RIETI Discussion Paper Series 23-E-044

Impact of E-commerce Development on Income Inequality: Evidence from rural China based on cross-county panel data

MA, Xinxin
Hosei University

KOMATSU, Sho
Asian Growth Research Institute

RIETI
Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 23-E-044

June 2023

Impact of E-commerce Development on Income Inequality: Evidence from rural China based on cross-county panel data

Xinxin Ma (Faculty of Economics, Hosei University)*

Sho Komatsu (Asian Growth Research Institute)

Abstract

Information and communications technology (ICT) is rapidly developing worldwide. Some studies argue that ICT increases income inequality in developed countries; however, evidence on the relationship between progress in ICT and income inequality in developing countries is scarce. Using an original cross-county panel data from 2011 to 2018, we investigated the impact of e-commerce development on income inequality in rural China while considering endogeneity issues. We found that the effect of e-commerce on income inequality differed by region: e-commerce development could expand income inequality in developed counties, while reducing it in less-developed ones; the total effect of e-commerce on the income inequality was insignificant. Additionally, this effect was greater in counties with a higher level of agricultural modernization. Furthermore, the decomposition results indicated that differences in e-commerce accessibility and income return of e-commerce usage contributed to widening the income inequality between developed and less-developed rural counties.

Keywords: e-commerce; rural areas; income inequality; Taobao Village; cross-county panel

JEL classification: C23, D31, L81, O14

* This research was supported by the Japan Society for the Promotion of Science (JSPS KAKENHI, Grant Numbers 20H01512). It was conducted as the Chinese economy part (Title: Empirical Research on the Changing Chinese Economy Upgrading, Expansion, Structural Reform) of the Project "Studies on Transformations of International Systems and Their Impact on Japan's Mid- & Long-term Competitiveness" undertaken at the Research Institute of Economy, Trade and Industry (RIETI). The author is grateful for the helpful comments and suggestions from Masayuki MORIKAWA (RIETI), Shujiro URATA (RIETI), Masato MIZUNO (RIETI), Asei ITO (University of Tokyo), Kai KAJITANI (Kobe University), Mariko WATANABE (Gakushuin University), Zhixiong GUAN (Nomura Institute for Capital Market Research), and the Discussion Paper seminar participants at RIETI.

1. Introduction

Information and communications technology (ICT) is rapidly developing worldwide. Numerous empirical studies have found that ICT may affect individuals' behaviors and outcomes, such as educational performance (Caldarulo et al., 2023), quality of life (Valentín-Sívico et al., 2023), employment (Luo et al., 2022), productivity (LoPiccalo, 2022), and income levels (Ma, 2022).

Furthermore, some studies have argued that ICT has contributed to the growth of income inequality in developed countries since the 1970s through what is known as skill-biased technology change (Balcilar et al., 2021; Card & Lemieux, 2001; Krueger, 1993). This refers to the phenomenon in which advancements in technology disproportionately increase the demand for skilled labor, leading to an income gap between skilled and unskilled laborers. On the one hand, Ghosh (2020), Lloyd-Ellis (1999), Pradhan et al. (2016), and Philip et al. (2017) hold that Internet usage has increased productivity and the income of rural residents in less-developed regions, thus reducing income inequality. By contrast, Dimaggio and Bonikowski (2008) and Furuholt and Kristiansen (2007) found the opposite effect in developed countries. However, direct evidence on the relationship between ICT progress and income inequality in developing countries is scarce.

Electronic Commerce (herein after e-commerce) is a modern business model that involves the trading of goods or services using ICT such as the Internet and computer technology. Since the early 2000s, e-commerce has experienced rapid global growth. However, despite numerous studies on the impact of ICT on income levels or income inequality (You and Zhang, 2018; Zeng, et al., 2018; Li et al., 2019; Chen, 2020; Li et al., 2021; Ma, 2022, etc.), there is limited evidence specifically focusing on e-commerce (Zeng et al., 2017; Li et al. 2021). Therefore, this study aims to investigate the effects of e-commerce development on income inequality in rural areas of China. China, being a developing country that has witnessed significant e-commerce growth and an increase in income inequality in recent years (Li et al., 2008; Sicular et al., 2020), serves as an ideal context for this research. By examining the relationship between e-commerce and income inequality, this study aims to broaden the scope of research on ICT and provide a deeper understanding of how ICT-driven e-commerce business models can shape economic activities.

Regarding the issue of income inequality in China, the Chinese government has undertaken market-oriented reforms, leading to remarkable economic growth since 1978. However, with the progress of economic transitions, income inequality has expanded since the 1990s. According to the National Statistics Bureau, the Gini coefficients for 2003–2021 ranged between 0.491 and 0.473. Studies investigating the determinants of income inequality in China have reported that the income inequality between rural and urban areas contributes to nationwide income inequality (Li et al., 2008; Sicular et al., 2020).

To reduce this rural-urban inequality, the policy agenda of "Alleviating Poverty through E-Commerce" has been featured annually in China's No. 1 Central Document since 2014 (Couture et al., 2018). In addition, the government announced the expansion of e-commerce in rural areas as a priority in its national policy. Alibaba Group, China's largest e-commerce company, launched its "Rural Taobao Program" in 2014. Taobao is an online trading platform founded by Alibaba that provides rural residents with better access to the Internet and helps them earn more by selling agricultural products directly to urban consumers via online platforms. The number of Taobao Villages[1] has been increasing annually since 2013, which has accelerated the rapid development of e-commerce and has played a crucial role in rural China.

Empirical evidence on the association between ICT (including e-commerce) and the income inequality in rural China is scarce. Qi et al. (2019) and Zeng et al. (2017) reported that e-commerce may increase rural farmers' income levels in some eastern coastal regions. Leng et al. (2020) found that ICT adoption had a positive

[1] The identification criteria of Taobao Village by Ali Research and Alibaba Group's research unit include the following: (1) trading place—located in a rural area—with the administrative village as a unit, (2) trading volume: the annual trading volume of e-commerce amounts to over RMB10 million, and (3) scale of online merchants: active online stores amount to over 100 or account for over 10% of local households.

and significant effect on income diversification among rural households in China. The studies by Li et al. (2021) and Zeng et al. (2018) are closely related to this study; however, certain limitations must be addressed.

Our contributions to the literature are as follows: First, we constructed original cross-county-level panel data for five provinces to investigate the impact of e-commerce development on the income inequality in rural areas.[2] This distinguishes our study from previous works (e.g., Li et al., 2021; Zeng et al., 2018) that only used cross-province-level panel data. In China, county-level data comprise the smallest unit of available regional data from the Provincial Statistical Yearbook published by the Provincial Bureau of Statistics. Therefore, using small-unit regional data is preferable, as it may considerably reduce the regional aggregation data bias compared to the large-unit regional data (province-level data) used in previous studies.

Second, to the best of our knowledge, this study is the first to focus on issues including developed and less-developed regions in rural areas. Only two studies (Li et al., 2021; Zeng et al., 2018) focused on this issue. However, they used data on e-commerce in developed regions (Zhejiang and Jiangsu provinces in the eastern region with high income levels). Conversely, we used data that included counties in both developed (Zhejiang and Jiangsu) and less-developed (Henan, Hunan, and Ningxia) provinces. We also compared the differences in the effects of e-commerce on these regions.

Third, we broke down the effects of e-commerce on the income inequality between developed and less-developed counties into two components: the endowment effect (i.e., the difference in access to e-commerce) and the price effect (i.e., the difference in the magnitude of e-commerce's effect on income levels), based on the Blinder-Oaxaca (B-O) decomposition approach (Blinder, 1973; Oaxaca, 1973). This study is the first to calculate the contribution rates of these two

[2] The utilization of cross-county data in this study is in line with the World Bank & Alibaba (2019).

components, which may help us understand the impact of e-commerce development on the income inequality in rural areas. Additionally, because the policy implications differ based on these two components, the results may provide rich empirical evidence for policymakers.

Fourth, we tackled endogeneity problems that have not been adequately addressed in previous studies using the generalized difference-in-differences (GDD) and fixed effect (FE) models. We also applied the fixed effects instrumental variable (FE_IV) method as a robustness check.

Finally, the effects of e-commerce on the income inequality in rural areas may differ across heterogeneous groups. Therefore, we compared the differences in the effects of e-commerce based on the county's agricultural technology and education levels. This was done for the first time, and the results will enrich the available evidence on this issue.

We found that e-commerce development in rural areas tends to expand the income inequality, but to an insignificant degree. The effects differ more by region: e-commerce development may increase the income inequality in developed counties, while reducing the income inequality in less-developed ones. The effect of e-commerce on income levels was greater in counties with higher levels of agricultural technology. Furthermore, the decomposition results indicated that differences in e-commerce accessibility and income return of e-commerce usage contributed to widening the income inequality between developed and less-developed rural counties in rural areas.

2. Literature review

2.1 Empirical studies on the relationship between e-commerce development and income inequality

We will summarize the results related to this issue according to the data sources. First, using individual/household survey data, Chen (2020), Li et al. (2019), and You and Zhang (2018) found that ICT usage may increase rural income. Using panel data from 31 provinces from 2002 to 2013, Zhang and Han (2017) found that e-commerce development may increase the income levels of urban and rural residents. Additionally, using survey data from rural areas in the Gansu, Henna, and Shandong provinces, Ma et al. (2018) found that ICT (smartphone) usage may increase farm and nonfarm income. Finally, using data from the China Labor Force Dynamics Survey of 2016, Leng et al. (2020) investigated the association between ICT and the income diversification of rural residents. They found that ICT adoption may increase income diversification among rural Chinese households.

Second, several studies used regional (i.e., cross-province) data, but the empirical results were mixed. For instance, Zhang and Han (2017) found that e-commerce development expanded urban income inequality. Using panel data from 31 provinces from 2015 to 2019, He (2020) found that rural e-commerce development might expand income inequality. Using panel data from 28 provinces from 2012 to 2016, Zhang (2019) reported that Internet trade contributed to the expansion of income inequality. On the other hand, Li et al. (2021) found that e-commerce and rural income inequality have an inverted U-shaped relationship. Furthermore, using longitudinal household data from the China Family Panel Studies of 2014–2018 and the decomposition method, Ma (2022) found that differences in Internet accessibility might expand income inequality. In contrast, the difference in the magnitude of the effect of Internet usage on income level (income return on Internet usage) may reduce it.

To the best of our knowledge, no study has focused on the impact of e-commerce development on the income inequality in rural areas, including in developed and less-developed counties. In addition, no study has investigated how differences in accessibility and the magnitude of the effect of e-commerce on income levels affect the income inequality between developed and less-developed counties. This study is unique in that it provides new evidence on these issues.

2.2 Two channels explaining the association between e-commerce development and income inequality

The two components considered to affect the impact of e-commerce development on income inequality are as follows.

First is the endowment difference, which is the difference in the number of e-commerce platforms among regions, may contribute to income inequality. E-commerce is a new business model using human capital that positively affects economic growth. This reduces transaction costs, helps farmers obtain more business information, and sells more agricultural produce. This may create new opportunities to help farmers transition to nonfarm work. This may raise the income levels of individuals or households. Therefore, when the number of e-commerce platforms (Taobao Village) is higher in developed regions than in less-developed ones, there may be a gap in income levels due to differences in e-commerce accessibility. This is called the "endowment effect."

Second, it is assumed that the proportion of highly educated people in the population is higher. Hence, the spillover effects of new technology or the e-commerce usage skill levels are greater in developed regions than in less-developed ones, which may lead to a greater income increase in developed regions. We call this the "price effect" (i.e., the difference in the magnitude of e-commerce's effect on income level or the income premium of e-commerce). Therefore, the difference in the price effect of e-commerce also contributes to the formation of an income inequality.

The policy implications differ across these two channels. For example, the policy of e-commerce expansion in less-developed counties is expected to reduce the differences in e-commerce accessibility. However, when the difference in the price effect is the main component, policies focusing on improving e-commerce usage skill for less-educated and low-skilled individuals (most of them are in less-developed counties) are necessary. Therefore, from the perspectives of both academia and policymaking, it would be interesting to investigate how these two components can influence the effect of e-commerce on the income inequality.

3. Methodology

3.1 Data

We constructed original cross-county-level panel data for five provinces (Zhejiang, Jiangsu, Henan, Hunan, and Ningxia) between 2011 and 2018. The Chinese government publishes yearbooks annually. We sourced information from counties in five provinces on each factor (e.g., GDP and education) from the official government database, except for the number of Taobao Villages. The total number of samples was 2,800 (see Appendix A1). We constructed the following variables:

- *Indicator of Income inequality*

The dependent variable was the regional income inequality. We used two types of indicators to measure the income inequality:

The first was the dependent variable in the income inequality function. As the analyzed unit was the county, we measured the income inequality in each county. Unfortunately, we could not access the official data to calculate the Gini coefficients. We utilized the concept of the poverty gap based on the Foster–Greer–Thorbecke indicator (Foster et al., 2010). We used the real provincial per capita rural household income as the standard income line and calculated a new indicator to measure the income gap as follows. First, we calculated the gap between real per capita rural household income in a county (A) and the real per capita income of the province to which the county belongs (B) [gap=A-B]. We then calculated the gap ratio (ratio = [A-B]/B). Finally, we calculated the squared value of the ratio, which was used as an income gap indicator in this study. As the ratio can be positive for a rich county or negative for a poor county, we used the squared value of the ratio as a measure indicator that could address the direction issue (negative or positive value), and the measure increased the sensitivity of the income gap (Foster et al., 2010).

We used the squared value of the ratio to address the direct issues.

The second was the differences in average income levels between developed and less-developed counties. We defined developed counties as having an income level higher than the provincial standard income line and less-developed counties as having an income level lower than the provincial standard income line. We decomposed the income gap between the two county groups using the B-O method.

・*Indicator of e-commerce development*

The key independent variable in this study was the level of e-commerce development. We used the Taobao Village as the indicator for the following reasons: First, Alibaba, the largest e-commerce company in China, has been actively promoting the establishment of Taobao Villages in rural areas since 2000. Taobao Villages have significantly contributed to the overall purchases and sales volume of e-commerce in China. Additionally, Taobao Villages represent primary e-commerce purchasing patterns in rural regions. Therefore, the number of Taobao Villages serves as a representative indicator of e-commerce development in rural China. Second, because our study utilized cross-county data within each province, we were unable to access specific county-level information regarding the level of e-commerce development, except for the number of Taobao Villages and the Alibaba E-commerce Development Index. Finally, previous studies employed various indicators of e-commerce development such as online retail sales, Internet trade volume, and per capita delivery volume. However, Wang and Yang (2020) argued that a universally accepted measurement indicator for e-commerce development has not yet been established. Previous studies (Qi et al., 2019; Zeng et al., 2017) have used the number of Taobao Villages as an indicator of e-commerce development in rural areas.

Therefore, in this study, we used two indicators: (i) whether a county had at least one Taobao Village [Taobao Village] and (ii) the natural logarithm of the number of Taobao Villages in a county [Ln (the number of Taobao Villages)]. These two indicators allowed us to evaluate the impact of both the breadth and depth of e-commerce development in each county.

Information on the number of Taobao Villages was sourced from the Taobao

Village list published annually by Alibaba since 2012. Zhejiang and Jiangsu provinces in the eastern region are the top two provinces in terms of the number of Taobao Villages in China. Henan and Hunan provinces are located in the central region, whereas Ningxia province is located in the western region. We used the period 2011–2018 for our analysis, as Alibaba started counting the number of Taobao Villages[2] and first published the Taobao Village list in 2012. There was a large regional disparity in the number of villages. For example, Taobao Villages have recently increased in the western region. The number of provinces with at least one Taobao Village increased to 24 in 2017, 25 in 2019, and 28 in 2020. Zhejiang province had 1,757 Taobao Villages, and Ningxia, Hainan, and Gansu provinces had only one Taobao Village in 2020. In this study, we used this regional disparity in Taobao Villages to construct indices for e-commerce development levels and represent distinct levels of e-commerce development.

We also constructed the following control variables (see Appendix Table A1).

- *GDP [Ln (GDP per capita), Ln (GDP per capita squared)*]: Kuznets (1955) advocated an inverted U-shaped relationship between economic growth and income inequality.[3] We used the logarithm of the county's per capita GDP and its square as indicators of the economic growth level.

- *Fiscal expenditure [Gov]* evaluates the government's participation in economic activities. The fiscal expenditure indicator is the ratio of fiscal expenditure to GDP, calculated annually. Regional disparities in public fiscal expenditure are expected to affect the cross-county income inequality.

- *Financial development [Finance]*: This index is the ratio of annual loans extended by financial institutions to the GDP. Zhang and Guo (2011) and Zhang et al. (2013) reported that financial development significantly affects the

---

[3] The empirical results of testing the Kuznets's hypothesis were mixed for China (Zhang, Chen, & Zhang, 2012), and evidence from the county level is scarce (Cheng & Wu, 2017).

income inequality. For example, Zhang et al. (2013) found that the relationship between the level of financial development and the income inequality among rural households has an inverted U shape.[4] Therefore, we used this variable to control for the influence of financial development on the income inequality.

- *Industrial structure [Sec_gdp, Ter_gdp]*: This index is the ratio of the added value of the secondary or tertiary sectors to annual GDP. They reflect the development of non-agricultural industries. Zhu (1992) argued that the disparity in developing non-agricultural industries among counties has widened the income inequality in rural China. Leng et al. (2020) found that ICT adoption positively and significantly affects income diversification among rural households in China. We constructed secondary and tertiary industry dummies to control for the influence of nonfarm income on the income inequality.

- *Capital investment [Invest]*: Capital investment is an important factor affecting regional economic development and income inequality (Qi et al., 2019). Raychaudhuri and Prabir (2010) reported that investment in fixed assets, such as public infrastructure, will likely widen the urban-rural income inequality. Disparities in capital investment may also affect this gap among rural counties. Therefore, we constructed the capital investment variable as "the ratio of fixed asset investment to GDP."

- *Education [Ln(edu)]*: Human educational capital stock is usually measured by years of schooling. However, owing to data limitations, we used the logarithm of the ratio of the number of students enrolled in secondary schools to the total population of a county as the education indicator. According to the human capital theory (Becker, 1964; Mincer, 1974), education may affect the income inequality.

- *Health [Ln(bed)]*: Because health status also contains an element of human capital that may affect income, we constructed a logarithm of the ratio of the number of hospital beds to the total population. The number of hospital beds reflects the level of medical care services, which may affect people's health status.

- *Land [Ln(pc_land)]*: The total cultivated area per capita logarithm was used. In China, approximately 98% of farmers cultivate less than two hectares. Furthermore, these small farms predominantly depend on family labor (Rapsomanikis, 2015). Therefore, the scale of cultivated land per capita is related to agricultural productivity, which can influence agricultural income. Moreover, farmland consolidation is progressing faster in neighboring urban villages, which may affect farmers' household incomes in rural areas. Therefore, we constructed a variable to control for these effects.

- *Instrument Variable* (*IV):* We used the IV method to check the robustness of the model. Luo and Niu (2019) found that participation in e-commerce is not random in Taobao Villages and stated that although lagged endogenous variables are commonly used as instrumental variables, they are not strictly exogenous. Li et al. (2021) used the number of post offices and telephones in 1991 as an instrumental variable for the ratio of online retail sales to GDP. Referring to previous studies, we applied several tests (e.g., over-identification, weak IV tests, and endogenous tests) to prove the validity of the IVs (see Table 4). We selected the logarithm of the provincial number of websites *[Ln (website)]* and the provincial ratio of rural broadband to total broadband subscribers *[Rural_internet]* as IVs for the Taobao Village variables in our study. These IVs passed both weak identification and over-identification tests. In addition, both IVs were statistically significant in the first-stage regression, indicating that they were valid. Passing another criterion of the IV—the "exclusion restriction"—at least one of our IVs did not seem to affect the income gap in rural areas directly.

3.3 Model

We adopted two models to address potential endogeneity problems: DID with the multi-period method and the FE model. We also used a combination model of the FE and IVs to check robustness.

First, the DID model was constructed as follows:

$$Gap_{it} = \beta_0 + \beta_1 EC_{it} + \beta_2 Post\_EC_{it} + \beta_3 X_{it} + \lambda_t + \mu_i + \varepsilon_{it}, \quad (1)$$

where i represents a county, t represents the year, and Gap represents the indicator of the income inequality in each county. EC is the treatment dummy (a county with at least one Taobao Village in any given year always has a value of 1, even before the establishment of Taobao Villages). $Post$ denotes the period after a county has established the first Taobao Village, and $Post\_EC$ is the post-treatment period dummy (1 is in the treatment group after at least one Taobao Village was established; otherwise, it is 0). It should be noted that because the year of Taobao Village establishment differs among counties, the $Post$ period differs by county, which is different from the traditional DID method. X represents a series of observable control variables. λ and μ represent the time and county fixed effects, respectively, and ε represents the error term.

Second, we used the FE or random effects (RE) model to address the heterogeneity problem due to time-invariant individual specificity, as shown in Eq. (2):

$$Gap_{it} = \beta_0 + \beta_1 EC_{it} + \beta_2 X_{it} + u_i + v_{it}, \quad (2)$$

where u is the time-invariant individual specificity and v is the true error term. The Hausman specification test was used to assess the validity of the FE and RE models.

Third, although it is difficult to find a perfect IV, we further considered the potential endogeneity of the Taobao Village variables and used the IV method to provide a robustness check. In this study, we used the provincial number of websites and the provincial ratio of rural broadband to total broadband subscribers as IVs for the Taobao Village variables. Eqs. (3) and (4) express the FE_IV model:

$$EC_{it} = a + \gamma_1 Z_{1it} + \varphi X_{it} + \epsilon_{it}, \quad (3)$$
$$Gap_{it} = a + \beta_1 \widehat{EC}_{it} + \lambda X_{it} + \theta_1 Z_{1it} + +u_i + v_{it}, \quad (4)$$
$$corr(Z, \varepsilon) = 0 \text{ and } corr(Z, v) \neq 0$$

where Z represents IVs (the provincial number of websites and provincial ratio of rural broadband to total broadband subscribers). A set of tests was performed to assess the validity of the IVs.

Finally, we used the B-O decomposition method to decompose the determinants of the income gap between the developed and less-developed counties into two components: the endowment difference (e.g., the difference in the number of Taobao Villages between the two counties) and the price difference (i.e., the difference in the magnitude of e-commerce's effect on income level between the two counties), represented by Eqs. (5) and (6), respectively:

$$\overline{lnin}_H - \overline{lnin}_L = \sum \beta_{\mathrm{H}}(X_H - X_L) + \sum(\beta_{\mathrm{H}} - \beta_{\mathrm{L}})\, X_L, \quad (5)$$

$$\overline{lnin}_H - \overline{lnin}_L = \sum \beta_{\mathrm{L}}(X_L - X_H) + \sum(\beta_{\mathrm{L}} - \beta_{\mathrm{H}})\, X_H, \quad (6)$$

where H and L denote the high- and low-income counties (developed and less-developed counties), which are distinguished based on the county's per capita income; $\overline{lnin}_H - \overline{lnin}_L$ is the difference in the mean value of the income levels between the two groups; X represents the mean values of a set of variables (including e-commerce and control variables); and β is the coefficient calculated from the income function. $\sum \beta_{\mathrm{H}}(X_H - X_L)$ or $\sum \beta_{\mathrm{L}}(X_L - X_H)$ expresses the difference in the endowment effect; $\sum(\beta_{\mathrm{H}} - \beta_{\mathrm{L}})\, X_L$ or $\sum(\beta_{\mathrm{L}} - \beta_{\mathrm{H}})\, X_H$ denotes the difference in the price effect.

4. Empirical results

4.1 Basic results

(1) Results of the DID model

Table 1 presents the results of the DID model. Columns (1) and (2) present the results for all counties. Columns (3) and (4) show the results for less-developed (low-income) counties, where the per capita household income is less than the provincial average per capita income of rural residents. Conversely, Columns (5) and (6) show the results for developed (high-income) counties, where per capita household income is more than the provincial average per capita income of rural residents.

--Table 1 near here—

First, the coefficient of $Post\_EC$ was the DID estimator of the treatment effect (a county with at least one Taobao Village in the year after the establishment of Taobao Villages). The values were negative and statistically significant at the 5% level in Columns (3) and (4) and positive and statistically significant at the 1% and 5% levels in Columns (5) and (6), respectively. The results indicated that e-commerce development may widen the income inequality in developed counties while reducing it in less-developed ones.

The reasons for this are as follows. First, in developed counties, the advancement of e-commerce has led to higher levels of ICT compared to less-developed ones. This "skill-biased technology change" can increase the demand for skilled labor and widen the wage gap between skilled and unskilled workers, resulting in income inequality growth (Balcilar et al., 2021; Card & Lemieux, 2001; Krueger, 1993). Second, in less-developed counties, e-commerce can provide opportunities to create new jobs, particularly in online businesses. For example, farmers can easily access customer information to promote the sale of agricultural products, thereby stimulating agricultural development and increasing income (Vatsa et al., 2022). Because low-income households often face more severe information asymmetry than high-income ones, e-commerce can have a more significant impact on boosting the income of low-income households, consequently reducing the income inequality.

Second, regarding other relevant socioeconomic variables, the relationship between GDP, per capita, and the income inequality had an inverted U shape, which is consistent with previous studies (Kuznets, 1955). The ratios of fiscal expenditure to GDP and of loans from financial institutions to GDP expanded the income inequality. By contrast, the added value of the tertiary sector to GDP reduced this gap in developed counties. However, the reverse was true in less-developed counties.

(2) Results from the FE model

Table 2 lists the results of the FE model. As the results of the Hausman specification test indicated that the FE model was more appropriate than the RE model, we present only the results of the FE model in Table 2.

--Table 2 near here—

The coefficients of EC (Taobao Villages and the natural logarithm of the number of Taobao Villages) estimated the effects of e-commerce development. Columns (1)－(4) show the results for all counties, columns (5)–(8) show the results for less-developed counties, and columns (9)–(12) show the results for developed counties. Columns (1), (3), (5), (7), (9), and (11) show the development of rural e-commerce using the existence of Taobao Villages (breadth of the development of rural e-commerce), while columns (2), (4), (6), (8), (10), and (12) show the development using the number of Taobao Villages (depth of the development of rural e-commerce).

The coefficients of the e-commerce variables were negative and statistically significant at the 5% level for less-developed counties, and positive and statistically significant at the 5% and 1% levels for developed counties. These results were consistent with those of the GDD model.

4.2 Robustness check

First, considering that the definitions of developed and less-developed counties may affect the estimated results, we changed the definition of the income standard and reran the estimations. We used income terciles to divide the sample into high-, middle-, and low-income counties (high-, middle-, and less-developed). Table 3 lists the results obtained using the FE model. The coefficients of both the Taobao Village dummy and the number of Taobao Villages were positive and statistically significant at the 1% level for high-developed counties, as shown in columns (1) and (2). These results were consistent with those shown in Tables 1 and 2.

--Table 3 near here—

Second, we used the IV method to address the endogeneity issue caused by unobservable variables. Table 4 presents the results obtained using the FE-IV method. Again, we limited the results to those that passed both the weak and over-identification tests.

Columns (1) and (2) show the results for high-developed counties. The first-stage regression showed that IVs statistically significantly affected the potential endogenous Taobao Village variables at the 5% or 1% levels. The coefficients of the e-commerce variables were positive and statistically significant at the 1% level, indicating that e-commerce development tends to expand the income inequality in rural areas. These results were consistent with those obtained from the GDD and FE models. This reconfirmed the conclusion.

--Table 4 near here—

4.3 Heterogenous effects of e-commerce development

We considered the possible heterogeneous effects of e-commerce development in regions with different levels of agricultural modernization and human capital. Because the results were significant only for developed counties, they are presented in Table 5.

--Table 5 near here—

The coefficients of the e-commerce variables (both Taobao Villages and the number of Taobao Villages) and the total power of agricultural machinery were positive and statistically significant at the 1% and 5% levels. By comparison, the coefficients of the interaction terms were negative and statistically significant at the 5% and 10% levels. These results indicate that e-commerce development may widen the income inequality in counties with lower levels of agricultural modernization. However, this is the reverse of that in counties with a higher level of agricultural modernization, suggesting that the income inequality widening effect of e-commerce can be mitigated by agricultural modernization. In other words, policies to promote agricultural modernization can reduce income inequality expansion with the development of e-commerce.

Regarding the heterogeneous effect of human capital, the interaction terms between the education and e-commerce variables were mostly insignificant, suggesting that in developed counties, the difference in the impact of e-commerce on the income inequality between low and high human capital counties is small.

4.4 Decomposition results

The decomposition results of the income inequality between developed and less-developed counties are summarized in Tables 6 (using the dummy variable of having a Taobao Village) and 7 (using the number of Taobao Villages). Decompositions were performed based on the results of the FE model.

--Table 6 near here—

--Table 7 near here—

First, the price effect (101.0%) had a greater impact on the income inequality than the endowment effect (-1.0%). The results indicated that the difference in income determination mechanisms contributed to widen the income, inequality while the difference in the amount of each variable contributed to reduce the income inequality, and the effect was greater for the former than for the latter.

Second, regarding the effect of e-commerce, (1) in the endowment effect, the contribution rate of both having a Taobao Village and that of the number of Taobao Villages was 0.1%. This indicates that the income level was higher in a county with a Taobao Village or a county with more Taobao Villages than in its counterparts (a county without a Taobao Village or a county with fewer Taobao Villages). These results suggest that differences in e-commerce accessibility contributed to the expansion of the income inequality within rural areas.

(2) In the price effect, the contribution rate of having a Taobao Village was -0.2% and that of the number of Taobao Villages was 0.1%, indicating that the income returns of having a Taobao Village contributed to reduce the income inequality, inequality while the income returns of the number of Taobao Villages contributed to widen the income inequality.

In summary, the results indicated that the differences in both the accessibility and income return of e-commerce usage contributed to expanding the income inequality between developed and less-developed counties within rural areas.

5. Conclusions

We constructed an original cross-county panel dataset from 2011 to 2018 and used the GDD, FE, and FE_IV approaches to investigate the impact of e-commerce development on the income inequality in rural China after considering endogeneity problems. The following three main conclusions were drawn:

First, the effects differed by region. For example, e-commerce development could expand the income inequality in developed counties while reducing it in less-developed ones. As a result, the total effect of e-commerce in rural areas was insignificant.

Second, the effect of e-commerce on the income inequality differed among heterogeneous groups. For example, the income inequality expansion effect was smaller in counties with a higher level of agricultural modernization than in those with a low level. This suggests that progress in agricultural technology and capital-intensive industrial agglomeration in rural areas may reduce the income inequality.

Third, the decomposition results indicated that both the differences in accessibility and income return of e-commerce usage (especially the income return of Taobao Village number) contributed to income inequality expansion in rural areas.

The policy implications of these empirical findings can be considered as follows. Although the government enforced the development of e-commerce as an effective way to alleviate rural poverty and increase the income of rural residents, it may expand the income inequality in rural areas, which may widen income inequality nationwide. Reducing the nationwide income inequality in the digital era is a significant challenge for governments. As two components, the difference in accessibility and income return of e-commerce, contribute to widening the income inequality between developed and less-developed counties, policies to expand the e-commerce platform in less-developed counties, and policies to improve e-commerce usage skills such as providing more e-commerce or ICT training to rural residents in less-developed counties, are expected to reduce the income gap in rural areas. Moreover, the results indicated that agricultural modernization may reduce the income inequality with the development of e-commerce, suggesting that policies to promote agricultural modernization may reduce the income inequality.

This study had several limitations. First, although we used the GDD, FE, and FE_IV approaches to address endogeneity issues, we could not identify the underlying causality between e-commerce development and the income inequality in rural areas, which should be addressed through a more in-depth analysis. Second, there was insufficient information available to measure a county's income inequality. Therefore, we used the relative income inequality to develop and construct an indicator of the income inequality in cross-county panel data, which could lead to new challenges in the future. Finally, the establishment of Taobao Villages is primarily determined by corporations such as Alibaba and local government support. Therefore, the number of Taobao Villages serves as an indicator of the impact of e-commerce on income inequality from the perspective of e-commerce suppliers (corporations). However, it is important to note that increased e-commerce usage (more transactions conducted through e-commerce platforms) can have spillover effects that contribute to the establishment of Taobao Villages in nearby areas. To further understand the effects of e-commerce on income inequality, it is crucial to examine both the role of e-commerce corporations and the behavior of consumers (e-commerce platform users). This distinction between the two sides will be the focus of our future research.

However, despite its limitations, we believe that this study, which constructed original cross-county panel data based on statistical information from 2011 to 2018 from the eastern, central, and western regions of rural China, provides new insights into the association between e-commerce development and the income inequality among rural counties in China. This study is also the first to decompose the effects of e-commerce into two components: differences in accessibility, and the magnitude of the effect of e-commerce on income levels. We expect that the Chinese experience will provide valuable lessons for developing countries wishing to expand their e-commerce activities and accelerate economic development.

References

Balcilar, M., Gupta, R., Ma, W., & Makena, P. (2021). Income inequality and economic growth: A re-examination of theory and evidence. *Review of Economic Development*, 25(2), 737–757.

Becker, G. S.（1964）. *Human capital: A theoretical and empirical analysis, with special reference to education.* New York: Columbia University Press.

Caldarulo,M. , Mossberger, K., & Howell, A. (2023). Community-wide broadband adoption and student academic achievement. *Telecommunications Policy*, 47, Article 102445.

Card, D., & Lemieux, T. (2001). Can falling supply explain the rising return to college for younger men? A cohort-based analysis. *The Quarterly Journal of Economics*, 116(2),705–746.

Chang, H. H., & Just, D. R. (2009). Internet access and farm household income: Empirical evidence using a semi-parametric Assessment in Taiwan. *Journal of Agricultural Economics*, 60, 348–366.

Chen, Y., Liu, M., & Zhang, Q. (2010). Development of financial intermediation and the dynamics of urban-rural disparity in China, 1978–1998. *Regional Studies*, 44(9), 1171–1187.

Cheng, W., &Wu, Y. (2017). Understanding the Kuznets process: An empirical investigation of income inequality in China: 1978–2011. *Social Indicators Research*, 134, 631–650.

Couture, V., Faber, B., Gu, Y., & Liu, L. (2018). Connecting the countryside via e-commerce: Evidence from China. *NBER Working Paper* 24384.

Dimaggio, P., & Bonikowski, B. (2008). Make money surfing the web? The impact of internet use on the earnings of U.S. workers. *American Sociological Review*, 2, 227–250.

Foste,J., Greer, J., & Erik Thorbecke, E. (2010). The Foster–Greer–Thorbecke (FGT) poverty measures: 25 years later. *The Journal of Economic Inequality*, 8,491–524.

Furuholt, B., & Kristiansen, S. (2007). A rural-urban digital divide? Regional aspects of internet use in Tanzania. *The Electronic Journal of Information Systems in Developing Countries*,31 (1), 1–15.

Ghosh, S. (2020). Impact of economic growth volatility on income inequality. *Quality & Quantity*, 54 (3), 807–850.

He, Y. (2020). Analysis of the effect of rural e-commerce development on the urban-rural income gap in China. *Journal of Commercial Economics*, 16, 91–94 (in Chinese).

Hübler, M., and Hartje, R. (2016). Are smartphones smart for economic development? *Economics Letters*, 141, 130–133.

Jiang, Y., Shi, X., Zhang, S., and Ji, J. (2011). The threshold effect of high-level human capital investment on China's urban-rural income gap. *China Agricultural Economic Review*, 3(3), 297–320.

Krueger, A.B. (1993). How computers have changed the wage structure: Evidence from microdata, 1984- 1989. *The Quarterly Journal of Economics*, 108(1), 33–60.

Kuznets, S. (1955). Economic growth and income inequality. *The American Economic Review*, 45(1),

1–28.

Li, S., Sicular, T., & Gustafsson, B. (2008). *Income distribution in China III*. Beijing: Beijing Normal University Press (in Chinese).

Li, L., Zeng, Y., Ye, Z. & Guo, H. (2021). E-commerce development and urban-rural income gap: Evidence from Zhejiang Province, China. *Papers in Regional Science*, 100, pp.475–494.

Li, L., Zeng, Y., Ye, Z, & Guo, H. (2021). E-commerce development and urban-rural income gap: Evidence from Zhejiang Province, China. *Papers in Regional Science*, 100, 475–494.

Li, Q., Tang, Y., & Ren, X. (2019). The development of e-commerce, spatial spillover effects and peasants' income growth. *Journal of Agrotechnical Economics*, 119–131 (in Chinese).

LoPiccalo, K. (2022). Impact of broadband penetration on U.S. Farm productivity: A panel approach. *Telecommunications Policy*, 46, Article102396.

Luo, Q., Hu, H., Feng, D., & He, X. (2022) How does broadband infrastructure promote entrepreneurship in China: Evidence from a quasi-natural experiment. *Telecommunications Policy*,46, Article 102440.

Luo, X., & Niu, C. (2019). E-commerce participation and household income growth in Taobao Villages. *World Bank Poverty & Equity Global Practice Working Paper* 198.

Lloyd-Ellis, H. (1999). Endogenous technological change and wage inequality. *American Economic Review*, 89 (1), 47–77.

Ma, X. (2022). Internet use and income gaps between urban and rural residents in China. *Journal of the Asia Pacific Economy*. doi:10.1080/13547860.2022.2054133

Ma, W., Grafton, R. Q., & Renwick, A. (2018). Smartphone use and income growth in rural China: Empirical results and policy implications. *Electronic Commerce Research*, 20(4), 713–736.

Mincer, J. (1974). *Schooling, experience and earning*. New York: Columbia University Press.

Philip, L., Cottrill,C., Farrington, J., Williams, F., & Ashmore, F. (2017). The digital divide: Patterns, policy and scenarios for connecting the 'final few' in rural communities across Great Britain. *Journal of Rural Studies,*54, 386–398.

Pradhan, R. P., Arvin, M. B., Norman, N. R., & Bennett, S. E. (2016). Financial depth, internet penetration rates and economic growth: Country-panel evidence. *Applied Economics*,48 (4),331–343.

Qi, J., Zheng, X., & Guo, H. (2019). The formation of Taobao villages in China. *China Economic Review*, 53, 106–127.

Rapsomanikis, G. (2015). *The economic lives of smallholder farmers: An analysis based on household data from nine countries.* Food and Agriculture Organization of the United Nations. Rome.

Shin, H. B. (2016). Economic transition and speculative urbanization in China: Gentrification versus dispossession. *Urban Studies*, 53(3), 471–489.

Sicular, T., Li, S., Yue, X., & Sato, H. (2020). *Changing Trends in China's Inequality: Evidence, Analysis, and Prospects*. Oxford UK: Oxford University Press.

You, S., & Zhang, Q. (2018). The influence of rural e-commerce development on farmers' propensity to consume: Stimulating effect and income growth effect. *Luojia Management Review*, 3, 182–191 (in Chinese).

Valentín-Sívico, J., Canfield, C., Sarah A. Low, S. A., & Gollnick, C. (2023). Evaluating the impact of broadband access and internet use in a small underserved rural community. *Telecommunications Policy*, 47, Article 102499.

Vatsa, P., Li, J., Luu, P. Q., & Botero-R, J. C. (2022) Internet use and consumption diversity: Evidence from rural China. *Review of Development Economics*, Article 12935.

Wan, G. (2004). Accounting for income inequality in rural China：A regression based approach. *Economic Research Journal*, 8, 117–127.

World Bank and Alibaba (2019). *E-commerce development: Experience from China.* Washington DC: The World Bank.

Zeng, Y., Jia, F., Wan, L., & Guo, H. (2017). E-commerce in agri-food sector: a systematic literature review. *International Food and Agribusiness Management Review*, 20(4), 439–460.

Zeng, Y., Guo, H., & Jin, S. (2018). Does e-commerce increase farmers' income? Evidence from Shuyang County, Jiangsu Province, China. *China Rural Economy*, 2, 49–64 (in Chinese).

Zhang, B., Liu, D. & Zheng, B. (2013). Has rural financial development eased rural residents' income gap? Analysis from the threshold regression model based on the inter-province panel data in China. *China Rural Survey*, 3,19–29 (in Chinese).

Zhang, H., Chen, W., & Zhang, J. (2012). Urban-rural income disparities and development in a panel data set of China for the period from 1978 to 2006. *Applied Economics*, 44, 2717–2728.

Zhang, J. & Guo, P. (2011) Impact of financial development on income inequality in rural China: Evidence based on VAR model. *Agricultural Technology Economics*, 1, 34–41(in Chinese).

Zhang, L. and Han, L. (2017). Has e-commerce economy development expanded urban-rural residents income gap? *Research on Economics and Management*, 38(5), 3–13 (in Chinese).

Zhang, Y. (2019). Can internet trade reduce the income gap? A double heterogeneous model and the experience from China. *Inquiry into Economic Issues*, 6, 50–58 (in Chinese).

Zhu, L. (1992). Impact of non-farm industrial activities on the income distribution pattern of farm households. *Economic Research Journal*, 3, 23–30(in Chinese).

Table1 Results of the association between e-commerce development and income inequality from the GDD model

	Total		Less-developed county		develoepd county	
	(1)	(2)	(3)	(4)	(5)	(6)
Post_EC	0.002	0.002	-0.009**	-0.007**	0.013***	0.008**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
Ln (GDP per capita)	0.300**	0.044	-0.259**	-0.410***	1.071***	1.309***
	(0.143)	(0.146)	(0.105)	(0.118)	(0.286)	(0.292)
Ln (GDP per capita squared)	-0.015**	-0.004	0.010**	0.018***	-0.049***	-0.060***
	(0.007)	(0.007)	(0.005)	(0.006)	(0.013)	(0.013)
Gov	-0.049*	-0.030	-0.103***	-0.102***	0.021	0.101**
	(0.029)	(0.029)	(0.031)	(0.031)	(0.047)	(0.044)
Finance	-0.01	-0.007	-0.029***	-0.025***	0.022	0.046***
	(0.011)	(0.012)	(0.010)	(0.009)	(0.016)	(0.016)
Sec_gdp		0.224***		0.188***		-0.209
		(0.061)		(0.046)		(0.131)
Ter_gdp		0.116*		0.236***		-0.531***
		(0.064)		(0.054)		(0.128)
Invest	-0.005	-0.001	-0.011	-0.007	-0.004	0.011
	(0.005)	(0.006)	(0.008)	(0.007)	(0.007)	(0.008)
Ln (edu)	-0.012	-0.006	0.000	0.003	-0.030**	-0.014
	(0.008)	(0.007)	(0.005)	(0.005)	(0.012)	(0.011)
Ln (bed)		0.002		-0.005		0.001
		(0.006)		(0.006)		(0.009)

Ln (pc_land)	0.019*	0.021**	0.017***	0.016***	0.018	0.016
	(0.010)	(0.010)	(0.005)	(0.005)	(0.015)	(0.013)
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-1.370*	-0.129	1.610***	2.183***	-5.579***	-6.682***
	(0.730)	(0.740)	(0.534)	(0.594)	(1.520)	(1.550)
Observations	1,579	1,578	891	890	688	688
R-squared	0.199	0.226	0.291	0.323	0.341	0.431
Number of county_id	220	220	127	127	101	101

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2 Results of the association between e-commerce development and income inequality from the FE model

	Total				Less-developed county				Developed county			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Taobao Village	0.002		0.002		-0.009**		-0.007**		0.013***		0.007**	
	(0.003)		(0.003)		(0.003)		(0.003)		(0.004)		(0.003)	
Ln (the number of Taobao Villages)		0.003*		0.003*		-0.005**		-0.003		0.010***		0.007***
		(0.002)		(0.002)		(0.002)		(0.003)		(0.002)		(0.002)
Ln (GDP per capita)	0.299**	0.310**	0.043	0.048	-0.259**	-0.238**	-0.410***	-0.384***	1.063***	1.115***	1.306***	1.344***
	(0.142)	(0.142)	(0.146)	(0.146)	(0.105)	(0.104)	(0.118)	(0.117)	(0.285)	(0.292)	(0.292)	(0.297)
Ln (GDP per capita squared)	-0.015**	-0.015**	-0.004	-0.004	0.010**	0.009*	0.018***	0.016***	-0.048***	-0.050***	-0.060***	-0.061***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.005)	(0.005)	(0.006)	(0.006)	(0.013)	(0.014)	(0.013)	(0.014)
Invest	-0.005	-0.004	-0.001	0.000	-0.011	-0.010	-0.007	-0.007	-0.004	0.000	0.011	0.013*
	(0.005)	(0.005)	(0.006)	(0.006)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)
Gov	-0.048*	-0.052*	-0.03	-0.034	-0.103***	-0.107***	-0.102***	-0.108***	0.022	0.011	0.102**	0.089**
	(0.029)	(0.029)	(0.029)	(0.029)	(0.031)	(0.031)	(0.031)	(0.032)	(0.048)	(0.043)	(0.044)	(0.042)
Finance	-0.010	-0.009	-0.007	-0.006	-0.029***	-0.030***	-0.025***	-0.026***	0.023	0.027*	0.047***	0.048***
	(0.011)	(0.011)	(0.012)	(0.011)	(0.010)	(0.010)	(0.009)	(0.009)	(0.016)	(0.015)	(0.016)	(0.015)
Ln (edu)	-0.012	-0.011	-0.006	-0.005	0.000	0.001	0.003	0.004	-0.030**	-0.032**	-0.014	-0.016
	(0.008)	(0.008)	(0.007)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)	(0.012)	(0.012)	(0.011)	(0.011)
Ln (pc_land)	0.019*	0.020**	0.021**	0.022**	0.017***	0.018***	0.016***	0.017***	0.018	0.023	0.016	0.020
	(0.010)	(0.010)	(0.010)	(0.010)	(0.005)	(0.005)	(0.005)	(0.005)	(0.015)	(0.015)	(0.013)	(0.013)
Sec_gdp			0.224***	0.234***			0.188***	0.188***			-0.21	-0.206
			(0.061)	(0.061)			(0.046)	(0.046)			(0.131)	(0.130)
Ter_gdp			0.116*	0.133**			0.236***	0.236***			-0.533***	-0.511***
			(0.064)	(0.066)			(0.054)	(0.056)			(0.128)	(0.127)
Ln (bed)			0.002	0.001			-0.005	-0.004			0.001	0.001

			(0.006)	(0.006)			(0.006)	(0.006)			(0.009)	(0.009)
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-1.363*	-1.432**	-0.123	-0.171	1.610***	1.502***	2.183***	2.058***	-5.534***	-5.871***	-6.663***	-6.916***
	(0.728)	(0.725)	(0.739)	(0.736)	(0.534)	(0.528)	(0.594)	(0.587)	(1.515)	(1.554)	(1.549)	(1.576)
Observations	1,579	1,579	1,578	1,578	1,579	1,579	1,578	1,578	891	891	890	890
R-squared	0.232	0.232	0.241	0.241	0.199	0.202	0.226	0.23	0.291	0.286	0.323	0.317
Number of county_id	220	220	220	220	220	220	220	220	127	127	127	127

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3 Results of changing the definitions of developed and less-developed counties

	High-income county		Middle-income county		Low-income county	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Taobao Village	0.016***		-0.003		-0.007	
	(0.005)		(0.003)		(0.005)	
Ln (the number of Taobao Villages)		0.013***		-0.001		-0.004
		(0.003)		(0.001)		(0.003)
Ln (GDP per capita)	0.966**	1.058***	0.070	0.096*	-0.293**	-0.292**
	(0.374)	(0.389)	(0.063)	(0.049)	(0.135)	(0.132)
Ln (GDP per capita squared)	-0.044**	-0.048***	-0.004	-0.005**	0.012*	0.012*
	(0.017)	(0.018)	(0.003)	(0.002)	(0.007)	(0.006)
Invest	0.003	0.01	-0.005	-0.005	-0.013**	-0.012**
	(0.010)	(0.010)	(0.005)	(0.005)	(0.006)	(0.006)
Gov	0.113	0.087	-0.012	-0.013	-0.096***	-0.098***
	(0.189)	(0.180)	(0.009)	(0.010)	(0.032)	(0.032)
Finance	0.007	0.012	-0.003	-0.003	-0.033***	-0.033***
	(0.027)	(0.028)	(0.007)	(0.007)	(0.010)	(0.009)
Ln (edu)	-0.038***	-0.040***	-0.004	-0.004	-0.004	-0.005
	(0.014)	(0.014)	(0.004)	(0.004)	(0.006)	(0.006)
Ln (pc_land)	0.037**	0.042**	-0.002	-0.002	0.025***	0.026***
	(0.017)	(0.017)	(0.004)	(0.004)	(0.008)	(0.008)
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-4.947**	-5.514**	-0.270	-0.404	1.830***	1.829***
	(2.058)	(2.140)	(0.356)	(0.278)	(0.691)	(0.680)
Observations	515	515	518	518	546	546
R-squared	0.407	0.428	0.076	0.068	0.407	0.405
Number of county_id	76	76	85	85	83	83

Notes: The fixed effect model was used. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4 Results of the association between EC development and income inequality from the FE-IV method

VARIABLES	(1)	(2)
Taobao Village	0.108***	
	(0.033)	
Ln (Number of Taobao Villages)		0.054***
		(0.016)
Ln (GDP per capita)	1.794***	1.739***
	(0.590)	(0.518)
Ln (GDP per capita squared)	-0.081***	-0.077***
	(0.027)	(0.024)
Invest	0.034*	0.040**
	(0.018)	(0.019)
Gov	-0.168**	-0.142**
	(0.082)	(0.072)
Finance	0.004	0.033
	(0.028)	(0.027)
Ln (edu)	-0.009	-0.027
	(0.020)	(0.016)
Ln (pc_land)	0.033	0.054**
	(0.026)	(0.025)
Year fixed effect	Yes	Yes
Constant	-9.769***	-9.635***
	(3.249)	(2.836)
First stage regression		
IV		
Ln (website)	-0.352**	-0.731***
	(0.150)	(0.242)
Rural_internet	-1.678**	-3.293**
	(0.792)	(1.434)
Weak identification test (Cragg-Donald Wald F statistic)	21.560	29.506
Weak identification test (Kleibergen-Paap rk Wald F statistic)	11.723	14.295

Overidentification test (Hansen J statistic p-value)	0.546	0.547
Observations	688	688
Number of county_id	101	101
Within R-squared	-	-
Between R-squared	0.0077	0.0032
Overall R-squared	0.0001	0.0099

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5 Heterogeneous effects of e-commerce on income inequality in developed counties

	(1)	(2)	(3)	(4)
Taobao Village	0.131**		0.030	
	(0.051)		(0.103)	
Ln (Number of Taobao Villages)		0.072**		-0.007
		(0.035)		(0.061)
Ln (agri_machine)	0.001***	0.001**		
	(0.000)	(0.000)		
Ln (edu)	0.001	-0.033**	-0.030**	-0.032**
	(0.011)	(0.013)	(0.012)	(0.012)
Interaction term of Taobao Village and Ln (agri_machine)	-0.009**			
	(0.004)			
Interaction term of Ln (the number of Taobao Villages) and Ln (agri_machine)		-0.005*		
		(0.003)		
Interaction term of Taobao Village and Ln (edu)			-0.002	
			(0.012)	
Interaction term of Ln (the number of Taobao Villages) and Ln (edu)				0.002
				(0.007)
Ln (GDP per capita)	0.851***	1.049***	1.068***	1.113***
	(0.224)	(0.300)	(0.297)	(0.296)
Ln (GDP per capita squared)	-0.038***	-0.047***	-0.048***	-0.050***
	(0.010)	(0.014)	(0.014)	(0.014)
Invest	0.007	0.011	-0.004	0.000
	(0.007)	(0.011)	(0.007)	(0.007)
Gov	0.059	0.165	0.023	0.010
	(0.083)	(0.130)	(0.048)	(0.042)
Finance	0.034**	0.025	0.023	0.027*
	(0.014)	(0.017)	(0.016)	(0.015)
ln(pc_land)	0.027*	0.023	0.018	0.023
	(0.014)	(0.017)	(0.015)	(0.015)
Year fixed effect	Yes	Yes	Yes	Yes
Constant	-4.759***	-4.778***	-5.170***	-5.064***
	(1.232)	(1.248)	(1.239)	(1.239)
Observations	590	590	688	688
R-squared	0.346	0.398	0.34	0.362
Number of county_id	93	93	101	101

Notes: The fixed effect model was used. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6 Decomposition results of the income gap between developed and less-developed counties (using the dummy variable of having Taobao Village)

	Value		Percentage (%)	
	Explained	Unexplained	Explained	Unexplained
Total	**-0.004**	**0.442**	**-1.0%**	**101.0%**
Having Taobao village	**0.000**	**-0.001**	**0.1%**	**-0.2%**
Other factors	**-0.004**	**0.443**	**-0.90%**	**101.02%**
Gov	-0.019	-0.041	-4.3%	-9.4%
Finance	0.000	-0.028	0.1%	-6.5%
Sec_gdp	0.001	-0.002	0.1%	-0.5%
Ter_gdp	0.003	-0.005	0.8%	-1.1%
Invest	-0.005	0.004	-1.1%	1.0%
lnedu	-0.001	-0.495	-0.1%	-113.3%
lnbed	0.004	0.015	0.8%	3.3%
lnpc_land	0.015	0.068	3.4%	15.6%
Year	-0.003	-0.019	-0.8%	-4.4%
Constant	0.000	0.946	0.0%	216.4%

Notes: Blinder-Oaxaca decompositions based on the results from the FE model.

Table7 Decomposition results of the income gap between developed and less-developed counties (using the number of Taobao villages)

	Value		Percentage (%)	
	Explained	Unexplained	Explained	Unexplained
Total	**-0.004**	**0.442**	**-1.0%**	**101.0%**
Number of Taobao village	**0.000**	**0.001**	**0.1%**	**0.1%**
Other factors	**-0.004**	**0.441**	**-1.1%**	**100.9%**
Gov	-0.018	-0.046	-4.2%	-10.5%
Finance	0.001	-0.025	0.2%	-5.7%
Sec_gdp	0.002	0.028	0.4%	6.3%
Ter_gdp	0.003	0.025	0.8%	5.7%
Invest	-0.005	0.008	-1.1%	1.9%
lnedu	-0.001	-0.445	-0.1%	-101.7%
lnbed	0.004	0.015	0.9%	3.3%
lnpc_land	0.013	0.083	3.0%	18.9%
Year	-0.003	-0.024	-0.9%	-5.3%
Constant	0.000	0.823	0.0%	188.0%

Notes: Blinder-Oaxaca decompositions based on the results from the FE model.

Appendix Table A1 Definition and descriptive statistics of variables

Variable	Definition	Obs.	Mean	Std. Dev.	Min	Max
County id	county id	2,800	175.5	101.05	1	350
Year	Year	2,800	2014.5	2.29	2011	2018
Province id	province id	2,800	3.71	1.19	1	5
Income gap	the square of the difference of real per capita income of rural residents of the county and real per capita income of the province to which the county belongs divided by real per capita income of the province to which the county belongs	2,607	0.13	0.34	2.68E-10	6.51
Taobao Village	at least one Taobao Village in a county=1	2,800	0.12	0.33	0	1
Number of Taobao Villages	the number of Taobao Village in a county	2,800	1.12	6.79	0	134
Sec_gdp	the added value of secondary sector to GDP	2,778	0.47	0.13	0.01	0.89
Ter_gdp	the added value of tertiary sector to GDP	2,778	0.39	0.12	0.08	0.96
GDP per capita	real GDP per capita (Yuan)	2,778	44881.54	31107.10	5258.53	198552.50
Invest	investment in fixed asset/GDP	2,598	0.82	0.36	0.11	4.58
Finance	loans of financial institutions/GDP	1,752	0.67	0.41	0.11	2.04
Gov	local public financial expenditure/GDP	2,777	0.19	0.19	0.02	4.19
Agri_machine	total power of agricultural machinery (kw)	2,263	596063.80	459760.30	12.64	2829220.00
pc_land	cultivate area per capita	2,765	13.42	7.51	0.01	55.96
Edu	number of students enrolling in regular secondary school per 100,000 persons	2,782	5254.89	1522.88	1953.4	11321.54
Bed	number of beds in health institutions per 1,000 persons	2,770	4.56	2.51	0.95	37.47
Website	the number of websites in a province	2,800	140008.20	101132.90	3000.00	417000.00
Rural_internet	the provincial ratio of rural broadband subscribers to the total broadband subscribers	2,800	0.27	0.08	0.08	0.43

Source: Creation by authors.