Research on Data Theory of Value

Li Haijian[1], Zhao Li[2*]
[1] Institute of Quantitative & Technological Economics, China Academy of Social Sciences (CASS), Beijing, China
[2] Ningbo University, Ningbo, China

Abstract: *This paper explores the data theory of value along the line of reasoning epochal characteristics of data – theoretical innovation – paradigmatic transformation and, through a comparison of hard and soft factors and observation of data peculiar features, it draws the conclusion that data have the epochal characteristics of non-competitiveness and non-exclusivity, decreasing marginal cost and increasing marginal return, non-physical and intangible form, and non-finiteness and non-scarcity. It is the epochal characteristics of data that undermine the traditional theory of value and innovate the "production-exchange" theory, including data value generation, data value realization, data value rights determination and data value pricing. From the perspective of data value generation, the levels of data quality, processing, use and connectivity, data application scenarios and data openness will influence data value. From the perspective of data value realization, data, as independent factors of production, show value creation effect, create a value multiplier effect by empowering other factors of production, and substitute other factors of production to create a zero-price effect. From the perspective of data value rights determination, based on the theory of property, the tragedy of the private outweighs the comedy of the private with respect to data, and based on the theory of sharing economy, the comedy of the commons outweighs the tragedy of the commons with respect to data. From the perspective of data pricing, standardized data products can be priced according to the physical product attributes, and non-standardized data products can be priced according to the virtual product attributes. Based on the epochal characteristics of data and theoretical innovation, the "production-exchange" paradigm has undergone a transformation from "using tangible factors to produce tangible products and exchanging tangible products for tangible products" to "using intangible factors to produce tangible products and exchanging intangible products for tangible products" and ultimately to "using intangible factors to produce intangible products and exchanging intangible products for intangible products".*

Keywords: *Data theory of value, data value generation, data value rights determination, data value pricing*
JEL Classification Code: M00
DOI: 10.19602/j.chinaeconomist.2024.05.02

1. Introduction

In the course of time, the sources of value have gradually expanded and the creators of value have become increasingly diversified. In the early years of the industrial economy, the labor theory of value took shape because labor played a leading role in value creation and value realization, and extra profits

* CONTACT: Zhao Li, email: 936674463@qq.com.
Acknowledgement: This study is funded by "Management Model Innovation of Chinese Enterprises" Research Project, Institute of Industrial Economics, CASS (Grant No. 2019-gjs-06); Project under the Graduate Student Scientific and Research Innovation Support Program, University of Chinese Academy of Social Sciences (Graduate School) (Grant No. 2022-KY-118).

were obtained by prolonging labor time and increasing labor intensity. In the late years of the industrial economy, the capital theory of value took shape because capital played a leading role in value creation and value realization, and enterprises could use both tangible capital and intangible capital (such as technology, knowledge, and management) to produce goods (Li and Feng, 2004). In the era of digital economy, data, as key factors of production, play a leading role in value creation and value realization. Therefore, it is imperative to establish a data theory of value based on both the labor and the capital theories of value, which would serve to clarify the role of data factors in the production process, and the methods of data rights determination and pricing in the exchange process. Such theory would guide the allocation of resources to maximize efficiency and minimize cost. It should be noted that, during the focus shift from the labor capital theories of value to the data theory of value (Li and Feng, 2004), the creators of value have also changed from tangible factors of production, such as labor and capital, to intangible factors, such as data, and this shift has brought about a paradigmatic change in value generation of data factors and value exchange of data products.

Our literature review shows that the existing research on data value focuses on two aspects. The first aspect concerns the process to realize data value; Yin et al. (2022) sustained that after data quality is improved by "purification", data value can be realized by clarifying property rights, assessing their value, determining their price, promoting their circulation and exchange, increasing their application, and adding their value; this is an evolutionary pathway of "data resource utilization – data asset utilization – data value". The second aspect is related to the obstacles of data value, such as data rights determination (Shen, 2020), data pricing (Ouyang and Gong, 2022), data trading (Tian and Liu, 2020), and data profit distribution (Li et al., 2022). In summary, the existing literature mentions some fragmented research on the pathway and obstacles of realizing data value but it does not show systematic knowledge. Considering the epochal characteristics of data factors, the realization pathway and obstacles of data value and the "production-exchange" value theory, this paper seeks to construct a systematic theory of data value which breaks down data value into four parts: Data value generation, data value realization, data value rights determination, and data value pricing.

This paper makes a marginal contribution in two aspects. The first is a systematic comparison between soft and hard factors along the line of reasoning "epochal characteristics of data – theoretical innovation – paradigmatic transformation", and a summary of the epochal characteristics of data factors, which has undermined the underlying logics and deconstructed the "production - exchange" theory in the traditional theories of value. The second is a systematic research on the issues of data value generation, realization, rights determination, and pricing, which has reconstructed the "production-exchange" theory in the traditional theories of value and promoted the paradigmatical transformation of "production – exchange", namely, a transition from the first two tiers ("using tangible factors to produce tangible products and exchanging tangible products for tangible products" and "using intangible factors to produce tangible products and exchanging intangible products for tangible products") to the third tier ("using intangible factors to produce intangible products and exchanging intangible products for intangible products"). The logical framework of the research in this paper is shown in Figure 1 below.

2. Epochal Characteristics of Data Factors

2.1 Comparison Between Soft and Hard Factors (R_2 and R_1)

This section focuses on the comparison of soft factors such as technology, knowledge, management, and data (R_2) and the three traditional hard factors, namely, labor, land and capital (R_1). Soft factors have the following characteristics: (i) Non-competitiveness and non-exclusivity. Hard factors, due to their competitiveness, can be used by only one entity at a time and their value is consumed after their use. The exclusive use of hard factors inevitably leads to their low retention rate, limited use, static nature, low

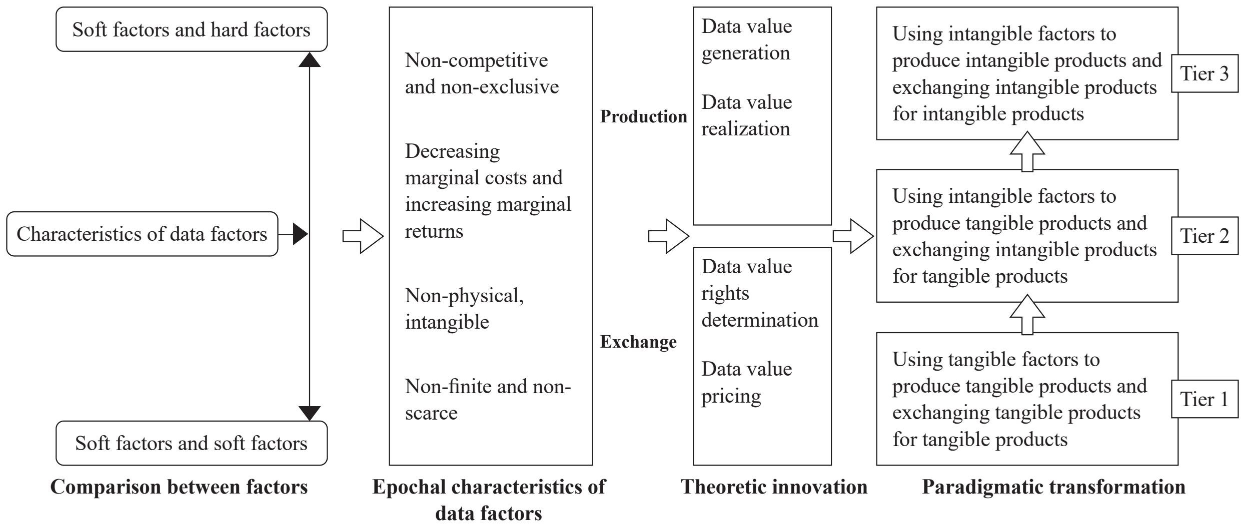

Figure 1: Logical Framework of Research on Data Theory of Value

penetration, interconnectivity and substitutability between different hard factors, and low externalities. In contrast, the non-competitiveness of soft factors allows them to be repeatedly used by several entities within the same period and their value is not consumed after their use. In addition, the non-exclusive use of soft factors inevitably leads to their high retention rate, unlimited use, dynamic nature, ability to penetrate, connect and substitute hard factors (Sun, 2013), and high externalities of data factors. (ii) Marginal costs decrease while marginal returns increase. If being used, hard factors will be depreciated and devalued, be restricted to a certain number of times, and become less and less in quantity and value. In other words, marginal costs increase while marginal returns decrease. In contrast, if being used, soft factors will increase and appreciate, not be restricted to a certain number of times, and increase in quantity and value. In other words, marginal costs decrease while marginal returns increase. Therefore, compared with hard factors, soft factors are more non-competitive and non-exclusive and their marginal costs decrease while their marginal returns increase.

2.2 Comparison of Soft Factors (R_{22} and R_{21})

The comparison of soft factors is made between data factors (R_{22}) and the three modern factors, namely, technology, knowledge, and management (R_{21}). They differ in that data factors have more "non-physical, intangible" characteristics in comparison to the three modern factors. These intangible qualities of data factors (R_{22}) are: (i) Pervasive presence. With the development of digital technology from the consumer internet to the industrial internet and then the internet of things, the internet of minds, and the metaverse, data penetration expands from the consumer side to the producer, thing and scenario sides, and more and more subjects, things and scenarios generate data in a coordinated way at any time and any place. (ii) Always online. Supported by computing power and algorithm, data can generate data, make analysis through modeling, update and optimize themselves, provide flexible feedback in real-time to work out dynamic solutions and enable data to reduce the delays, achieve high alignment and be permanently online. (iii) Extremely large scale. In the era of digital economy, each person, machine, thing and scenario is a data generator and data can be generated regardless of the spatial and temporal constraints and see an explosive growth. (iv) Extremely cheap. Under the influence of the Moore's Law, the costs of data analysis and processing have declined sharply, which is helpful for reducing the fixed cost of digital infrastructure. Moreover, the fact that data are highly reusable, widely shared and easily reproduced has significantly reduced the marginal cost of data, even to zero. (v) High added value. Compared with the three modern factors, data factors have the characteristics of high penetration, mobility, integration, recyclability, sharing, and enabling, which inevitably leads to the high added value of data factors. Therefore, data factors have more "non-physical, intangible" characteristics than the three modern factors and are high-ranking resources and factors, *i.e.* $R_{22}>R_{21}$.

2.3 Intrinsic Characteristics of Data Factors

In addition to the characteristics shared by soft factors, data factors have their own characteristics which are noticeable for their "non-finiteness and non-scarcity", specifically, "four infinite features": (i) Infinite reuse. Running in the form of bits, data can be infinitely copied and used and be infinitely cleaned, processed and applied to multiple scenarios by different enterprises based on the original data to achieve the effect of "one-time production, infinite reuse". (ii) Infinite combination. Unlimited processing, unlimited connection and unlimited matching can be realized between different data to the effect of infinite combination. The essence of infinite data combination is to connect data islands, push down data chimneys, drop data barriers, bridge data gaps, and combine fragmented and disorderly data into systematic and orderly data. (iii) Infinite iteration. Data can be infinitely perfected and configured in the virtual space to achieve zero-cost data iteration and trial. (iv) Infinite supply. Infinite data production, reuse, combination and iteration lead to infinite supply of data factors, and overcome the constraints caused by the scarcity of traditional factors.

3. Data Value Generation

3.1 Different Levels of Data Quality Generate Different Values

The data quality ensures the accuracy of prediction, and the positive externalities of high-quality data benefit data consumers and a broader ecosystem. However, incorrect data tabs, imbalanced data sets, and other data quality issues are difficult to identify on a large scale and result in a gap in the predictive data models and output. In addition, data quality issues not only affect the internal decision making of an enterprise, but also cause negative externalities to data stakeholders due to the compound effect of "data cascades". In particular, with the wide application of those abnormal data outputs to specific scenarios, the scope of data cascades expands infinitely, their cycle infinitely extends and their iteration cost infinitely increases (Sambasivan et al, 2021).

3.2 Different Levels of Data Processing Generate Different Values

The essence of data processing is to extract valuable data from disordered data, namely, to realize the transformation of unstructured low-value data to structured high-value data in data processing (Xu, 2022). The inconsistency between the structures, concepts, calibers, and logics of different types of data is likely to cause disorder, low applicability, and other problems. To solve these problems, it is necessary to establish common standards and languages and a unified open processing system when processing various types of data. According to the different levels of data processing, data can be divided into raw data and value-added data. Raw data are unprocessed data with fragmentation and low value density; value-added data are processed and integrated based on raw data, with large scale and high value density.

3.3 Different Levels of Data Use Generate Different Values

As data are used more often, broadly, and deeply, their value increases. (i) Frequency of use. The exponential growth of data means that as more data are fed into the intelligent learning system, the data used for training becomes more accurate, the machine learning ability becomes stronger, the output of the system becomes more exact, and the value of the data becomes higher. (ii) Breadth of use. Data are not only used in the new economy to realize digital industrialization, but also empower the R&D, manufacturing, marketing, and operation of enterprises to promote digital industrial transformation. The extensive use of data factors in such digital industrialization and industrial digitization brings about an increase in both breadth of use and the value of data. (iii) Depth of use. As data-intensive products and services continue to descend to various fields, links, and entities, data allow to connect them with each other.

3.4 Different Levels of Data Connectivity Generate Different Values

Digital infrastructure connectivity is the foundation for realizing the value of data applications. (i) Connections between digital infrastructures. From the consumer internet, the industrial internet, the internet of things, the internet of minds to the metaverse, the objects to be connected include consumer services, production and manufacturing. In extensive connection of digital devices, embedded connectivity components are used to strengthen the connection and interaction between products and users and provide users with individualized and intelligent products and services (Cao et al., 2022). In addition, through strengthening the connection and interaction between products, expanding the range of products, and optimizing their functions, a digital platform that integrates multiple products and functions can be established and an intelligent product system can be formed to provide intelligent life solutions (Li and Chen, 2011). (ii) Connection between data. The intelligent interconnection between products and users and between products in digital infrastructure connection is actually the connection between product data and user data and between product data. Stronger data connectivity brings about deeper interaction between users, products, services, and content, broader aggregation, more blurred boundaries between enterprises, and higher value of data application.

3.5 Different Scenarios of Data Application Generate Different Values

Data application scenarios mainly include head scenarios and tail scenarios. Head scenarios are the main scenario, in which enterprises collect, analyze and use head data to optimize most of the production decisions and transaction decisions of enterprises. Tail data scenarios focus more on marginalized and individual scenarios, in which enterprises collect users' tail data based on their tail behaviors and optimize their algorithmic models to increase the accuracy of data prediction. In the differentiated tail scenarios, with stronger, wider long-tail coverage, more long-tail data collected and stronger long-tail learning effect, the ability to optimize marginalized and individualized scenarios will become stronger (Iansiti, 2021), and optimize data value which decreases or increases with the tail scenarios (Wang et al., 2022).

3.6 Different Levels of Data Openness Generate Different Values

Depending on the degree of data openness, data systems can be categorized into data walls and data sharing systems. Data walls mean that data are restricted to internal use and are not allowed to be used by external entities, in which case the data resources are "privately owned". While enabling enterprises to accumulate their internal data assets, optimize business processes and maximize the efficiency of internal allocation of data resources, the effect of "once produced, one user" brought by data walls also increases the exclusivity of data and creates significant barriers for external entities to use the data. Data openness means that data are not restricted to internal use, and more incentives and tolerance are granted to allow more external entities to use them. While empowering the internal business process of enterprises, the effect of "once produced, several users" brought by data sharing also removes data walls and boundaries, and realizes end-to-end data interconnection and sharing, maximizes the efficiency of social allocation of data resources through the market-based data opening and circulation, and enhances the non-exclusivity of data. Therefore, the difference in data values resulting from different levels of data openness is manifested as the difference between the maximized efficiency of social allocation of data resources brought about by data openness and the maximized efficiency of internal allocation of data resources brought about by data walls.

4. Data Value Realization

4.1 Value Creation Effect of Data as Independent Factors of Production

Data play their own role as factors of production in addition to labor and capital. In other words, before data become independent factors of production, the production function is Y = F (K, L)[1]; after they become independent factors of production, the production function is Y = F (K, L, D). The computing scenario and algorithmic ecosystem formed by “data + computing power + algorithm” are supported by labor, capital and other factors of production. The “non-competitive and non-exclusive” characteristics of data can create a constructive interaction between data, labor and capital, mitigate the market failure caused by asymmetric and inconsistent information and enable data to generate the precision, real time, network and expectation effects.

4.1.1 Precision effect

If we compare the era of industrial economy and the era of digital intelligent economy, the data search method has changed from “people looking for data” to “data looking for people”, and the search method has become more precise. For example, when users go shopping in offline stores, the traditional data search method is that users take the initiative to search for store locations, product attributes, prices and other information according to their consumption needs. In this process, the data are passively collected, transmitted and analyzed, forming the search model of “people looking for data”, which has a narrow search scope, slow search speed and inaccurate search results. In contrast, when users go shopping in online stores, helped by the strength of artificial intelligence, data are automatically collected, transmitted and analyzed, and can provide a multi-dimensional “user profile” by describing their consumption habits, preferences and frequency, which realizes the intelligent search model of “data looking for people”. This model has a more comprehensive search scope, faster search speed and more accurate search results and can effectively shorten the search time, reduce the search cost, enable a precise matching between data and people, all of which create the data precision effect.

4.1.2 Real-time effect

The ongoing iterative optimization of data, computing power and algorithms guarantees the real-time processing, analysis, and feedback of data to make the best decision, realize instant supply, all of which makes up the real-time effect (Li and Zhao, 2021). Data can create the real-time effect depending on the following factors: (i) Instant data. Considering the time-sensitive nature of data, historical data should be filtered out, and more attention should be paid to current and future data to achieve data “optimization” and “denoising”. For example, when enterprises provide mobile search advertising services based on user location data, geographic location data are highly time-dependent data, in which case the immediate and continuous user data flow has more economic value than a large amount of historical data. (ii) Instant computing power. Computing power is the basis for real-time data processing. Ongoing iterative optimization of an enterprise’s computing power, computing load, and computing speed can improve data processing capabilities and efficiency, meet the demand for real-time data processing, and provide safe, reliable, and real-time computing support for data-driven decision making.

[1] For the purpose of this paper, the production function is simplified to include only two factors of production, labor and capital. However, in reality, data play the role of independent factor of production, empowering and substituting other factors of production, and the production function includes not only labor and capital but also other factors of production, such as land, technology, knowledge and management. The purpose of such simplification is only to demonstrate the independent, empowering and substituting effect of data.

(iii) Instant algorithm. The computing speed, memory space, accuracy, readability, and robustness of algorithms need to be perfected on an ongoing basis, so the computing time lag can be shortened. By selecting models, accumulating weights, adjusting parameters, and optimizing the construction of models, enterprises will select algorithms with optimal results and faster response to provide the synergistic real time effect of "data + computing power + algorithm".

4.1.3 Network effect

The data network effect is the root cause of an enterprise's cross-border competitive advantage; it can be unilateral or bilateral: (i) Unilateral network effect means that the larger the number of product or service users, the more information a platform can obtain from the user data it has collected, and the more value such data will create for each user, in which case the data network effect takes shape. (ii) Bilateral network effect. It means that the utility and value of one side of the market is affected by the scale of the other side of the market. Take DiDi Chuxing as an example. A larger number of passengers attracts more drivers and increases the platform's "capacity". Meanwhile, on the strength of "data + computing power + algorithm", the platform can analyze the interaction data between passengers and drivers on a real-time basis, promote faster, more accurate matching between supply and demand, and realize an increase in both data-driven value creation and user experience value. While providing matching services to users on both supply and demand sides, the platform enterprise continues to integrate fragmented and scattered user data in the value network, generating a larger and stronger data ecosystem. As the scope of the data ecosystem continues to expand, the boundaries of the enterprise continue to expand, and its data advantages are gradually transformed into cross-border competitive advantages.

4.1.4 Expectation effect

With the support of data mining and data analysis, the evolutionary path of "data-information-knowledge-wisdom" spirals upward for description, diagnosis, and prediction, which reduces uncertainty, alleviates the limitations of human cognition, enhances rationality in decision-making, and thus forms rational expectations. The expectation effect of data is reflected in the following aspects: (i) Enhancing the market's intelligent decision-making capability. With data mining and data analysis, market players can effectively optimize their behaviors, fully grasp the market operation mechanism, including price mechanism, supply and demand mechanism, competition mechanism and risk mechanism, and generate accurate forecasts for decision-making. In this way, efficiency of resource allocation can be improved, and a prototype of "strong market" in the era of digital and intelligent economy will take shape. (ii) Enhancing the ability of national macro-control ability. On the one hand, with data mining and data analysis, the government can conduct predictive analysis of macroeconomic indicators such as GDP, unemployment rate and inflation rate, master the regular pattern of macroeconomic operation, and improve the accuracy and effectiveness of macroeconomic policies; on the other hand, in response to the supply-demand mismatch and resource misallocation, the technological combination of "data + arithmetic + algorithms" can carry out predictive analysis of aggregate supply and aggregate demand, and improve the efficiency of resource allocation. On the other hand, in order to cope with the mismatch between supply and demand and the misallocation of resources, the government can use the technical combination of "data + computing power + algorithm" to conduct predictive analyses of total supply and demand to achieve its macroeconomic goals, and improve the efficiency of resource allocation and form a prototype of "strong government" in the era of digital and intelligent economy.

4.2 Data Empower Other Factors of Production to Create a Value Multiplier Effect

The fact that data can empower other factors of production means that, in addition to becoming a factor of production in the combination of factors, data can, relying on their "non-physical and intangible" characteristics, penetrate and empower other factors of production to improve the resource allocation efficiency of other factors of production and realize the value multiplier effect. In other words, before data empower other factors of production, the production function is Y = F (K, L); after data empower other factors of production, a new production function emerges: Y_D = F (K(D), L(D)), in which K(D) and L(D) are the capital and labor after data has been employed in the production system. At this moment, data indirectly take part in production and the "new" factors of production are the updated and advanced version of the "old" factors of production as well as the result of the improved efficiency of various factors of production. In this case, factors of production (capital and labor) undergo only a change in form rather than in substance and they still participate in production in their traditional roles.

4.2.1 Data can improve R&D efficiency

Data factors can empower the R&D process, forming a user-centered "crowdfunding" and "crowdsourcing" mechanism for product development. With the support of "data + computing power + algorithm", enterprises can use the data-based R&D platform of "modeling analysis + virtual simulation" to continuously interact with users and generate feedback of "big data" and "small data". "Big data" mean generic data generated when users use products, such as age data, behavioral data, and geographical data; "small data" mean anomalous data generated when users use products, such as data on complaints and bad comments (Long et al., 2021). Enterprises need to fully leverage the combination of "big data + small data" to form a systematic user perception and to improve R&D problems, such as cost, quality, and cycle, therefore helping to reduce the "friction" between products and users and deliver better products through user data collection, iterative testing and continuous design optimization.

4.2.2 Data can improve manufacturing efficiency

Data factors can empower the manufacturing process. Sensors installed in plants, machine tools and other production equipment can collect production data on a real-time, dynamic basis, open the black box of the production process, and achieve data-based, transparent production process. Furthermore, enterprises can use the technical combination of "data + computing power + algorithm" to process and analyze the production process data, and according to user needs, use data flow to make real-time adjustments to production materials, production processes and core components to enable data flow to lead the flows of materials, capital and technology. The data flow leads the material flow, capital flow and technology flow. Third, data application. Through data collection, enterprises can dynamically monitor equipment operation, transmit machine operation data on a real-time basis, generate the "operation profile" of machines, give instant warning and feedback of production failures, identify potential equipment failures, generate failure solutions in advance, and develop elastic, agile and individualized resource allocation programs to improve internal and external production efficiency.

4.2.3 Data can improve marketing efficiency

Data factors can empower the marketing process. First, data reshape the traditional marketing model. With the development of "digital technology group", sensors can capture users' browsing "traces" at any time and generate "three-dimensional" user preference portrait. This means that a distributed full-time interaction is established between enterprises and users and reduces the cost incurred from real-time "tracking" of user preferences. Meanwhile, enterprises make use of the "user profile" generated at any time to change the advertising mode from the "thousand persons, one face" mode in an abstract, fixed and uniform consumption scenario to the "one person, one face" mode in a specific, customized and individualized consumption scenario, and optimize the efficiency of long-tail marketing channels to achieve precision marketing, fine marketing and real-time marketing. Second, data connect online and offline marketing channels. Under the influence of online virtual stores which are in the space-time interactive model of "full-time operation, platform aggregation", offline physical stores are characterized by "fragmented convenience" and "centralized experience", effectively rectify the shortcomings of "delayed delivery and inadequate experience" of the products in online stores, connect "online and offline" channels and complement each other's strengths to realize the "full-channel" marketing covering both online and offline markets and improve internal and external marketing efficiency.

4.2.4 Data can improve operation efficiency

Data factors can empower the operation process. First, data effectively reduce the risk of operation within an enterprise. Digital technology with data factors as its core effectively connects, integrates, configures an enterprise's R&D, manufacturing, marketing, and operation processes and makes full use of the high mobility of data to alleviate the information asymmetry and inconsistency between various processes, reduces the "friction" between system and system, software and hardware, and equipment and staff and reduces the risk of internal operation of the enterprise. Second, data can effectively alleviate the friction of external operations. With the support of data, an enterprise can effectively connect external entities, reduce the "friction" between enterprises and users, enterprises and suppliers, enterprises and the government, balance the demands of the enterprise and its external stakeholders, cope with the zero-sum game between different entities, realize the tolerance of different incentives of such entities, and form a synergistic coexisting system.

4.3 Data Create a Zero-Price Effect by Replacing Other Factors of Production

The "non-physical and intangible" nature of data factors enables data to fully penetrate and substitute other factors of production, optimizes the systematic allocation of factor resources, and maximizes efficiency. While replacing other factors of production, data change the allocation ratio between factors of production. This means that the proportion of traditional factors of production represented by capital and labor declines, and the proportion corresponding to data rises. As a result, data may partially or even completely replace other factors of production, to produce digital and data products. Their production function changes from $Y_D = F(K(D), L(D))$ to $Y_D = F(D_K, D_L)$, in which D_K and D_L, as substitutes for "new" capital and labor, achieve equivalent function and output with fewer physical factors and more data factors and create the zero-price effect based on the characteristics of data factors that "marginal costs diminish while marginal returns increase".

The fact that data create the zero-price effect by replacing other factors of production is shown by the following facts: (i) Dematerialization of factors of production. As the digital twin technology connects physical space and virtual space, the materialized factors of production in physical space are mapped to the virtual space, and digital, data-based factors of production are substituted for physical factors of production. (ii) Dematerialization of product forms. The continuous development of the "digital technology cluster" accelerates the informatization, digitization and intelligence of product forms. From the era of industrial economy to the era of digital and intelligent economy, the relationship between product function and physical media changes from "1+1" (one product function, one physical medium) to "N+1" (multiple product functions, one physical medium), "N+1/M" (multiple product functions, one minus physical medium), and "N+0" (multiple product functions, no physical medium).

In the era of industrial economy, enterprises use products to make money, in which case, making money is incompatible with free products or services; in the era of digital and intelligent economy, enterprises use data to make money, in which case, making money is compatible with free products or services. Traditional factors are low-ranking, physical factors, and high-value investment in technological innovation and product research and development leads to high initial cost of physical products. The exclusivity and competitiveness of such factors and the finiteness of resources make it inevitable that the traditional factors will become fewer after being used but their marginal cost will rise. According to the principle that the price of a product is equal to its marginal cost, the price of physical products also rises.

Data factors are high-ranking, virtual factors. While the fixed and initial costs of digital infrastructure are high, the non-exclusivity and non-competitiveness of data factors inevitably lead to the infinite supply of data factors and extremely low or even almost zero marginal cost of reproducing data products. According to the principle that the price of a product is equal to its marginal cost, the price of data products will become lower, ultimately leading to a zero-price effect on data products.

The zero-price effect of data products emphasizes that the price of data products is extremely low or even close to zero, not that the value of data products is zero. On the contrary, in the era of digital and intelligent economy, data products have great value, especially producer surplus and consumer surplus. (i) Producer surplus. Thanks to the innovation of general technology in the "new technology cluster", the continuous interaction among users on digital platforms has gradually increased the number of "user-generated content". In this process, users, as free digital workers, produce many free data products. This saves the labor cost of enterprises, and meanwhile, data are introduced into the production and operation activities of enterprises as free factors. In particular, the gradual increase of data is helpful to stimulate the "learning effect" of machines, optimize the machine learning model, improve the model matching efficiency, and realize the unlimited reuse and appreciation of data in multiple scenarios with zero cost, and thus enhances the residual value of the platform enterprise. (ii) Consumer surplus. In the process of providing digital products or services, platform enterprises leverage their data mining and data analysis capabilities to form a self-organized technology system that automatically adapts to, adjusts, recycles and provides feedback on user data on an ongoing basis to continuously optimize the platform's products or services, meet the diversified, individualized and experience-based needs of their users, increase their users' value perception, and generate huge consumer surplus. It is difficult to directly account for such hidden consumer surplus in the GDP outcome, but consumer well-being is greatly improved.

5. Data Value Rights Determination

5.1 Based on the Theory of Property, the Tragedy of the Private Outweighs the Comedy of the Private with Respect to Data[2]

The theory of property holds that clearly defined property is the foundation and prerequisite for market circulation and transactions. The clearly defined property means that property is exclusive; moreover, the property system can make full use of the market competition mechanism to maximize the efficiency of resource allocation (Tang, 2021). The theory of property mainly emphasizes the exclusivity and competitiveness of physical factors. Different from physical factors, data factors are non-physical factors characterized by non-exclusivity and non-competitiveness. If the property rights of personal data are assigned to individuals or enterprises according to the theory of property, it will be difficult to balance an individual's demand for privacy and an enterprise's demand for data application, and it is even more difficult to maximize efficiency in data resources allocation. In this case, with respect to data, the tragedy of the private outweighs the comedy of the private.

5.1.1 Data property rights assigned to individuals

Assigning data property rights to individuals means fully respecting the privacy of personal data and protecting the personal rights of individuals. In accordance with the General Data Protection Regulation (GDPR) issued by the European Commission in 2016, individuals have rights such as the right of information, the right of access, the right to rectification, the right to portability, the right to erasure, the right to restriction of processing, the right to object, and the right to decision. The GDPR is considered to

[2] This is based on Garrett James Hardin's "tragedy of the common" as tragedy (comedy) of the common (private), which means, based on the economic theory of property rights. When data property rights are assigned to individuals, the private tragic effect of data is greater than the private comic effect. Based on the theory of sharing economy, the comedic effect of data on the Commons is greater than the tragic effect of data on the Commons.

be the most stringent personal data protection regulation. While the personal data right is fully respected and protected, there are also issues such as excessive data restriction and excessive data supervision, which could trigger the "tragedy of the private" (Qi and Liu: 2020). When data property rights are assigned to individuals, on the one hand, the heterogeneity and dispersion of personal data will increase with the exponential growth of data, leading to a sharp rise in the costs of negotiation, transaction and rights confirmation between enterprises and individuals on a "one-to-one" basis; on the other hand, assigning data property rights to individuals will exacerbate the friction between enterprises and individuals in terms of data collection, data circulation and data application, and to some extent hinder the growth of benefit brought about by the scenario-based and large-scale application of data on the part of enterprises. Therefore, when data property rights are assigned to individuals, the increased cost and decreased benefit as a result of defining data property rights will increase the negative externalities of data rights determination, weaken the incentives for enterprises to invest in, develop and use data resources, reduce the efficiency of data resource allocation, and indicate a room for Pareto improvement (Fei et al., 2018).

5.1.2 Data property rights assigned to enterprises

While assigning data property rights to enterprises can optimize the efficiency of data resource allocation, it is difficult to ensure the full respect for individuals' personality rights. In particular, the following two issues need to be resolved:

(i) Data monopoly issue. Based on the path dependence and cyclic inertia of data, when used, data will increase in their quantity, accuracy, uniqueness, and value, and will be more likely to form a monopoly. Data monopoly is mainly divided into: a. Monopoly within the industry. Taking DiDi Chuxing as an example, "DiDi Chuxing" involves DiDi's businesses including Flash, Express, Hitchhiking, Designated Driving, Car Rental, Premier and Freight, which is deeply committed to the mobility industry and undermines the traditional mobility business. By using the technological combination of "data + computing power + algorithm", it fully connects the supply side and the demand side and improves the efficiency of matching between them. With strong "transportation capacity", it attracts both drivers and passengers, leveraging cross-side network effect to generate large-scale and individualized mobility data, including not only the basic information of personal accounts such as user name, gender, cell phone number, and bank card number, but also the user's travel data as well as occupational data, preference data, and consumption ability, which are analyzed accordingly, forming a data-radiated monopoly map with the mobility industry as the core and multiple business segments as branches. b. Cross-industry monopoly. Some enterprises are not only involved in one industry, but also in multiple industries. For example, Alibaba Group not only forms a data-radiated monopoly map with the retail industry as the core, and Tmall, Taobao, Juhuasuan, Freshippo and other business segments as branches, but also has the data-radiated monopoly map of the retail industry infinitely connected with those of the mobility, logistics, and financial industries, which together "weave" a crisscrossing data eco-network and facilitate the formation of monopoly data. On the one hand, by relying on the first-mover advantages in monopoly data, technology and capital, such enterprises set discriminatory prices through "data-driven exploitation", deprive of consumer surplus, create monopoly profits, and generate data rents; on the other hand, by enhancing "data enclosure" behavior, such enterprises set industry entry barriers, deny potential competitors' access to data and restrict the development and innovation of start-ups so as to achieve "data blockade" and accelerate the agglomeration of data, technology and capital to top enterprises, thus forming a data-monopolized circular system and leading to the inefficient state of data factor allocation.

(ii) Data privacy issue. Assigning data property rights to enterprises will certainly improve the efficiency of data use, but it has caused problems such as data leakage and privacy violation. a. Uninformed collection. Before an individual uses a product or service provided by an enterprise, the enterprise needs to inform the user through an explicit or implicit user agreement to obtain consent or authorization for the use of personal data. However, the “notice-consent” model obtains only “ostensible consent” from individuals. User agreements are complex, lengthy and highly professional, and individuals have little knowledge of the specific operating rules of enterprises, such as the purpose, scope, and extent of collection, and the technology and algorithms used by enterprise. Furthermore, individuals who reject to sign user agreements have to relinquish the products or services offered by enterprise. In other words, enterprises use an ostensibly “informed consent” in exchange for individuals’ “forced consent” to the collection of data by enterprises. Consequently, this breaks the balance between enterprises and clients in terms of their individual rights. b. Illegal transactions. Data trading is an effective way to realize the commercial value of data. However, it is difficult to trade personal privacy data which have huge use value and exchange value at a designated trading place, and it is easy to trigger a large number of gray and black transactions and make personal privacy data “transparent” in the underground data trading market, which will exacerbate the data security crisis and violate individuals’ rights. c. Non-transparent application. Uninformed collection and illegal transactions result in non-transparent application of data. After being analyzed and integrated, the preference data, behavioral data and other data at a high level of granularity still lack detailed application field, scope and scenario, resulting in the application of data in a “black box” mode.

5.2 Based on the Theory of Sharing Economy, the Comedy of the Commons Outweighs the Tragedy of the Commons with Respect to Data

Data generators, collectors and annotators as well as data users and operators are involved in the different stages of the data lifecycle; therefore, data generation is an ecosystem in which different right holders collaborate with each other. The fact that the ownership of property rights for data factors involve various entities and dynamically evolve makes it difficult to assign data property rights to a specific entity based on the traditional theory of property. However, the theory of sharing economy emphasizes the right of use rather than ownership. In addition, the non-exclusive and non-competitive characteristics of data factors confront the traditional mentality that emphasizes ownership to “private property”. Instead, we should consider the claims of diverse data subjects and then strengthen data connection, sharing and openness, promote large-scale data mining, analysis and application and market-based circulation and trading, and fully leverage the “public” characteristics of data to maximize the efficiency of data resource allocation while ensuring the compatibility of incentives for diverse data subjects. If data property rights are assigned based on the theory of sharing economy, the comedy of the commons will outweigh the tragedy of the commons with respect to data. The comedy of the commons requires that the goal of data rights determination be changed from data property rights to the rights and interests of data subjects. Specifically, the rights and interests of data subjects include:

5.2.1. Individuals emphasize the personality right to data privacy

The uninformed collection, illegal trading and non-transparent application of personal data by enterprises increase the negative externalities of privacy breaches in data collection and data transfer, and it is imperative to safeguard individuals’ right of personality. The Personal Information Protection Law (PIPL), which took effect in China in 2021, clearly provides that individuals have the right to information, decision, access and duplication, correction and supplementation, erasure, the right to request for explanation, and judicial protection in personal information processing activities; it also clearly states that “sensitive personal information means personal information that, once leaked or illegally used, may easily lead to harm to the personality or dignity of natural persons or harm to personal or property security”. The implementation of the PIPL effectively protects personal privacy in the form of law and strengthens individuals’ effective “control” over data through “authorization” and “permission”. In addition, it is not allowed to excessively collect personal information through “package authorization" or mandatory unification, so as to promote the reasonable use of personal information. During the data collection process, individuals can implement "one authorization for one scenario" based on their privacy risk preference (Liu et al., 2023).

5.2.2. Enterprises emphasize property rights in data collection and data circulation

After obtaining authorization and permission from individuals, enterprises may engage in data application and data circulation on the premise of using digital technology in desensitization and decryption processing to protect individual privacy. Data can be used for two purposes: Self-use and circulation or trading. In the case of data self-use, there is input of digital technology, digital capital and intellectual labor in the course of data processing. In this process, enterprises enjoy the "right to use data for processing". In addition, based on individual authorization, data can also be used for market circulation or trading. In this process, enterprises enjoy the "right to operate data products". Out of respect for capital and labor and considering the contribution made by enterprises to value appreciation in data processing, it is reasonable that enterprises should enjoy the property rights of commercially derived data and data generated from deep processing. In this case, the market should fully play its role in the primary distribution and the mechanism of "evaluating contribution by the market and determining rewards based on contributions" should be followed. In protecting the data property rights of enterprises, it is necessary to fully leverage the trustworthiness and traceability of blockchain, record the "value creation" in each stage of data collection, processing, use and circulation by enterprises, so as to identify the value added by enterprises and leverage the foundational power of technological regulation. To protect the property rights of enterprises in the process of data collection and circulation, China issued the *Opinions on Building a Fundamental Data System to Play a Better Role of Data Factors* (the "Twenty Articles on Data"), which put forward a data property rights system framework of "separation of three rights" – the right to hold data resources, the right to use data for processing, and the right to operate data products.

5.2.3. Governments emphasize the security and core interests of public data

Any leakage of the key data and sensitive data collected and generated by the governments in performing their public management functions and providing public services following the law will seriously undermine the authority and overdraw the credit of the governments. Therefore, it is necessary to fully protect the security and core interests of governmental data. Governmental data has public attributes. As the holder of public data, governments have the responsibility to disclose data to society, improve the transparency of public data, promote the opening and sharing of public data, and increase the supply of high-value data factors, so as to form a public data pool, in which the data application resources are supplied based on demand. Governments also need to encourage and support citizens and enterprises to use open public data to conduct scientific research and innovative development, promote the integration and interoperability of public data with personal data and enterprise data, remove data barriers and data silos, unleash the economic and social value of public data, so as to increase public interests, improve public well-being, and realize the social sharing of the dividends of public data.

6. Data Value Pricing

6.1 Pricing According to Physical Product Attributes

6.1.1 Supply and demand for data products determine the formation of equilibrium price

According to the utility theory, the highest price that demanders are willing to pay is determined by the marginal utility of data value, data quality and user-perceived value. In different application scenarios, the strong "subjectivity" and "sensitivity" of the demanders of data products and data services lead to differentiated and individualized maximum prices. The lowest price data suppliers are willing to accept is determined based on the data processing costs of data collection, annotation, searching, correlation, and storage, historical costs, and operation and maintenance costs. The static pricing method is applied to standardized data products. The maximum and smallest prices from frequent transactions and standardized data products jointly form the supply and demand curves, and their intersection is the equilibrium data product quantity and price.

6.1.2 Data market structure influences data product pricing

The supply and demand of data products provide competition in the data market, which in turn influences the pricing of data products. (i) In a fully monopolized data market, individualized, diversified and networked scenario data are aggregated into a single enterprise, and the supply-oriented monopolistic enterprise adopts the strategies of "1,000 people, 1,000 prices", "number of sales" and "market separation" to segment consumers and the market, implements multilevel price discrimination, sets the monopoly data price, and squeezes the surplus of consumers, so as to enjoy the profit from data monopoly. (ii) In an oligopoly data market, oligopolistic enterprises use their data encryption technology, desensitization technology, and intelligent analysis technology to enhance their data mining, analysis, integration, desensitization, and confidentiality capabilities, and enable primary processed products and in-depth processed products to be aggregated into several leading enterprises. This means that the data trading market stays supply-oriented, and oligopolistic enterprises have competitive advantages and pricing advantages. (iii) In a fully competitive data market where many suppliers provide similar data products, it is easy for suppliers to enter and exit from the data market, and there is no threshold restriction. In addition, such market is highly transparent, and the price of data products is equal to the marginal cost. Although data products have low marginal costs, their high fixed costs bring about high capital and high technical thresholds for data processors; thus, the fully competitive data market exists only in the ideal state or in public data disclosed by the government. In general, the price of standardized data products is set based on data market supply and demand and market competition.

6.1.3 Data trading modes are the foundation of data product pricing

Data trading modes can be divided into two categories according to trading venues:

(i) Over-the-counter trading. The supply side and the demand side negotiate prices on a "one-on-one" and dynamic basis within the price range between the highest price and the lowest price and data prices are variable. However, the "one-on-one" over-the-counter trading mode may cause high search cost, negotiation cost and premium cost and add data uncertainty; moreover, in the absence of third-party supervision, the data trading process may lead to non-compliance with respect to data encryption, data anonymization, and data trading standards, which may further lead to other problems, such as data leakage that breaches privacy (Koutroumpis, Leiponen and Thoma, 2020). This also proves that in over-the-counter data trading, it is difficult to formulate "informal", "highly subjective" standard rules by relying on the contractual relationship; instead, "win-win benefits" or "external reputation" may facilitate the management of the negotiations between data trading parties. (ii) On-exchange trading. Like general commodities, data trading faced many problems in the early days, so "atomized" over-the-counter trading began to converge on fixed, centralized trading venues. Data trading centers (data exchanges) are important venues for on-exchange trading. Data exchange trading centers include several advantages: First, trading venues can reduce high trading cost; as special commodities, data are highly specialized and heterogeneous; they have great value for specific groups in specific scenarios at specific times, while in other cases the value is smaller and highly volatile, making it more difficult to match supply and demand in the short term. In contrast, fixed and concentrated trading venues can significantly reduce the time, money and energy spent in searching information and maximize the efficiency of data resource allocation. Second, trading venues can reduce information asymmetries; as special commodities, data face greater information asymmetry, that is, sellers have all the information, buyers have no information, and data commodities are difficult to “inspect”. Fixed and centralized trading venues can provide compliance constraints for both parties through “credibility”, “after-sale” and “accountability” to avoid market collapse as a result of “bad money driving out good”. Third, trading venues can promote the deep specialization in division of labor. To achieve the long-term healthy development of the data industry, it is necessary to form a specialized, ecosystem-like “data business system”. At the early stage of data commodity development, the data trading party handled the whole process of data trading chain, such as data trading parties search, data contract formulation, data quality audit, and data delivery. In the process of data trading, the non-specialized trading method and high opportunity cost limited the expansion of the data market. However, a fixed and centralized trading venue is conducive to the emergence of “data dealers” specializing in data trading, and the supply and demand sides only need to trade with “data dealers” who play a role in transaction matching, quality assessment and asset assessment. By breaking the whole chain of data trading, a trading venue will form a centralized bilateral or multilateral market. This will prompt the data supply side to focus more on data collection and data processing, while “data dealers” will greatly improve the trading efficiency and productivity by accurately matching data supply and demand sides and promote the prosperous development of the data industry and create a win-win situation for all parties by relying upon the “Smithian dynamics”. The on-exchange pricing method means that “data dealers”, through auction pricing and game pricing, use mechanism design and contractual relationship to reveal the true data valuation of both supply and demand sides, and mitigate the information asymmetries between supply and demand sides (Koutroumpis et al., 2020).

At present, in the domestic data market, the trading volume between the supply and demand sides is small, the trading frequency is low, standardization of data products is at a low level, and scenarios are limited; as a result, prices generated between the supply and the demand sides greatly deviate from the value and it is difficult to form a fair data pricing mechanism (Huang et al., 2022). Therefore, data exchanges/data trading platforms and other third-party institutions need to play a role in generating data prices. The data valuation model and system need to be set up by introducing third-party institutions and considering the elements that influence data value. In data valuation, the business value, cost value, application value, and brand value of data should be comprehensively considered and elements that influence data valuation should be identified in details based on a level-1 indicator established for data valuation. A level-2 indicator should also be set up, including the survey cost, development cost, operation and maintenance cost, management cost, connectivity, granularity, affiliation, reusability, completeness, compliance, timeliness, security and complementarity. A weight should be given to each indicator to form a multidimensional and full-range data valuation system.

6.2 Pricing Based on Virtual Product Attributes

6.2.1 Pricing based on data scenarios

Data are highly dependent on scenarios, and data pricing that ignores scenarios will distort the “value-price” correlation (Ouyang and Gong, 2022). Therefore, it is advisable to adopt a dynamic pricing method, namely “one scenario, one price”. Scenario-based pricing can be divided into following categories: (i) Application scenario-based pricing. Data price is highly dependent on the heterogeneity of quality, connectivity, privacy and granularity. For instance, Zhang and Beltrán (2020) divide data pricing into two categories according to data attributes: One is data granularity-based pricing. Based on data granularity, this data pricing method can be further divided into dataset-based pricing and search-based pricing. In particular, in the case of search-based pricing, the principle of no arbitrage and no discount should be strictly followed and the specific methods include image pricing and minimum unit pricing. The other is data privacy-based pricing. In the process of pricing, “noise” may be added to the original data or search results to realize “scenario justice” that considers both privacy protection and data circulation and to maximize the accuracy of data search results subject to budget or minimize data payment subject to fixed accuracy targets. (ii) Trading scenario-based pricing. The heterogeneity of trading scenarios mainly refers to that of the supply and demand sides in terms of quantity in trading. When the number of data demanders is relatively small, data are highly exclusive and can be priced through auction so that those who offer the best price can obtain data products. When the number of data demanders is relatively large, data are less exclusive and can be priced in the cost-plus method; this method can satisfy the need of data demanders while increasing the revenue of data suppliers and further increasing the total surplus value. (iii) Pricing based on self-production and self-use scenario. Data self-production and self-use is the use of an enterprise' own data to empower the operational decision-making and production process and improve the individualization and experience of its products and services. In this process, although data self-production and self-use is different from external use and does not generate monetized revenue directly, data factors can effectively improve corporate performance and productivity by empowering the operational decision-making and production process. On the one hand, due to data self-production and self-use, enterprises can adopt the cost-plus method in data pricing based on data processing cost. Such data pricing method is relatively "static"; however, in reality, if data are accumulated within an enterprise and are not actively used, the cost-plus pricing will overestimate data value, or if data are often used within an enterprise, the cost-plus pricing will underestimate data value. Given the limitations of the cost-plus method in pricing, the revised historical cost method may be used to comprehensively consider the effect of data quality and use on data cost (Hu and Xu, 2022). On the other hand, as data are self-produced and self-used, internal use of data within enterprises can improve the expected performance and productivity. Therefore, the difference between performances before and after enterprises use data is the price of such data. As such, enterprises can adopt the income method to measure the value of self-produced and self-used data.

6.2.2 Pricing based on financial product attributes

With the increasing volume of data factors and the gradual generation of data assets, profit-seeking financial capital has gradually penetrated all aspects of the data industry (Jin and Chen, 2022). In particular, through financial innovation, financial capital creates a new form of "data factors + mortgage loans" and "data factors + assets backed securities" to pursue the maintenance and increase of data asset value, enjoy the "premium" bonus of data assets and realize the transformation from "data assets" to "data capital". Pricing based on financial product attributes means that the potential risk and expected return of data products should be fully considered in their pricing. (i) Data mortgage. Data mortgage is an effective channel for data monetization by using enterprise data as collateral for fundraising. In a debt contract, data mortgage is different from physical asset mortgage. In the era of industrial economy, enterprises create mortgage over tangible assets, such as plants, and their scale, operation and management capabilities and credit standing are considered financing indicators and constraints. In the era of digital and intelligent economy, enterprise create mortgage over data and other intangible assets, and the types, economies of scale and network of their data are considered key financing indicators and constraints (Zou et al., 2016). (ii) Data securities. Cao Shuo et al. (2021) defined data asset securitization as securities issued by using data assets as underlying assets and using expected cash flow as payment source. By "packaging" and "stratifying" underlying data assets, data asset securitization can form standardized data assets with stable data quality, value and return, reduce the risks in individually heterogeneous data, realize "risk segregation", "risk decentralization" and "sensitivity consistency" among "data packages" and "data levels" and fully capitalize on the "risk decentralization" effect. In addition, data asset securitization is also conducive to alleviating the information costs arising from differences in the quality, value, and returns of data assets, reducing the "lemon discount rate" and mitigating incentive distortions (Zou et al., 2014). It can be concluded that the price of data mortgage and data securities is a collection of discounted prices for their potential risks and expected returns.

7. Conclusions

The "production-exchange" paradigm consists of three tiers: (i) The use of tangible factors to produce tangible products and the exchange of tangible products for tangible products. In the early years of industrial economy, enterprises used labor, land, capital, and other tangible factors of production to produce tangible products. Tangible products were exchanged in the form of currency of equal value. This is the "production-exchange" paradigm of "using tangible factors to produce tangible products and exchanging tangible products for tangible products". (ii) The use of intangible factors to produce tangible products and the exchange of intangible products for tangible products. In the late years of industrial economy, enterprises used technology, knowledge, management, and other intangible factors of production to "produce" tangible products. Technology products, knowledge products, management products, and other intangible products were exchanged for currency. This is a "production-exchange" paradigm of "using intangible factors to produce tangible products and exchanging intangible products for tangible products". (iii) The use of intangible factors to produce intangible products and exchanging intangible products for intangible products. In the era of digital and intelligent economy, enterprises use technology, knowledge, management, data, and other intangible factors to "produce" technology products, knowledge products, management products, data products, and other intangible products and exchange such intangible products for technology products, knowledge products, management products and data products. In this case both factors and products are intangible, which is a "production-exchange" paradigm of "using intangible factors to produce intangible products and exchanging intangible products for intangible products".

From the era of the industrial economy to the era of the digital and intelligent economy, the "production-exchange" paradigm will move from the first two tiers to the third tier. Tiers 1 and 2 are the foundation of tier 3 and tier 3 is the advanced form of tiers 1 and 2. The emergence and development of intangible factors and products, such as data factors and products, have led to a revolution in the "production-exchange" paradigm at tiers 1 and 2 and a transition to the tier-3 "production-exchange" paradigm. It should be noted that the new paradigm is not opposed to the old one but is the result of a systematic transformation and all-round upgrading of the old one, in the process of which, data become new factors of production as well as part of the production function. The new paradigm contains, optimizes, and surpasses the old one and constitutes one of the characteristics of this era. CE

References:

Cao S., Liao C., Zhu Y. Y. Research on the Securitization Path of Data Elements ——Based on DAITs Model[J]. Securities Market Herald, 2021(10): 44-51.

Cao X., Ouyang T. H., and Huang J. M. Research on How Smart Connected Products Reshape Enterprise Boundaries: The Case of Xiaomi[J]. Journal of Management World, 2022(4): 125-142.

Fei F. Y., Yan Z. X., Chen Y. W. Data Nature, Property Rights and Competition in the Digital Economy[J]. Research on Financial and Economic Issues, 2018(2): 3-21.

Hu Y. Y., and Xu X. C. Research on the Statistical Measurement of Business Data Assets[J]. Statistical Research, 2022(9):3-18.

Huang Q. Q., Wang J. D, Chen D. et al. Research on the Data Price Generation Mechanism under the Hyperscale Data Factor Market System[J]. E-Government, 2022(2):21-30.

Iansiti M. The Value of Data and Its Impact on Competition[R]. Harvard Business School Working Paper, 2021(w22-002).

Jin C. L., Chen R. D. Data Valuations and Its Derived Financial Attributes: Formation Logic and Future Challenges[J]. Journal of

Quantitative & Technological Economics, 2022 (7):69-89.

Koutroumpis P., Leiponen A., Thomas L. D. W. Markets for data[J]. Industrial and Corporate Change, 2020(3): 645-660.

Li B., Sun K., Sun G. J. The Participation of Data Elements in Income Distribution: Theoretical Analysis, Factual Basis and Practical Path[J]. Reform, 2022(3):66-76.

Li H. J., Chen X. Y. Study on the Boundaryless Development of Firm——Based on Case Analysis[J]. China Industrial Economics, 2011(6):89-98.

Li H. J., Feng L. Study on Theories of Value Origin of Corporation[J]. China Industrial Economics, 2004(3):52-60.

Li H. J., Zhao L. Data Becomes a Factor of Production: Characteristics, Mechanisms, and the Evolution of Value Form[J]. Shanghai Journal of Economics, 2021(8):48-59.

Liu T. X., Li R. F., Rong K. The Theory of Data Rights and the Hierarchical Authorization Based on Data Generation Scenarios[J]. Journal of Management World, 2023(2):22-39.

Long D., Zhao C. M., Fu S. Q. Digital Engine -- Action Guide for Enterprise Digital Transformation[M]. Beijing: CITIC Press, 2021.

Ouyang R. H., Gong W. The Pricing Mechanism of Data Elements Based on Data Element Value and the Contribution of Market Assessment[J]. Reform, 2022(3):39-54.

Qi Y. D., Liu H. H. A Study on the Factor Property and the Market-Oriented Allocation Mechanism of Data in Digital Economy[J]. Economic Review Journal, 2020(11):63-76+2.

Sambasivan N., Kapania S., Highfill H. et al. “Everyone Wants to Do the Model Work, Not the Data Work”: Data Cascades in High-Stakes AI[C]//Yokohama: The 2021 CHI Conference on Human Factors in Computing Systems, 2021.

Shen W. X. On Usufruct of Data [J]. Social Sciences in China, 2020(11):110-131+207.

Sun F. E. Study on the Rent Allocation of Modular Network Organization[J]. China Industrial Economics, 2013(11):109-121.

Tan Y. J. Economic Analysis of Data Property Rights[J]. Social Science Journal, 2021(1):98-106+209.

Tian J. T., Liu L.Y. Transaction Mode, Rights Definition and Data Element Market Cultivation[J]. Reform, 2020(7):17-26.

Wang C. X, Zhang W. D., Yan M. Is More Data Always Better——An Interdisciplinary Analysis of the Nature of Returns to Data[J]. China Industrial Economics, 2022(7):44-64.

Xu X. C., Zhang Z. W., Hu Y. Y. Research on Deriving Measures of Data Assets in National Accounts[J]. Journal of Management World, 2022(2):16-30+2.

Yin X. M., Lin Z. Y., Chen J. et al. Research on the Dynamic Value Creation Process of Data Element[J]. Studies in Science of Science, 2022(2):220-229.

Zhang M., Beltrán F. A Survey of Data Pricing Methods[J]. SSRN, 2020, 3609120.

Zou X. M., Zhang M., Gao B. The Supply and Demand of Securitization: A Literature Review[J]. Chinese Review of Financial Studies, 2014(4):99-111+125-126.

Zou Z. F, Zou S. Q., Zhang P. Research on Data Pledge Supply Chain Financing under Big Data Environment[J]. Science and Technology Management Research, 2016(20):201-205+233.