

2020

## Internet penetration and consumption inequality in China

Shu Zhang

Feng Li

Jing Jian Xiao

*University of Rhode Island, jjxiao@uri.edu*

Follow this and additional works at: [https://digitalcommons.uri.edu/hdf\\_facpubs](https://digitalcommons.uri.edu/hdf_facpubs)

**The University of Rhode Island Faculty have made this article openly available.  
Please let us know how Open Access to this research benefits you.**

This is a pre-publication author manuscript of the final, published article.

### Terms of Use

This article is made available under the terms and conditions applicable towards Open Access Policy Articles, as set forth in our [Terms of Use](#).

### Citation/Publisher Attribution

Zhang, S, Li, F, Xiao, JJ. Internet penetration and consumption inequality in China. *Int J Consum Stud.* 2020; 00: 1– 16. doi: 10.1111/ijcs.12575

Available at: <https://doi.org/10.1111/ijcs.12575>

This Article is brought to you for free and open access by the Human Development and Family Science at DigitalCommons@URI. It has been accepted for inclusion in Human Development and Family Science Faculty Publications by an authorized administrator of DigitalCommons@URI. For more information, please contact [digitalcommons@etal.uri.edu](mailto:digitalcommons@etal.uri.edu).

---

# **Internet Penetration and Consumption Inequality in China**

(Published in *International Journal of Consumer Studies*, DOI: 10.1111/ijcs.12575)

Shu Zhang, Feng Li, Jing Jian Xiao

## **Abstract**

Growing research shows that information technology accelerates economic growth and development, but the effect of internet penetration on inequality is less well documented, especially about consumption inequality. On the one hand, internet lowers transaction costs and offers equal access to online products especially beneficial for remote and poor populations, seemingly reducing inequality. On the other hand, uneven access to the internet may increase divergences. This study examines the relationship between internet penetration and consumption inequality. Using data from 155 counties available from 2010-2016 China Family Panel Studies, this study examines whether internet penetration potentially impacts consumption inequality considering regional heterogeneity. Based on fixed-effect models and the two-stage least squares regressions, results suggest the internet penetration may increase consumption inequality measured by the Gini index. Furthermore, higher education and over a certain internet penetration rate buffer the positive impact of the internet. In some cases, the internet has smaller positive or even negative impacts on consumption inequality in regions with higher education levels and over threshold penetrations.

**Key Words:** Gini index, consumption inequality, internet penetration

# 1. Introduction

Despite the sustained high economic growth in China, the share of household consumption in GDP is declining and inadequate (Aziz and Cui, 2007). The household consumption ratio (% of GDP) has decreased from 46.9% in 2000 to 35.4% in 2010. Although increasing slightly to 38.7% in 2018, it is still lower than most of the countries, including western countries like the US (67%) and Germany (52%), Asian countries like Japan (56%) and Korea (49%), and developing countries like India (59%) and Brazil (64%) (World Bank, 2019).

Besides, widening consumption disparities are one of the most remarkable social concerns in recent years. In China, consumption inequality has an increasing wave-like trend over last three decades (Ding & He, 2018; Xia, Li, & Song, 2017) and is more severe within inland areas, poor areas, and higher-education groups (Qu & Zhao, 2008; Zhao et al., 2017). Research on the reduction of the disparities is necessary for balanced economic developments.

The diffusion of the internet has led to an increasing but possibly unbalanced economic development. With the accelerated information spreading and improved job matching, information technology facilitates economic growth (Koutroumpis, 2009; Czernich et al, 2011) as well as per capita consumption. On the one hand, reducing dramatically costs of economic activities in isolated areas, it enables suppliers to reach consumers and obtain profits and productivity over long distances (Forman et al., 2005), and offering job opportunity especially for job seekers in poor and rural areas (Aker & Mbiti 2010; Kuhn & Mansour, 2014). On the other hand, the widespread use of

technology generates wage growth only in well-off regions with high incomes and large population sizes, causing greater interregional divergences (Forman et al., 2012).

Besides, due to the existence of the digital divide, not all residents have the same access to and skilled use of information and communications technologies (ICT). The digital divide has been identified in various dimensions such as among countries, regions, and sociodemographic (race, age, education, income, etc.) groups (Song, 2008; Xia & Lu, 2008; Lin et al., 2017; Hwang & Nam, 2017). Geographically, technologies like the internet are invented and diffused from the large metropolitan areas with dominant shares of telecommunication infrastructure, and will reinforce these centralizing tendencies (Hwang, 2004). Thus, whether the internet reinforces or reduces consumption inequality still needs to be examined.

Unlike previous research mostly concentrating on income inequality, this study attempts to investigate the relationship between internet penetration and consumption inequality. We use the data of 155 counties available from 2010-2016 China Family Panel Studies (CFPS) and China Statistical Yearbook. To solve the problem of potential endogeneity, fixed-effect models and the two-stage least squares (2SLS) regressions are utilized. Following Gao, Zang, and Sun (2018), we explore the change of provincial internet penetration rate, caused by an exogenous policy shock (Cuncun Tong project), as the instrumental variable of county-level internet penetration. The Cuncun Tong (CT) project, with the purpose of providing villages with the internet accessibility, is planned by the provincial governments and implemented by state-owned companies. The 2SLS estimations suggest that the positive impact of the internet on consumption inequality is buffered by regional demographics and penetration level. Regions with higher education

levels show fewer inequality increases caused by the internet. Besides, fewer inequalities are generated in regions with a higher internet penetration rate. When the buffering effect is strong enough, the internet causes small consumption inequality increases in some advanced counties. Finally, the outcomes of a robustness check are consistent with the main results.

This study is similar to the work by Hounghonon and Liang (2018), which documents the negative effect of broadband internet on income inequality in counties of the United States. Compared with their research, this study has unique contributions to the literature. First, we focus on consumption inequality rather than income inequality, as consumption is a better indicator than income for measuring the welfare (Meyer & Sullivan, 2003; Attanasio & Pistaferri, 2016). Second, we discuss conceptually advantages and disadvantages (digital divide) of internet diffusion that may co-exist, which may cause bidirectional impacts of the internet on consumption inequality. Empirical analyses based on the data from the emerging market (China) are conducted. Different from the negative effect from the work by Hounghonon and Liang (2018), our results suggest internet penetration may increase consumption inequality within a region at the early stage of internet development. Third, we consider the buffering effects of regional factors, especially education. In developed regions with higher education levels and higher penetration rates, the internet may result in smaller consumption inequality increases.

The rest of the paper is organized as follows. Section 2 presents the literature review and hypotheses. Section 3 describes the data and empirical strategies and section 4 presents basic results. Section 5 focuses on heterogeneity and section 6 concludes.

## **2. Literature Review and Hypotheses**

### **2.1 The Internet and E-Commerce in China**

Since connected to the global network in 1994, the internet of China has been taken off phenomenally. The number of internet users has increased more than 480 times, from 620,000 in 1997 to 298 million in 2008 (CNNIC, 2009), the largest of the world. However, despite the rapid growth, the internet penetration rate, 22.6%, was far below the advanced economies (Li & Shiu, 2012), and the severe digital divide exists (Song, 2008). For example, the coastal provinces have high internet penetration rates, while the rates of the central and western provinces are lower than the national average. Besides, the rural areas have less access to the internet and slower diffusion rates (Xia & Lu, 2008; Lin et al., 2017).

In 2006, the central government promulgated “Information Development Strategy 2006-2020” and the 11<sup>th</sup> Five-year Plan to guide the development of the internet. The internet services are mainly provided by the largest state-owned telecommunication companies (China Telecom, China Unicom and China Mobile), which accounted for over 90% of the internet user population in 2018.

Since 2005, the Cuncun Tong (CT) project, an exogenous policy shock aimed to provide villages with public infrastructures, has focused on the broadband internet accessibility. At the beginning of the year, provincial governments made the plan about the internet connection, which was then implemented by the three state-owned companies. The government plans and company establishment are the primary determinants of the change of internet availability at provincial level, making the CT project an exogenous shock (Gao et al., 2018). The counties within provinces are also

subject to this policy. Due to the lack of statistics about counties, we utilize the change of provincial internet penetration as the instrument variable of county internet level. The validity of the IV will be explained in section 4.2.

The internet development has brought about the rise of e-commerce. In China, the nominal value of online sales increased from RMB 10,400 billion in 2013 to RMB 29,160 billion in 2018, with an annual growth rate of 29.4% (CNNIC, 2019). Around 73.6% of internet users have purchased online, contributing to 18.4% of total retail sales. Besides, residents from smaller cities spend more disposable income online (Dobbs et al., 2013), and the growth rate of online shoppers is larger in rural (19.8%) than urban (12.9%) in 2015 (CNNIC, 2016).

## **2.2 Internet and Income Inequality**

Generally, the advent of internet technology has transformed the economy and may impact inequality within a country. Providing that improving access to information with lower costs, the diffusion of the internet would enhance economic development (Koutroumpis, 2009; Czernich et al., 2011; Billon et al., 2018) and wage growth (Forman et al., 2012). Reducing dramatically the costs of economic activities in isolated areas, it enables suppliers to reach consumers and gain profits and productivity over long distances (Forman et al., 2005), and also offers job opportunities especially for job seekers in poor and rural areas (Aker & Mbiti, 2010; Kuhn & Mansour, 2014). Besides, the internet exhibits increasing returns due to network externalities (Koutroumpis, 2009), strengthens social capital and ICT skills for job seekers to alleviate poverty (Blanco & López Bóo, 2010; Richmond & Triplett, 2017).

Empirical studies present mixed evidence on how the internet impacts income inequality. An empirical study exhibits a significantly positive relationship between the ICT stock capital share with income inequality in 51 countries (Jaumotte et al., 2013). However, other researchers provide opposite evidence. Ningsih and Choi (2018) contend that internet penetration reduced income inequality in Southeast Asian countries. Using data from EU-27 countries, Asteriou et al. (2014) find that increasing ICT spending share of GDP is associated with decreasing income inequality, but only significant in high technology countries. Richmond and Triplett (2014) conclude that the association conditionally depends on the type of access; Internet and mobile use are negatively, while broadband penetration is positively linked with income inequality. Thus, the effect of the internet on income inequality seems mixed and conditional on regional characteristics.

### **2.3 Internet, E-Commerce and Consumption Inequality**

The consumption inequality closely tracks income inequality in China (Ding and He, 2018), and the internet also matters for the distribution of consumption. With emerging e-commerce changing consumer behavior substantially, several mechanisms through which e-commerce may benefit consumers. Firstly, e-commerce platforms provide consumers with greater varieties of products (Brynjolfsson et al. 2003), including products unavailable offline and products of new companies entered. Secondly, the introduction of the internet or e-commerce might improve the match quality between consumers and products (Glenn & Ellison, 2014) and reduce prices (Brown & Goolsbee, 2002). Furthermore, the increased competition may cause reduced price dispersions in online retailing: Yet, there are persisting substantial price dispersions across online stores



(Clay et al., 2002). Affluent consumers can avoid trips to physical stores and gain convenience benefits from online shopping (Dolfen et.al, 2019).

For the impact of the internet, previous literature used to pay attention to the general dispersions of countries, while more recent literature points to consumption divergences across geographic and social groups within regions, including counties. Geographically, remote areas far from the markets benefit more from e-commerce. Combing e-commerce sales data with city-level characteristics in China, Fan et al. (2018) manifest that consumers in the smallest and remote cities spend the greatest proportions of purchases online and gain the largest welfare (real wage growth). Based on household data, Couture et al. (2018) conclude that the introduction of e-commerce in villages leads to a reduction in the cost of living especially for households far from the township centers. Also, Luo et al. (2019) find that the relationship between consumption growth and county e-commerce development is stronger for rural and inland households. Among socioeconomic groups, poor households have the largest consumption growth (Luo et al., 2019), while younger and richer households obtain higher consumer gains (Couture et al., 2018). Based on the empirical studies, we extend the study on the relationship of the internet and the consumption inequality to county-level analysis.

Internet may affect consumers through two mechanisms: wage growth and equal access to products. For low quintile group, the poor and remote workforce may reap higher wage growths with the internet thus narrowing wage gaps (Aker & Mbiti 2010; Kuhn and Mansour, 2014), then resulting in purchasing power convergence. Besides, the internet provides consumers with equal access to products far away. In contrast to the high penetration of large, big-box retail stores in developed countries, China's traditional

retail industry is less developed and centered in metropolitan areas (Fan et al., 2018). To make up the fixed costs of opening new stores and transportation costs along with, big retailers are more likely to gain profits entering big markets with large demands. Thus, regions with fewer population densities suffer from limited product varieties. The equal accessibility to products through the internet might break down the unequal distribution of retail infrastructure and unleash the pent-up demands. However, despite the widespread use of the internet, the digital divide is a serious phenomenon within regions (Salemink et al., 2017). Absent or unqualified internet users may benefit less for wage growth, missing their spots in the “peer-to-peer” economy associated with Uber, Airbnb and other websites (Robinson, 2015), increasing consumption inequality in the information era. Thus, we propose the following alternative hypotheses:

**Hypothesis1a:** Internet penetration increases consumption inequality.

**Hypothesis1b:** Internet penetration decreases consumption inequality.

Since the digital divide and regional heterogeneity exist, effects of the internet may differ among regions. Usually, innovations such as information technology are invented in and diffused from advanced areas to peripherals, causing the digital divide. When the internet starts to spread, the positive externalities kick in when reaching a certain threshold and the effect grows with increasing adoption (Koutroumpis,2009; Galperin & Vicens, 2017). The internet investment seems only cause wage growth in advanced counties where per capita income, population size, and IT usage are salient factors before the internet taking off (Forman et al., 2012). In practice, these advanced counties have the advantage of internet adoption and benefit first from the threshold effect and reap fast growth. Due to the coordination and market forming of the internet

economy, the advantaged counties are centered by information industries and absorb much more workforces. It is even harder for less advantaged counties to develop and narrow the growth gap. Besides the technological intervention, the effect of ICT also interacts with economic and political factors that are combined to determine the distribution of consumption (Bauer, 2018). Further, the development of e-commerce is supported by logistics, highways, and other related infrastructures, resulting in substantial incremental effects in advanced areas.

These advanced areas may be salient in some pertinent regional factors. First, GDP per capita is associated with technology changes, labor market transitions, and product supplements, probably exacerbating disparities (Seven and Coskun, 2016; Richmond and Triplett, 2018). Second, regions with larger population sizes have larger labor markets for technology adoption, productivity improvement (Forman et al., 2012) as well as consumer benefits. Besides, the age structure is also one of the main factors for inequality (Houngbonon & Liang, 2018), as the young generation tends to utilize technology and skills more than the elder. Education can improve the accessibility and skilled usage of the internet including information searching (Ertiö & Räsänen, 2019). Not only does the education reduce the digital divide (Li & Shiu, 2012), but it also narrows the wage gap, which is essential for poverty alleviation. Finally, the threshold effect of the internet means that the high penetration rate may exhibit positive externalities. The developed areas, which are in the top quartile in these characteristics, may benefit more and gain more from the increasing internet penetration. Thus, we propose the following hypothesis:

**Hypothesis 2:** The effects of internet penetration on consumption inequality are moderated by per capita GDP, population size, age structure, education, and high internet level.

## **3. Method**

### **3.1 Data and Variables**

The data used in this study come from the China Family Panel Studies (CFPS) and China Statistical Yearbook. The CFPS is a nationally representative micro-household survey conducted by the China Social Science Research Center of Peking University. The CFPS survey sample covers 14,798 households in 162 counties among 25 provinces across the country. The households were chosen randomly at village level (Xie & Hu, 2014). By collecting individual, household, and community data, this dataset provides information for academic research and public policy analyses. At present, CFPS has been carried out four waves in the year 2010, 2012, 2014 and 2016. In this study, county-level panel data from all four waves were used.

The CFPS surveys consist of household and personal level questionnaires, thus data processing is necessary to compile county-level data. First, we combine personal level with household level data to collect information about internet accessibility. Second, variables based on household level data are averaged at the county level. Third, we only keep the counties investigated from the first wave, obtaining samples who participated in all waves. Finally, there are 155 counties and 569 observations from the year 2010 to 2016<sup>1</sup>.

---

<sup>1</sup> More details about data processing are shown in the Appendix.

*Consumption inequality* is the dependent variable at the county level.

Consumption inequality can be measured by the Gini coefficient which is widely used in previous research. Followed by Gini (1921) and Yao (1999), the definition of the Gini coefficient is presented as follows:

$$\text{Gini} = 1 - \sum_{i=1}^n p_i (2 \sum_{k=1}^i Q_k - w_i) \quad (1)$$

Within each county,  $w_i$  and  $p_i$  represent, respectively, the consumption share and relative population frequency of the household  $i$  ( $i = 1, 2, \dots, n$ ). Besides,  $Q_i = \sum_{k=1}^i w_k$  is the cumulative consumption share up to  $i$ . Before calculation, households should be arranged to follow the ascending order of mean household consumption. According to China Statistical Yearbook, the Gini coefficients of income inequality were 48.1, 47.4, 46.9, and 46.5 for the year 2010, 2012, 2014 and 2016, respectively. From Table 1, the mean values of consumption inequality of counties were 45.8, 47.7, 41.5 and 44.0, which shows a similar decline trend from 2012 to 2016.

*Internet* is the key factor in this study. We use the log of the number of households with internet connections per 100 households in one county to represent the degree of internet availability at the county level. In the dataset, internet users may have access to the internet through either smartphones or laptops. From Table 1, internet penetration keeps growing during the period.

Control variables include county and provincial level variables. The county-level variables including education are shown in Table 1, and provincial control variables in Table 2. Education is measured by the ratio of high school graduates to the total population in each county and may reduce consumption disparities (Jaumotte et al., 2008; Asteriou et al., 2014). Economic development is controlled by GDP per capita and

industrial structure, as industrialization and service sector development can boost economic growth. The industrial structure is measured by the compositions of secondary and tertiary industries in provinces' GRP separately. Previous studies show that GDP per capita (Seven & Coskun, 2016; Richmond & Triplett, 2018) and shares of agriculture and manufacturing industry (Jaumotte et al., 2008; Asteriou et al., 2014) may widen social inequality. Per capita consumption is used to measure the average consumption level. Age structure is represented by the ratio of the elder generation with age above 65 to the total provincial population, following Hounghonon and Liang (2018). All variables are adjusted with the CPI index as a deflator. Urban rate is measured by the proportion of the urban population that may impact inequality at the national and city levels (Kanbur & Zhuang, 2013; Li et al., 2019; Wang, Shao and Li, 2019). To control regional agglomerative economies' effects, we include population size, measured by the logarithm of the provincial population. Unemployment rate and foreign direct investment (logged form) are influential factors on wage distribution as well as consumption. Unemployment and openness may reinforce disparities (Lee & Kim, 2016).

### 3.2 Models

We focus on the internet penetration and consumption inequality among counties. Based on hypothesis 1, internet penetration might have an impact on consumption disparity. Based on the model from Hounghonon and Liang (2018), we formulate the fixed effect regression model:

$$\ln Gini_{it} = \alpha + \beta internet_{it} + \gamma C_{it} + v_t + \mu_i + \varepsilon_{it} \quad (1)$$

The outcome variable  $\ln Gini_{it}$  captures consumption inequality in county  $i$  in year  $t$ . The key variable  $internet_{it}$  is measured by the log of the number of households

per 100 households connected with internet in county  $i$  in year  $t$ .  $C_{it}$  represents control variables measuring regional conditions including education, GDP per capita, industrial structure, consumption, population, age structure, FDI and unemployment rate. Last,  $v_t$  and  $\mu_i$  are the dummy variables for year fixed effects and county fixed effects.

However, this model has some problems to account for the causal relationship. First, reverse causality may exist, causing biased estimators. Besides, some omitted variables may simultaneously impact both internet penetration and consumption inequality. Thus, we add more fixed effects to control for unobserved fixed factors, including province-year fixed effects. Furthermore, we use an instrumental variable and 2SLS method to identify potential causality. Following Gao et al. (2018), the change of provincial internet is utilized as the instrumental variable. The test of causality and justification of the IV will be discussed in section 4.2.

## **4. Results**

### **4.1 Baseline model**

The results of the baseline regression are shown in Table 3. Column 1 measures the association between internet penetration and inequality with county and time fixed effects, showing a significantly positive correlation. After controlling province-year fixed effects in Column 2, the association is still significant. The estimated coefficient 0.0718 means counties with one percent increase of internet penetration rate, will experience 0.0718 percent increase of Gini index of consumption inequality. Based on the estimated mean value of internet penetration rate and Gini index, we can calculate the real percentage variations. Ten percentage points increase of internet penetration rate are associated with 0.738 ( $10 \times 0.0718 \times 0.0718 / 44.795$ ) change in Gini index of consumption

inequality, a little larger than estimated result 0.36 of income inequality estimated by Hounghonon and Liang (2018).

The results imply that more households' accesses to the internet may result in higher consumption inequality among counties. In column 3, we include all control variables, provincial GDP per capita significantly increase consumption inequality, while consumption per capita reduces it. As the endogeneity problem exists, the IV is utilized to demonstrate the causality in the following section.

## 4.2 Evidence on the Causal Direction

As mentioned before, internet penetration may be endogenous as advanced regions invest more in broadband infrastructure. We use the instrument variable and the 2SLS regressions to address the problem of the measurement-error bias and omitted variables bias (Angrist & Pischke, 2009). The regression models used can be written as follows:

$$internet_{it} = \alpha + \beta \Delta PI_{it} + \gamma C_{it} + v_t + \mu_i + \varepsilon_{it} \quad (2)$$

$$\ln Gini_{it} = v + \lambda \widehat{internet}_{it} + \phi C_{it} + v_t + \mu_i + \varepsilon_{it} \quad (3)$$

where  $\Delta PI_{it}$  denotes the change of the provincial internet penetration rate, and  $\widehat{internet}_{it}$  is the predicted internet penetration rate at the county level. Equation (2) and (3) are first-stage and second-stage estimations, respectively. In the first stage, we use the change of provincial internet ( $\Delta PI_{it}$ ) as the IV to predict county-level internet penetration rate ( $\widehat{internet}_{it}$ ). In the second stage, we explore the impact of the predicted internet ( $\widehat{internet}_{it}$ ) on consumption inequality. The estimated outcomes of two stages are shown in column 1 and 2 of Table 4. Before we discuss estimated results, several related tests



should be processed to address the qualification of the IV, including the correlation, independence, and exclusion.

First, the IV should be correlated with the endogenous variable. From column 1 of Table 4, the change of provincial internet has a close association with county-level internet penetration rate and the explained variance is adequate (R-square value=0.719). Besides, the correlation justification is supported by the under-identification test (P-value=0.0016) and the weak identification test. Second, the IV should be independent of omitted variables. The correlation of the IV with the residuals estimated from equation (1) is statistically insignificant (P-value=0.256). Third, the exclusion condition means the IV can only impact consumption inequality through the internet. As emphasized before, the change of provincial internet penetration rate is exogenous policy shock and has no direct influence on county-level consumption distribution. Thus, the IV is qualified to solve the potential endogeneity problem.

In column 2 of Table 4, the  $\widehat{internet}$  represents the predicted internet penetration rate at the first stage. The estimated coefficient of the impact of predicted internet on consumption inequality is 0.587, higher than previous baseline regressions in Table 3. One potential explanation is that measurement error leads to a downward bias of baseline regressions (Zhang, Liu and Yung, 2007). The internet increases consumption inequality significantly, supporting hypothesis 1a.

## **5. Heterogenous Effect and Robustness Check**

### **5.1 Heterogeneity Caused by Regional Factors**

We investigate the heterogeneous effects of internet penetration among four regional conditions: GDP per capita, education, population size, and age structure.

Specifically, we interact the predicted internet penetration rate with these heterogeneous items separately, shown in Table 5. The increasing effect of internet penetration diminishes in counties with higher education levels, where equal technology diffusion exists, and the advantages of the internet is unleashed. Education is vital to avoid the digital divide and narrow consumption gap. However, the moderate effects of regional GDP per capita, population size and age structure are negatively but insignificant. One possible reason is that provincial control variables have limited explanations for consumption distribution. In conclusion, these counties with more education levels enable lower social classes to make full use of the internet for income-generating and consumption beneficial activities.

## 5.2 Nonlinear Effect of the Internet

Previous research shows the existence of the non-linear effect of the internet with a certain threshold, which is identified as 10% (Czernich et al., 2011) and 30% (Houngbonon & Liang, 2018). We test the nonlinear relationship of the internet on consumption inequality using the following model:

$$\ln Gini_{it} = \alpha + \beta_1 \widehat{internet}_{it} + \beta_2 \widehat{internet}_{it} * D_{it} + \delta D_{it} + \phi C_{it} + v_t + \mu_i + \varepsilon_{it} \quad (4)$$

In the equation above,  $D_{it}$  is a dummy variable when the internet penetration is above a certain level. After testing threshold effects, we find that a significant threshold effect exists when internet penetration above 46%. Suggested in column 6 in Table 6, the interaction between internet and high penetration rate status may reduce consumption inequality.

## 5.3 Heterogeneity of Subgroups

Since the buffering effect of education and high internet penetration, we wonder whether there is potential heterogeneities where the internet may reduce inequality. Thus, we define a well-developed county of which the internet penetration is over 46% and the education ranks among the top 5% counties. As shown in equation 5,  $H_{it}$  equals one when the county is defined as well developed. The estimation of the interaction between  $H_{it}$  and  $internet_{it}$  is listed in column 6 of Table 6, which is negatively significant. Compared to other counties, the internet causes smaller consumption inequality increases in the highly developed counties. With higher education and technology penetration, the internet can strength its advantages and benefit consumers more evenly causing small increases in consumption inequality. However, for places lack of education and technology, the penetration of the internet would generate a larger gap and make the situation worse.

$$\ln Gini_{it} = \alpha + \beta_1 \widehat{internet}_{it} + \beta_2 \widehat{internet}_{it} * H_{it} + \delta H_{it} + \phi C_{it} + v_t + \mu_i + \varepsilon_{it} \quad (5)$$

## 5.4 Robustness Check

For robustness check, we use two methods separately. First, a Logit regression model is established to test the propensity to be a higher inequality group. The dependent variable is a binary variable which equals one when a county has higher than average consumption inequality. In table 6, counties with higher internet show the significant propensity of being in a high Gini group. The estimated results of the 2SLS regressions exhibit the causal relationship and buffering effects of regional characteristics, consistent with previous results.

The second robustness check is based on the cross-sectional data at county-level from CFPS in the year 2010, which is the only wave provided with detailed county-level

control variables. The OLS regression model is used and the estimated results of baseline regression are shown in Table 7, presenting a significant positive impact of internet penetration. The buffering effect of high education (the top 5%) and high penetration rate (over 62.88%)<sup>2</sup> is presented in Table 8. Furthermore, the advanced counties are re-defined as those with the top 5% education level and high penetration rate. Compared to other counties, there is a significantly negative association between internet and consumption inequality in advanced counties, suggesting internet penetration may decrease consumption inequality.

## **6. Conclusions**

This study examines the potential impact of internet penetration on consumption inequality. With county-level data from 2010-2016 CFPS, results suggest that internet penetration may have a significantly positive effect on consumption inequality. However, the buffering effects exist where high education rates and high internet penetration rate (over 46%) will reduce the positive impact of internet on consumption inequality. Improving education and raising the penetration rate of the internet may result in more people benefited, causing smaller positive impact on consumption inequality. Furthermore, based on a cross sectional data from wave 2010, the OLS estimation shows that there is a negative association between internet and inequality in advanced counties. That is, the internet may reduce consumption inequality for counties with the high education level and high internet penetration rate.

---

<sup>2</sup> The threshold effect is re-tested based on data from the year 2010.

These results shed light on the internet and equal development. Besides massive broadband infrastructure investment, the adoption and skilled usage of the internet also matter to reduce the digital divide. Education not only reduce inequality but also buffer the positive effect of the internet on consumption inequality. Thus, internet literacy and related education should be enhanced to narrow digital divide. Meanwhile, speeding up the construction of the high-speed broadband network and reducing related fees, as General Office of the State Council suggested (2015), will benefit rural consumers' well-beings and bring balanced economic developments. To take full advantage of ICT, related policies should be made to support the development of internet-based industries such as e-commerce, and allocate resources among production, logistics, and distribution sectors. Provided with a better economic environment for internet, consumers would be more equal to establish and maintain close relationships with firms in the era of information economy.

Several limitations exist in this study. More waves with detailed data should be included and examined for further research. For example, the impact of the internet on categories of consumption, such as housing expenditure, can be explored. Besides, detailed geographic information will allow studies on the spatial autocorrelation. Also, further research could examine the urban-rural consumption gap. In this paper, we assume the internet may impact consumers through two mechanisms by wage growth and equal access to e-commerce. However, these two mechanisms still need to be tested with more specific variables. For examples, on what conditions low quintiles will gain more wage growth from internet adoption? Besides, which subgroups of consumers benefit more from e-commerce or online shopping? The mechanisms behind the relationship

between internet penetration and consumption inequality can be addressed in future research.

## Reference

- Aker, J. C., & Mbiti, I. M. (2010). Mobile phones and economic development in Africa. *Journal of Economic Perspectives*, 24(3), 207-32.
- Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.
- Asteriou, D., Dimelis, S., & Moudatsou, A. (2014). Globalization and income inequality: A panel data econometric approach for the EU27 countries. *Economic modelling*, 36, 592-599.
- Attanasio, O. P., & Pistaferri, L. (2016). Consumption inequality. *Journal of Economic Perspectives*, 30(2), 3-28.
- Aziz, M. J., & Cui, L. (2007). *Explaining China's low consumption: the neglected role of household income*. International Monetary Fund Working Paper no. 07/181.
- Bauer, J. M. (2018). The Internet and income inequality: Socio-economic challenges in a hyperconnected society. *Telecommunications Policy*, 42(4), 333-343.
- Billon, M., Crespo, J., & Lera-López, F. (2018). Educational inequalities: do they affect the relationship between Internet use and economic growth?. *Information Development*, 34(5), 447-459.
- Blanco, M., & López Bóo, F. (2010). Ict skills and employment: A randomized experiment. *IZA Discussion Paper No. 5336*.
- Brown, J. R., & Goolsbee, A. (2002). Does the Internet make markets more competitive? Evidence from the life insurance industry. *Journal of political economy*, 110(3), 481-507.
- Brynjolfsson, Erik, Yu, Hu, Smith, Michael D., 2003. Consumer surplus in the digital economy: estimating the value of increased product variety at online booksellers. *Manag.Sci.* 49 (11), 1580–1596.
- China Internet Network Information Center (CNNIC), 2009. The 23th China Statistical Report on Internet Development.  
[http://www.cnnic.net.cn/hlwfzyj/hlwzbg/index\\_4.htm](http://www.cnnic.net.cn/hlwfzyj/hlwzbg/index_4.htm)

- China Internet Network Information Center (CNNIC), 2016. The report on China's rural Internet development in 2015.  
<http://www.cnnic.net.cn/hlwfzyj/hlwxyzbg/ncbg/201608/P020170907348967498375.pdf>
- China Internet Network Information Center (CNNIC), 2019. The 43rd China Statistical Report on Internet Development.  
<http://www.cnnic.net.cn/hlwfzyj/hlwxyzbg/hlwtjbg/201902/P020190318523029756345.pdf>
- Clay, K., Krishnan, R., Wolff, E., & Fernandes, D. (2002). Retail strategies on the web: Price and non-price competition in the online book industry. *The Journal of Industrial Economics*, 50(3), 351-367.
- Couture, V., Faber, B., Gu, Y., & Liu, L. (2018). E-Commerce Integration and Economic Development: Evidence from China. *National Bureau of Economic Research*, No. w24384.
- Czernich, N., Falck, O., Kretschmer, T., & Woessmann, L. (2011). Broadband infrastructure and economic growth. *The Economic Journal*, 121(552), 505-532.
- Ding, H., & He, H. (2018). A tale of transition: An empirical analysis of economic inequality in urban China, 1986–2009. *Review of Economic Dynamics*, 29(3), 106-137.
- Dobbs, Richard, Chen, Yougang, Orr, Gordon, Manyika, James, Chui, Michael, Chang, Elsie, 2013. China's E-tail Revolution: Online Shopping as a Catalyst for Growth. The McKinsey Global Institute.
- Dolfen, P., Einav, L., Klenow, P. J., Klopock, B., Levin, J. D., Levin, L., & Best, W. (2019). Assessing the Gains from E-Commerce. *National Bureau of Economic Research Working Paper*, No. w25610.
- Ellison, G., & Ellison, S. F. (2018). Match quality, search, and the Internet market for used books. *National Bureau of Economic Research*, No. w24197.
- Ertiö, T., & Räsänen, P. (2019). Consumerism in online health information search and self-tracking devices. *International Journal of Consumer Studies*, 43(3), 245-252.
- Fan, J., Tang, L., Zhu, W., & Zou, B. (2018). The alibaba effect: Spatial consumption inequality and the welfare gains from e-commerce. *Journal of International Economics*, 114, 203-220.

- Forman, C., Goldfarb, A., & Greenstein, S. (2005). How did location affect adoption of the commercial Internet? Global village vs. urban leadership. *Journal of urban Economics*, 58(3), 389-420.
- Forman, C., Goldfarb, A., & Greenstein, S. (2012). The Internet and local wages: A puzzle. *American Economic Review*, 102(1), 556-75.
- Galperin, H., & Fernanda Viacens, M. (2017). Connected for Development? Theory and evidence about the impact of Internet technologies on poverty alleviation. *Development Policy Review*, 35(3), 315-336.
- Gao, Y., Zang, L., & Sun, J. (2018). Does computer penetration increase farmers' income? An empirical study from China. *Telecommunications Policy*, 42(5), 345-360.
- General Office of the State Council. (2005). Guiding Opinions on Accelerating the Construction of High-speed Broadband Network to Increase Internet Speed and Cut Service Charges (Guo Ban Fa [2015]. No 41).
- Gini, C. (1921). Measurement of inequality of incomes. *The Economic Journal*, 31(121), 124-126.
- Houngbonon, G. V., & Liang, J. (2018). Broadband Internet and Income Inequality. *Social Science Research Network*, No. 2963860.
- Hwang, H., & Nam, S. J. (2017). The digital divide experienced by older consumers in smart environments. *International journal of consumer studies*, 41(5), 501-508.
- Hwang, J. S. (2004). Digital divide in Internet use within the urban hierarchy: The case of South Korea. *Urban Geography*, 25(4), 372-389
- Jaumotte, F., Lall, S., & Papageorgiou, C. (2013). Rising income inequality: technology, or trade and financial globalization?. *IMF Economic Review*, 61(2), 271-309.
- Kanbur, R., & Zhuang, J. (2013). Urbanization and inequality in Asia. *Asian Development Review*, 30(1), 131-147.
- Koutroumpis, P. (2009). The economic impact of broadband on growth: A simultaneous approach. *Telecommunications policy*, 33(9), 471-485.
- Kuhn, P., & Mansour, H. (2014). Is Internet job search still ineffective?. *The Economic Journal*, 124(581), 1213-1233.
- Lee, J. H., & Kim, J. (2016). Do Free Trade Agreements Affect Income Inequality?: An Empirical Investigation. *Journal of International Trade & Commerce*, 12(6), 53-



- Li, J., Cheong, T. S., Shen, J., & Fu, D. (2019). Urbanization and Rural–Urban Consumption Disparity: Evidence from China. *The Singapore Economic Review*, 64(04), 983-996.
- Li, R., & Shiu, A. (2012). Internet diffusion in China: A dynamic panel data analysis. *Telecommunications Policy*, 36(10-11), 872-887.
- Liang, J. (2018). Broadband Infrastructure and Income Inequality \*, (July), 1–30.
- Lin, J., Yu, Z., Wei, Y. D., & Wang, M. (2017). Internet access, spillover and regional development in china. *Sustainability*, 9(6), 946.
- Meyer, B. D., & Sullivan, J. X. (2003). *Measuring the well-being of the poor using income and consumption* (No. w9760). National Bureau of Economic Research.
- Ningsih, C., & Choi, Y. J. (2018). An Effect of Internet Penetration on Income Inequality in Southeast Asian Countries. *22nd ITS Biennial Conference Paper*.
- Qu, Z. F., & Zhao, Z. (2008). Urban-rural consumption inequality in China from 1988 to 2002: Evidence from quantile regression decomposition. *Institute for the Study of Labor (IZA) Discussion Papers*.
- Richmond, K., & Triplett, R. E. (2018). ICT and income inequality: a cross-national perspective. *International Review of Applied Economics*, 32(2), 195-214.
- Robinson, L., Cotten, S. R., Ono, H., Quan-Haase, A., Mesch, G., Chen, W., ... & Stern, M. J. (2015). Digital inequalities and why they matter. *Information, communication & society*, 18(5), 569-582.
- Salemink, K., Strijker, D., & Bosworth, G. (2017). Rural development in the digital age: A systematic literature review on unequal ICT availability, adoption, and use in rural areas. *Journal of Rural Studies*, 54, 360-371.
- Seven, U., & Coskun, Y. (2016). Does financial development reduce income inequality and poverty? Evidence from emerging countries. *Emerging Markets Review*, 26, 34-63.
- Song, W. (2008). Development of the Internet and digital divide in China: A spatial analysis. *Intercultural Communication Studies*, 12(3), 20–43.
- Wang, X., Shao, S., & Li, L. (2019). Agricultural inputs, urbanization, and urban-rural income disparity: Evidence from China. *China Economic Review*, 55, 67-84.

- World Bank, 2019. "Households and NPISHs final consumption expenditure (% of GDP)." Webpage. <https://data.worldbank.org/indicator/NE.CON.PRVT.ZS>.
- Xia, J., & Lu, T. J. (2008). Bridging the digital divide for rural communities: The case of China. *Telecommunications Policy*, 32(9-10), 686-696.
- Xia, Q., Li, S., & Song, L. (2017). Urban consumption inequality in China. *Institute for the Study of Labor (IZA) Discussion Paper Series*, No. 11150.
- Xie, Y., & Hu, J. (2014). An introduction to the China family panel studies (CFPS). *Chinese sociological review*, 47(1), 3-29.
- Yao, S. (1999). On the decomposition of Gini coefficients by population class and income source: a spreadsheet approach and application. *Applied economics*, 31(10), 1249-1264.
- Zhang, J., Liu, P. W., & Yung, L. (2007). The Cultural Revolution and returns to schooling in China: Estimates based on twins. *Journal of Development Economics*, 84(2), 631-639.
- Zhao, D., Wu, T., & He, Q. (2017). Consumption inequality and its evolution in urban China. *China Economic Review*, 46, 208-228.

## Tables

**Table 1. Descriptions of County-Level Variables**

	<b>Description</b>	<b>2010</b>	<b>2012</b>	<b>2014</b>	<b>2016</b>
Num. of counties	The number of counties	150	142	140	137
Sum household number	The mean number of households included in counties	81.893	81.338	79.629	79.044
Gini	Gini index of consumption within counties	45.828	47.745	41.502	43.973
Internet	logarithm of the number of households with internet per 100 households in each county	3.365	3.655	3.833	4.102
Education	The ratio of high school graduates' share of county population (%)	14.137	17.644	21.287	21.327

**Table 2 Descriptions of Provincial Level Variables**

<b>Variable</b>	<b>Description</b>	<b>Obs</b>	<b>Mean</b>	<b>Std. Dev.</b>	<b>Min</b>	<b>Max</b>
GDP	The logarithm of provincial GDP per capita	100	10.427	0.479	9.303	11.374
consumption	The logarithm of provincial consumption per capita	100	9.380	0.443	8.526	10.517
population	The logarithm form of Population size	100	8.422	0.508	7.169	9.306
urban rate	Urbanization rate (%)	100	56.209	13.853	33.810	89.600
age structure	The ratio of elder generation (age>65) (%)	100	9.739	1.616	6.560	14.080
unemployment rate	Unemployment ratio at provincial level (%)	100	3.452	0.649	1.200	4.500
secondary industry share	The share of the secondary industry in provincial GRP (%)	100	47.285	7.485	19.738	59.045
tertiary industry share	The share of the tertiary industry in provincial GRP (%)	100	42.669	9.607	29.265	79.653
FDI	logarithm of foreign direct investment	100	6.580	1.191	3.721	9.082
Provincial internet change	The change in provincial internet penetration (%)	100	8.439	3.859	2.294	20.460

**Table 3. The Effect of Internet Penetration on Consumption Inequality**

	(1) <b>lnGini</b>	(2) <b>lnGini</b>	(3) <b>lnGini</b>
Internet	0.0718*** (2.94)	0.0555* (1.83)	0.0631** (2.09)
education			-0.000432 (-0.20)
GDP			0.703*** (2.68)
consumption			-0.545** (-2.40)
population			-0.511 (-0.77)
urban rate			-0.00191 (-0.26)
age structure			-0.0102 (-0.94)
unemployment rate			-0.0230 (-0.54)
secondary industry share			-0.0153* (-1.75)
tertiary industry share			-0.00668 (-0.71)
FDI			-0.0422 (-0.77)
_cons	3.572*** (44.16)	3.689*** (36.63)	7.364 (1.19)
FE of year	Y	Y	Y
FE of county	Y	Y	Y
FE of year and province		Y	
<i>N</i>	569	569	569
adj. <i>R</i> <sup>2</sup>	0.170	0.225	0.176

Note: *t* statistics in parentheses

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

**Table 4. The Causality of Internet Penetration and Consumption Inequality**

	(1) First-stage internet	(2) Second-stage lnGini
Provincial Internet change	0.0189*** (2.85)	
$\widehat{internet}$		0.587** (2.22)
education	0.0130*** (2.85)	-0.00745 (-1.64)
GDP	0.538 (1.23)	0.293 (0.74)
consumption	-0.315 (-0.80)	-0.313 (-0.94)
population	0.780 (0.59)	-0.550 (-0.53)
urban rate	0.0479*** (4.12)	-0.0281* (-1.74)
age structure	0.0205 (1.52)	-0.0161 (-1.21)
unemployment rate	-0.0208 (-0.27)	-0.00895 (-0.19)
secondary industry share	-0.0430** (-2.24)	0.00664 (0.40)
tertiary industry share	-0.0452** (-2.49)	0.0154 (0.88)
FDI	-0.189 (-0.96)	0.0565 (0.63)
_cons	-3.798 (-0.32)	6.855 (0.71)
FE of time	Y	Y
FE of county	Y	Y
the corr. of residuals and iv (P value)		0.256
LM test for under-identification test (P-value)		9.906 (0.0016)
Cragg-Donald Wald F statistic for weak identification test (15% maximal IV size)		9.806 (8.96)
<i>N</i>	569	569
adj. <i>R</i> <sup>2</sup>	0.719	

Note: *t* statistics in parentheses

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

**Table 5. Heterogeneity of Regional Factors**

<b>lnGini</b>	<b>(1) Provincial per capita GDP</b>	<b>(2) Education</b>	<b>(3) Populatio n size</b>	<b>(4) age structure</b>	<b>(5) Nonlinear effect</b>	<b>(6) Highly developed county</b>
<i>internet</i>	0.771* (1.72)	0.565*** (2.91)	1.070** (2.08)	0.607*** (2.91)	0.520*** (2.67)	0.540*** (2.86)
<i>internet</i> *GDP	-0.0208 (-0.49)					
<i>internet</i> *education		-0.00493*** (-2.69)				
<i>internet</i> *population			-0.0539 (-1.08)			
<i>internet</i> *consumption						
<i>internet</i> *age structure				-0.00478 (-0.44)		
<i>internet</i> *Penetration above 46%					-0.103*** (-2.69)	
Penetration above 46%					0.404*** (2.79)	
<i>internet</i> *Highly developed county						-0.415*** (-5.51)
Highly developed county						1.735*** (5.43)
_cons	6.535 (1.05)	9.941 (1.46)	6.219 (1.09)	6.830 (1.11)	9.268 (1.39)	8.839 (1.49)
FE of time	Y	Y	Y	Y	Y	Y
FE of county	Y	Y	Y	Y	Y	Y

Control variables	Y	Y	Y	Y	Y	Y
<i>N</i>	569	569	569	569	569	569
adj. <i>R</i> <sup>2</sup>	0.184	0.202	0.186	0.184	0.201	0.207

---

Note: *t* statistics in parentheses

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$



**Table 6. A Robustness Check with Logit Regression Model**

High Gini Group	(1) Correlation	(2) Causality	(3) Provincial per capita GDP	(4) Education	(5) Population size	(6) Nonlinear effect	(7) Highly developed county
$\widehat{internet}$	1.312** (2.52)						
$\widehat{internet}$		2.505** (2.26)	14.54* (1.79)	6.471* (1.76)	9.788 (1.00)	5.768 (1.57)	6.083* (1.65)
$\widehat{internet}$ *GDP			-0.928 (-1.14)				
$\widehat{internet}$ *education				-0.0787** (-2.51)			
$\widehat{internet}$ *population					-0.390 (-0.38)		
$\widehat{internet}$ *consumption							
$\widehat{internet}$ * Penetration above 46%						-1.508** (-2.27)	
$\widehat{internet}$ * High developed county							-5.042 (-1.60)
FE of time	Y	Y	Y	Y	Y	Y	Y
Fe of county	Y	Y	Y	Y	Y	Y	Y
Control variables			Y	Y	Y	Y	Y
<i>N</i>	446	446	446	446	446	446	446
pseudo <i>R</i> <sup>2</sup>	0.145	0.141	0.210	0.226	0.206	0.226	0.219

Note: *t* statistics in parentheses

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

**Table 7 The OLS Regression Based on data from 2010 Wave**

	(1)	(2)
	<b>lnGini</b>	<b>lnGini</b>
Internet	0.0758*** (2.69)	0.0614** (1.99)
county consumption	-0.117*** (-3.06)	-0.122** (-2.53)
county GDP per capita		-0.0139 (-0.61)
county population		0.0418** (2.51)
county employment rate		0.245 (1.30)
average education year at county level		0.0211 (1.13)
age structure at county level		0.00840 (1.57)
_cons	4.566*** (17.03)	3.823*** (11.11)
<i>N</i>	150	150
adj. <i>R</i> <sup>2</sup>	0.059	0.097

Note: *t* statistics in parentheses

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

From the 2010 wave of the CFPS dataset with detail county-level variables, the only one variable related to education is “Average education year at county level”, which is then used for controlling education for the Table 7 and 8. “Internet” represents the county level internet penetration rate. The descriptions of the control variables are shown in Table 3 of the Appendix.

**Table 8. The Buffering Effect of Regional Factors**

	(1)	(2)	(3)	(4)
	lnGini	lnGini	lnGini	lnGini
Internet	-0.116 (-0.71)	0.0722** (2.24)	0.0558* (1.77)	0.0644** (2.07)
Internet*county GDP per capita	0.0192 (1.02)			
Internet*top 5% education level		-0.403** (-2.26)		
Internet*penetration ratio over 62.88%			-0.793** (-2.32)	
penetration ratio over 62.88%			3.445** (2.34)	
Internet*advanced county				-1.263*** (-4.63)
advanced county				5.502*** (4.64)
_cons	3.821*** (10.01)	3.984*** (10.54)	3.986*** (10.73)	3.943*** (11.03)
Control variables	Y	Y	Y	Y
<i>N</i>	150	150	150	150
adj. <i>R</i> <sup>2</sup>	0.096	0.106	0.115	0.124

*t* statistics in parentheses

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

## Appendix

Table 1 Numbers of Households Used to Measure Inequality in Each County

year	Mean	Std. Dev.	min	max
2010	81.893	13.828	39	105
2012	81.338	16.060	40	107
2014	79.629	16.133	43	107
2016	79.044	17.056	39	107

We drop the bottom 5% and the top 5% sample based on the rank of numbers of households in each county to minimize the size gap.

Table 2 Detail statistics about county-level variable

county id	Mean of Gini index	std. dev. of Gini index	Mean of internet	std. dev. Of internet	Mean of education	std. dev. of internet
1	0.4492	0.0501	0.3453	0.1735	0.1074	0.0750
2	0.4360	0.0578	0.5223	0.1815	0.1935	0.0767
3	0.4988	0.0551	0.3685	0.1002	0.1226	0.0638
4	0.4245	0.0388	0.3244	0.2214	0.1073	0.0270
5	0.5418	0.0687	0.3776	0.0000	0.0684	0.0204
6	0.4552	0.0611	0.4537	0.1410	0.1160	0.0310
7	0.4254	0.0207	0.6837	0.0856	0.2546	0.0458
8	0.4162	0.1202	0.7193	0.0424	0.3454	0.0527
9	0.4327	0.0463	0.4479	0.1600	0.1581	0.0459
10	0.4483	0.0521	0.7100	0.0855	0.2566	0.0462
11	0.4865	0.0744	0.5713	0.1065	0.1181	0.0325
12	0.4646	0.0461	0.4247	0.0716	0.1197	0.0391
13	0.4333	0.0409	0.5550	0.1887	0.1799	0.0447
14	0.4877	0.0895	0.4300	0.1773	0.0999	0.0425
15	0.4081	0.0624	0.3016	0.2123	0.0879	0.0367
16	0.4999	0.1261	0.3623	0.1976	0.1089	0.0410
17	0.4458	0.0372	0.3144	0.1693	0.1045	0.0161
18	0.4778	0.0593	0.3212	0.1146	0.0405	0.0185
19	0.4459	0.0263	0.4040	0.1793	0.0785	0.0384
20	0.3979	0.0501	0.4604	0.0835	0.2274	0.0160
21	0.4942	0.0404	0.3711	0.1676	0.0979	0.0525
22	0.4721	0.0173	0.2289	0.1552	0.0770	0.0558
23	0.5741	0.0507	0.3927	0.2048	0.1138	0.0531
24	0.4797	0.0158	0.4328	0.1668	0.1507	0.0329
25	0.4740	0.0487	0.3471	0.1955	0.0931	0.0471
26	0.4485	0.0399	0.3766	0.2041	0.0991	0.0379
27	0.4949	0.0860	0.4797	0.1716	0.0768	0.0403
28	0.4637	0.0573	0.4180	0.1527	0.1051	0.0450
29	0.3967	0.0942	0.8310	0.0685	0.5295	0.0025
30	0.3961	0.0512	0.5565	0.1695	0.3176	0.0397
31	0.4430	0.0577	0.3812	0.2215	0.1598	0.0559
32	0.4855	0.0168	0.5183	0.1502	0.1429	0.0293
33	0.4902	0.0103	0.2093	0.0699	0.0765	0.0381
34	0.5025	0.1114	0.5628	0.1394	0.2390	0.0338
35	0.5319	0.0700	0.5047	0.0783	0.2126	0.0626
36	0.5184	0.1126	0.6215	0.0976	0.2446	0.0614
37	0.4064	0.0333	0.7206	0.0975	0.2512	0.0464
38	0.4890	0.0411	0.3494	0.1449	0.1365	0.0394
39	0.4461	0.0299	0.3912	0.1339	0.1497	0.0281
40	0.3962	0.0355	0.3101	0.1127	0.0989	0.0617

---

41	0.4728	0.0753	0.2880	0.0786	0.1134	0.0573
42	0.4358	0.0406	0.4661	0.1039	0.0902	0.0351
43	0.4587	0.0400	0.2021	0.0520	0.1102	0.1317
44	0.3795	0.0586	0.2270	0.1606	0.2381	0.1951
45	0.4213	0.0740	0.6968	0.0821	0.4741	0.0151
46	0.4375	0.0000	0.4556	0.0000	0.1293	0.0000
47	0.4860	0.1082	0.4143	0.2050	0.0818	0.0302
48	0.4948	0.0370	0.4431	0.0923	0.1757	0.0407
49	0.4142	0.0375	0.4777	0.1845	0.1147	0.0206
50	0.4463	0.1053	0.3117	0.1405	0.1617	0.0333
51	0.4435	0.0857	0.5921	0.1336	0.3837	0.0533
52	0.4630	0.0773	0.3804	0.1251	0.1384	0.0327
53	0.4557	0.0466	0.2389	0.1008	0.1483	0.0195
54	0.3892	0.0624	0.4822	0.1274	0.2793	0.0340
55	0.4065	0.0184	0.4821	0.1065	0.1972	0.0223
56	0.3788	0.0445	0.2903	0.1779	0.1333	0.0357
57	0.3799	0.0753	0.4309	0.0912	0.2079	0.0332
58	0.4320	0.1643	0.5212	0.0545	0.4056	0.0566
59	0.4041	0.0551	0.4375	0.1039	0.2737	0.0246
60	0.4984	0.0881	0.3574	0.1257	0.1794	0.0615
61	0.4291	0.0294	0.4644	0.1567	0.1552	0.0395
62	0.3948	0.0491	0.4044	0.0888	0.1370	0.0430
63	0.4734	0.0853	0.4192	0.2014	0.1858	0.0421
64	0.4347	0.0567	0.4080	0.0891	0.1443	0.0639
65	0.5012	0.0596	0.3452	0.1151	0.2003	0.0370
66	0.4334	0.0355	0.4248	0.1657	0.1331	0.0324
67	0.4372	0.0170	0.3598	0.1499	0.1786	0.0595
68	0.3311	0.0658	0.1224	0.0000	0.0747	0.0478
69	0.4060	0.0495	0.3604	0.1514	0.1112	0.0339
70	0.4649	0.0559	0.6856	0.0714	0.1701	0.0146
71	0.4010	0.0228	0.4878	0.1427	0.1459	0.0344
72	0.4906	0.0265	0.4623	0.1660	0.1901	0.0330
73	0.4110	0.0329	0.5621	0.1873	0.1307	0.0323
74	0.4516	0.0179	0.2799	0.1459	0.1513	0.0485
75	0.4932	0.0842	0.5441	0.0826	0.3138	0.0402
76	0.4293	0.0329	0.5958	0.1273	0.2326	0.0465
77	0.4245	0.0195	0.6252	0.1091	0.1499	0.0472
78	0.4649	0.0486	0.3565	0.1226	0.1388	0.0340
79	0.4246	0.0448	0.6816	0.0561	0.3004	0.0392
80	0.3945	0.0000	0.3913	0.0000	0.0727	0.0000
81	0.3861	0.1358	0.4885	0.0421	0.4137	0.0259
83	0.5425	0.1766	0.6560	0.0374	0.3409	0.0043
85	0.5404	0.0000	0.6308	0.0000	0.4270	0.0000
87	0.4243	0.0000	0.8293	0.0000	0.5349	0.0000

---

---

88	0.3450	0.0262	0.7021	0.0931	0.4200	0.0316
89	0.4521	0.0815	0.7202	0.0533	0.3598	0.0093
90	0.4854	0.0822	0.6718	0.1020	0.3801	0.0299
91	0.4301	0.0676	0.4827	0.0628	0.2418	0.0412
93	0.3494	0.0207	0.5086	0.0070	0.1658	0.0557
94	0.4683	0.0734	0.4443	0.0514	0.2130	0.0820
95	0.4277	0.0720	0.5918	0.1353	0.2622	0.0399
97	0.4747	0.0396	0.5255	0.0830	0.2960	0.0087
98	0.4419	0.0208	0.3550	0.0141	0.1209	0.0304
99	0.3841	0.1127	0.7181	0.0978	0.3558	0.0237
100	0.4517	0.1038	0.3389	0.1766	0.1030	0.0513
101	0.4912	0.0756	0.5171	0.1194	0.1610	0.0418
102	0.5013	0.0862	0.4651	0.1535	0.0903	0.0329
103	0.5170	0.1135	0.4348	0.1879	0.1051	0.0596
104	0.4878	0.0034	0.4628	0.0739	0.1080	0.0172
105	0.4173	0.0548	0.6729	0.1367	0.3767	0.0326
106	0.3954	0.0583	0.7062	0.1071	0.4146	0.0147
107	0.4679	0.0406	0.5706	0.0883	0.1983	0.0273
108	0.4616	0.0462	0.4232	0.1764	0.1390	0.0320
109	0.5277	0.0811	0.3931	0.1727	0.1251	0.0458
111	0.5125	0.0979	0.3904	0.1263	0.1323	0.0610
112	0.4627	0.0501	0.4097	0.1292	0.0703	0.0321
113	0.4903	0.0790	0.3298	0.1326	0.1078	0.0470
114	0.4857	0.0643	0.5248	0.1617	0.1701	0.0448
115	0.5067	0.0400	0.7014	0.1708	0.3319	0.0805
116	0.4353	0.1123	0.6200	0.1416	0.2115	0.0848
117	0.4759	0.0328	0.3731	0.1646	0.0953	0.0504
118	0.3728	0.0305	0.7776	0.0415	0.3105	0.0399
119	0.3653	0.0749	0.7136	0.0602	0.2141	0.0571
120	0.4423	0.0237	0.7794	0.0508	0.2958	0.0441
121	0.4850	0.0487	0.3052	0.0943	0.1081	0.0656
122	0.5839	0.0332	0.5769	0.1632	0.1534	0.0510
123	0.4993	0.0765	0.4152	0.0232	0.1379	0.0392
124	0.5082	0.0618	0.2867	0.1422	0.1532	0.0704
125	0.4277	0.0918	0.3421	0.1753	0.0929	0.0446
126	0.4847	0.1315	0.3837	0.1632	0.1022	0.0453
127	0.4374	0.0327	0.7452	0.1115	0.2724	0.0358
128	0.3564	0.0544	0.5021	0.1050	0.1138	0.0279
129	0.5041	0.0974	0.3623	0.1688	0.0869	0.0462
130	0.4432	0.0535	0.3755	0.2028	0.0918	0.0389
131	0.4337	0.1039	0.6992	0.1214	0.3076	0.0319
132	0.4327	0.0739	0.4065	0.1460	0.1564	0.0379
133	0.5231	0.0891	0.4697	0.1333	0.2492	0.0332
134	0.4537	0.0269	0.4534	0.2305	0.1828	0.0720

---

---

135	0.4303	0.0613	0.4528	0.1332	0.1988	0.0521
136	0.4123	0.0177	0.2971	0.1816	0.0820	0.0583
137	0.5034	0.0474	0.4128	0.2388	0.1193	0.0939
138	0.4336	0.0159	0.4624	0.1265	0.1529	0.0541
139	0.4283	0.0719	0.3627	0.0996	0.1286	0.0510
140	0.5488	0.1407	0.4875	0.2213	0.1281	0.0290
141	0.4892	0.0606	0.4548	0.1667	0.1241	0.0599
142	0.4648	0.0723	0.3193	0.1652	0.1387	0.0570
143	0.4881	0.0534	0.2330	0.1218	0.1309	0.0722
144	0.4094	0.0470	0.2854	0.2441	0.1039	0.1064
145	0.3931	0.0058	0.3660	0.0370	0.1012	0.0116
146	0.4979	0.1169	0.2251	0.1674	0.0770	0.0467
147	0.3486	0.0209	0.5312	0.1147	0.3924	0.0595
148	0.4189	0.0507	0.3069	0.0990	0.1553	0.0454
149	0.4019	0.0929	0.5795	0.0584	0.4787	0.0235
150	0.5496	0.0039	0.3520	0.0769	0.2013	0.0688
151	0.2926	0.0701	0.5869	0.1027	0.3567	0.0464
152	0.3462	0.0981	0.5362	0.1287	0.3179	0.0869
153	0.3736	0.0375	0.4642	0.0541	0.2720	0.0427
154	0.4330	0.0063	0.5682	0.1094	0.2617	0.0442
155	0.4277	0.0833	0.4040	0.0633	0.1934	0.0279
156	0.4085	0.0497	0.4988	0.0448	0.1244	0.0322
157	0.3689	0.0417	0.3834	0.2214	0.0901	0.0452
158	0.4457	0.0712	0.3738	0.0620	0.0849	0.0451
159	0.4423	0.0493	0.3377	0.2029	0.0831	0.0376
160	0.4776	0.0715	0.2447	0.1676	0.0879	0.0334
161	0.4097	0.0510	0.6255	0.0963	0.3671	0.0377

---



**Table 3. County Level Control Variables from the Wave 2010 of CFPS**

<b>Variable</b>		<b>Obs</b>	<b>Mean</b>	<b>Std. Dev.</b>	<b>Min</b>	<b>Max</b>
County GDP per capita	The logarithm of GDP per capita at county level	150	10.101	0.998	8.062	12.676
County population	The logarithm of population at county level	150	13.074	0.742	10.711	15.914
County employment	employment ratio (%) at county level	150	0.664	0.102	0.430	0.900
Average education year at county level	Average education year at county level	150	9.076	1.431	4.830	12.780
Age structure at county level	The elder generation ratio (age $\geq$ 65) (%)	150	9.124	2.222	2.350	17.970
County consumption	The logarithm of consumption per capita at county level	150	8.625	0.493	7.531	9.861

Note: Employment ratio (%) is defined as the ratio of employed adult on total adult population (age $\geq$ 16), directly attained from the CFPS dataset.