The Impact of E-Commerce on Physical Retail: An Empirical Study of Beijing, Wuhan, and Hangzhou

Tianqi Zheng
University of Edinburgh (BSC Applied Mathematics), Old College, South Bridge, Edinburgh EH8 9YL, United Kingdom

Abstract: With China's fast digital transformation, e-commerce has grown very quickly and is bringing more and more structural challenges to traditional industries. This paper uses data from three representative cities between 2018 and 2024 to study the impact of e-commerce on local physical industries. The study uses local statistical reports and yearbooks to build a regression model. The empirical results show that the regression coefficient of e-commerce share is significantly negative at the 1% level. This means that for every 10% increase in e-commerce penetration, physical retail sales drop by about 3.7%, especially the effect is even stronger in cities with a more developed digital economy. This provides empirical evidence for understanding the relationship between e-commerce development and the urban consumption structure, and also offers some policy insights on how to promote the integration of online and offline retail.

Keywords: E-Commerce; Physical Retail; Total Retail Sales of Consumer Goods

1. Introduction

After the millennium, e-commerce in China experienced rapid growth and was widely seen as reshaping the local consumption landscape. During this process, the conflict between e-commerce and local physical industries has become increasingly clear. However, there is still a lack of systematic and quantitative evidence at the city level to understand this issue. This paper takes this as an opportunity and uses city-level data from 2018 to 2024. Three representative Chinese cities are selected: Beijing, Hangzhou, and Wuhan. These cities represent the political, cultural, and economic center of China, a super first-tier city, and a first-tier city, respectively. The study tests whether the development of e-commerce has had a significant impact on local physical retail industries. The annual sales of physical industries are set as the dependent variable, while e-commerce sales are used as the main independent variable. GDP, population, household income, and other macro factors are included as control variables. A multiple regression model is built, and consistency tests are then conducted based on this model. The core conflict is whether, and to what extent, the city's physical economy is negatively affected by the rapid growth of the digital economy.

2. Literature Review

International research usually looks at the relationship between online and offline channels. The main question is whether e-commerce is a substitute or complement. Studies show when retailers add an online channel, it often reduces store sales a lot. Only when the channel works together better, the overall performance maybe get better. Also, some field experiments with causal methods show that the total effect of online and offline is decided by things like customer travel cost and how attractive the store is. In the domestic market, data from the National Bureau of Statistics show that in 2022, total online retail sales were about 13.8 trillion RMB. [9] The growth of online sales is much faster than total retail sales, and it keeps growing in 2023, [10] showing that the move of consumers to digital channels is still going on. From an academic view, most domestic studies use panel data at the province level or higher. They study the relationship between e-commerce development and things like consumption upgrade, income, or employment. These studies show that e-commerce has a clear impact on household spending and the regional economy, but research at the city level is still quite limited. Also, many samples stop around 2019, so they do not fully cover the structural changes caused by the pandemic. The comparison of how prosperous different cities are is also not very complete. To address these gaps, this paper uses

a multiple log-linear regression model with data from 2018 to 2024 for three representative cities. It tests the substitution effect of e-commerce on physical retail and the differences between cities, adding more evidence to the study of how e-commerce impacts physical retail in China.

3. Research Method

In theory, the lower marginal cost and higher product variety of e-commerce, together with China's well-developed logistics network and big data search, can create a certain level of substitution for the offline physical economy. But it is also possible that, through targeted promotion using big data and supply expansion from other places, e-commerce may lead to a "net growth in demand" for the physical economy. Without clear data to support it, there is a lot of uncertainty. Therefore, after collecting the data, we use regression to test two hypotheses:

After controlling for macro factors, the coefficient of e-commerce sales is negative. If this is true, it means e-commerce development is strongly negatively related to physical retail sales.

This negative effect is stronger in cities where the e-commerce industry is more active, has wider coverage, the economy is more developed, and consumers have higher spending power.

Based on the above research background and significance, this paper uses panel data from major Chinese cities to conduct an empirical analysis, testing the impact of e-commerce development on the local physical economy. The main data come from the National Bureau of Statistics and the statistical yearbooks of each province and city. Other parts come from the official annual reports of different e-commerce platforms. The model and its variables are set as follows. The dependent variable is the development level of the physical economy, measured by the physical retail part of the total retail sales of consumer goods. This is used to show the size of the physical consumption market. The main independent variable is the level of e-commerce development, measured by the share of e-commerce sales in the total retail sales of consumer goods. Specifically, "e-commerce sales share" is defined as the share of online retail sales of physical goods by above-designated-size units in the total retail sales of consumer goods, following the National Bureau of Statistics statistical caliber. [8][10] For example, statistics show that in 2020, national online retail sales reached 11.76 trillion RMB, of which 9.76 trillion RMB were from online retail, about one-fourth of the total retail sales. [9] This share is then used for analysis. For the control variables, several indicators of social and economic characteristics of the cities are included. This helps to control for differences between cities. We estimate a two-way fixed-effects panel model (city and year fixed effects) to control for time-invariant heterogeneity across cities and common time shocks. The model formula is:

$$Y= \alpha+\beta X+\gamma Z+\varepsilon \quad (1)$$

Here, Y is the dependent variable, which represents the development level of the physical economy. X is the share of e-commerce development. Z is a vector of control variables. α is the constant term, meaning the expected level of the physical economy in a city if there were no e-commerce. β and γ are the estimated coefficients, and ε is the random error term. All specifications include city and year fixed effects as stated above.

4. Empirical Analysis

The analysis shows that between 2018 and 2024, Beijing, Hangzhou, and Wuhan all show a similar trend in their consumption structure. E-commerce expanded quickly, while the share of physical retail kept decreasing. Looking at the data in detail, in Beijing the share of online retail in total retail sales rose from less than 20% to nearly 40% around 2022. [1][2] This shows the continuous deepening of online growth in major Chinese cities. In Hangzhou, online retail sales passed the total local retail sales as early as 2019, [3] showing the strongest e-commerce dominance. From then until the end of the analysis in 2024, it stayed at a high level and is expected to keep this trend in the future. [4][5] In Wuhan, after the impact of the pandemic, the share of online retail stayed stable at around 30%. [6][7] At the same time, physical retail sales in the three cities showed a small decline in 2020 due to the uncontrollable impact of the pandemic. [2] Although there was some recovery later, the speed of recovery moved in the opposite direction to the deepening local penetration of e-commerce, showing a clear trade-off pattern. Beijing and Wuhan show the pattern of "total sales growing, but the online share keeps rising." Hangzhou, on the other hand, shows a trend closer to "total sales growing

faster, with the structure shifting to online even more quickly." These descriptive facts provide a clear prior for the following regression analysis. The coefficient for physical retail sales is expected to be negative. If this negative relationship is stronger in cities like Hangzhou, where e-commerce is highly concentrated, it would further confirm that the effect comes from structural substitution rather than just general macroeconomic fluctuations, as shown in Table 1.

Table 1. Regression Results for the Effect of E-Commerce Share on Physical Retail

Variable	Coefficient estimates	Standard error	T-value	Significance level
Log(percentage)	-0.371	0.029	-12.6	p<0.01
Log(GDP)	-0.305	0.093	-3.264	p<0.01
Log(Population)	0.368	0.111	3.319	p>0.01
Log(Income)	0.034	0.032	1.043	p<0.01
α	0.870	0.316	2.756	p<0.01
R^2	0.977			

5. Results

The regression results show that, after controlling for GDP, population, and income factors, the growth of e-commerce is strongly and significantly negatively related to the share of physical retail in the economy. The log analysis result of -0.371 shows an elastic relationship between the two. For every 10% increase in the share of e-commerce sales, the share of physical retail decreases by about 3.71%. This finding is consistent with the earlier expectations: The share of online retail in the total retail sales of consumer goods has risen quickly. For example, in 2019, the national share of online retail for physical goods reached 20.7%. [8] In more developed cities like Beijing, the share was as high as 27.4% in the same year, while the share of offline physical retail dropped to around 70% of its previous level. Similarly, in 2022, the national share of online retail for physical goods rose further to 27.2%. [9] E-commerce has become an important force driving consumption growth and upgrading. Therefore, overall, the results confirm the original estimation. The development of e-commerce has a clear negative impact on the level of physical retail, and this effect is significant at the 1% level. For the coefficient γ of the control variables, the regression shows some meaningful results. The coefficient of GDP is -0.305 and significantly negative at the 1% level. This means that, with the same e-commerce share, cities with larger economic size tend to have a lower share of physical retail. This may be because, in a more developed economy with ongoing consumption upgrading, consumers are more likely to choose new forms of online consumption. The coefficient of permanent population is +0.368, showing that cities with larger populations have a slightly higher share of physical retail. This may be because big cities have more dense physical retail networks, so a large part of consumer demand can still be met offline, keeping a certain share even under the impact of e-commerce. Lastly, The coefficient of per capita disposable income is positive but not significant, with a p-value around 0.30. This may be because income and GDP are highly correlated, making it hard to separate its independent effect from that of GDP. In this model, the marginal effect of per capita income on the share of physical retail does not pass the significance test.

The model's goodness of fit is 0.98, and the adjusted value is 0.97. This shows that the selected variables can explain about 98% of the variation in the share of physical retail. This reflects that the model fits well within the sample. Next, we evaluate it through residual analysis and graphical checks. As shown in Figure 1 (QQ plot of regression residuals), the graph compares the theoretical normal quantiles with the actual residual quantiles. From the QQ plot, we can see that the data points mostly fall near the diagonal line, which means that after the log transformation, the data roughly follow a normal distribution. There is some slight deviation at the tails, which may indicate mild skewness or a few outliers, but this is within a normal range. Therefore, overall, taking the log of the Y variable for linear regression is reasonable.

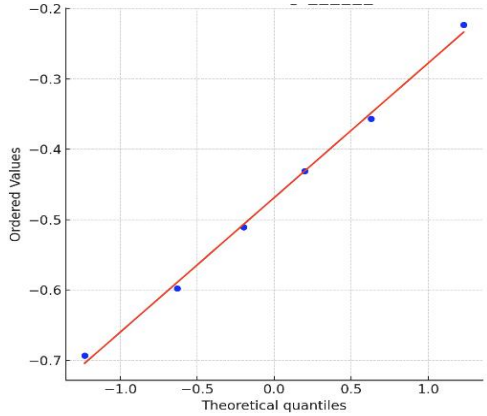

Figure 1. Normal Q–Q Plot of Regression Residuals.

6. Conclusion

In this paper, annual data from 2018 to 2024 for Beijing, Hangzhou, and Wuhan are used as the sample. Focusing on the core question of whether e-commerce development affects the physical economy, a log-linear multiple regression model is built for empirical testing. The results are consistent with the observed trends and support the research hypothesis of a "structural substitution effect" of online retail on offline retail. From the city comparison, the substitution effect is more evident in cities with a stronger e-commerce foundation, such as Hangzhou and Beijing. This suggests that the intensity of the impact varies across different city levels.

After testing with clustered robust standard errors and residual diagnostics (including the QQ plot), the model shows reliable performance in both fitting and inference.

The main contributions of this study are as follows:

First, it applies the international framework of "online–offline substitution/complementarity" to comparable data at the prefecture-level city scale in China, filling the gap left by previous studies that mostly focused on national or provincial levels. Second, it covers a new sample period that fully includes the pandemic years, showing the latest picture of rising e-commerce penetration and the redistribution of offline retail share. Based on this, the policy implications of this paper mainly include: promoting the digital transformation of physical retail, optimizing the layout of commercial outlets and last-mile logistics, and encouraging data collaboration between online platforms and offline business districts, so as to move from a "zero-sum substitution" pattern toward "mutual growth."

However, this study has some limitations. The sample is limited to three cities, and some years have missing data, so caution is needed when generalizing the conclusions. In addition, city-level e-commerce indicators may differ, such as "online retail sales above a designated size" versus "total online retail sales," which could lead to measurement errors. Future research can be expanded in three directions.

First, include more city samples and introduce city-time fixed effects to improve identification strength.

Second, combine major events, such as policy changes or shifts in consumer behavior, and use methods like difference-in-differences or regression discontinuity for better causal inference.

Third, collect micro-level store data or conduct field studies to enhance the theoretical credibility of the findings.

Overall, this paper provides quantitative evidence at the city level: during the study period, e-commerce penetration significantly squeezed the share of physical retail. This structural change deserves continued tracking with richer data and more advanced methods.

References

[1] Beijing Municipal Bureau of Statistics. Statistical Bulletin on Beijing's National Economic and Social Development in 2018 [EB/OL]. March 20, 2019. Accessed: August 19, 2025.

[2] Beijing Municipal Bureau of Statistics. Beijing's Total Retail Sales of Consumer Goods Declined by 8.9% in 2020 (Including Online Retail Sales of 442.33 Billion RMB) [EB/OL]. January 20, 2021. Accessed: August 19, 2025.

[3] Hangzhou Municipal Government. Report on the Implementation of Hangzhou's National Economic and Social Development Plan for 2019 and the Draft Plan for 2020 (Excerpt: Total Retail Sales 621.5 Billion RMB, Online Retail Sales 727.77 Billion RMB) [EB/OL]. May 29, 2020. Accessed: August 19, 2025.

[4] Hangzhou Municipal Government. Report on the Implementation of Hangzhou's National Economic and Social Development Plan for 2021 and the Draft Plan for 2022 (Excerpt: Total Retail Sales 674.4 Billion RMB, Online Retail Sales 995.15 Billion RMB) [EB/OL]. April 12, 2022. Accessed: August 19, 2025.

[5] Hangzhou Municipal Bureau of Commerce / Hangzhou Daily. Total Retail Sales of Consumer Goods in Hangzhou Reached 729.4 Billion RMB in 2022 [EB/OL]. February 15, 2023. Accessed: August 19, 2025.

[6] Hubei Provincial Government Offices Administration (Reprint from Wuhan Municipal Bureau of Statistics). Statistical Bulletin on Wuhan's National Economic and Social Development in 2018 (Excerpt: Total Retail Sales 684.39 Billion RMB) [EB/OL]. July 12, 2019. Accessed: August 19, 2025.

[7] Wuhan Municipal Bureau of Statistics. Statistical Bulletin on Wuhan's National Economic and Social Development in 2019 (Excerpt: Total Retail Sales 744.964 Billion RMB; Online Retail Sales above Designated Size 75.634 Billion RMB) [EB/OL]. March 29, 2020. Accessed: August 19, 2025.

[8] National Bureau of Statistics of China. Total Retail Sales of Consumer Goods Increased by 8.0% in 2019 (Among Them: Online Retail Sales of Physical Goods 8.5239 Trillion RMB, Share 20.7%) [EB/OL]. January 17, 2020. Accessed: August 19, 2025.

[9] State Council Client / National Bureau of Statistics of China. China's National Economy Withstood Pressure and Reached New Heights in 2022 (Among Them: Total Retail Sales 43.9733 Trillion RMB; Online Retail Sales of Physical Goods 11.9642 Trillion RMB, Share 27.2%) [EB/OL]. January 17, 2023. Accessed: August 19, 2025.

[10] National Bureau of Statistics of China. Statistical Bulletin on the National Economic and Social Development of the People's Republic of China in 2023 (Among Them: Online Retail Sales of Physical Goods 13.0174 Trillion RMB, Share 27.6%) [EB/OL]. February 29, 2024. Accessed: August 19, 2025.

