Data Products Pricing Mechanism: A Harmonious and Mutual-beneficial Perspective

Li Xiong* and Huixian Zheng*
School of Management, Shanghai University, Shanghai, China

*Corresponding author e-mail: 446270731@qq.com, xiongli8@163.com

Abstract. Data as a new production material has great potential value, and its value is reflected after the utility generated. So it is difficult to set prices for data products. This paper analyzes the characteristics of data products, builds a data product pricing model from a harmonious and mutual-beneficial perspective, and verifies the operability of the model in the simulation pricing strategy. The results of this research can help the construction and development of the data transaction market, and provide ideas for future research on data pricing.

1. Introduction
Data is large-scale and complex with essential characteristics such as rapid growth, diversity, value sparseness and authenticity [1]. Compared to traditional products, data is significantly different in terms of storage methods, analytical methods, and value returns. Data has enormous commercial value, and data will be a strategic resource for economic and social development like land, oil and capital [2–4].

As a new type of production factor, data has the same tradable attributes as traditional production factors. However, the value of data products does not generate direct value like land, raw materials, equipment, and capital. Instead, it refines information from data, transforms information into knowledge. Then, knowledge guides production and management activities or governance practices to improve efficiency, improve accuracy, reduce costs, and reduce deviations. Data ultimately produces direct or indirect economic and social values [5–6]. Therefore, data transactions are more dependent on new-generation information technologies such as analysis and mining. At this stage, data aggregation has achieved scale results, but there is still a lack of cross-industry and cross-domain mobility, showing an “islanding effect”; to achieve the scale effect and strategic value of data, it is necessary to break through the “data island” and promote the free flow of data. Data transaction is a key link in data production and application [7].

Data product pricing is difficult to operate based on general commodity pricing [8]. From the perspective of the seller, the data products cannot be priced according to the microeconomic principles of price equal to the marginal cost. From the perspective of the buyer, the actual price and psychological expected value are compared when purchasing the goods, and then make a purchase decision. Therefore, the pricing of data products can be discussed from the estimated value of both parties. This paper expounds the current status of data transactions, analyzes the dilemma of data product pricing, and on this basis, focuses on the estimated value of data transaction parties to propose a new data products pricing mechanism.
2. Game Analysis of Data Transaction Pricing
This section addresses the solution to the pricing problem of data products. First, establish a two-party pricing model. Both parties have the behavior of buying and selling data products. Both parties will consider the total utility and total value. Assuming that the values of the two data products are different, the pricing formula is obtained through analysis and calculation (Figure 1).

![Data product pricing mechanism based on the perspective of harmonious and mutual-beneficial.](image)

First, suppose that there are only two people in the market to conduct mutual data transactions (codes A and B). Both parties have data products, and there is a willingness to purchase the other party's data products, that is, there are behaviors for buying and selling data products. Since both of them have data products, it is necessary to consider the case where the two data products are equal in value or one side is higher than the other, and the game return matrix in three cases is obtained.

(1) When the value of the data products of the two parties is equal, the strategy set of A is \( S_A = \{Y_A, N_A\} \); the same, the policy set of B is \( S_B = \{Y_B, N_B\} \). At this time, the evolutionary game return matrix between A and B is Table 1.

**Table 1.** The return matrix of A and B when the values of the data products of both parties are equal.

|       | \( Y_A \) | \( N_A \) |
|-------|-----------|-----------|
| \( Y_A \) | \( a_1, a_2 \) | \( b_1, b_2 \) |
| \( N_A \) | \( c_1, c_2 \) | \( d_1, d_2 \) |

When the value of the data products of the two parties is not equal, there will be two situations. One is that the data product value of A is greater than the data product value of B, and the other is that the data product value of B is greater than the data product value of A. At this time, the evolutionary game return matrix between A and B is Table 2 and Table 3, respectively.

**Table 2.** A and B’s return matrix when A’s data product value is high.

|       | \( Y_A \) | \( N_A \) |
|-------|-----------|-----------|
| \( Y_A \) | \( a_1, a_2 \) | \( b_1, b_2 \) |
| \( N_A \) | \( c_1, c_2 \) | \( d_1, d_2 \) |
Table 3. A and B’s return matrix when B’s data product value is high.

| Y_a | Y_b | N_a | N_b |
|-----|-----|-----|-----|
| a_1, a_2 | b_1, b_2 | c_1, c_2 | d_1, d_2 |

Where: $a_1, a_2$ respectively indicate the profit of each of A and B when A and B select the “transaction” strategy when the value of the data products of both parties is equal; $b_1, b_2$ respectively indicate that when the value of the data products of the two parties are equal, A selects the “transaction” strategy, and B selects the “no transaction” strategy, the respective benefits of A and B; $c_1, c_2$ respectively indicate that when the value of the data products of the two parties are equal, A selects the “no transaction” strategy, and when B selects the “transaction” strategy, the respective benefits of A and B; $d_1, d_2$ respectively indicate that when both sides of the data products are equal in value, both A and B select the “no transaction” strategy, and the respective benefits of A and B. Similarly, $a_1, a_2, b_1, b_2, c_1, c_2, d_1, d_2$ respectively indicates that when A’s data product value is high, A and B choose different strategies, A and B respectively have their own benefits.

(2) Suppose that the probability that player A adopts "transaction" strategy is $x$, the probability that player B adopts "transaction" strategy is $y$, and the probability of adopting "no transaction" strategy is $1-x, 1-y$.

When the value of the digital products of the two parties is equal, then the expected return function ($E_A^{11}$) of the player A to select the “transaction” decision is:

$$E_A^{11} = y \cdot a_1 + (1 - y) \cdot b_1 \quad (1)$$

Gamer A selects the expected return function ($E_A^{12}$) of the “no transaction” decision as:

$$E_A^{12} = y \cdot c_1 + (1 - y) \cdot d_1 \quad (2)$$

The overall average expected return ($E_A^1$) of Game Party A is:

$$E_A^1 = x \cdot E_A^{11} + (1 - x) \cdot E_A^{12} \quad (3)$$

Similarly, the expected return function ($E_B^{11}$) of the game party B selecting the “transaction” decision is:

$$E_B^{11} = x \cdot a_2 + (1 - x) \cdot c_2 \quad (4)$$

Game Party B selects the expected return function ($E_B^{12}$) of the “no transaction” decision as:

$$E_B^{12} = x \cdot b_2 + (1 - x) \cdot d_2 \quad (5)$$

Then the overall average expected return ($E_B^1$) of player B is:

$$E_B^1 = y \cdot E_B^{11} + (1 - y) \cdot E_B^{12} \quad (6)$$

In summary, the replication dynamic equations of the game A and B evolution models are:
\[
\begin{align*}
\frac{dx}{dt} &= x \cdot (E_i - E_{i+1}) \\
\frac{dy}{dt} &= y \cdot (E_j - E_{j+1})
\end{align*}
\]

(7)

Finished up:

\[
\begin{align*}
\frac{dx}{dt} &= x \cdot (E_i - E_{i+1}) = x \cdot (1 - x) \left[ (b_i - d_i) + y \cdot (a_i - b_i - c_i + d_i) \right] \\
\frac{dy}{dt} &= y \cdot (E_j - E_{j+1}) = y \cdot (1 - y) \left[ (c_j - d_j) + x \cdot (a_j - b_j - c_j + d_j) \right]
\end{align*}
\]

(8)

(9)

Through the above equation, the two-game system has five replicated dynamic equilibrium points on the plane \( m = \{(x, y) | 0 \leq x, y \leq 1\} : (0, 0), (1, 0), (1, 1), (0, 1) \) and \( \left\{(a_i - b_i - c_i + d_i)/(d_i - b_i), (a_j - b_j - c_i + d_j)/(d_j - c_j)\right\} \), if and only if \( 0 \leq (a_i - b_i - c_i + d_i)/(d_i - b_i) \) and \( 0 \leq (a_j - b_j - c_i + d_j)/(d_j - c_j) \) is established.

3. Game pricing model

3.1. Seller's perspective

The source of data products is basically the resources deposited by the daily operations of enterprises, governments and scientific research institutions. It can be said to be renewable resources with almost no consideration of production costs. To provide high-value data requires more data processing costs. However, the cost of copying and distributing data products is negligible [9].

Since only two sellers conduct mutual data transactions, the pricing of their data products is affected by the value of the other's data products. If the two parties have the same value, the price will not be much different. If the two parties have different values, the party with the higher value will set a higher price than the one with the lower value. Therefore, it can be concluded that:

\[
q_A = q_B \Rightarrow p_A = p_B
\]

(10)

\[
q_A > q_B \Rightarrow p_A > p_B
\]

(11)

\[
q_A < q_B \Rightarrow p_A < p_B
\]

(12)

3.2. Buyer's perspective

Buyers in the data market have personal preferences that make their purchase decisions based on their requirements, preferences, and prices through a self-selection process [10]. This self-selection is described by the buyer's utility function, which can be expressed in terms of the value of the data product and the willingness to purchase: \( w(\theta, q) = \theta q \), where \( \theta \) represents the buyer's preference coefficient and \( q \) represents the value of the data product the buyer wants to purchase. Set \( \theta \in [0, \theta_{\max}] \), \( q \in [0, q_{\max}] \), where \( \theta_{\max} \) and \( q_{\max} \) are the corresponding maximum values. The utility obtained by the buyer under the price \( p \) is:

\[
w(\theta, p, q) = \theta q - p
\]

This function assumes that the willingness to purchase is a linear function of the value of the data product and reflects the difference in buyer value preferences. These upper and lower thresholds are referred to as saturation and reserve value levels, respectively. Adopt this model and assume that the buyer \( j \) (\( j = A \) or \( B \)) has an \( \theta \) preference coefficient and the purchase intention for the data product \( i \) (\( i = A \) or \( B \)) is \( w_i \):
\[ w_j = \begin{cases} 
0, & \text{if } q_j < q_j^i \\
\theta_j(q_j - q_j^i), & \text{if } q_j^i < q_j < q_j^i \\
\theta_j q_j^i, & \text{if } q_j > q_j^i 
\end{cases} \quad (i=A \text{ or } B) \tag{13} \]

The utility obtained from purchasing data product \( i \) can be expressed as:

\[ u_j(q_j, q_j, \theta_j) = \begin{cases} 
0, & \text{if } q_j < q_j^i \\
\frac{w_j}{\theta_j} - p_j, & \text{otherwise} 
\end{cases} \tag{14} \]

### 3.3. Two-party pricing model

Considering the total utility and total value of the two parties, a two-party game pricing model was established to solve this problem. First, as a seller, you can determine the sales price and value level of the data product. Then, as a buyer, you can make your own purchase decisions through self-selection. Players pursue maximum profit and maximum utility, and make optimal decisions based on known asymmetric information.

The profit function of player A is: \( \pi_A = p_A x \); the utility function of game side A is: \( u_A = (w_A - p_B)y \); the profit function of game party B is: \( \pi_B = p_B y \); the utility function of game side B is: \( u_B = (w_B - p_B)x \).

Trading decision as a seller

\[ \max \pi(q_j, p_j) = \pi_A + \pi_B = p_A x + p_B y \] (15)

Trading decision as a seller

\[ \max u(q_j, \theta_j) = u_A + u_B = (w_A - p_B)y + (w_B - p_B)x \] (16)

### 4. Data product pricing strategy

Considering the interests of both parties in the transaction, the pricing model of data products is established from the perspective of harmony and mutual-benefit. In this part, the possible transaction situations of data products are analyzed and the optimal pricing decision of data products is made.

Since the decision tree can analyze multiple situations and make feasible and effective results, this paper will use the decision tree method to simulate the data product pricing process, which shows that the pricing model is suitable for most data product pricing and easy to operate (such as Figure 2).

![Figure 2. Data product pricing decision tree based on the perspective of harmonious and mutual-beneficial.](image-url)
Where, $\theta$ is the buyer's preference coefficient, which can be used as the probability of occurrence of various natural states, and $Z$ is the total transaction benefit under the scheme. By substituting the standardized desensitization data after cleaning into this model, various results are obtained for various schemes. Under the condition, the calculation and comparison of the profit and loss value can obtain the solution with the largest expected value and promote the completion of data transaction.

5. Conclusion
The rapid development of the Internet in various industry sectors has made the role of data in production and operation increasingly important. As a new production material, the scarcity and uncertainty of data values make the pricing of data products more difficult. This paper fully considers the psychology of both parties, and constructs the data product pricing mechanism from the perspective of harmony and mutual-beneficial, and provides ideas for future research on data pricing.

This paper proposes that data product pricing should consider the interests of the transaction entity, and the value evaluation model is established based on the unique properties of data products, and use this model to measure the price. In the perspective of a harmonious and mutual-beneficial, it is necessary to consider the buyer's estimated value and the cost of the seller to balance the benefits of the buyer and seller, which is operability and rationality. At present, this paper only establishes the pricing of data under the game between the two parties. Future research can add more trading entities, which is closer to the reality.

References
[1] George G, Haas MR, Pentland A. Big Data and Management: From the Editors [J]. Academy of Management Journal, 2014, 57(2):321-326.
[2] Gupta S.-Kar-A.-K.-Baabdullah-A.-&-Al-Khowaiter-W.-A.. Big data with cognitive computing: A review for the future [J]. International Journal of Information Management, 2018, 42:78-89.
[3] Hashem IAT, Chang V, Anuar NB, et al. The Role of Big Data in Smart City [J]. International Journal of Information Management, 2016, 36(5):748-758.
[4] Ragini JR, Anand PMR, Bhaskar V. Big Data Analytics for Disaster Response and Recovery Through Sentiment Analysis [J]. International Journal of Information Management, 2018, 42(10):13-24.
[5] Maxime c. cohen, Ruben lobel, Georgia perakis. Dynamic Pricing through Data Sampling [J]. Production and Operations Management, 2018, 27(6):1074-1088.
[6] Liang F, Yu W, An D, et al. A Survey on Big Data Market: Pricing, Trading and Protection [J]. IEEE Access, 2018, 6(99):15132-15154.
[7] Mauro femminella, Matteo pergolesi, Gianluca reali. IoT, big data, and cloud computing value chain: pricing issues and solutions [J]. Springer International Publishing, 2018, 73(7):511-520.
[8] Shasha liu, Bingjia shao, Yuan gao, et al. Game theoretic approach of a novel decision policy for customers based on big data [J]. Springer US, 2018, 18(2):225-240.
[9] Acharya A.-Singh-S.-K.-Pereira-V.-&-Singh-P.. Big data, knowledge co-creation and decision making in fashion industry [J]. International Journal of Information Management, 2018, 42:90-101.
[10] Erevelles S, Fukawa N, Swayne L. Big Data Consumer Analytics and the Transformation of Marketing [J]. Journal of Business Research, 2016, 69(2):897-904.