

The impact of Data assets on price markup

Yitong Liu

School of Tourism management, Yanshan University, Qinhuangdao 066004, China

lyt312227@163.com

Abstract. In the digital economy, data assets have become a critical driver of competitive advantage, yet existing literature remains scarce on how these assets influence firms' market power. Using a sample of China A-share listed firms from 2007 to 2024, this study applies the De Loecker and Warzynski production function approach to estimate firm-level price markups and systematically empirically examines the impact, mechanisms, and economic consequences of data assets on price markups. The findings reveal that data assets significantly elevate firms' price markups, and this relationship withstands a battery of robustness checks. Mechanism analysis demonstrates that data assets enhance markups through three channels: raising capacity utilization, alleviating financing constraints, and reducing agency costs. Heterogeneity tests show that these effects are more pronounced in non-state-owned enterprises, industries with higher competitive intensity, and regions with more developed data factor market institutions. Finally, the economic consequence analysis confirms that the markup gains driven by data assets substantially boost total factor productivity, thereby empowering firms to pursue high-quality development. This study enriches the microeconomic literature on the role of data as an input factor and provides new empirical evidence on the value-creation mechanisms of data assets, offering important insights for firms' digital transformation strategies and for improving data factor market governance.

Keywords: data assets, price markup, capacity utilization, financing constraints, agency costs.

1. Introduction

The rapid development of the digital economy is profoundly reshaping global industrial structures and competitive dynamics among firms. As the most strategically valuable resource in this era, data has been established as a core factor of production alongside land, labor, capital, and technology. The Opinions on Building a More Perfect Institutional Mechanism for Market-Oriented Allocation of Factors, issued by the Central Committee of the Communist Party of China and the State Council in 2020, was the first to elevate data to the same status as traditional production factors, explicitly calling for the accelerated development of a data factor market. The Opinions on Building Basic Data Systems to Better Leverage the Role of Data Factors, released in 2022, further delineated the foundational institutional framework for data property rights, circulation and trading, and benefit distribution. In 2023, the Ministry of Finance promulgated the Interim Provisions on Accounting Treatment of Enterprise Data Resources, which formally incorporated eligible data resources into corporate balance sheets for recognition and measurement. Collectively, this series of institutional arrangements marks the progression of data factors from macro-level policy articulation to enterprise-level practices of recognition, valuation, and management, while also providing academic research with a new institutional setting and empirical foundation.

At the micro level, the accumulation and utilization of data assets are increasingly becoming a core driver in reshaping firm competitiveness. Through the continuous collection, integration, and deep mining of production and operational data, market data, and consumer behavioral data, firms can achieve intelligent optimization of production processes, precision matching in supply chains, and personalized customization of products and services, thereby establishing differentiated competitive advantages in product markets. Yet a critical question remains: how exactly do data assets translate into market power and pricing capability? What are the underlying transmission channels and boundary conditions? To date, theoretical explanations and empirical evidence addressing these questions remain insufficient, calling for further in-depth investigation.

Price markup serves as a core indicator for measuring firms' market power and pricing capability. A higher markup implies that a firm can set product prices well above marginal cost without significant customer attrition, constituting a direct manifestation of competitive advantage and market position. The traditional literature has explained cross-sectional differences and temporal variations in firm-level markups primarily from the perspectives of R&D innovation, market structure, trade openness, and managerial efficiency, laying a solid foundation for understanding how pricing power is formed. In the digital economy era, however, data as a novel production input and competitive instrument has yet to be incorporated into mainstream analytical frameworks of firm pricing capability. Data assets possess a dual function—reducing marginal costs while strengthening demand-side premium extraction—and should theoretically exert a positive effect on firm markups. Nonetheless, this logical inference still requires rigorous empirical verification.

Against this backdrop, the present study takes China's A-share listed firms over the period 2007-2024 as the research sample, constructs the core explanatory variable using data asset information disclosed in corporate financial reports, and measures firm-level price markups as the dependent variable following the production function approach of De Loecker and Warzynski (2012)^[1]. It then systematically examines the effect of data assets on firm markups, the underlying transmission mechanisms, and heterogeneous patterns. The potential marginal contributions of this paper are threefold. First, by incorporating data assets into the analytical framework of firm market power, this study extends the literature on the determinants of price markups from a digital economy perspective. While existing research has focused predominantly on traditional factors such as R&D innovation and market structure, this paper is among the first to treat data assets as a core explanatory variable, providing fresh empirical evidence on the sources of pricing capability in the digital era. Second, this study identifies two specific transmission channels—capacity utilization and financing constraints—through which data assets influence price markups, effectively opening the “black box” between the two. Third, by exploring heterogeneity along the dimensions of firm ownership and industry competition intensity, this study delineates the boundary conditions under which data asset effects operate, offering meaningful empirical guidance for firms seeking to formulate differentiated digital transformation strategies.

2. Organization of the Text

2.1. Literature review

In the digital economy era, data assets have gained increasing prominence as a critical factor of production, prompting extensive scholarly inquiry into their conceptual delineation, economic effects, and measurement approaches.

Regarding conceptual delineation, Yu, Niu, and Yuan (2024)^[2] define data assets as data resources recorded in physical or electronic form that are owned or controlled by a firm and capable of generating future economic benefits. Together with Xu, Zhang, and Hu (2022)^[3], they further note that the formation of data assets follows a value chain process of “data collection, storage, analysis, and application,” clarifying their nature as a factor of production. In terms of economic effects, Yu et al. (2024) identify data assets as an effective driver of firm value enhancement; Li, Dong, Zhang, and Tao (2023)^[4] further demonstrate that data assets can significantly boost corporate innovation investment by alleviating financing constraints and strengthening inter-firm R&D collaboration. With respect to measurement approaches, Yu et al. (2024) and Li et al. (2023) employ text analysis methods to quantify the level of data assets, whereas Xu et al. (2022) propose an “adjusted cost approach” from a macroeconomic accounting perspective. Additionally, Wang and Yang (2024)^[5] examine the transformation of the Chinese-style management accounting system in the context of data as a factor of production, while Lu et al. (2023)^[6] provide empirical evidence on the relationship between data assets and firm development.

Price markup is a core indicator for gauging firms' market power and pricing capability. The traditional literature has approached this topic primarily from the perspectives of cost shocks, market

structure, trade openness, and managerial efficiency. Regarding cost shocks, Sheng and Liu (2017)^[7] and Zhao . (2018)^[8] show that upward cost pressures arising from minimum wage legislation and exchange rate fluctuations significantly erode firm markups. From a market competition perspective, Wang et al. (2020)^[9] find that the establishment of administrative approval centers intensifies competition by lowering market entry barriers, thereby depressing markups; Zhao et al. (2019) demonstrate the adverse effects of spatial agglomeration on markups from a clustering standpoint. With respect to firm heterogeneity, Bessen and Righi (2019)^[10] show that large firms' investment in proprietary software can raise markups; Jiang (2021)^[11] finds that increases in firm age, labor-to-capital ratio, and scale all contribute to higher markups; while Zhu and Zhang (2015)^[12] document that non-exporting firms in China generally exhibit higher markups than their exporting counterparts. In addition, Li and Chen (2018)^[13] provide evidence from the perspective of market segmentation, and Jensen and Meckling (1976)^[14] lay the theoretical foundation linking agency costs to corporate governance.

In the domain of the digital economy and servicification of inputs, Bai and Yu (2021)^[15] demonstrate that digital economy development significantly reduces firm markups through competition effects and cost pressure channels, whereas Cheng and Sun (2020)^[16] find, from a global value chain perspective, that servicification of manufacturing inputs promotes markup growth via cost reduction and technological innovation channels; Yu et al. (2025)^[17] further confirm that digital human capital investment can significantly raise firm markups by improving product quality and enhancing production efficiency. In terms of measurement methodology, the production function approach proposed by De Loecker and Warzynski (2012)—commonly referred to as the DLW method—has become the mainstream choice given its independence from specific market structure assumptions. Beyond this, a body of internationally influential scholarship, including Brynjolfsson and McElheran (2016)^[18], Varian (2018)^[19], Goldfarb and Tucker (2019)^[20], Autor et al. (2020)^[21], and Hall (2018)^[22], has examined such themes as data-driven decision-making, the economics of artificial intelligence, and the rise of superstar firms, collectively offering important reference points for understanding the evolution of market power in the digital economy.

2.2. Theoretical Analysis and Hypothesis Development

On the theoretical front, data assets refer to data resources that are accumulated, owned, or controlled by a firm in the course of its production and business activities, recorded and stored in electronic form, and capable of generating future inflows of economic benefits. Compared with traditional physical assets, data assets possess several distinctive properties: non-rivalry, increasing returns to scale, strong complementarity, and time sensitivity. Non-rivalry means that the same data can be repeatedly deployed across multiple business scenarios without depletion. Increasing returns to scale implies that the value of data grows non-linearly as its volume expands and its dimensions become richer. Strong complementarity indicates that data assets must be deeply integrated with algorithms, computing power, and application contexts before their value can be fully unlocked. Time sensitivity suggests that the economic value of data may decay over time, necessitating continuous updating and maintenance.

These properties collectively determine that the value creation mechanism of data assets differs fundamentally from that of conventional assets, endowing data with a distinctive set of pathways through which it shapes firm competitive advantage. Non-rivalry drives the marginal cost of reusing data across multiple business scenarios toward zero, directly underpinning improvements in capacity utilization. Strong complementarity requires deep fusion of data with algorithms and computing infrastructure; the high technological threshold inherent in this integration erects information barriers that strengthen a firm's ability to command demand-side premiums. Time sensitivity, in turn, compels firms to engage in sustained digitalization investment to maintain their competitive edge; this imperative of continuity positions data assets as a critical signal through which firms can alleviate financing constraints and secure long-term capital support.

The impact of data assets on firm markups can be analyzed in depth along two dimensions: the cost side and the demand side. From the cost side, production and operational data accumulated by a firm can support applications such as intelligent production scheduling, predictive maintenance, and real-time quality monitoring, effectively reducing equipment downtime and defect rates while improving the continuous operating efficiency of production lines, thereby significantly lowering the marginal cost of production per unit of output. As emphasized by Bloom et al. (2014)^[23], the diffusion of information technology can substantially enhance productivity; data assets, serving as the core vehicle in this process, spread the data service cost per unit of output through economies of scale. Meanwhile, data assets enable firms to accurately forecast and dynamically match upstream and downstream supply-and-demand information, reducing inventory overstocking, stockout losses, and logistics waste attributable to information asymmetry, which in turn compresses marginal costs in procurement and warehousing. Furthermore, data assets exhibit near-zero marginal replication costs—once the underlying data infrastructure is in place, the incremental data processing cost required for each additional unit of output is extremely low. As the scale of data assets continues to expand, the data service cost per unit of output is progressively diluted, giving rise to data-based economies of scale. The resulting decline in marginal cost directly drives markup expansion: holding the product price constant, a reduction in marginal cost raises the ratio of price to marginal cost. Even if a firm chooses to pass part of the savings on to consumers, so long as the price reduction is smaller in magnitude than the cost reduction, the markup can still increase.

From the demand side, data assets can substantially strengthen a firm's premium-pricing capability, enabling it to set product prices at higher levels without incurring significant customer attrition. By leveraging user behavioral and preference data, firms achieve personalized product customization and precision-driven iteration, thereby improving the fit between products and consumer needs and enhancing product distinctiveness. According to Porter's (1985)^[24] theory of competitive advantage, such an increase in differentiation reduces the price elasticity of demand, allowing firms to sustain higher markups. The greater the degree of differentiation, the lower consumers' willingness to switch to substitutes, the smaller the price elasticity of demand a firm faces, and the larger the markup space it can maintain. Moreover, data assets support the implementation of differentiated pricing strategies based on consumer profiling, enabling firms to precisely identify willingness to pay across distinct market segments and set prices closer to consumers' reservation prices. Through the continuous accumulation of user data, firms build personalized service ecosystems—including recommendation systems, customized services, and loyalty programs—that effectively raise consumer switching costs. When migrating to a competitor entails forfeiting an established personalized experience, consumer sensitivity to price changes diminishes and the firm's pricing latitude expands accordingly. At the same time, the accumulation of data assets exhibits pronounced path dependence and first-mover advantage characteristics; firms that are first to amass large-scale, high-quality data can erect data barriers that competitors find difficult to replicate in the short run, thereby securing an informational advantage in the market and obtaining stronger market power and pricing authority.

It warrants emphasis that cost reduction does not constitute a sufficient condition for markup expansion. Under perfectly competitive market conditions, cost savings are readily dissipated through intensified price rivalry among incumbents. Accordingly, the markup-enhancing effect of data assets is fundamentally contingent upon the synergistic interaction between cost-side efficiency gains and demand-side differentiation. Data assets simultaneously compress marginal production costs and augment product distinctiveness through personalized customization, thereby attenuating the price elasticity of demand. This dual mechanism furnishes the requisite market power foundation that enables firms to appropriate, rather than fully transmit to consumers, the rents generated by cost efficiencies, sustaining elevated markup levels in equilibrium.

Synthesizing the foregoing analysis from both the cost side and the demand side, data assets simultaneously reduce firms' marginal production costs and strengthen their premium-pricing capability; these two channels jointly drive the expansion of price markups. Based on this reasoning, the following hypothesis is proposed:

H1: Data assets exert a significant positive effect on firm price markups.

Building upon the baseline analysis, this study further examines the transmission mechanisms through which data assets influence price markups along two pathways: capacity utilization and financing constraints.

Capacity utilization reflects the ratio of a firm's actual output to its potential maximum output and serves as a core indicator of the efficiency with which production equipment and resources are employed. An increase in capacity utilization implies that a firm has achieved a higher level of effective output under existing resource constraints, with fixed and quasi-fixed costs allocated to each unit of product correspondingly reduced, thereby effectively lowering average costs. Data assets can drive improvements in capacity utilization through multiple channels. First, by leveraging historical sales data, market intelligence, and external environmental data to construct demand forecasting models, firms can substantially enhance the accuracy of demand anticipation, mitigating capacity idleness or shortfalls attributable to demand misjudgment and enabling precise alignment between production schedules and actual demand. As Acemoglu and Restrepo (2018)^[25] note, technological progress can enhance production efficiency by optimizing human-machine collaboration; data assets operate precisely through this mechanism, maximizing equipment utilization time while meeting delivery requirements and reducing changeover losses and idle running. Through continuous monitoring and analysis of equipment operational data, firms can establish fault early-warning models that identify anomalous signals before serious malfunctions occur and schedule preventive maintenance, thereby averting disruptions to production continuity caused by unexpected downtime and compressing average costs. When a firm achieves higher effective output with equivalent resource inputs, the fixed costs borne by each unit of product are substantially diluted and the total cost curve shifts downward; provided that product prices remain unchanged or decline by less than the cost reduction, price markups expand accordingly. Based on this reasoning, the following hypothesis is proposed:

H2: Data assets enhance firm price markups through the channel of improving capacity utilization.

Financing constraints refer to the situation in which firms, due to information asymmetry in external capital markets, insufficient collateral, or limited credit history, are unable to obtain adequate external financing at reasonable cost and are consequently forced to forgo investment opportunities with positive net present value. Financing constraints restrict firms' capital allocation toward R&D innovation, market expansion, and capacity upgrading, impeding their ability to build and sustain competitive advantages through continuous investment and thereby limiting the enhancement of their pricing power.

Data assets can effectively alleviate the financing constraints confronting firms through multiple channels. First, a firm's possession of abundant data assets signals a high degree of digitalization and informatization in its business operations, enabling it to provide external investors and creditors with more comprehensive, timely, and verifiable operational information, thereby significantly reducing the degree of information asymmetry between borrowers and lenders. When financial institutions can more accurately assess a firm's operational performance, cash flow prospects, and risk profile, their demand for uncertainty-related risk compensation declines, and both lending willingness and credit limits increase correspondingly. Second, according to the credit rationing theory of Stiglitz and Weiss (1981)^[26], the fundamental obstacle to external financing lies in information asymmetry. As data asset ownership certification and valuation systems progressively mature, data assets are increasingly gaining recognition from financial institutions as a novel form of collateral. Firms can broaden their financing channels through mechanisms such as data asset pledging and data trusts, providing asset-light technology firms with financing alternatives beyond traditional real estate and equipment collateral. Third, a substantial stock of data assets transmits positive signals to capital markets, indicating that a firm possesses strong digitalization capabilities and future growth potential, which facilitates equity investment attraction and lowers the risk premium on debt financing.

The alleviation of financing constraints enables firms to sustain investment at lower capital costs—in R&D innovation to strengthen product differentiation, in capacity upgrading to reduce production costs, and in marketing to reinforce brand premiums. These investments ultimately translate into stronger core competitiveness and higher price markups. Based on this reasoning, the following hypothesis is proposed:

H3: Data assets enhance firm price markups through the channel of alleviating financing constraints.

The principal-agent problem constitutes a critical factor constraining the efficiency of internal resource allocation within firms. The separation of ownership and control in modern corporations gives rise to goal incongruence and information asymmetry between shareholders and management; as agents, managers may deviate from the objective of shareholder value maximization and engage in self-interested behaviors such as perquisite consumption, overinvestment, and “empire building.” The resulting agency costs inflate period expenses and operational costs, distort resource allocation structures and strategic decision-making, and ultimately erode firms’ competitiveness and pricing power in product markets. As Jensen and Meckling (1976) observe, agency costs represent a significant impediment to the efficiency of internal resource allocation.

Data assets can effectively reduce internal agency costs through multiple dimensions. First, once a firm establishes a data governance framework covering core business functions including finance, operations, procurement, and sales, managerial decision-making behavior and resource allocation processes are systematically recorded, stored, and presented in structured form. Shareholders and boards of directors can leverage digitalized monitoring tools such as operational analytics dashboards and anomaly alert systems to obtain timely and verifiable information on managerial conduct at costs substantially lower than those associated with traditional governance mechanisms, thereby compressing the scope for managerial opportunism and reducing principals’ monitoring costs. Second, the sufficient accumulation of data assets enables major business decisions—including investment evaluation, pricing strategy formulation, and capacity planning—to be substantiated through historical data analysis and quantitative modeling, rendering the decision-making process increasingly standardized and traceable; in the absence of adequate data-based justification, managers find it difficult to advance projects that deviate from value-maximizing objectives, effectively constraining their discretionary latitude. Furthermore, data assets provide the informational foundation for constructing multidimensional performance evaluation systems, enabling firms to integrate financial indicators organically with non-financial metrics such as customer satisfaction, product quality, and innovation progress, thereby measuring managers’ genuine operational contributions with greater precision. The enrichment of evaluation dimensions and cross-validation among indicators increase the difficulty for managers to embellish overall performance through selective manipulation, strengthening the effectiveness of incentive contracts.

Once agency costs are reduced through the aforementioned mechanisms, the transmission effects on price markups manifest along two dimensions. On the cost side, the reduction of perquisite consumption, the curtailment of redundant investment, and the improvement of operational efficiency directly lower administrative expense ratios and unit operating costs, widening the gap between product price and cost given unchanged output prices. On the strategic side, as managerial decision-making realigns with value-maximizing objectives, firms become more inclined to invest in differentiation-oriented R&D and brand equity development—strategic initiatives with long-term returns that reinforce product premium capacity from the demand side. These two effects operate synergistically, jointly driving the enhancement of price markups. Based on this reasoning, the following hypothesis is proposed:

H4: Data assets enhance firm price markups through the channel of reducing agency costs.

3. Research Design

3.1. Sample Selection and Data Sources

This study takes Chinese A-share listed firms over the period 2007-2024 as the initial research sample and applies the following screening procedures: excluding firms in the financial and insurance industries; excluding observations with missing variable data; excluding firms designated as ST, *ST, or PT; and to mitigate the influence of outliers, winsorizing all continuous variables at the 1st and 99th percentiles. Firm-level data are obtained from the China Stock Market and Accounting Research (CSMAR) database.

3.2. Key Variable Definitions

(1) Dependent Variable: Firm Price Markup (*Markup*). Following the methodology established in De Loecker and Warzynski (2012) and related seminal studies, this paper employs the DLW approach to measure firm-level price markups. Grounded in the cost-minimization conditions of the firm's production process, this method defines the price markup as the ratio of the output elasticity of a variable input to that input's share in total operating costs. Compared with traditional financial indicators such as gross profit margins or accounting profit ratios, the DLW method more precisely isolates distortions arising from changes in returns to scale and shifts in factor input structures, thereby providing a more rigorous and objective reflection of a firm's true pricing power and premium-capturing ability in product markets.

(2) Core Explanatory Variable: Data Assets (*DA*). Given that data assets, as an emerging asset category, remain in the progressive stages of institutional and practical implementation, neither academia nor industry has yet converged on a unified measurement standard. The prevailing approaches in the literature involve either computing data assets from financial statement figures or quantifying them through textual information disclosed in annual reports. Considering the various limitations inherent in text-based analysis, this study follows the approach of Lu Zheng et al. (2023) and measures firm-level data assets (*DA*) as $\text{Ln}(\text{market value} - \text{fixed assets} - \text{financial assets} - \text{intangible assets})$.

(3) Mechanism Variables. Based on the preceding theoretical analysis and research hypotheses, this study selects the following three mechanism variables in accordance with the core logic of mediation testing. The first is capacity utilization (*CU*), measured following existing literature as the ratio of actual output to potential frontier output; a higher value indicates greater synergistic efficiency in activating idle capacity and reducing front-end sunk costs. The second is financing constraints, for which classical indices include the SA, WW, and KZ indices; in line with the research objectives, this study adopts the widely applied KZ index (*KZ*) to capture the severity of financing constraints, where a larger value signifies stronger external financing constraints and greater difficulty in accessing financial resources. The third is agency costs; given that internal agency frictions in modern corporate governance can severely erode operational efficiency and pricing strategies, this study employs the administrative expense ratio (*Mfee*) as the proxy for agency costs, where a higher ratio indicates more severe internal managerial inefficiency and rent-seeking costs.

(4) Control Variables. To mitigate estimation bias arising from potential omitted variables, this study follows the established research paradigm on firm price markups and incorporates a set of control variables that may influence firms' pricing power into the econometric model. These include firm size (*Size*), leverage ratio (*Lev*), return on total assets (*Roa*), Tobin's Q (*Tbq*), board size (*Board*), shareholding proportion of the largest shareholder (*Top1*), proportion of independent directors (*Indep*), and firm age (*Age*). Detailed definitions and measurement methods for all variables are presented in Table 1.

Table 1. Variable Definitions

Type	Variable Name	Symbol	Measurement
Dependent variable	Price markup	<i>Markup</i>	Estimated via the DLW method
Explanatory variable	Data assets	<i>DA</i>	$\ln(\text{market value} - \text{fixed assets} - \text{financial assets} - \text{intangible assets})$
Mechanism variables	Capacity utilization	<i>CU</i>	Ratio of actual output to frontier output
	Financing constraints	<i>KZ</i>	KZ index measuring financing constraints
	Agency costs	<i>Mfee</i>	Administrative expenses / operating revenue
Control variables	Firm size	<i>Size</i>	Natural logarithm of total assets at year-end
	Leverage ratio	<i>Lev</i>	Total liabilities / total assets at year-end
	Return on assets	<i>Roa</i>	Net profit / total assets at year-end
	Market value	<i>Tbq</i>	Tobin's Q
	Ownership concentration	<i>Top1</i>	Shareholding proportion of the largest shareholder
	Board size	<i>Board</i>	Natural logarithm of the number of board directors
	Independent director ratio	<i>Indep</i>	Number of independent directors / number of board directors
	Firm age	<i>Age</i>	Natural logarithm of the difference between the observation year and the listing year

3.3. Model Specification

To empirically examine the impact of data assets on firm price markups, the following regression model is constructed:

$$Markup_{it} = \alpha_0 + \alpha_1 DA_{it} + \alpha_h \sum Control_{it} + FirmFE + YearFE + \varepsilon_{it} \quad (1)$$

where the dependent variable is the price markup (*Markup*) of firm *i* in period *t*, the explanatory variable is data assets (*DA*) of firm *i* in period *t*, *Control_{it}* denotes the set of control variables, *FirmFE* represents firm fixed effects, *YearFE* represents year fixed effects, and ε_{it} is the random disturbance term. If the estimated coefficient α_1 is significantly positive, it indicates that data assets can enhance firm price markups, and vice versa.

Meanwhile, following the methodological guidance of Jiang Ting (2022)^[27] on mechanism testing, the mediation analysis is conducted through the following regression model:

$$Med_{it} = \beta_0 + \beta_1 DA_{it} + \beta_h \sum Control_{it} + FirmFE + YearFE + \varepsilon_{it} \quad (2)$$

where *Med_{it}* represents the mechanism variable, and all other variables are defined consistently with the baseline regression model above.

4. Empirical Results

4.1. Descriptive Statistics

Table 2 reports the descriptive statistics for the main variables. The dependent variable, price markup (*Markup*), has a mean of 1.1578, a median of 1.0672, and a standard deviation of 0.3772, indicating substantial heterogeneity in premium-capturing ability across firms. The core explanatory variable, data assets (*DA*), has a mean of 22.7057, a median of 22.5372, and a standard deviation of 1.1455, suggesting that data asset allocation among Chinese listed firms exhibits considerable imbalance. Regarding the control variables, firm size (*Size*) has a mean of 22.4134, leverage ratio (*Lev*) a mean of 0.4379, return on assets (*Roa*) a mean of 0.0347, and Tobin's Q (*Tbq*) a mean of 1.9641. The statistical properties of the remaining corporate governance variables are consistent with their intended role as controls in the research design and are broadly aligned with those reported in the existing literature, indicating that the sample selection in this study possesses satisfactory representativeness. Furthermore, the variance inflation factors (*VIF*) for all major variables fall well below the conventional threshold of 10, confirming that the regression model does not suffer from serious multicollinearity.

Table 2. Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Median	Min	Max
<i>Markup</i>	36379	1.1578	0.3772	0.6989	1.0672	3.2521
<i>DA</i>	36379	22.7057	1.1455	20.6230	22.5372	26.2797
<i>Size</i>	36379	22.4134	1.3051	20.0568	22.2127	26.4258
<i>Lev</i>	36379	0.4379	0.2004	0.0615	0.4339	0.8792
<i>Roa</i>	36379	0.0347	0.0579	-0.2052	0.0338	0.1962
<i>Tbq</i>	36379	1.9641	1.2168	0.8297	1.5745	7.9338
<i>Board</i>	36379	2.2792	0.2523	1.6094	2.3026	2.8904
<i>Top1</i>	36379	0.3377	0.1503	0.0809	0.3140	0.7418
<i>Indep</i>	36379	0.3837	0.0741	0.2500	0.3636	0.6000
<i>Age</i>	36379	2.2995	0.7362	0.6931	2.3979	3.4012

4.2. Baseline Regression Results

Table 3 reports the baseline regression results for the impact of data assets on firm price markups. In the specification that includes only the core explanatory variable, the estimated coefficient on data assets is 0.0802, statistically significant at the 1% level. This provides preliminary evidence that data assets exert a positive effect on firm price markups. To further enhance the reliability of parameter estimates and to rule out potential confounding influences of other financial characteristics or governance structures on pricing power, a subsequent specification incorporates a series of control variables including firm size, leverage ratio, return on assets, and corporate governance attributes. The results indicate that, after adequately controlling for potential confounding factors, the estimated coefficient on data assets remains statistically significant at the 1% level. Hypothesis H1 is thus supported by the empirical evidence.

Table 3. Baseline Regression Results

Variable	(1) <i>Markup</i>	(2) <i>Markup</i>
<i>DA</i>	0.0802*** (0.0052)	0.0833*** (0.0107)
<i>Size</i>		-0.0057 (0.0119)
<i>Lev</i>		-0.2387*** (0.0259)
<i>Roa</i>		1.1561*** (0.0558)
<i>Tbq</i>		-0.0200*** (0.0046)
<i>Board</i>		0.0033 (0.0074)
<i>Top1</i>		0.0324 (0.0380)
<i>Indep</i>		-0.0090 (0.0201)
<i>Age</i>		0.0226** (0.0107)
<i>Constant</i>	-0.6629*** (0.1171)	-0.5702*** (0.1363)
<i>Firm FE</i>	YES	YES
<i>Year FE</i>	YES	YES
<i>N</i>	36379	36379
<i>adj. R²</i>	0.7475	0.7743

Note: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Cluster-robust standard errors are reported in parentheses. The same applies to all subsequent tables.

4.3. Robustness Checks

(1) Instrumental Variable Approach. Given that reverse causality and omitted variable bias may inevitably introduce endogeneity concerns into the baseline estimates, this study employs an instrumental variable (*IV*) approach to address potential endogeneity. Specifically, the one-period lagged value of the core explanatory variable is adopted as the instrument for current-period data assets. In terms of relevance, a firm's prior-period data asset investment and accumulation constitutes the foundation for its current stock of data factors, ensuring a strong correlation between the instrument and the endogenous variable. In terms of exogeneity, prior-period data assets represent sunk historical stocks that do not directly intervene in a firm's current pricing decisions for end products, thereby satisfying the exclusion restriction required of a valid instrument.

Columns (1) and (2) of Table 4 report the first-stage and second-stage regression results of the instrumental variable estimation, respectively. The first-stage results in column (1) show that the estimated coefficient on the instrument is 0.2403, positive and statistically significant at the 1% level, confirming that the selected instrument satisfies the relevance condition. Furthermore, the p-value of the underidentification test (Kleibergen-Paap rk LM statistic) is 0.0000, and the weak identification test statistic (Kleibergen-Paap Wald rk F statistic) reaches 711.45, substantially exceeding the critical value of 16.38 established by Stock and Yogo at the 10% maximal IV size threshold. These diagnostics collectively rule out concerns of underidentification and weak instrumentation. The second-stage results reported in column (2) indicate that, after correcting for potential endogeneity bias, the estimated coefficient of data assets on price markups remains positive and statistically significant at the 1% level.

(2) Alternative Clustering of Standard Errors. Building upon the original firm-level clustering, this study elevates the clustering level of standard errors to the industry level to correct for potential downward bias in standard error estimation. The results reported in column (3) of Table 4 show that, after switching the clustering level, the estimated coefficient on data assets remains 0.0833 and statistically significant at the 1% level, further corroborating the baseline regression findings.

(3) Subsample Regression. The public health crisis that emerged in 2020 inflicted severe short-term exogenous shocks on global supply chains, logistics networks, and the normal resumption of business operations among domestic firms. The resulting extreme imbalance between market supply and demand may have prompted firms to passively reduce prices to clear inventories, or alternatively generated abnormal price premiums attributable to material shortages. To eliminate the noise introduced by this black swan event into normal market pricing mechanisms, this study re-estimates the baseline model after excluding observations from 2020. The results reported in column (4) of Table 4 indicate that, after removing the influence of this macroeconomic exogenous shock, the core explanatory variable remains statistically significant, and the main conclusions are not materially altered.

Table 4. Robustness Check Results

Variable	(1) <i>DA</i>	(2) <i>Markup</i>	(3) <i>Markup</i>	(4) <i>Markup</i>
<i>DA</i>		0.1674*** (0.0265)	0.0833*** (0.0203)	0.0849*** (0.0109)
<i>IV</i>	0.2403*** (0.0090)			
<i>Size</i>	0.7021*** (0.0120)	-0.0823*** (0.0251)	-0.0057 (0.0203)	-0.0061 (0.0121)
<i>Lev</i>	0.2245*** (0.0250)	-0.2590*** (0.0307)	-0.2387*** (0.0496)	-0.2404*** (0.0265)
<i>Roa</i>	0.8113*** (0.0400)	1.0798*** (0.0586)	1.1561*** (0.1720)	1.1640*** (0.0606)
<i>Tbq</i>	0.3053*** (0.0051)	-0.0508*** (0.0089)	-0.0200** (0.0075)	-0.0206*** (0.0046)
<i>Board</i>	0.0150* (0.0086)	0.0045 (0.0077)	0.0033 (0.0074)	0.0044 (0.0076)
<i>Top1</i>	-0.0670* (0.0399)	0.0441 (0.0428)	0.0324 (0.0453)	0.0334 (0.0388)
<i>Indep</i>	-0.0054 (0.0229)	-0.0063 (0.0208)	-0.0090 (0.0276)	-0.0100 (0.0213)
<i>Age</i>	0.0044 (0.0132)	0.0385** (0.0163)	0.0226 (0.0229)	0.0227** (0.0109)
<i>Constant</i>	0.7822*** (0.1480)		-0.5702*** (0.1589)	-0.5999*** (0.1371)
<i>Firm FE</i>	YES	YES	YES	YES
<i>Year FE</i>	YES	YES	YES	YES
<i>N</i>	29416	29416	36379	33702
<i>adj. R²</i>	0.9756	0.1288	0.7743	0.7688

5. Further Analysis

5.1. Mechanism Analysis

The baseline regression and robustness checks presented above have established that data assets significantly enhance firm price markups. To further explore the underlying transmission channels, this study draws upon the theoretical analysis developed earlier and conducts mechanism tests along three dimensions: improving capacity utilization, alleviating financing constraints, and reducing agency costs.

(1) Enhancing Capacity Utilization. The core value of data assets lies in breaking down information silos and enabling efficient coordination and precise matching of market supply and demand. Column (1) of Table 5 shows that the estimated coefficient of data assets on capacity utilization (*CU*) is positive and statistically significant at the 1% level. This suggests that the accumulation of data assets effectively empowers both the front-end sales and back-end manufacturing functions of firms, substantially mitigating the “bullwhip effect” in supply chains. Leveraging big data analytics and flexible manufacturing systems, management can dynamically adjust production scheduling in response to real-time fluctuations in market demand, thereby avoiding resource misallocation arising from blind expansion. Such effective activation of idle capacity and maximal dilution of fixed costs solidify the efficiency foundation on the production-supply side that underpins firms’ ability to capture higher price markups.

(2) Alleviating Financing Constraints. Sustained and sufficient financial resources are indispensable for firms seeking to elevate market pricing power through product innovation iteration and brand differentiation. In column (2) of Table 5, the estimated coefficient of data assets on the KZ

index of financing constraints is -0.1343 , negative and statistically significant at the 1% level. This strongly corroborates the signaling and certification effects of data assets. On one hand, the enhancement of digital infrastructure improves corporate financial transparency; on the other hand, substantial data asset reserves convey positive signals to capital markets regarding a firm's robust growth potential. The resulting reduction in information asymmetry significantly lowers external investors' required risk premium, enabling firms to cross the threshold of financing constraints at a lower cost of capital, which in turn facilitates higher price markups.

(3) Reducing Agency Costs. The separation of ownership and control in modern corporations is frequently accompanied by internal agency frictions. This study employs the management expense ratio as a proxy variable for agency costs. The results reported in column (3) of Table 5 show that the estimated coefficient of data assets on the management expense ratio is -0.0068 , highly significant at the 1% level. This finding reveals that data assets effectively compress the rent-seeking space available to management, thereby reducing non-productive internal dissipation of administrative expenses. The resulting improvement in governance efficiency enables firms to channel limited resources precisely toward value-creating activities, which in turn supports a sustained upward shift in product premium-capturing capacity.

Table 5. Mechanism Analysis Results

Variable	(1) <i>CU</i>	(2) <i>KZ</i>	(3) <i>Mfee</i>
<i>DA</i>	0.0206*** (0.0035)	-0.1343*** (0.0453)	-0.0068*** (0.0016)
<i>Size</i>	-0.0359*** (0.0041)	-0.1304** (0.0510)	-0.0056*** (0.0018)
<i>Lev</i>	0.0912*** (0.0092)	6.0751*** (0.1178)	-0.0440*** (0.0049)
<i>Roa</i>	0.4491*** (0.0156)	-11.0520*** (0.2779)	-0.2427*** (0.0093)
<i>Tbq</i>	-0.0036** (0.0014)	0.4900*** (0.0223)	0.0058*** (0.0008)
<i>Board</i>	0.0039 (0.0026)	-0.0348 (0.0422)	0.0021 (0.0014)
<i>Top1</i>	0.0024 (0.0142)	-0.3361* (0.1779)	-0.0059 (0.0065)
<i>Indep</i>	-0.0003 (0.0072)	0.0794 (0.1272)	-0.0077* (0.0041)
<i>Age</i>	0.0090*** (0.0032)	0.5853*** (0.0524)	-0.0069*** (0.0018)
<i>Constant</i>	0.8688*** (0.0545)	2.7025*** (0.6542)	0.3939*** (0.0249)
<i>Firm FE</i>	YES	YES	YES
<i>Year FE</i>	YES	YES	YES
<i>N</i>	36379	36379	36379
<i>adj. R²</i>	0.7269	0.7522	0.7208

5.2. Heterogeneity Analysis

The preceding analysis has established that data assets exert a positive effect on firm price markups, yet it remains unclear whether this effect exhibits asymmetry across different firm types and external environments. This study further explores heterogeneity along three dimensions: ownership structure, industry competition intensity, and the development of external data factor markets.

(1) Ownership Structure Heterogeneity. China's conomy features a pronounced dual ownership structure, with firms of different ownership types diverging markedly in their market-oriented operational imperatives and resource acquisition capabilities. State-owned enterprises (SOEs)

typically enjoy higher market entry barriers and implicit government endorsement, and their product pricing is often constrained by macroeconomic regulation and public interest objectives. In contrast, non-state-owned enterprises confront intense market competition, and their survival and growth are heavily contingent upon genuine market-based profitability. Consequently, these firms possess stronger incentives to deeply exploit the commercial potential of data assets, leveraging precision marketing and incremental product innovation to construct pricing moats. The subsample regression results reported in columns (1) and (2) of Table 6 corroborate this logic: the enhancing effect of data assets on price markups is considerably larger in the non-state-owned enterprise group than in the state-owned enterprise group, indicating that non-state-owned firms capture the pricing dividends enabled by data assets with greater sensitivity and efficiency.

(2) Industry Competition Heterogeneity. The competitive landscape of an industry constitutes a critical external constraint that compels firms to seek digital breakthroughs. This study employs the Herfindahl-Hirschman Index (*HHI*) to characterize industry concentration and divides the sample into two groups: low competition (high *HHI*) and high competition (low *HHI*). In highly competitive industries, severe product homogeneity compresses profit margins to extreme levels, rendering any mechanism capable of achieving differentiated pricing or cost reduction a vital lifeline for firm survival. This intense selection pressure forces firms to deeply integrate data assets into their R&D and sales processes in order to capture pricing power within niche market segments. By contrast, in monopolistic or low-competition industries, firms can effortlessly sustain high markups by virtue of natural market entry barriers, and their urgency to pursue digital transformation is comparatively limited. The results presented in columns (3) and (4) of Table 6 align closely with this theoretical conjecture: the estimated coefficient on data assets is larger in highly competitive industries than in less competitive ones, underscoring the catalytic role of competitive market environments in activating the empowering effects of data assets.

(3) Data Factor Market Development Heterogeneity. As a novel factor of production, the release of data's micro-level commercial value cannot be divorced from the macro-level foundation of factor market infrastructure. Well-developed data factor markets effectively dismantle "data silos" between firms, reduce the acquisition costs of high-quality external data, and generate multiplier effects when a firm's internal data assets interact with the broader external data ecosystem. This study introduces a dummy variable capturing data factor market development to conduct subsample tests. The results in columns (5) and (6) of Table 6 reveal that, in the subsample where firms are located in regions that have initiated data factor market construction, the marginal empowering effect of data assets on price markups is slightly higher than in the group without such development. This finding uncovers a pronounced synergy: the improvement of external data factor market infrastructure effectively smooths macro-level information linkages, amplifying the specialized advantages of firms' internal data assets in demand insight and product pricing.

Table 6. Heterogeneity Analysis Results

Variable	(1) <i>Markup</i>	(2) <i>Markup</i>	(3) <i>Markup</i>	(4) <i>Markup</i>	(5) <i>Markup</i>	(6) <i>Markup</i>
<i>DA</i>	0.0535*** (0.0165)	0.0976*** (0.0121)	0.0562*** (0.0142)	0.1076*** (0.0155)	0.0881*** (0.0171)	0.0810*** (0.0114)
<i>Size</i>	0.0212 (0.0188)	-0.0151 (0.0138)	-0.0018 (0.0162)	-0.0066 (0.0174)	-0.0206 (0.0203)	-0.0036 (0.0126)
<i>Lev</i>	-0.2567*** (0.0564)	-0.2211*** (0.0257)	-0.1481*** (0.0345)	-0.2961*** (0.0349)	-0.2809*** (0.0426)	-0.1740*** (0.0283)
<i>Roa</i>	1.6181*** (0.1342)	1.0020*** (0.0584)	1.2945*** (0.0701)	0.9812*** (0.0826)	0.9067*** (0.0841)	1.2621*** (0.0617)
<i>Tbq</i>	-0.0112 (0.0072)	-0.0220*** (0.0054)	-0.0132* (0.0068)	-0.0242*** (0.0063)	-0.0262*** (0.0088)	-0.0180*** (0.0049)
<i>Board</i>	0.0131 (0.0124)	-0.0057 (0.0093)	0.0072 (0.0100)	-0.0068 (0.0111)	-0.0174 (0.0118)	0.0114 (0.0083)
<i>Top1</i>	0.1249* (0.0654)	0.0280 (0.0461)	-0.0056 (0.0430)	0.0578 (0.0532)	0.0402 (0.0782)	0.0125 (0.0433)
<i>Indep</i>	-0.0572* (0.0308)	0.0202 (0.0266)	0.0303 (0.0259)	-0.0258 (0.0283)	0.0350 (0.0314)	-0.0309 (0.0246)
<i>Age</i>	0.0032 (0.0222)	0.0037 (0.0129)	0.0125 (0.0156)	0.0060 (0.0147)	0.0374** (0.0183)	0.0097 (0.0127)
<i>Constant</i>	-0.5425** (0.2409)	-0.6127*** (0.1745)	-0.1014 (0.1880)	-0.9855*** (0.2090)	-0.2268 (0.2585)	-0.6264*** (0.1499)
<i>Firm FE</i>	YES	YES	YES	YES	YES	YES
<i>Year FE</i>	YES	YES	YES	YES	YES	YES
<i>N</i>	13662	22627	17775	18007	13530	22574
<i>adj. R²</i>	0.7690	0.7868	0.7826	0.8052	0.8152	0.7730

5.3. Economic Consequence Analysis

A natural question arises: does the observed increase in firm-level price markups merely reflect short-term rent extraction through market dominance—a zero-sum appropriation of consumer surplus—or does it genuinely contribute to productive efficiency and long-run value creation? To address this question, we extend the analysis to a core indicator of high-quality firm development, namely total factor productivity (TFP), and examine whether the pricing power driven by data assets translates into substantive efficiency gains. Following Wang et al. (2024)^[28], we specify the following model to investigate this economic consequence:

$$ESG_{it} = \theta_0 + \theta_1 \gamma CU_{it} + \theta_h \sum Control_{it} + YearFE + IndFE + \varepsilon_{it} \quad (3)$$

As reported in Table 7, the estimated coefficients of data asset-driven price markups on contemporaneous TFP and one-period-ahead TFP are 0.2571 and 0.2354, respectively, both statistically significant at the 1% level. This finding confirms that the markup-enhancing effect of data assets extends well beyond short-term profit improvement, and instead serves as a durable source of long-run competitive advantage. By generating above-normal profits through data-enabled capabilities, firms substantially alleviate the financial constraints that would otherwise impede R&D investment and capital upgrading. These surplus resources are subsequently channeled into flexible production line restructuring, core technology advancement, and managerial process reengineering, ultimately yielding a pronounced improvement in firm-level total factor productivity and fostering high-quality development.

Table 7. Economic Consequence Analysis Results

Variable	(1) TFP	(2) F.TFP
∇DA	0.2571*** (0.0221)	0.2354*** (0.0253)
Size	0.3993*** (0.0223)	0.3386*** (0.0251)
Lev	0.4857*** (0.0468)	0.5338*** (0.0532)
Roa	2.1814*** (0.0786)	1.5619*** (0.0779)
Tbq	-0.0562*** (0.0074)	-0.0190** (0.0083)
Board	0.0175 (0.0134)	0.0207 (0.0155)
Top1	0.0090 (0.0742)	0.0842 (0.0855)
Indep	-0.0164 (0.0370)	0.0123 (0.0438)
Age	0.0147 (0.0167)	-0.0025 (0.0202)
Constant	-4.9616*** (0.2907)	-3.1954*** (0.3369)
Firm FE	YES	YES
Year FE	YES	YES
N	35613	28813
adj. R2	0.9273	0.9189

6. Conclusions

Drawing on data from Chinese A-share listed firms over the period 2007-2024, this paper systematically examines the effect of data assets on firm-level price markups, the underlying transmission mechanisms, and the heterogeneous patterns across different subsamples. The main conclusions are as follows:

First, data assets significantly enhance firm-level price markups. Across both baseline fixed-effects specifications and instrumental variable estimations, the coefficient on data assets remains robustly positive and economically meaningful. This conclusion continues to hold after a battery of robustness checks, including alternative clustering of standard errors, exclusion of the pandemic year, and substitution of markup measurement approaches, indicating that data assets constitute a substantive source of market power for firms.

Second, data assets promote firm-level price markups through three distinct channels: improving capacity utilization, alleviating financing constraints, and reducing agency costs. The capacity utilization channel operates primarily from the cost side—data assets optimize production scheduling and resource allocation, effectively minimizing idle capacity and unplanned downtime, thereby lowering unit production costs. The financing constraint channel functions mainly from the investment side—data assets improve the external financing environment faced by firms, enabling them to secure the capital required for sustained investment at lower costs. The agency cost channel works by compressing managerial rent-seeking space, reducing nonproductive overhead embedded in administrative expenses. These three pathways complement one another and collectively constitute the micro-level mechanism through which data assets create value.

Third, the effects of data assets exhibit significant heterogeneity across firm ownership types, industry competitive structures, and the degree of external data factor market development. Non-

state-owned enterprises, subject to stronger market-based incentives and equipped with more agile decision-making mechanisms, are able to more fully convert data assets into pricing power. Highly competitive industries, where profit margins are severely compressed, create intense survival pressure that compels firms to deeply embed data assets into operations, thereby magnifying the markup-enhancing effect. In addition, well-established regional data factor market infrastructure generates synergies that further amplify the specialized advantages of firms' internal data assets in demand sensing and product pricing.

Fourth, the markup enhancement driven by data assets carries far-reaching economic consequences. Empirical analysis demonstrates that the data asset-driven increase in price markups is not confined to short-term profit gains; rather, it translates into a pronounced improvement in total factor productivity (TFP). This finding further corroborates the long-run role of data assets in facilitating high-quality firm development.

The findings of this study carry several policy and managerial implications. For firms, the accumulation and development of data assets should be incorporated into long-term competitive strategic planning. Beyond data collection and storage, particular emphasis ought to be placed on the deep application of data across production optimization, demand forecasting, and product innovation, thereby leveraging data-enabled capabilities to achieve comprehensive improvements spanning both cost advantages and differentiation advantages. For policymakers, advancing institutional frameworks for data property rights, valuation standards, and circulation mechanisms can lower the barriers and transaction costs associated with data asset conversion, thereby unlocking the full economic value of data as a factor of production. At the same time, close attention should be paid to the potential rise in market concentration that may accompany data asset accumulation; efforts to facilitate the efficient flow of data factors must be accompanied by safeguards to preserve a fair and competitive market environment.

This study is subject to several limitations. First, given that data asset disclosure remains in its infancy, the identification and measurement of data resource-related items in corporate annual reports may be affected by noise and potential underestimation. As accounting standards are implemented and disclosure practices mature, the precision of data asset measurement is expected to improve. Second, while this paper examines three transmission channels—capacity utilization, financing constraints, and agency costs—additional pathways through which data assets may influence pricing power, such as R&D innovation, supply chain bargaining power, and customer stickiness, remain unexplored due to constraints of scope and variable availability. Third, the sample is confined to Chinese A-share listed firms, and the external validity of these findings for small and medium-sized enterprises or for institutional contexts in other countries warrants further investigation. Future research may exploit the quasi-natural experiment created by the formal implementation of accounting standards for enterprise data resources beginning in 2024, constructing cleaner causal identification strategies based on pre- and post-policy variation. In addition, scholars may extend the analysis to spillover effects along industrial supply chains and to emerging issues surrounding cross-border data flows.

References

- [1] J. De Loecker, F. Warzynski, Markups and firm-level export status, *Am. Econ. Rev.* 102 (2012) 2437-2471.
- [2] X. Yu, B. Niu, Z.M. Yuan, Data assets, human capital upgrading, and firm value, *J. Zhongnan Univ. Econ. Law* (2) (2024) 109-122.
- [3] X.C. Xu, Z.W. Zhang, Y.R. Hu, Research on data asset statistics and accounting, *Manage. World* 38 (2) (2022) 16-30.
- [4] J. Li, X.F. Dong, J.L. Zhang, et al., The impact of data assets on corporate innovation investment, *Foreign Econ. Manage.* 45 (12) (2023) 18-33.

- [5] Y. Wang, D. Yang, The transformation of Chinese management accounting system: From data factors to data assets, *Manage. World* 40 (10) (2024) 171-189.
- [6] Z. Lu, T. Zhou, L. Wang, et al., Data assets and firm development: Empirical evidence from Chinese listed companies, *Ind. Econ. Res.* (4) (2023) 128-142.
- [7] D. Sheng, Z.Q. Liu, Exchange rate movements, processing trade, and markups of Chinese firms, *J. World Econ.* 40 (1) (2017) 3-24.
- [8] R.L. Zhao, C.R. Sun, Y.B. Chen, Minimum wages and firm price markups, *J. World Econ.* 41 (2) (2018) 121-144.
- [9] L. Wang, Q.F. Wu, D. Luo, Market barriers, administrative approval, and firm price markups, *China Ind. Econ.* (6) (2020) 100-117.
- [10] J.E. Bessen, C. Righi, Identifying the impact of software on firm performance, Boston University School of Law, Law and Economics Research Paper, 2019.
- [11] G.H. Jiang, Financing constraints and the export mode choice of Chinese firms, *Finance Trade Econ.* (5) (2016) 106-118.
- [12] S.J. Zhu, P.H. Zhang, Domestic technology content of China's manufacturing exports and its determinants, *Stat. Res.* 30 (6) (2013) 58-66.
- [13] X.P. Li, K. Chen, The impact of market segmentation on firm price markups: Evidence from Chinese manufacturing firms, *Econ. Manage.* 40 (11) (2018) 140-156.
- [14] M.C. Jensen, W.H. Meckling, Theory of the firm: Managerial behavior, agency costs and ownership structure, *J. Financ. Econ.* 3 (4) (1976) 305-360.
- [15] P.W. Bai, L. Yu, Digital economy development and firm price markups: Theoretical mechanisms and empirical evidence, *China Ind. Econ.* (11) (2021) 59-77.
- [16] L.H. Cheng, T.Y. Sun, The impact of service inputs in manufacturing on firm price markups: A global value chain perspective, *Reform* (3) (2020) 111-125.
- [17] D.F. Yu, Y. Zhang, Q. Song, Digital human capital inputs and firm price markups: Evidence from recruitment big data, *China Ind. Econ.* (7) (2025) 62-80.
- [18] E. Brynjolfsson, K. McElheran, The rapid adoption of data-driven decision-making, *Am. Econ. Rev.* 106 (5) (2016) 133-139.
- [19] H.R. Varian, Artificial intelligence, economics, and industrial organization, NBER Working Paper No. 24839, 2018.
- [20] A. Goldfarb, C. Tucker, Digital economics, *J. Econ. Lit.* 57 (1) (2019) 3-43.
- [21] D. Autor, D. Dorn, L.F. Katz, C. Patterson, J. Van Reenen, The fall of the labor share and the rise of superstar firms, *Q. J. Econ.* 135 (2) (2020) 645-709.
- [22] R.E. Hall, New evidence on the markup of prices over marginal costs and the role of mega-firms in the US economy, NBER Working Paper No. 24574, 2018.
- [23] N. Bloom, R. Sadun, J. Van Reenen, The diffusion of IT and productivity, *Am. Econ. Rev.* 102 (3) (2014) 167-170.
- [24] M.E. Porter, *Competitive Advantage: Creating and Sustaining Superior Performance*, Free Press, New York, 1985.
- [25] D. Acemoglu, P. Restrepo, The race between man and machine: Implications of technology for growth, factor shares, and employment, *Am. Econ. Rev.* 108 (6) (2018) 1488-1542.
- [26] J.E. Stiglitz, A. Weiss, Credit rationing in markets with imperfect information, *Am. Econ. Rev.* 71 (3) (1981) 393-410.
- [27] T. Jiang, Mediating and moderating effects in empirical studies of causal inference, *China Ind. Econ.* (5) (2022) 100-120.
- [28] M.B. Wang, T. Ye, D.M. Kong, Green manufacturing and corporate environmental information disclosure: A policy experiment based on China's green factory initiative, *Econ. Res. J.* 59 (2) (2024) 116-134.