	E_{21}	Payoff for markets (Strict monitoring)	
	E_{22}	Payoff for markets (Lenient monitoring)	
Exogenous Variables	E_{31}	Payoff for government (Strong regulation)	Derived from the Payoff Matrix
	E_{32}	Payoff for government (Regular regulation)	
	E_{41}	Payoff for demanders (Cautious analysis)	
	E_{42}	Payoff for demanders (Direct trust)	
Parameters	The 22 external parameters (e.g., R_p , C_{p1} , F_p , ...)		

Table 6. Simulation initial values.

C_{p1}	C_{p2}	C_{p3}	R_p	λF_p	L_p	C_{e1}	C_{e2}	R_e	F_e	L_e	C_{g1}	C_{g2}	βF_g	L_g	R_g	αL_d	L_s	R_d	R_s	η
4	2	1.5	10	5	10	1	5	1.2	10	10	8	2	5	10	10	5	10	10	10	0.5

Keeping other parameters unchanged, adjust $Cp_1 = 10$, $Ce_2 = 12$, and $C_d = 6$ to satisfy the conditions of Scenario 1. The simulation results are shown in Figure 7(a). When the government regulatory department adopts the “strong regulation” strategy, the system evolution results in $(0,0,1,0)$, which means (provide low-quality data products, lenient monitoring, strong regulation, direct trust). Keeping other parameters unchanged, adjust $Cp_1 = 6$, $Ce_2 = 12$ to satisfy the conditions of Scenario 2. The simulation results are shown in Figure 7(b). When the government regulatory department adopts the “strong regulation” strategy, the system evolution results in $(1,0,1,0)$, which means (provide high-quality data products, lenient monitoring, strong regulation, direct trust).

Similarly, keeping other parameters unchanged, adjust $Cp_1 = 14$, $Cp_1 = 0.5$, $L_p = 5$, $Ce_2 = 12$, $F_e = 5$, and $L_e = 5$ to satisfy the conditions of Scenario 3. The simulation results are shown in Figure 7(c). When the government regulatory department adopts the “strong regulation” strategy, the system evolution results in $(0,0,1,1)$, which means (provide low-quality data products, lenient monitoring, strong regulation, cautious analysis). In summary, the simulation results are consistent with the theoretical derivation results, proving the validity of the model.

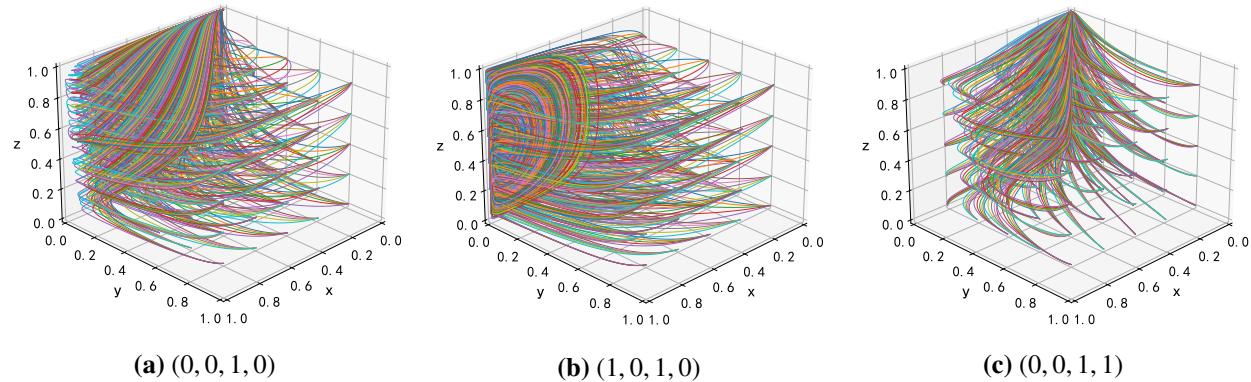


Figure 7. The results of the strategy evolution process of the four-player game agents.

When the data provider chooses to provide high-quality data products and the government regulatory agency enforces strong regulation, the stability of the data transaction market is relatively high. In this scenario, the data market and consumers can opt to save costs rather than take additional actions. At this point, the strategy combination $(1, 0, 1, 0)$ becomes the stable strategy combination in the system’s evolution.

However, even when the government regulatory agency enforces strong regulation, some risk-seeking data providers may engage in rent-seeking with data markets to pursue high profits. As the end users of data products, mature data demanders will choose cautious analysis, while naive data demanders will not take additional actions. This is the potential reason why strategy combinations $(0, 0, 1, 0)$ and $(0, 0, 1, 1)$ may become stable strategy combinations. This situation corresponds to the initial stage of the development of the data element transaction market, where data resources are gradually being recognized as an important economic element. At this stage, the government plays a crucial role in policy guidance, strategic planning, standard setting, and education and outreach. However, due to issues such as the unclear definition of data ownership and usage rights, lack of

supporting laws and regulations, and low standardization, it is necessary for relevant government departments to choose strong regulatory measures to establish a more robust data transaction market environment and improve the overall quality level of data products.

5.3. Sensitivity analysis

5.3.1. The impact of regulatory agency strategies

The government's strong regulatory strategy helps maintain a healthy, safe, and sustainable environment amid the rapid development of the data market. To further verify the effectiveness and feasibility of this regulation, simulations were conducted by setting different regulatory states: $z = 0$, $z = 0.3$, and $z = 0.9$, which represent different levels of government regulation. A simulation analysis of the evolutionary process of various initial strategies of the data provider, data market, and data demander was performed in three-dimensional space. The simulation results are shown in Figure 8.

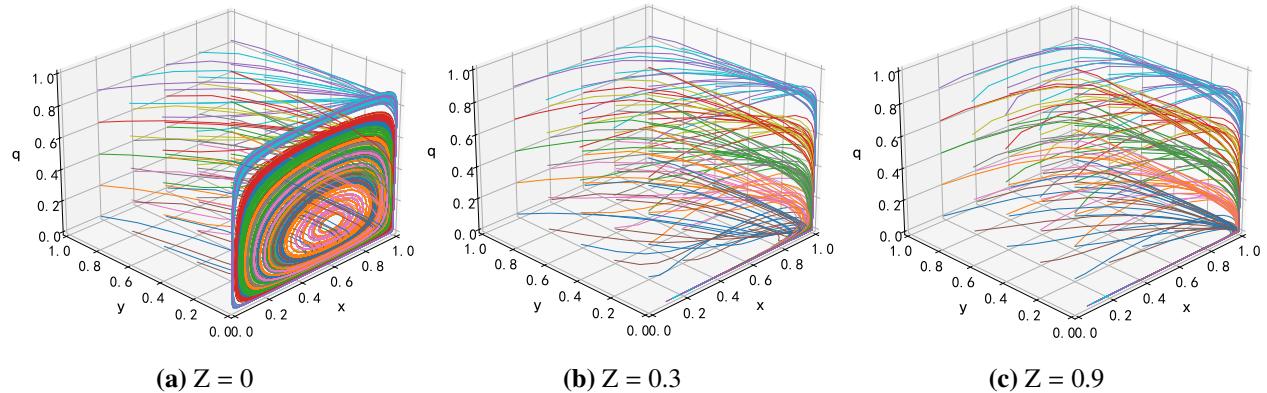


Figure 8. Simulation results.

From Figure 8(a), it can be seen that when the government regulatory agency does not implement any regulatory measures ($z = 0$), the system exhibits a complex and unstable state. The strategy evolution paths show various periodic changes, indicating that the system lacks stability. In the absence of regulatory measures, the behavior of the involved parties struggles to reach a stable strategic equilibrium. The strategies of the parties continuously change, and the overall system displays unstable dynamic characteristics.

From Figure 8(b), it is evident that when the government regulatory agency opts for a moderate level of regulation ($z = 0.3$), the system's state improves. The strategy evolution paths of the parties involved in the game tend to stabilize, although some fluctuations and changes still persist. In this scenario, the introduction of regulatory measures reduces some of the strategic uncertainty, but the relatively low level of regulation still allows the parties considerable room to adjust their strategies, leading to dynamic changes in the system's state.

From Figure 8(c), it is clear that when the government regulatory agency implements strong regulatory measures ($z = 0.9$), the system shows a marked convergence trend. The strategies of the parties gradually stabilize, and the system achieves a dynamic equilibrium. This indicates that a strong regulatory mechanism can significantly enhance the strategic stability of the parties involved in

the game, leading them to adopt stable strategies that comply with regulatory requirements. The overall system reaches a more predictable and stable state.

In summary, when there is no regulation or the regulatory intensity is weak, the system exhibits obvious instability. In contrast, stronger regulatory measures (such as $z = 0.9$) can effectively reduce system instability, prompting the strategies of all parties to evolve toward stability. Strong regulatory measures not only help stabilize the strategies of the game participants but also contribute to enhancing the overall stability of the market. This implies that in the quality management of the data trading market, increasing the intensity of government regulation can more effectively control the behavior of market participants and promote the healthy development of the market.

5.3.2. The impact of strong regulation costs Cg_1

Let Cg_1 take the values 10, 8, and 2. The strategy evolution process and results for the four-party game participants are seen in Figure 9, where the strong regulation costs are high, and the probability of strategies adopted by data providers exhibits significant cyclical fluctuations. At certain points in time, data providers significantly reduce product quality, while at other times they maintain high product quality. This fluctuation may be due to the high supervision costs making it difficult for the government to continuously enforce high-intensity regulation on data providers. Data providers take advantage of this by lowering product quality during periods of weak regulation to reduce costs.

As the strong regulation cost decreases to 2, as shown in Figure 9(c), the system finally stabilizes at (1,1,1,0). The reduction in strong regulation costs not only improves the stability of government regulation but also enhances the overall health of the market, leading to more consistent and predictable behavior among data providers, data markets, and data consumers. At this point, data providers' strategies stabilize in providing high-quality data products, as they are evidently unwilling to risk lowering product quality. This is because low supervision costs mean that the government can conduct regulation more frequently, indicating that low supervision costs effectively increase the frequency and intensity of regulation, thereby encouraging data providers to consistently produce high-quality products. The strategy of government regulatory bodies also shows stability under these circumstances, maintaining a strong regulatory stance. Similarly, the strategies of third-party institutions and consumers tend to stabilize, with consumers more inclined to trust high-quality products and no longer frequently conducting additional analyses and evaluations. This reflects the stability of the market environment and the consistency of product quality.

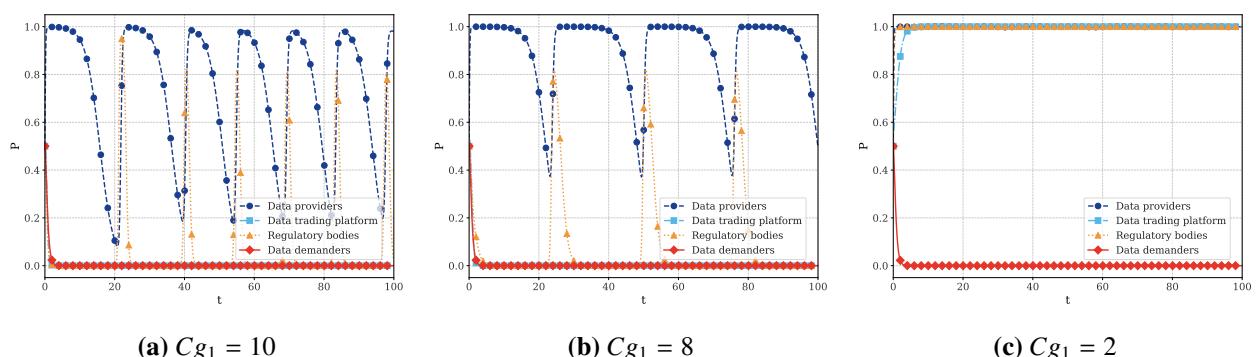


Figure 9. The results of the strategy evolution process of the four-player game agents.

In summary, under high, strong regulation costs, government regulatory bodies find it difficult to maintain high-intensity regulation, leading to cyclical fluctuations in data providers' strategies and instability in product quality. In contrast, with low, strong regulation costs, the government can maintain effective regulatory pressure, encouraging data providers to provide high-quality data products. Therefore, by reducing strong regulation costs, the government can effectively enhance its control over the behavior of data providers, reducing the likelihood of suppliers cutting corners on product quality. This helps to establish a more robust data trading market environment and improve the overall quality level of data products.

5.3.3. The impact of fines λF_p on data providers

To verify the impact of fines imposed on data providers in the regulation of data transaction quality on the evolutionary game process and outcomes, the fines for data providers were set at $\lambda F_p = 5$, $\lambda F_p = 20$, and $\lambda F_p = 80$. As shown in Figure 10, when the fines are relatively low, the strategies of data providers exhibit noticeable cyclical fluctuations. At regular intervals, the probability of data providers' strategies drops rapidly, indicating that data providers periodically provide low-quality data products. This fluctuation may be due to the fines being insufficient to consistently motivate data providers to provide high-quality data products. When data providers find the penalties relatively mild, they may be inclined to risk lowering the quality of their data products to pursue short-term profits. Similarly, the strategies of government regulatory bodies also show cyclical fluctuations, particularly when the quality of data providers' products decreases, leading to a reduction in regulatory intensity. This may be because regulatory bodies, with limited resources, are unable to maintain high-intensity regulation continuously, resulting in delayed or insufficient responses to changes in data providers' strategies.

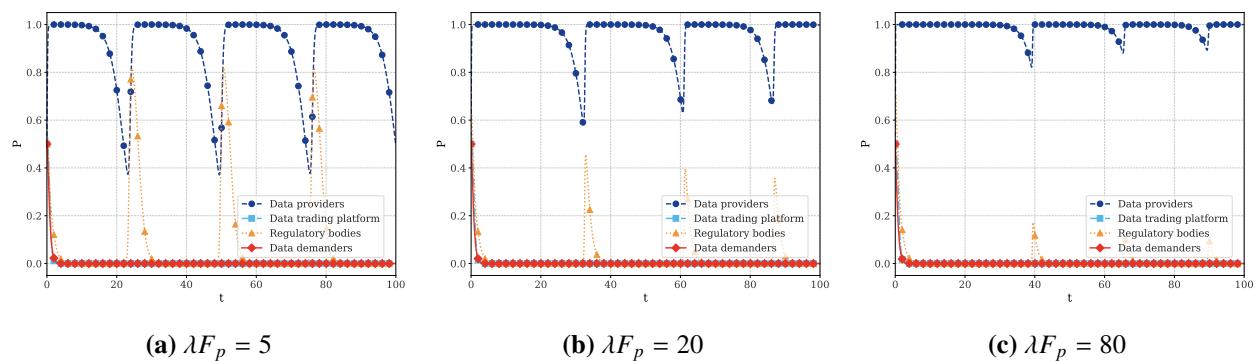


Figure 10. The results of the strategy evolution process of the four-player game agents.

However, when the fine amount is increased to 80, as shown in Figure 10(c), the frequency of fluctuations in the data providers' strategies significantly decreases, and the amplitude of these fluctuations is also notably reduced. Data providers are more inclined to maintain a higher product quality, indicating that higher fines effectively suppress the opportunistic behavior of data providers. The increased fines substantially raise the risk cost of lowering quality, thereby reducing the motivation for data providers to compromise on quality. Under higher fine amounts, the volatility in

the strategies of government regulatory bodies decreases, demonstrating more stable regulatory behavior. This may be because higher fines improve the compliance of data providers, thereby reducing the regulatory burden on the government. As the strategies of data providers and government regulatory bodies stabilize, the responses of data markets and data demanders become more consistent and stable, ultimately converging to zero. In a high-quality environment, data markets may reduce the intensity of their inspections, and data demanders' confidence in high-quality products increases, leading to a decrease in their vigilance.

These cyclical fluctuations reflect the real-world risk of *intermittent enforcement*. As market compliance improves, cost-sensitive regulators often succumb to *regulatory fatigue* and relax supervision. Crucially, this retraction inadvertently creates *quality windows*—temporary periods of lax oversight. Dishonest providers exploit these windows to opportunistically release low-quality data before enforcement retightens, effectively turning the governance process into a speculative cat-and-mouse game. In summary, first, lower fines are ineffective in curbing the behavior of data providers providing low-quality data products, leading to cyclical fluctuations in the strategies of market participants. In contrast, higher fines significantly increase the motivation for data providers to provide high-quality data, reducing the instability of the entire system. Second, higher fines not only stabilize the strategies of data providers but also enhance the overall stability of the market system, including the behavior of government regulators, data markets, and the confidence of data demanders. Finally, relying solely on government regulation is not always effective, especially when fines are low. However, by increasing fines and strengthening regulation, the behavior patterns of data providers can be significantly improved, thereby enhancing product quality and the overall health of the market.

5.3.4. The impact of regulatory fines F_e

Let F_e take the values 0, 10, and 20. The strategy evolution process and results of the four-party game participants are shown in Figure 11. When the fine amount is 0, the strategy of data providers exhibits significant cyclical fluctuations. This indicates that in the absence of fines imposed on data markets, there may be rent-seeking between data providers and data markets, leading to the choice of low-cost, low-quality production strategies due to the lack of punitive pressure. This unstable strategic behavior reflects a tendency among data providers to engage in short-term speculative behavior to pursue high profits in an environment lacking effective regulation, without fearing detection by data markets. Since data markets are not constrained by fines, they lack sufficient motivation to conduct thorough inspections, resulting in their strategy stabilizing at 0, indicating lax inspections. The strategy of data demanders also shows significant cyclical fluctuations at this time. Due to the rent-seeking and lax inspection behaviors between data markets and suppliers in the market, data demanders lack confidence in product quality, leading to frequent changes in their strategy choices. As the fines increase, when $F_e = 20$, the strategies of data providers and demanders gradually stabilize at “providing high quality data products” and “direct trust,” respectively. This is because higher fines significantly increase the cost of rent-seeking behavior between data markets and data providers, and the presence of fines encourages data markets to strengthen their inspections, allowing supervisory work to be effectively carried out. Additionally, higher fines enhance market transparency and compliance, thereby increasing the confidence of data demanders in product quality.

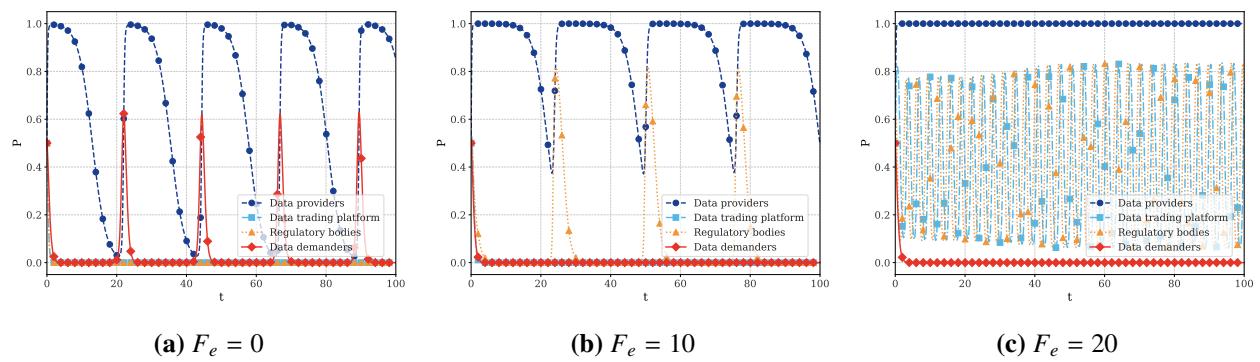


Figure 11. The results of the strategy evolution process of the four-player game agents.

The analysis demonstrates that raising direct rent-seeking costs mitigates the motivation for data providers to engage in bribery. Nevertheless, if the potential returns from low-quality data are substantial, providers may still resort to risk-taking speculation despite high bribery costs. Furthermore, our findings suggest that penalizing the platform is generally more effective than solely increasing the cost of bribery. High penalties compel the platform to implement rigorous inspections, effectively driving the probability of success down to near zero and severing the rent-seeking mechanism at its root. Overall, under high, strong regulation costs, government regulatory bodies find it difficult to maintain high-intensity regulation, leading to cyclical fluctuations in data providers' strategies and instability in product quality. In contrast, with low, strong regulation costs, the government can maintain effective regulatory pressure, encouraging data providers to provide high-quality data products. Therefore, by reducing strong regulation costs, the government can effectively enhance its control over the behavior of data providers, reducing the likelihood of suppliers cutting corners on product quality. This helps to establish a more robust data trading market environment and improve the overall quality level of data products.

6. Conclusions and recommendations

6.1. Conclusion

An evolutionary game analysis was conducted on the issue of quality regulation of data trading products in the data element market. The analysis specifically explored the aspects of limited rationality of decision-makers, collaborative quality regulation, and the impact of key elements on strategy evolution. The main research conclusions are as follows:

(1) The quality of data products is not the result of a single linear factor, but the result of the interaction of multiple factors. Reducing data production and processing costs, increasing reputation penalties for low-quality data providers, and establishing a comprehensive punishment mechanism can significantly improve data quality. Helping data providers select a suitable cost-benefit evaluation system can mitigate uncertainties in the data trading market and reduce the negative effects of limited rationality and decision bias on behavior.

(2) The strategy choices among data providers, regulatory agencies, data markets, and data consumers are interrelated, and a stable state is achieved through repeated dynamic games. When regulatory agencies choose strict supervision, data markets adopt rigorous inspections, and data

consumers opt for cautious analysis, thereby data providers are incentivized to provide higher quality data, ultimately achieving market stability. It should be noted that strengthening data market regulation, enhancing enforcement supervision by regulatory agencies, improving data consumer identification capabilities, and perfecting the quality feedback mechanism are key to ensuring that decisions do not deviate and to fully exert the effectiveness of the collaborative quality regulation system.

(3) The stability and orderliness of the data market are closely related to the government's emphasis on data quality and the quality awareness of data demanders. The importance attached to quality by regulators, data providers, and data markets, as well as the enhancement of quality awareness by data demanders, can effectively ensure the stability of the data market.

6.2. *Recommendations*

Based on the above conclusions, the following suggestions are put forward:

First, government regulatory bodies, industry associations, and other non-governmental organizations, as the builders of the market regulatory environment, should actively create a favorable business atmosphere and development environment for data suppliers by implementing targeted reward and punishment strategies, improving industry entry standards, providing proper guidance to data suppliers, and timely disclosing "black and red lists" of data suppliers, among other effective measures. A one-size-fits-all regulatory approach should be abandoned in favor of establishing a dynamic credit-based regulatory system. Credit profiles could be created based on data providers' historical compliance records, with those achieving high credit scores placed on a "whitelist" to reduce the frequency of random inspections. A "blacklist" should be established for data providers with lower credit scores, allowing for increased inspection frequency within reasonable limits. This approach maintains deterrence while substantially reducing overall societal regulatory costs, enabling the government to sustain an effective regulatory equilibrium over the long term. As market participants and implementers, data suppliers should actively improve the production process of data products, strengthen quality control, and strictly monitor the quality of data products. With a strong external regulatory framework as a constraint and positive internal controls as detailed norms, the collaboration between external regulation and internal control will help ensure product quality and safety. Future research will focus on comparative analysis with the data transaction market regulatory frameworks of other jurisdictions, drawing on regulatory experiences from different regions to enhance the universality of the research conclusions and their policy reference value.

Second, to achieve high-level quality governance, it is essential to continuously strengthen the positive interaction between government regulatory bodies and data suppliers. Government regulators should innovate regulatory methods based on the specific regulatory environment, employing a combination of soft and hard approaches, as well as reward and punishment strategies, to conduct categorized supervision. Emerging technologies such as artificial intelligence and blockchain should be utilized to tackle difficult points in the regulatory process, reduce regulatory costs, and improve efficiency. To further overcome regulatory resource constraints, artificial intelligence technologies may be leveraged. Through natural language processing and machine learning algorithms, these can automatically verify compliance documentation, metadata standards, and privacy agreements for data, thereby directly reducing the intensive regulatory costs within the model. A Blockchain's distributed ledger technology can be employed to permanently and immutably record the entire process of data

production and communication. Should any falsification of quality occur, this tamper-proof record would substantially increase the reputational risk for data providers. This would render the long-term costs of falsification far exceeding any short-term gains, compelling them to adopt high-quality strategies. At the same time, in addition to optimizing the regulatory framework and models at the macro level, the micro-level data quality regulatory mechanisms are equally crucial and serve as the cornerstone for ensuring the healthy and sustainable development of the data transaction market. As emphasized in this paper, improving data quality regulatory mechanisms should become a key direction in the future construction of the regulatory system. Specifically, to effectively improve data quality standards, efforts must be made both at the technical and non-technical levels in a coordinated manner.

Finally, in terms of technological expansion, current research primarily focuses on areas such as strategy selection and equilibrium analysis, with limited exploration of the specific application of emerging technologies like artificial intelligence and blockchains in the regulation of data trading markets. We recognize that integrating these cutting-edge technologies into regulatory framework design holds immense potential and forward-looking value. Future research will focus on exploring how to leverage artificial intelligence, blockchains, and other technologies to enhance the intelligence, transparency, and efficiency of data trading market regulation, thereby contributing to the development of a more comprehensive and advanced regulatory system for data trading markets.

Author contributions

All authors contributed to the study conception and design. All authors read and approved the final manuscript.

Use of Generative-AI tools declaration

The authors declare that they did not utilize any artificial intelligence (AI) tools in the creation of this article.

Acknowledgments

This work was supported by the National Social Science Fund of China (Grant No. 23BJY205).

Conflict of interest

The authors declare that they have no competing interests.

References

1. L. Theodorakopoulos, A. Karras, A. Theodoropoulou, G. Kampiotis, Benchmarking big data systems: Performance and decision-making implications in emerging technologies, *Technologies*, **12** (2024), 217. <https://doi.org/10.3390/technologies12110217>
2. H. Taherdoost, A systematic review of big data innovations in smart grids, *RINENG*, **22** (2024), 102132. <https://doi.org/10.1016/j.rineng.2024.102132>

3. S. Latif, N. Javaid, F. Aslam, A. Aldegeishem, N. Alrajeh, S. H. Bouk, Enhanced prediction of stock markets using a novel deep learning model PLSTM-TAL in urbanized smart cities, *Heliyon*, **10** (2024), 6. <https://doi.org/10.1016/j.heliyon.2024.e27747>
4. J. Kennedy, P. Subramaniam, S. Galhotra, R. Castro Fernandez, Revisiting online data markets in 2022: A seller and buyer perspective, *SIGMOD Rec.*, **51** (2022), 30–37. <https://doi.org/10.1145/3579371>
5. J. Yang, C. C. Zhao, C. X. Xing, Big data market optimization pricing model based on data quality, *Complexity*, **2019** (2019), 5964068. <https://doi.org/10.1155/2019/5964068>
6. H. Sun, M. J. Xiao, Y. Xu, G. J. Gao, S. Zhang, Crowdsensing data trading for unknown market: Privacy, stability, and conflicts, *IEEE Trans. Mob. Comput.*, **23** (2024), 11719–11734. <https://doi.org/10.1109/TMC.2024.3399816>
7. S. M. Xiong, P. C. Chen, S. S. Ge, Q. Ni, SFOM-DT: a secure and fair one-to-many data trading scheme based on blockchain, *IEEE Trans. Inf. Forensics Secur.*, **19** (2024), 10029–10042. <https://doi.org/10.1109/TIFS.2024.3475816>
8. Y. Lv, J. Yang, X. N. Sun, H. F. Wu, Evolutionary game analysis of stakeholder privacy management in the AIGC model, *Oper. Res. Perspect.*, **14** (2025), 100327. <https://doi.org/10.1016/j.orp.2025.100327>
9. B. Mi, Z. W. Han, K. Y. Chen, Multi-party data pricing for complex data trading markets: a rubinstein bargaining approach, *arXiv preprint arXiv:2502.16363*, 2025. <https://doi.org/10.48550/arXiv.2502.16363>
10. M. Rasouli, M. I. Jordan, Data sharing markets, *arXiv preprint arXiv:2107.08630*, 2021. <https://doi.org/10.48550/arXiv.2107.08630>
11. F. Liang, W. Yu, D. An, Q. Y. Yang, X. W. Fu, W. Zhao, A survey on big data market: Pricing, trading and protection, *IEEE Access*, **6** (2018), 15132–15154. <https://doi.org/10.1109/ACCESS.2018.2806881>
12. S. Karkošková, Data governance model to enhance data quality in financial institutions, *Inf. Syst. Manage.*, **40** (2023), 90–110. <https://doi.org/10.1080/10580530.2022.2042628>
13. M. C. Cohen, R. Lobel, G. Perakis, Dynamic pricing through data sampling, *Prod. Oper. Manag.*, **27** (2018), 1074–1088. <https://doi.org/10.1111/poms.12854>
14. J. R. Wang, Y. Liu, P. G. Li, Z. X. Lin, S. Sindakis, S. Aggarwal, Overview of data quality: Examining the dimensions, antecedents, and impacts of data quality, *J. Knowl. Econ.*, **15** (2024), 1159–1178. <https://doi.org/10.1007/s13132-022-01096-6>
15. T. S. Sun, Z. Yuan, C. X. Li, K. F. Zhang, J. Xu, The value of personal data in internet commerce: A high-stakes field experiment on data regulation policy, *Manage. Sci.*, **70** (2024), 2645–2660. <https://doi.org/10.1287/mnsc.2023.4828>
16. R. Zhao, L. Peng, Y. L. Zhao, Y. B. Feng, Coevolution mechanisms of stakeholder strategies in the green building technologies innovation ecosystem: An evolutionary game theory perspective, *Environ. Impact Assess. Rev.*, **105** (2024), 107418. <https://doi.org/10.1016/j.eiar.2024.107418>
17. T. Hendershott, M. Rysman, R. Schwabe, Stock exchanges as platforms for data and trading, *J. Financ. Mark.*, **6** (2025), 100986. <https://doi.org/10.1016/j.finmar.2025.100986>

18. Y. L. Chen, H. Y. Yin, Y. C. Zhang, W. Ren, Y. Ren, A privacy-aware and fair self-exchanging self-trading scheme for IoT data based on smart contract, *Int. J. Grid. Util. Comput.*, **16** (2025), 191–201. <https://doi.org/10.1504/IJGUC.2025.146277>

19. J. Wang, S. H. Ma, Q. Lv, Q. Li, Demographic forecast modelling using SSA-XGBoost for smart population management based on multi-sources data, *PLoS One*, **20** (2025), e0320298. <https://doi.org/10.1371/journal.pone.0320298>

20. L. H. Huang, M. K. Guo, Z. Q. Shao, X. Qin, Q. F. Tang, Thoughts on national unified data asset registration system, *Bull. Chin. Acad. Sci.*, **37** (2022), 1426–1434. <https://doi.org/10.16418/j.issn.1000-3045.20220623001>

21. T. Peixoto, B. Oliveira, Ó. Oliveira, F. Ribeiro, Data quality assessment in smart manufacturing: a review, *Systems*, **13** (2025), 243. <https://doi.org/10.3390/systems13040243>

22. F. Ridzuan, W. M. N. W. Zainon, A review on data quality dimensions for big data, *Pro. Comput. Sci.*, **234** (2024), 341–348. <https://doi.org/10.1016/j.procs.2024.03.008>

23. Z. Sun, W. Ying, K. N. Shen, X. Liu, Assetization-oriented data governance: an exploratory case study, *ICEB 2025 Proc.*, **6** (2025). <https://aisel.aisnet.org/iceb2025/24>

24. Y. R. Qiu, Z. M. Hu, Data governance and open sharing in the fields of life sciences and medicine: A bibliometric analysis, *Digital Health*, **11** (2025), 20552076251320302. <https://doi.org/10.1177/20552076251320302>

25. Y. Gao, M. Janssen, C. Zhang Understanding the evolution of open government data research: towards open data sustainability and smartness, *Int. Rev. Adm. Sci.*, **89** (2023), 59–75. <https://doi.org/10.1177/0020852321100955>

26. L. Huang, Y. Dou, Y. Liu, J. Wang, G. Chen, X. Zhang, R. Wang, Toward a research framework to conceptualize data as a factor of production: the data marketplace perspective, *Fundam. Res.*, **5** (2021), 586–594. <https://doi.org/10.1016/j.fmre.2021.08.006>

27. S. Gathu, High-performance computing and big data: Emerging trends in advanced computing systems for data-intensive applications, *J. Adv. Comput. Syst.*, **4** (2024), 22–35.

28. A. Liu, M. Osewe, Y. Shi, X. Zhen, Y. Wu, Cross-border e-commerce development and challenges in China: A systematic literature review, *J. Theor. Appl. Electron Commer. Res.*, **17** (2021), 69–88. <https://doi.org/10.3390/jtaer17010004>

29. L. B. Bai, Q. Xie, J. C. Lin, S. Y. Liu, C. S. Wang, L. Wang, Dynamic selection of risk response strategies with resource allocation for construction project portfolios, *Comput. Ind. Eng.*, **191** (2024), 110116. <https://doi.org/10.1016/j.cie.2024.110116>

30. X. T. Qi, F. L. Bu, A system dynamics-based model for the evolution of environmental group events, *Sci. Rep.*, **14** (2024), 9711. <https://doi.org/10.1038/s41598-024-59283-1>

31. X. G. Zhao, Y. Z. Zhang, The system dynamics (SD) analysis of the government and power producers' evolutionary game strategies based on carbon trading (CT) mechanism: A case of China, *Sustainability*, **10** (2018), 1150. <https://doi.org/10.3390/su10041150>

- 32. L. F. Cheng, R. B. Sun, K. Wang, F. Yu, P. G. Huang, M. Y. Zhang, Advancing sustainable electricity markets: evolutionary game theory as a framework for complex systems optimization and adaptive policy design, *Complex Intell. Systems*, **11** (2025), 320. <https://doi.org/10.1007/s40747-025-01916-7>
- 33. C. Hu, J. Yang, P. Gao, Y. Zhang, Evolutionary game analysis of three stakeholders in big data trading market, *J. Ind. Manage. Optim.*, **20** (2024), 2135–2152. <https://doi.org/10.3934/jimo.2023159>
- 34. T. Stefes, Data quality of different modes of supervision in classroom surveys, *Educ. Sci.*, **14** (2024), 299. <https://doi.org/10.3390/educsci14030299>



AIMS Press

©2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>)