From a Way of Life to Ways of Earning a Living

The Impact of E-commerce on Occupational Choices*

Wei Li†
HKUST

Yao Amber Li‡
HKUST

This Version: July 2023

Abstract

While e-commerce has dramatically changed the landscape of the economy in many countries, we still lack a comprehensive understanding of how e-commerce affects job and income distribution. This paper links regional e-commerce development, measured by its extensive and intensive margins, to labor market outcomes. The oligopoly telecommunications industry in China and an Internet policy launched in 2015 allow us to identify the causal effect by applying an instrumental variable identification. We establish that e-commerce increases average employment rates, individual wages, and household incomes, but older workers may lose. People exposed to active e-commerce markets tend to quit the agricultural sector and get employed in enterprises of the manufacturing and service sectors. Moreover, with the creative destruction brought by e-commerce, people, especially older ones, tend to switch occupations and have to handle new tasks. Interestingly, e-commerce is not a technological-biased change as it significantly benefits less-educated workers. Based on firm-level data, we confirm that e-commerce facilitates more related firms to be created, and running e-commerce-related businesses is indeed associated with more employment of less-educated workers. These results suggest that e-commerce can serve as a promising driver of structural transformation.

* We are grateful for comments from Robin Kaiji Gong, Ohchan Kwon, Kira Matus, Jin Wang, and Xu Zhang. Any errors are our own.

† The Hong Kong University of Science and Technology, email: wlick@connect.ust.hk.

‡ The Hong Kong University of Science and Technology, email: yaoli@ust.hk.

1. Introduction

Will e-commerce kill jobs and render people unemployed? With the advances in digital technology and modern logistics, e-commerce, a new business model that allows consumers or firms to trade online, emerges and dramatically changes people's way of life. While some observers have expressed concerns that e-commerce may shut down offline stores and reduce job opportunities,[1] e-commerce may create new jobs to hire workers with the application of new technology (Acemoglu and Restrepo 2019; Autor et al. 2021; Braxton and Taska 2023). The net effect of e-commerce on employment and wages remains ambiguous. Moreover, who benefits or loses in the e-commerce era remains unclear.

Generally, e-commerce can be subsumed into business-to-business (B2B), business-to-consumer (B2C), and consumer-to-consumer (C2C) activities. The B2C and C2C e-commerce increases consumer gains by reducing prices (Lieber and Syverson 2012; Goldfarb and Tucker 2019; Jo et al. 2019; Couture et al. 2021), adding product varieties (Brynjolfsson et al. 2003; Randall et al. 2006; Brynjolfsson et al. 2022; Dolfen et al. 2023), and mitigating the distance and time restrictions of shopping (Sinai and Waldfogel 2004; Pozzi 2013; Lendle et al. 2016; Fan et al. 2018; Relihan 2022).

Beyond consumer gains, e-commerce affects the activities of producers and sellers. It leads to total market expansion (Duch-Brown et al. 2017; Fan et al. 2018), lowers firms' entry costs (e.g., costs of setting up storefronts and obtaining information), and encourages entrepreneurship and competition (Huang et al. 2021; Jin and Sun 2021). Furthermore, two competing forces emanating from information technology reshape the market structure: the "winner-take-all" (or "superstar") effect and the "long-tail" effect. Whereas the winner-take-all effect leads to the domination of "big" companies in some industries (Goldmanis et al. 2010), the long-tail effect of e-commerce allows the survival of "small" firms catering to special demands of rare and niche products (Brynjolfsson et al. 2003; Brynjolfsson et al. 2011; Brynjolfsson et al. 2022; Dolfen et al. 2023).

The changes in market structure induced by e-commerce have implications for the labor market: It can be a disaster for some while a windfall for others. For example,

[1] See articles published in media outlets, for example, Ben Unglesbee, "E-commerce Could Kill 30K Stores and Half a Million Jobs by 2025," *Retail Dive*, https://perma.cc/ZCM3-3G9M (accessed on 25 May 2023), and David Sax, "What We Would Miss in an All-Amazon-Shopping World," *The New Yorker*, https://perma.cc/B6U6-D4KD (accessed on 25 May 2023).

Chava et al. (2022) find that the coming of e-commerce fulfillment centers undermines the employment and wages of offline retailers. On the other hand, e-commerce may bring other positive effects to the labor market. E-commerce boosts demand from low-skilled jobs (e.g., couriers) to high-skilled positions (e.g., programmers). It even creates numerous occupations in new sectors, from information-based services to e-commerce training businesses (Bakos 2001). Additionally, the long-tail effect of e-commerce implies that people can work for the nitch market (Bakos 2001; Agrawal et al. 2015). Evidence also shows that e-commerce can encourage entrepreneurship (Huang et al. 2021; Jin and Sun 2021), implying the creation of jobs. Moreover, more entry of firms induced by e-commerce can increase real wages by lowering the price index and raising nominal wages (Fan et al. 2018).

However, with these labor market implications of e-commerce, we still lack direct evidence on e-commerce's impact on the labor-market structure: How e-commerce affects individual occupational choices? How does e-commerce relocate the workforce across sectors? What are the resultant changes in wages and incomes? This paper aims to fill this gap in the literature.

In this paper, we use Chinese data to identify the causal impact of e-commerce on labor market outcomes. China provides an ideal setting to study this question. Through decades of growth since the 1990s, China has become the largest online market in the world by 2020: Its e-commerce market value reached 1,344 billion USD, compared to 538 billion USD in the United States, 461 billion USD in Europe, and 513 billion USD in the rest of the world.[2] It is also worth noting that there is significant heterogeneity in e-commerce adoption within China. Most e-commerce activities concentrate on coastal areas.[3] With the largest global e-commerce market and enormous internal heterogeneity, China has a full spectrum of development stages of e-commerce.

More importantly, the institutional context in China provides me with a unique chance to implement an instrumental-variable strategy. It is difficult to identify the causal impact of e-commerce development as it is adopted endogenously by people and firms. One of the pivotal premises of e-commerce penetration is the Internet. At least before 2015, the oligopoly telecommunications market in China hindered the quality and affordability of Internet services. In 2015, the Chinese government launched a

---

[2] Sofia Zavialova, "E-commerce Report 2021," *Statista Digital Market Insights*, https://www.statista.com/study/42335/ecommerce-report/ (accessed on 20 November 2021).

[3] In 2020, 53% of the online retail sales were from three coastal provinces (Guangdong, Zhejiang, and Shanghai), and 49% of online shops were located in five eastern provinces (Guangdong, Zhejiang, Jiangsu, Shandong, and Henan) (Ministry of Commerce 2021).

policy named “Speed up and Cheapen the Internet” to force those state-owned telecommunications operators to improve Internet services. The stringency for implementing the policy is left to local governments’ discretion and depends on negotiations between local governments and telecommunications operators. Therefore, policy stringency varies across regions. Utilizing a stringency measure based on media contents in Communist-Party-of-China newspapers, we can extract variations in e-commerce development induced by the exogenous improvement in Internet services.

Tidying labor market outcomes from two nationally representative household surveys, we establish that e-commerce increases average employment rates, individual wages, and total household incomes. People exposed to more developed e-commerce markets are likelier to quit the agricultural sector and get employed in enterprises of the manufacturing and service sectors. In addition, evidence suggests that e-commerce involves creative destruction, and people who remain employed tend to switch occupations and handle new tasks. These results are robust under various checks, including using alternative e-commerce measures, clustering the standard errors at the individual and community-by-year levels, and controlling province-level Internet development and individual-level Internet usage.

Heterogeneity analysis shows that e-commerce mainly increases employment rates of the young- and middle-aged workforce (18-40 years of age). Older workers are the vulnerable group in the e-commerce era as they tend to lose their jobs and face more dramatic changes in job tasks. For gender heterogeneity, e-commerce encourages more females to start their own businesses. Interestingly, e-commerce seems not to be a technological-biased change: It mainly increases employment rates of less-educated workers by relocating them to jobs in enterprises with higher wages from the agricultural sector. By examining the migrant status of workers, we confirm the importance of labor mobility to ensure a smooth transformation induced by e-commerce.

Finally, we use administrative and survey data on firms to investigate mechanisms. Based on the firm registration database, we document that e-commerce penetration differs across industries. The online consumption market facilitates the establishment of firms running e-commerce-related businesses, implying the creation of jobs. Drawing on two firm surveys conducted by the National Bureau of Statistics, we document that both province-level e-commerce market development and running e-commerce-related businesses on firms’ own are associated with more employment in the manufacturing sector. In addition, evidence from economic census data confirms a tendency to employ less-educated workers in e-commerce-related businesses of the service sector.

Beyond the literature on e-commerce's socioeconomic effect, this paper speaks to studies on the labor-market impacts of information and communications technology (ICT). Research has established that the Internet can improve job matching (Autor 2001; Kuhn and Mansour 2014; Bhuller et al. 2023) and differently affect employment and wages for people with different socioeconomic backgrounds or skill levels (Forman et al. 2012; Kolko 2012; Akerman et al. 2015; Hjort and Poulsen 2019; Zuo 2021; Chiplunkar and Goldberg 2022). This paper furthers our understanding by showing how ICT facilitates economic transformation when it combines with commerce and modern logistics.

This paper also contributes to the extensive literature on technological change, wage structure, and job polarization. Scholars have long been intrigued by the evolution of technological progress and inequality. Whereas *the skill-biased technological change* complements non-routine, cognition-intensive tasks and increases demand for high-skilled workers, *the routine-biased technological change* substitutes routine tasks of many traditional middle-skilled jobs but has a limited impact on low-skilled non-routine occupations. Therefore, we have seen widening wage inequality and job polarization in recent decades (Goldin and Katz 2007; Acemoglu and Autor 2011; Forman et al. 2012; Goos et al. 2014; Michaels et al. 2014; Acemoglu and Restrepo 2019). Moreover, other channels, such as the education race, the automation-reinstatement race, and artificial intelligence, have transformed the landscape of labor markets (see a review by Autor 2022). What roles does e-commerce play in inequality issues? What are the regularities of e-commerce regarding skill-biased or routine-biased characteristics? This paper answers these questions by showing that e-commerce does not bring skill-biased change and benefits less-educated workers.

This paper speaks to a larger body of literature on the structural transformation of an economy. Structural transformation refers to the relocation of resources between sectors, which is usually measured by employment shares, value-added shares, and consumption expenditure shares across sectors (Herrendorf et al. 2014). Research has identified some drivers of structural change, such as human capital (Porzio et al. 2022), trade liberalization (Cravino and Sotelo 2019), and technological advances (Bustos et al. 2016; Hjort and Poulsen 2019). Whereas previous research has contributed to the impact of e-commerce on consumption and industry structure, how e-commerce transforms employment share across sectors remains unclear. This paper establishes that e-commerce significantly relocates the labor force from the agricultural sector to others, serving as a promising driver of structural transformation.

The findings of this paper are of profound policy implications. Many countries use e-commerce as a policy tool to develop the regional economy. For example, there are Rural E-commerce Demonstration Counties in China (since 2014), Digital India in India (since 2015), and E-Commerce Development Master Plan (2021-2025) in Vietnam. There is no warrant for the success of these programs unless we have a thorough understanding of the interplay between e-commerce and each agent. For instance, is the transformation brought by e-commerce efficient? Should the government intervene in the process? Who needs external help? What kind of help is needed? Characterizing empirical regularities of e-commerce can help us design appropriate public policies.

2. Empirical Strategy

2.1. Identification strategy

To examine the impact of e-commerce on labor market outcomes, we estimate the following equation,

$$y_{ipt} = \alpha_0 + \alpha_1 EC_{p,t-1} + \gamma_i + \delta_t + \varepsilon_{ipt}, \tag{1}$$

where i, p, and t denote the individual, province, and year, respectively. y_{ipt} is the outcome of interest, including various indicators of occupational choices and incomes. $EC_{p,t-1}$ is a measure of e-commerce development over the province p. Given that the impact of e-commerce on occupations needs time to be observed (e.g., there may be friction in market structure transformation), we lag the e-commerce measurement by one year. γ_i is an individual fixed effect absorbing time-invariant personal characteristics such as innate intelligence. δ_t is a year fixed effect capturing macro time trends. The standard error ε_{ipt} will be clustered at the individual level.

Equation (1) may suffer from endogenous problems. Even conditioning on individual fixed effects and year fixed effects, some factors varying at the individual-by-year level can simultaneously affect labor-market outcomes and e-commerce development. To tackle the endogenous concerns, we instrument the e-commerce measurement with an Internet policy launched in 2015 named “Speed up and Cheapen the Internet” (henceforth, the SUCI policy).

The SUCI policy provided an exogenous push to the oligopoly telecommunications industry in China. Three state-owned enterprises have dominated the telecommunications industry since 2008 (Xia 2017).[4] In such an oligopoly market, telecommunications prices were sticky, and the improvement of Internet services was incredibly slow. The public had long complained about the poor Internet services.[5] In responding to such complaints, the State Council of China launched a policy in May 2015 to increase Internet speed and lower its prices.[6] Measures to fulfill the policy goals can be summarized into two aspects: (1) increasing investment in telecommunications infrastructure, including fiber-optic networks and base stations; (2) upgrading or launching favorable telecommunications packages for consumers.[7] According to government statistics, compared to the rates in 2015, the average price of Internet services (including fixed-line and mobile Internet) has decreased by more than 95% by 2021, and the fiber-optic, 4G, and 5G networks are increasingly prevalent.[8] However, looking into the implementation process, the stringency and timing of the policy varied by province, as local governments had the discretion to propose schedules and measures to urge those local telecommunications operators. Such variations allow me to identify the effect of the SUCI policy on Internet usage and e-commerce development across provinces.

The Chinese context makes the SUCI policy an ideal instrumental variable (IV) for e-commerce development. First, telecommunications oligopolies have no motivation to improve their services due to a lack of competition, and the SUCI policy gave an exogenous push to these telecommunications operators. As Internet availability is a necessary condition for e-commerce, better Internet services may facilitate the development of e-commerce, given other advantages that online shopping provides (e.g., lower prices and abundant goods varieties). Second, the three telecommunications operators are all state-owned enterprises, meaning they have restricted leeway to downsize. Hence, although the SUCI policy reduces the markup of state-owned telecommunications operators, there is hardly a direct labor demand shock induced by

[4] The three operators are China Mobile, China Telecom, and China Unicom.

[5] For example, as early as 2011, a news article published in the Beijing News asserted that the Internet speed was lower than what the operators had promised. Moreover, fees per Mbps in China are three times the costs of similar services in Vietnam, four times in the United States, 29 times in South Korea, and 469 times in Hong Kong, China (see https://perma.cc/WUM8-SENR, in Chinese, accessed on 28 June 2023).

[6] The State Council, "Guideline to improve Internet speed, lower prices," https://perma.cc/566W-PZTQ (accessed on 18 May 2023).

[7] For example, see a proposal by China Telecom (China Telecom, "Ten measures to speed up and cheapen the Internet by China Telecom (in Chinese)," https://perma.cc/X3G3-QCUM, accessed on 22 June 2023). The proposal's implementation required the supervision and cooperation of the local government and CPC.

[8] Han Xin, "China continues to increase Internet speed, quality, and lower rates," *People's Daily*, https://perma.cc/N2CN-NV7B (accessed on 18 May 2023). http://en.people.cn/n3/2021/0429/c90000-9845006-2.html.

this policy. Finally, some may concern that the SUCI policy can directly affect occupational choices due to Internet accessibility (Hjort and Poulsen 2019; Zuo 2021; Chiplunkar and Goldberg 2022) rather than the development of e-commerce. By controlling for individual and year fixed effects in Equation (1), we can capture most individual-specific channels through which the Internet can exert an impact. In Section 3.3, we will provide additional evidence on the exclusion restriction by conducting tests following the idea of zero-first-stage tests (Bound and Jaeger 2000) and controlling other potential time-variant confounders regarding Internet development. In summary, the SUCI policy should satisfy the relevance and exclusion requirements for a valid IV, allowing me to extract exogenous variations in e-commerce development.

Formally, the first-stage specification for the two-stage least squares estimation is,

$$EC_{p,t-1} = \pi_0 + \pi_1 P_{p,t-1} + \gamma_i + \delta_t + \nu_{ipt}, \tag{2}$$

where $P_{p,t-1}$ is a measure of policy stringency over a region, and ν_{ipt} is the standard error.

The explanatory variable and IV vary by province and year. One implicit assumption underlying the measures is that the province boundaries form segmented markets for employment, e-commerce, and telecommunications services. Chinese context makes it a reasonable assumption. First, the household registration system (the so-called *Hukou*) still imposes significant barriers to labor mobility for migrant workers (Song 2014; Zhang et al. 2020).

Second, local governments in China perform local protectionism by issuing discriminatory policies to favor local firms, resulting in segmented markets within China (Barwick et al. 2021; Fang et al. 2022). In addition, stemming from local officials' incentives to be promoted, local governments may compete in development policies in their jurisdictions (Li and Zhou 2005; Zhou 2007). Depending on local leaders' goals and focus of attention, e-commerce policies show a significant regional disparity across provinces.[9]

Third, the telecommunications market is segmented due to the institutional context. Before September 2017, users should pay domestic roaming fees when they leave the local service area, hindering arbitrage opportunities across provinces.[10] Although the long-distance and roaming fees were waived later, the telecommunications operators

[9] By retrieving records of regulations and laws related to e-commerce from 2009 to 2021 in the PKULaw database, we found that more than 3,000 official documents were issued in Zhejiang and Guangxi Province, respectively, compared to less than 600 issued in Tibet, Tianjin, Xinjiang, and Hainan Province.

[10] The State Council, "End of era for long-distance, roaming charges," https://perma.cc/L9VD-WRRK (accessed on 18 May 2023).

can use price discrimination based on the market conditions across regions, especially after the telecommunications reform in 2014 that allowed operators to differentiate prices based on market demand.[11] As Internet services are one of the premises of e-commerce, the segmented telecommunications markets divide e-commerce consumer markets in terms of penetration rates.[12]

2.2. Data and measurements

2.2.1. E-commerce

As it is the B2C and C2C e-commerce that dramatically changes people's lives, this paper measures e-commerce development based on the consumer side. Two complementary e-commerce measures are adopted: a percentage measure to capture extensive margins and a volume measure to capture intensive margins.

The percentage measure is constructed from the China Household Finance Survey (CHFS). The CHFS is a biennially longitudinal survey executed by the Southwestern University of Finance and Economics (see Appendix I for more information on the data set). One prominent feature of the CHFS is that it has gleaned representative data at the province level since 2013, which is rare compared to other nationally representative household surveys in China. The CHFS asks households, "*How much did your family shop online last year?*"[13] we first define a dummy variable equaling one if a household had positive online shopping expenditure. Next, we calculate the mean of the dummy variable using sampling weights by province, yielding a percentage measure for 2012, 2015, 2016, and 2018. A higher measurement reading indicates a higher percentage of the population adopting online shopping in a province. In addition, we also aggregate households' monthly online expenditures in the same vein. A higher value represents a higher monthly online expenditure per household by province. Due to potential

[11] Ministry of Industry and Information Technology, "The government lifted the price supervision of telecommunications services (in Chinese)," https://perma.cc/TGA6-TCHC (accessed on 18 May 2023).

[12] Measuring e-commerce development at the provincial level is also due to data availability. As the exact information on respondents' residential places is only available at the province level due to privacy policies, we can only merge external source measurements with individual observations based on province names.

[13] CHFS 2015 is an exception that this year the survey asked a household, "*How much did you spend on online shopping last month*." A household may use online shopping but happen to buy nothing online during the month before the survey (the CHFS usually administers surveys from July to August). To avoid underestimating the e-commerce penetration in this wave, we utilized other information to overcome the limitation. For example, for households stating that their usage of online banking services is mainly for online shopping, we will also define them as online-shopping adopters. Nevertheless, underestimation is still possible. In the regression analysis, survey year fixed effects can mitigate the measurement error of this kind.

measurement errors (e.g., recall bias in past spending), we use this volume measure in the robustness check.

The volume measure in the benchmark analysis is generated based on the “online retail transaction volumes” published by the National Bureau of Statistics and the Ministry of Commerce (see Appendix I for more information on constructing the time series). The statistic is a firm-based measure involving online retail trading of goods (including virtual goods) and services. Although there are other statistics on online transactions, such as e-commerce sales and e-commerce procurement volume, we choose the retail transaction volumes to emphasize final consumption rather than firms’ intermediate inputs. Excluding transactions of intermediate goods also makes the volume measure more comparable to my percentage measure, which is constructed from consumer behavior. In the benchmark analysis, we normalize the online retail transaction volumes by province population. The usage of other mathematical transformations of the statistic, such as the logarithm form, is presented in Section 3.3 for sensitivity checks.

It is worth noting one nuanced disparity between the two measures. The percentage measure reflects the behavior of online consumers, capturing the development of the regional e-commerce consumer market. In contrast, the original “online retail transaction volumes” describe the behavior of online sellers, including their transactions with dealers within and outside the firms’ located province. By applying the IV strategy that utilizes variations in local Internet services, we can extract the transactions realized by consumers from the same province as the sellers, which differs from the scope captured by its original variations.

2.2.2. Employment

Variables from the CHFS

The data on labor market outcomes comes from two nationally representative household surveys. The first data set is the CHFS, which provides respondents’ employment status, positions, and characteristics of employment providers in 2011, 2013, 2015, 2017, and 2018. we first generate a dummy variable equal to one if a respondent reported that he/she had a job. As the CHFS only documents coarse employment categories, we create indicators corresponding to three positions: employer or self-employed worker, employee, and farmer. Although some research points out the difference between employers and self-employed workers (Levine and Rubinstein 2018), the division between them is only available starting from CHFS 2017.

we provide results of separating the two positions in the appendix. In addition, we can observe the characteristics of employment providers. we define three types of providers: public institutions, enterprises (including state-owned, domestic private, and foreign-owned), and individual businesses or others. All the dummy variables regarding positions and employment providers are defined over the analysis sample and are NOT conditional on being employed.

Besides occupations, we also utilize the information on individuals' wages and household incomes, which has important implications for the welfare effect of e-commerce.

Variables from the CFPS

Although the CHFS allows me to observe a longer period of individual data, it only provides coarse information on occupations. To overcome the limitation, we utilize data from the China Family Panel Studies (CFPS), a biennially longitudinal survey maintained by Peking University since 2010 (see Appendix I for more information on the data set). One salient feature of the CFPS is that it provides detailed categories and affiliated industries of occupations.[14]

The occupational category in the CFPS is a five-digit coding system with eight general categories and more than 400 subdivisions. Figure A1 illustrates the coding system. In addition, the CFPS assigns an industry out of twenty categories for each occupation. Utilizing the longitudinal data, we can identify whether individuals changed occupations or working industries in successive survey waves. Specifically, we define whether individuals change their occupations at five-, three-, or one-digit levels (from lower to higher coding levels).[15] Switches at higher digit levels indicate more considerable changes in job tasks. However, we should note that the change in occupational categories is not necessarily due to job-hopping across companies. The category codes will change if a respondent gets promoted or demoted within the same company. In addition, the measurements may underestimate the job-hopping rate as we can only measure their jobs at the category level while people can hop jobs across firms but perform the same tasks (i.e., such cases will be coded as maintaining the same occupational category).

[14] As of writing this paper, the detailed categories of occupations in CFPS 2020 are unavailable.

[15] The CFPS does not systematically code occupations in the agricultural sector at the lowest level. For example, the CFPS 2012 assigns the majority of occupations in the agricultural sector to the highest level of "agricultural and fishery workers – 5" without providing subdivisions. As our goal is to identify a structural change from the agricultural sector to other sectors, we recode all the occupations in the agricultural sector as 50000 across waves but retain lower level coding of occupations in other sectors. In addition, we consolidate the category "armed forces" to the type "others" due to its rare cases.

Sample restriction

As my purpose is to examine the impact of e-commerce on occupational choices, we restricted the individual-level data to respondents aged 18-55 and excluded full-time students, minimizing the selection problems due to education or retirement.

2.2.3. The SUCI policy

An ideal way to measure the stringency of the SUCI policy is to gauge the degree of improvement in Internet services, such as higher speed and lower real prices, induced by the policy. Unfortunately, public data on these dimensions are unavailable. Alternatively, this paper uses the number of news articles published in the official newspapers of the Communist Party of China (henceforth, CPC newspapers) mentioning the SUCI policy to measure the policy stringency.

Newspapers in China can be roughly divided into two types: the CPC newspapers serving as mouthpieces of the Party and government and the commercialized newspapers aiming to cater to market demand (e.g., entertainment). A pivotal role of CPC newspapers is to inform the public about the Party's and government's policies or decisions (Zhao 1998). Since the journalism reform in 2003, the trend in product differentiation has been strengthened: CPC newspapers focus more on propaganda content and leave other content to their commercialized counterparts (Qin et al. 2018). Therefore, contents in the CPC newspapers can serve as ideal indicators of policies.[16] Although it is the government's attitudes to certain issues or policies that news articles directly measure, it is reasonable to assume that if a policy is repeated frequently in official outlets, the government implements the policy more stringently, which is also supported by empirical evidence in Table A3 (elaborated in Section 2.3). In addition, news articles can capture regional disparities since the local CPC committees and governments can decide what articles to be published (Zhao 1998; Xu 2011).

We retrieved news articles mentioning the SUCI policy's Chinese name from the WiseNews database, the largest Chinese media database in the world (see Appendix I for more information on the database).[17] We obtained 28,622 unique records published in more than 600 print newspapers. To identify the types of newspapers (i.e., CPC papers or commercialized ones), we retrieved newspapers' registered information from the National Press and Publication Administration, a national regulator providing

[16] Baker et al. (2016) and Huang and Luk (2020) use the contents of news articles to construct economic policy uncertainty indices.

[17] We use the Chinese characters "*ti su jiang fei*" (meaning the SUCI policy) combined with "*wang luo*" (meaning the Internet) to locate and retrieve the news records.

registered information for all circulative newspapers in China.[18] Next, we merged the registered records with the articles from the WiseNews database and kept media outlets directly run by province or prefecture CPC committees.[19] The exercise yielded 12,762 articles published in 224 CPC newspapers from 2009 to 2020, accounting for 45% of the total records. Examining the contents, we find that they either declare the implementation or propagate the achievement of the SUCI policy, which justifies the validity of my measure.

With some exceptions, the scope of circulation of the CPC newspapers parallels the hierarchy of the CPC administrative system: nation, province, and prefecture (Stockmann 2013).[20] Accordingly, local policies' sphere of influence should correspond to newspapers' circulation areas. Explicitly speaking, policy articles published in a prefecture-level newspaper should measure the desires or achievements of the prefectural administration. In contrast, articles published in a province-level newspaper should measure the impact of the policy affecting the whole province. Given the different implications of articles published at each level of CPC newspapers, we use the percentage of a jurisdiction's population over a province, based on the 2010 census, as weights to calculate a weighted sum of news records by province and year.[21] Finally, we calculate the accumulative number of the weighted sum and take the logarithm form of the measurement.

2.2.4. Firm-level measurements

To examine the mechanisms of e-commerce's occupational impact from the perspective of firms, we draw on firm registration data from the State Administration for Market Regulation (the former State Administration for Industry and Commerce). The database provides information on firms' names, establishment dates, ownership,

[18] https://www.nppa.gov.cn/nppa/publishing/paper.shtml.

[19] We use the ownership information to determine the type and administrative level of a newspaper. For example, *Nanfang Daily*, directly run by the CPC Guangdong Provincial Committee, is a province CPC paper, while *Guangzhou Daily*, directly run by the CPC Guangzhou Prefectural Committee, is a prefecture CPC paper. Compared to the demarcations used in Qin et al. (2018), i.e., Party Dailies, Party Evenings, and Subsidiaries, my method subsumes the former two into one category and leaves the remaining newspapers as commercialized ones.

[20] The complete hierarchy of Chinese administration should include "county" as the lowest administrative unit. However, the central government of China shut down most county-level CPC newspapers in 2003.

[21] For example, in the 2010 census, the population is 12,701,948 in Guangzhou City (the provincial capital of Guangdong Province) and 104,320,459 in Guangdong Province. Articles published in the CPC newspaper of Guangzhou City will be assigned a weight of 12,701,948/104,320,459≈0.12. In contrast, articles published in the CPC newspaper of Guangdong Province will be given a weight of 1. We use the 2010 census data to weight the measurement for two reasons: First, the census has high data accuracy. Second, the constant weight can alleviate concerns about endogenous weighting in terms of the population.

locations, industries, and businesses since 1949 when the People's Republic of China was founded. It covers over 120 million records of firm registration (including individual businesses), with around 80 million in the recent five years (i.e., from 2015 to 2019, 67% of the total). As firm exit may not be updated timely for the data in recent years, identifying surviving firms each year may suffer from measurement errors. Hence, we only calculate flow measures for firm creation.

In addition, related to this research, we identify a firm running e-commerce-related businesses if their names or business descriptions hit one of a list of e-commerce-related keywords (Table A2 itemizes the list). "E-commerce-related" is defined to encompass various direct or indirect e-commerce activities, from online selling services to derivative businesses such as e-commerce training services. In the analysis, we aggregate the data to obtain the total number and the percentage of newly registered firms running e-commerce-related businesses by city, industry, and year.

As the Firm Registration Database does not contain time-variant information, such as employment, we draw on two firm surveys maintained by the National Bureau of Statistics: The 2008 Economic Census and the Annual Survey of Industrial Enterprises (ASIE) from 2002 to 2012 (more details regarding the AISE are provided in the appendix). The two firm surveys supplement each other. Whereas the ASIE documents the performance of manufacturing firms, the Economic Census covers the universe of firms in the service sector.[22] Whereas the ASIE provides time-series data of firms, such as annual average employment, the Economic Census has cross-sectional but detailed information on employment composition at the end of 2008 by employees' gender and educational background. Moreover, we merge the firm registration records with firms in the ASIE and the 2008 Economic Census, identifying whether a firm runs e-commerce-related businesses. Therefore, we can investigate the association between firms' e-commerce-related businesses and their employment decisions.

However, it is worth noting that the number of employees could be under-reported in the 2008 Economic Census and ASIE for two reasons. First, as the taxes and fees paid to the local government are proportional to the total employment, firms have incentives to report fewer formal employment. Second, firms in China usually hire temporary workers through a third party (the so-called "*laowu paiqian*" in Chinese).

---

[22] The sectors covered by the 2008 Economic Census include: construction; transportation and storage; information and communication; wholesale and retail trade; accommodation and food service; leasing and business services; professional, science, and technology activities; hydraulic engineering, environment and public facilities management; resident services and other services; education; human health and social work; arts, entertainment and recreation; public administration and social organization.

The data does not document such temporary workforce usage (Brandt et al. 2014). Hence, the employment statistics in these surveys may be a lower bound.

2.3. Summary statistics

Before discussing regression results, we provide summary statistics on e-commerce development, labor market outcomes, and the newspaper-based IV.

Figure 1 illustrates the trend and regional disparity in e-commerce development. The left panel displays the yearly estimates with 95% confidence intervals for the percentage of the population using online shopping. Over each year, each bar represents the statistic of eastern, middle, and western regions, respectively.[23] The right panel shows the trend in online retail transaction volume per capita (10,000 CNY, 2009 price) and its national growth rate. Over each year, the bars for eastern, middle, and western regions are stacked, showing the difference among regions.

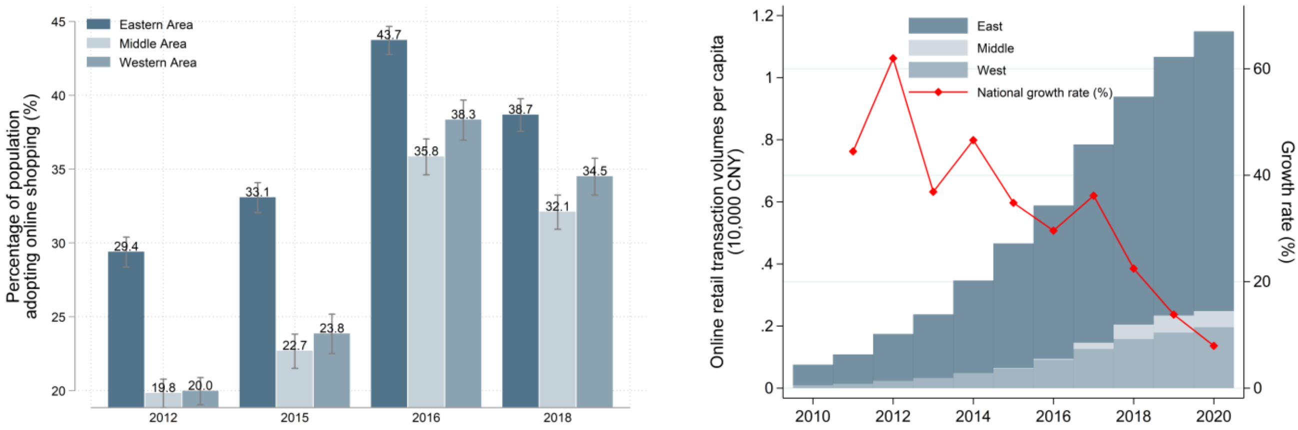
As shown in the left panel of Figure 1, after experiencing a surge in 2016, the adoption rate of e-commerce in each region shrunk. For example, the adoption rate in the eastern region grew to 43.7% in 2016 but decreased to 38.7% in 2018.[24] Given that the penetration rate is less than half, there may still be ample room for e-commerce to expand. Whereas the percentage measurement observed a decline, the online retail transaction volume per capita has increased since 2009, even after counting inflation rates. The growth rate during the decade remained positive, with a tremendous average rate of around 37%. Overall, the statistics imply that the intensive margin is a more vital driver in the recent e-commerce development in China.

The e-commerce development is unbalanced among regions. Whereas the adoption rate in the eastern area has been at the top across the years, the differences between regions narrowed down gradually. In contrast, the difference in transaction volumes between regions enlarged. Intriguingly, while the adoption rate in the middle region is lower than that in the western region, the transaction volume per capita in the middle region is higher. Overall, the considerable heterogeneity of e-commerce within China allows me to identify its impact on labor market outcomes across provinces.

---

[23] The demarcation of eastern, middle, and western regions, drawn from the National Bureau of Statistics, is a coarse geographical division based on economic development. For example, in 2018, the average GDP of provinces in the eastern region was 45,444 ten million, compared to 28,134 in the middle and 15,762 in the western.

[24] Although the discussion about the decrease in e-commerce usage is out of this paper's scope, one possible explanation is that China has passed the peak of e-commerce promotion. In such a case, discounts in the online market have been reduced, making online shopping less attractive to marginal consumers.



Percentage of the population adopting online shopping   Online retail transaction volumes and annual growth rate

Figure 1 The development of e-commerce

Notes: This figure shows the development of e-commerce by year and region. Two e-commerce measures are adopted: the percentage of the population adopting online shopping, serving as a measure of extensive margins, and the online retail transaction volumes per capita (10,000 CNY, 2009 price), serving as a measure of intensive margins.

Data sources: The percentage measurement (left panel) is estimated from the CHFS 2013 – 2019. The volume measurement (right panel) is from the National Bureau of Statistics.

Table 1 presents summary statistics on labor market outcomes from the CHFS. We can only use data from respondents observed at least twice in an individual fixed effect model. Therefore, we first present the mean over the full and traced samples. It turns out that the means across the two samples are comparable, alleviating the concern that the successfully traced sample is self-selective in labor market outcomes. In the traced sample of the CHFS data set, around 78% of the respondents aged 18-55 had jobs. Approximately 10% of the respondents ran their businesses, 46% got employed, and 18% worked as farmers.[25] Regarding employment providers, a majority of them worked in enterprises (25%) or individual businesses and others (14%), compared to a relatively small fraction working for public institutions (9%).

Next, we divide the traced sample, year by year, into two groups based on whether the respondents live in provinces with online retail per capita lower than the sample median. We then show the means and their differences across subsamples. The results show that respondents residing in areas with more e-commerce activities had better labor market outcomes, including higher employment rates, more people relocating to the secondary and tertiary sectors, and higher wages or household incomes.

[25] As the occupation outcomes are not conditional on being employed, the sum of the percentage of each postion is not equal to 100%.

Table 1 Summary statistics: outcomes from the CHFS

	Full sample	Traced sample	High online retail per capita	Low online retail per capita	Diff.
Individual observations					
Have a job	0.77	0.78	0.78	0.77	0.01***
	(0.42)	(0.42)			
Employer or self-employed worker	0.10	0.10	0.11	0.09	0.02***
	(0.30)	(0.30)			
Employee	0.48	0.46	0.51	0.39	0.12***
	(0.50)	(0.50)			
Farmer	0.16	0.18	0.13	0.26	-0.13***
	(0.36)	(0.38)			
Work for public institutions	0.09	0.09	0.09	0.09	0.00***
	(0.29)	(0.29)			
Work for enterprises	0.27	0.25	0.29	0.19	0.11***
	(0.44)	(0.44)			
Work for individual businesses or others	0.14	0.14	0.14	0.13	0.01***
	(0.34)	(0.34)			
Ln(Wages)[26]	9.86	9.89	9.97	9.72	0.25***
	(1.48)	(1.40)			
Observations	247,898	170,309			
Household observations					
Ln(Total household incomes)	10.09	10.06	10.21	9.78	0.43***
	(2.22)	(2.17)			
Observations	148,392	110,571			

Notes: This table reports the mean of variables. Standard deviations are shown in parentheses. The individual sample is restricted to respondents who are 18 – 55 years of age and not full-time students. The division of high or low online retail per capita is defined by whether the number is lower than the sample median.

Significant levels: $*** \, p < 0.01, \; ** \, p < 0.05, \; * \, p < 0.1$

Data sources: CHFS 2011, 2013, 2015, 2017, and 2019

Table 2 reports summary statistics on labor market outcomes from the CFPS. Similarly, statistics across the full and traced samples are comparable. The employment rate is around 77% in the traced sample, almost the same as in the CHFS. The traced sample indicates that for people holding jobs in successive waves, 42% of them switched jobs (at the 5-digit level).

[30] The number may underestimate the job-hopping rate as we can only measure their occupations at the category level while people can hop jobs within the same category. Among the job-switchers, more than half of them changed jobs at the one-digit level (0.28/0.42), indicating that a change in jobs tends to accompany a significant shift in tasks. As for industries, approximately 25% of the respondents relocated to different industries. Next, we divided the traced sample into two groups based on

[26] Based on data availability, observations of wages are 116,461 in the full sample and 62,284 in the traced sample.

[30] An individual should be observed at least three times to enter the analysis sample of job switching.

regional e-commerce activities. The difference in mean shows that respondents exposed to more active e-commerce markets had higher employment and job-switching rates.

Table 2 Summary statistics: outcomes from the CFPS

	Full sample		Traced sample		High online retail per capita	Low online retail per capita	
	Mean	Obs.	Mean	Obs.	Mean	Mean	Diff.
Have a job	0.76	113024	0.77	104077	0.79	0.76	0.03***
	(0.42)		(0.42)				
Change in occupations (1-digit)	0.29	50955	0.28	43832	0.32	0.24	0.07***
	(0.45)		(0.45)				
Change in occupations (3-digit)	0.38	50950	0.37	43825	0.42	0.31	0.11***
	(0.48)		(0.48)				
Change in occupations (5-digit)	0.43	50162	0.42	43042	0.48	0.35	0.13***
	(0.50)		(0.49)				
Change in industries	0.26	47119	0.25	39775	0.28	0.22	0.06***
	(0.44)		(0.43)				

Notes: This table reports the mean of variables. Standard deviations are shown in parentheses. The sample is restricted to respondents who are 18-55 years of age and not full-time students. The division of high or low online retail per capita is defined by whether the number is lower than the sample median. The change in occupations can only be identified for respondents who remain employed in successive survey waves.

Significant levels: $***\, p < 0.01,\ **\, p < 0.05,\ *\, p < 0.1$

Data sources: CFPS 2010, 2012, 2014, 2016, and 2018

Finally, we present some evidence on the justification for using news articles to measure the stringency of the SUCI policy. Figure 2 shows the number of news articles mentioning the SUCI policy by year and newspaper type. The number surged in 2015 when the central government announced the SUCI policy, and it gradually faded after 2019. Although articles in the CPC papers only take up a small portion of the total number, the trends in the prefecture and province CPC papers are mostly parallel to the overall pattern. In the analysis below, we use the accumulative weighted sum of news articles by province and year, as elaborated in Section 2.2.3, to instrument the e-commerce measurements.

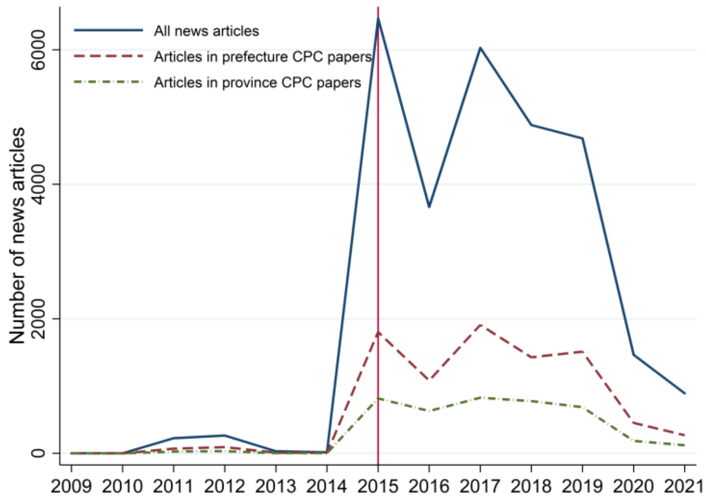


Figure 2 Trends in the number of news articles mentioning the Internet policy

Notes: This figure illustrates the number of news articles mentioning the "Speed up and Cheapen the Internet" policy by year and newspaper type. The verticle red line over 2015 indicates the official beginning year of the policy.

Data sources: the WiseNews Database

Suppose the number of news articles serves as a measure of policy stringency. In that case, more SUCI-related news articles should significantly decrease household Internet costs and increase the Internet speed people can enjoy, increasing Internet usage. Although there is no public data on Internet speed, we provide suggestive evidence of the news measure's effect on Internet cost and users. The CHFS gleaned households' communications costs, including telephone, Internet, and postage expenses. We deflate households' communications costs into the 2009 price and take its logarithm form as the dependent variable. While the variable does not solely measure the Internet costs, we still expect a negative coefficient of the news measure if there is no systematic change in the composition of communications costs and the decrease in Internet expenses dominates the total costs. Columns 1 – 2 of Table A3 show such an investigation. In addition, we collected provincial-level series on the number of Internet users from 2014 to 2020 and checked whether the news measurement is positively associated with it. Columns 3 – 5 of Table A3 present the results.

In line with our expectation, conditioning on the household workforce, residential place (rural or urban), and province and year fixed effects, the coefficient of the news-article measurement is negatively significant at the 1% level (column 1 of Table A3). A ten percentage point increase in the accumulative weighted number of SUCI-related news articles will lead to a 0.9 percentage points decrease in household communications costs. Although controlling the city fixed effects makes the coefficient insignificant, the negative sign retains (column 2 of Table A3). Similarly, conditioning on province and year fixed effects, a ten percentage point increase in the news-article measurement increases the Internet users in a province by more than one percentage point, with the significance at the 5% level (column 3 of Table A3). The estimates remain robust after controlling the logarithm form of province population and GDP (column 4 of Table A3) and clustering the standard errors at the province level (column 5 of Table A3, a wild-bootstrap adjusted p-values is provided to account for the small number of clusters). The evidence gives us more confidence that the more news articles mentioning the SUCI policy capture the more stringent implementation and significant improvement in Internet services.

Some may concern that the policy stringency is endogenous to the local market. For example, people in more economically developed areas demand better Internet services, and the policy stringency will respond to such appeals. To examine the possibility, we run regressions of my news-article measure on several one-year-lagged demographic and economic characteristics. They include the percentage of the population aged 0-14 and 15-64 (proxies for demographic structure), GDP, GDP in the primary, secondary, and tertiary sectors, consumption per capita, and disposable income per capita. The measurements are in logarithm form except for the percentage measures, and all monetary variables are deflated into the 2009 price. The results are shown in Table A4. In the first nine columns, we examine each variable separately and, in the final column, include all the predictive variables in one regression. The standard errors are clustered at the province level, and wild-bootstrap adjusted p-values are provided. It shows that all the coefficients are insignificant, except the population in the logarithm form has a fragile significant result (at the 10% level). These results suggest that the stringency of policy implementation is mostly independent of socioeconomic conditions in a province.

3. Main Results: Individual-level Evidence

In this section, we first investigate the impact of e-commerce development on various occupational outcomes, household incomes, and individual wages. Next, we provide validity and robustness checks for the main results. Finally, we examine the heterogeneous effects of e-commerce.

3.1. The impact of e-commerce on occupational choices

3.1.1. Employment status

We start by investigating the impact of e-commerce on employment rates in Table 3. The CHFS and CFPS both collect information on respondents' employment status. We define a dummy variable equal to one if respondents report that they have jobs. For comparison, columns 1 – 2 present the results based on the CHFS, and columns 3 – 4 present the results based on the CFPS. Two e-commerce measures at the province level are used: the percentage of the population adopting online shopping, serving as a measure of extensive margins, and the online retail transaction volumes per capita, serving as a measure of intensive margins. Panel A of Table 3 shows OLS estimates and Panel B shows IV estimates. For IV estimates, *p*-values from weak-IV-robust Anderson-Rubin tests (henceforth, the AR test) of the explanatory variable, first-stage estimates, and K-P *F* statistics are reported. All the regressions control individual and year fixed effects. Standard errors are clustered at the individual level.

Estimates show a positive effect of e-commerce on the probability of having a job. Most of them are significant at the 1% level. The IV estimates are larger than their OLS counterparts for both data sets, which may reflect a local average treatment effect (Angrist and Pischke 2008). While the K-P *F* statistics outstrip 10 (the rules-of-thumb criterion for a weak IV), the AR *p*-values double assure the statistical significance of the IV estimates. In line with the expectation, the coefficients in the first-stage estimations, i.e., the effect of the SUCI policy stringency on e-commerce development, are positive.

Taking the IV estimates as the preferred results, if the percentage of the population adopting online shopping increases by ten percentage points, employment rates among the labor force aged 18-55 will increase by six to seven percentage points. On the other hand, if the online retail volumes per capita increase by 1000 CNY (2009 price), the employment rate will increase by around two percentage points. These results establish that e-commerce development in a province increases employment rates on average.

Table 3 The impact of e-commerce on employment status

	(1)	(2)	(3)	(4)
	Have a job			
Data source	CHFS		CFPS	
Panel A: OLS estimates				
Percentage of population adopting online shopping	0.256***		0.496***	
	(0.026)		(0.043)	
Online retail transaction volumes per capita		0.055***		0.004
		(0.005)		(0.006)
R-squared	0.60	0.58	0.60	0.50
Panel B: IV estimates				
Percentage of population adopting online shopping	0.566***		0.983***	
	(0.115)		(0.152)	
Online retail transaction volumes per capita		0.180***		0.205***
		(0.037)		(0.025)
Anderson-Rubin test	0.00	0.00	0.00	0.00
First stage	0.039 (0.001)	0.114 (0.003)	0.043 (0.001)	0.168 (0.003)
K-P *F* statistics	2346	1404	1531	2659
Observations	155954	170309	58631	104077
Individual fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes

Notes: This table reports OLS and IV estimates of e-commerce's impact on employment rates. Observations are at the individual-by-year level. The sample is restricted to respondents who are 18 – 55 years of age and not full-time students. Two e-commerce measures are adopted: the percentage of the population adopting online shopping, serving as a measure of extensive margins, and the online retail transaction volumes per capita (10,000 CNY, 2009 price), serving as a measure of intensive margins. Both measurements vary by province and year. For IV estimates, the e-commerce measures are instrumented with a stringency measure of an Internet policy. The p-values from the weak-IV-robust Anderson-Rubin tests of the explanatory variable, first-stage estimates of the policy stringency measure, and K-P F statistics are reported. All the regressions control individual and year fixed effects. Standard errors in parentheses are clustered at the individual level.

Significant levels: $*** \, p < 0.01,\ ** \, p < 0.05,\ * \, p < 0.1$

Data sources: CHFS 2013 – 2019 and CFPS 2014 – 2018 are used when using the percentage measure, and CHFS 2011 – 2019 and CFPS 2010 – 2018 are used when using the volume measure.

3.1.2. Position

Given that e-commerce development in a region boosts employment, how does it affect labor force distribution across jobs? In the CHFS, respondents reported their roles in the workplace. We define dummy variables to indicate three positions: employer or self-employed worker, employee (i.e., work for firms or other entities), and farmer. Table 4 reports the results of these outcomes, with a similar presentation logic to Table

3.

The first finding from Table 4Table 3 is that e-commerce significantly pushes people to work as employees (columns 3 – 4) and cease to be farmers (columns 5 – 6), but it imposes no impact on being an employer or self-employed worker (columns 1 – 2). Again, the IV estimates are larger than their OLS counterparts. The K-P *F* statistics and AR tests confirm the precision of the estimates. If the percentage of the population adopting online shopping increases by ten percentage points, the probability of being an employee among the respondents aged 18-55 will increase by 13 percentage points, and the probability of being a farmer will decrease by five percentage points. If the online retail volumes per capita increase by 1000 CNY (2009 price), the probability of being an employee will increase by 3.5 percentage points, and the probability of being a farmer will decrease by one percentage point.

The OLS estimates of the percentage and volume measurements have opposite signs for the indicators "being an employer or self-employed worker" and "being a farmer." The opposite signs may imply the importance of local and external e-commerce transactions in affecting these two positions. As online retail transaction volumes are counted based on firms running e-commerce businesses, the statistic includes trading volumes within and outside a firm's home province. Online trading with other provinces may involve additional macro bilateral confounders, complicating the estimates' implications. Using the local SUCI policy stringency to instrument the volume measurement, we extract variations from consumers residing in the same province as those firms, making the variations more comparable to the percentage measurement. As expected, for the farmer indicator, the IV estimate of the volume measure has the same sign as the percentage measure, and both are negatively significant at the 1% level. While the opposite signs of the two measures remain in the IV estimation for the employer indicator, both e-commerce estimates are insignificant.

Some research has pointed out differences in abilities and activities between entrepreneurs and self-employed workers (Levine and Rubinstein 2018). Treating the employer and self-employed worker as one identity may be inappropriate. However, the division between the two roles is unavailable before the CHFS 2017. We check whether we can observe significant results by separating the two roles using the recent two waves of data (i.e., the CHFS 2017 and 2019). The results are shown in Table A5. However, none of the coefficients has significant IV estimates.

Table 4 The impact of e-commerce on positions

	(1)	(2)	(3)	(4)	(5)	(6)
	Employer or self-employed worker		Employee		Farmer	
Panel A: OLS estimates						
Percentage of population adopting online shopping	0.049***		0.598***		-0.344***	
	(0.017)		(0.029)		(0.019)	
Online retail transaction volumes per capita		-0.014***		0.058***		0.013***
		(0.004)		(0.006)		(0.004)
R-squared	0.66	0.64	0.69	0.67	0.72	0.71
Panel B: IV estimates						
Percentage of population adopting online shopping	0.076		1.260***		-0.493***	
	(0.082)		(0.129)		(0.095)	
Online retail transaction volumes per capita		-0.002		0.353***		-0.092***
		(0.025)		(0.041)		(0.030)
Anderson-Rubin test	0.35	0.93	0.00	0.00	0.00	0.00
First stage	0.039 (0.001)	0.114 (0.003)	0.039 (0.001)	0.114 (0.003)	0.039 (0.001)	0.114 (0.003)
K-P F statistics	2346	1404	2346	1404	2346	1404
Observations	155954	170309	155954	170309	155954	170309

Notes: This table reports OLS and IV estimates of e-commerce's impact on positions. Observations are at the individual-by-year level. The sample is restricted to respondents who are 18-55 years of age and not full-time students. Two e-commerce measures are adopted: the percentage of the population adopting online shopping, serving as a measure of extensive margins, and the online retail transaction volumes per capita (10,000 CNY, 2009 price), serving as a measure of intensive margins. Both measurements vary by province and year. For IV estimates, the e-commerce measures are instrumented with a stringency measure of an Internet policy. The p-values from the weak-IV-robust Anderson-Rubin tests of the explanatory variable, first-stage estimates of the policy stringency measure, and K-P F statistics are reported. All the regressions control individual and year fixed effects. Standard errors in parentheses are clustered at the individual level.

Significant levels: $*** \, p < 0.01$, $** \, p < 0.05$, $* \, p < 0.1$

Data sources: CHFS 2013 – 2019 are used when using the percentage measure, and CHFS 2011 – 2019 are used when using the volume measure.

As the classification of positions provided by the CHFS remains rough, we supplement the analysis using data on occupational categories from the CFPS. We generate dummy variables for each general category of occupations over the whole workforce (not conditional on being employed). Figure 3 displays the IV estimates of e-commerce's impact on the probability of working in each occupation. Each point in Figure 3 represents an estimate from a regression, with a line representing the 95% confidence interval and capped spikes representing the 90% ones. The p-value obtained from the AR test is attached to each estimate. The left subgraph shows estimates of the percentage of the population adopting online shopping, and the right one shows that of the online retail transaction volumes per capita. As the sample used for each explanatory variable is the same across outcomes, the K-P F statistics and observation number are provided under the heading of each subgraph.

Overall, the results mirror those presented in Table 4 but provide more details on labor force relocation across occupations. The estimates' signs are the same between the percentage and the volume measures. Estimates suggest that e-commerce increases the probability of working as "legislators, senior officials, and managers," "workers in commerce and service industries," and "workers in production and equipment operation." In contrast, it decreases the probability of working as "professionals in science, engineering, economics, culture, and others," "staff in administration, security, post and telecommunications, fire department, and others," and "agricultural and fishery workers." However, only the estimates for "workers in commerce and service industries" and "agricultural and fishery workers" are statistically significant.

Taken together, we conclude that e-commerce development pushes the workforce to cease to be farmers and make a living as workers in the manufacturing or service sectors.

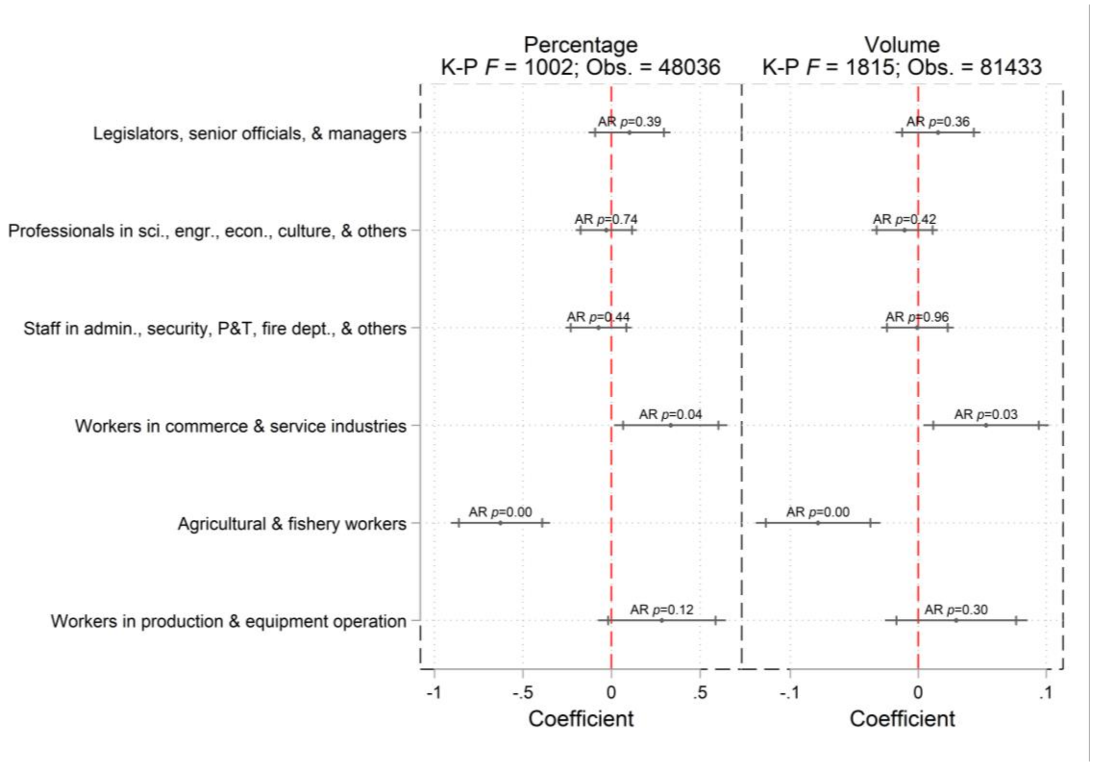


Figure 3 The impact of e-commerce on occupational categories

Notes: This figure illustrates IV estimates of e-commerce's impact on the probability of working in each occupation. For aesthetic reasons, two names of occupational categories adopt abbreviations, including "professionals in science, engineering, economics, culture, and others" (the second outcome) and "staff in administration, security, post and telecommunications, fire department, and others" (the third outcome). Each point represents an estimate from a separate regression. The line on each point represents the 95% confidence interval of the estimate, and the capped spikes specify the 90% one. The *p*-value obtained from the AR test is attached to each estimate. Two e-commerce measures are adopted: the percentage of the population adopting online shopping, serving as a measure of extensive margins, and the online retail transaction volumes per capita (10,000 CNY, 2009 price), serving as a measure of

intensive margins. Both measurements vary by province and year. The e-commerce measures are instrumented with a stringency measure of an Internet policy.
Data sources: CFPS 2014 – 2018 are used when using the percentage measure, and CFPS 2010 – 2018 are used when using the volume measure.

3.1.3. Employment providers

E-commerce can worsen the conditions of those low-productivity firms (Goldmanis et al. 2010; Chava et al. 2022). Investigating e-commerce's impact on types of employment providers can be conducive to linking the market structure to occupational choices. Based on the CHFS,we identify three types of employment providers: (1) public institutions, (2) enterprises, and (3) individual businesses or others. The results are shown in Table 5.

The IV estimates show that the development of e-commerce in a region significantly (at the 1% level) relocates the labor force to enterprises. If the share of the population adopting online shopping increases by ten percentage points, the probability of working for enterprises among the respondents aged 18-55 will increase by 11 percentage points. If the online retail volumes per capita increase by 1000 CNY (2009 price), the probability of working for enterprises will increase by 3.4 percentage points. The IV estimates of "working for the public institution" and "working for the individual business or others" are negative but insignificant.

These results are compatible with previous findings in the literature. As e-commerce boosts commercial activities and entrepreneurship, enterprises will demand more workers. On the other hand, most small offline retail stores serving the community are run by individuals or families in China. With the advent of e-commerce, such small businesses suffer from intense competition and are forced to exit (Chava et al. 2022).

Table 5 The impact of e-commerce on types of employment providers

	(1)	(2)	(3)	(4)	(5)	(6)
	Public institution		Enterprise		Individual business and others	
Panel A: OLS estimates						
Percentage of population adopting online shopping	0.025*		0.367***		0.061**	
	(0.014)		(0.029)		(0.026)	
Online retail transaction volumes per capita		-0.004		0.075***		-0.011*
		(0.004)		(0.007)		(0.006)
R-squared	0.75	0.73	0.63	0.61	0.47	0.45
Panel B: IV estimates						
Percentage of population adopting online shopping	-0.076		1.054***		-0.185	
	(0.061)		(0.129)		(0.125)	
Online retail transaction volumes per capita		-0.007		0.340***		-0.060
		(0.019)		(0.040)		(0.036)
Anderson-Rubin test	0.22	0.72	0.00	0.00	0.14	0.10
First stage	0.039 (0.001)	0.114 (0.003)	0.039 (0.001)	0.114 (0.003)	0.039 (0.001)	0.114 (0.003)
K-P F statistics	2346	1404	2346	1404	2346	1404
Observations	155954	170309	155954	170309	155954	170309

Notes: This table reports OLS and IV estimates of e-commerce's impact on types of employment providers. Observations are at the individual-by-year level. The sample is restricted to respondents who are 18-55 years of age and not full-time students. Two e-commerce measures are adopted: the percentage of the population adopting online shopping, serving as a measure of extensive margins, and the online retail transaction volumes per capita (10,000 CNY, 2009 price), serving as a measure of intensive margins. Both measurements vary by province and year. For IV estimates, the e-commerce measures are instrumented with a stringency measure of an Internet policy. The p-values from the weak-IV-robust Anderson-Rubin tests of the explanatory variable, first-stage estimates of the policy stringency measure, and K-P F statistics are reported. All the regressions control individual and year fixed effects. Standard errors in parentheses are clustered at the individual level.

Significant levels: $*** p < 0.01,\ ** p < 0.05,\ * p < 0.1$

Data sources: CHFS 2013 – 2019 are used when using the percentage measure, and CHFS 2011 – 2019 are used when using the volume measure.

I provide additional evidence on the characteristics of employment providers by examining the affiliated industries of individuals' occupations based on the CFPS. Following the same presentation logic as Figure 3, Figure 4 shows the results of e-commerce's impacts on labor force distribution across industries.

Consistent with previous results, e-commerce shows a significantly negative effect on the probability of working in "agriculture, forestry, and fishing" and a discernible positive effect on "manufacturing" and "wholesale and retail trade," no matter which measure of e-commerce is examined. Note that the positive effect on "wholesale and retail trade" does not contradict previous explanations regarding the decline of small offline retailers. The labor force may relocate from individual retail businesses to those large-scale online wholesale or retail firms. In addition, we observe that both e-commerce measures have significantly negative estimates for "human health and social work," and the percentage measure has a discernible positive estimate for "accommodation and food service." However, the magnitude of these estimates is smaller than those emphasized earlier.

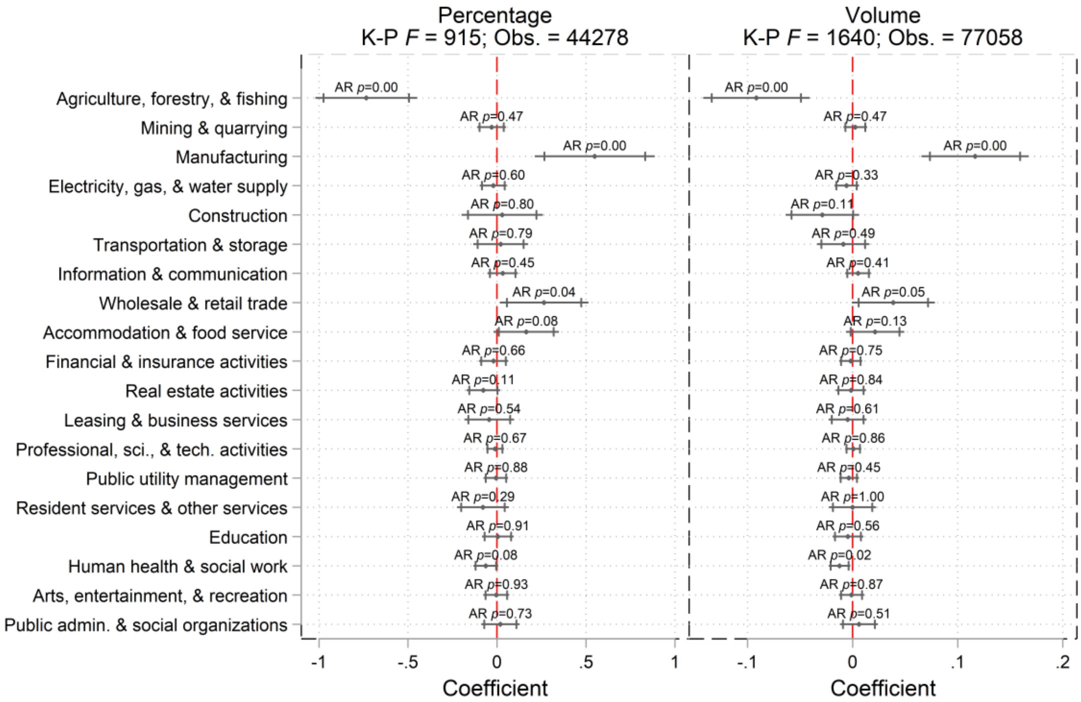


Figure 4 The impact of e-commerce on labor force distribution across industries

Notes: This figure illustrates IV estimates of e-commerce's impact on the probability of working in each industry. For aesthetic reasons, several names of industries adopt abbreviations, with "sci," "tech," and "admin" standing for science, technology, and administration, respectively. Each point represents an estimate from a separate regression. The line on each point represents the 95% confidence interval of the estimate, and the capped spikes specify the 90% one. Two e-commerce measures are adopted: the percentage of the population adopting online shopping, serving as a measure of extensive margins, and the online retail transaction volumes per capita (10,000 CNY, 2009 price), serving as a measure of intensive margins. Both measurements vary by province and year. The e-commerce measures are instrumented with a stringency measure of an Internet policy.

Data sources: CFPS 2014 – 2018 are used when using the percentage measure, and CFPS 2010 – 2018 are used when using the volume measure.

3.1.4. Job-switching

Does e-commerce bring the economy more creative destruction and induce more job-switching among the labor force?We construct four indicators to gauge an individual's job-switching behavior: change occupations at the one-digit level, three-digit level, and five-digit level (from high to low in degrees of the change), respectively, and change working industries. Table 6 reports the results.

Both OLS and IV estimates are positive, with the IV ones larger than their OLS counterparts, implying that e-commerce development will induce people to switch jobs more frequently. In line with expectations, the IV estimates of the two e-commerce measures increase as the digit level lowers, indicating that the changes are likelier among occupations with similar tasks. Such results make sense as the larger differences in job tasks, the higher costs and friction in job-switching (Lee and Wolpin 2006). Specifically, as the percentage of the population adopting online shopping increases by ten percentage points, the probability of changing occupations in successive waves will increase by 6.61%, 7.82%, and 10.16% at the one-digit, three-digit, and five-digit levels, respectively. All the estimates are significant at the 10% level or above. For the volume measure, if the online retail volumes per capita increase by 1000 CNY (2009 price), the probability will increase by 0.88%, 1.05%, and 1.46%, respectively, and the estimates of 1.05% and 1.46% are significant at the 10% level or above.

For the change in working industries, only the IV estimate of the percentage measure shows a weak significant result. The results make sense as changing the working industries is usually accompanied by a considerable change in tasks, and therefore, people may face a tremendous challenge in switching jobs across industries, at least in the short term.

In summary, Table 6 confirms that e-commerce, a new business model that transforms our lives, may involve creative destruction and require people to handle new tasks in their jobs if they want to remain employed.

Table 6 The impact of e-commerce on job-switching

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Change in occupations (1-digit)		Change in occupations (3-digit)		Change in occupations (5-digit)		Change in industries	
Panel A: OLS estimates								
Percentage of population adopting online shopping	0.066		0.109		0.113		0.088	
	(0.077)		(0.078)		(0.076)		(0.078)	
Online retail transaction volumes per capita		0.031**		0.026**		0.030**		0.016
		(0.012)		(0.012)		(0.012)		(0.013)
R-squared	0.56	0.52	0.61	0.58	0.65	0.62	0.59	0.55
Panel B: IV estimates								
Percentage of population adopting online shopping	0.661*		0.782**		1.016***		0.654*	
	(0.354)		(0.351)		(0.335)		(0.371)	
Online retail transaction volumes per capita		0.088		0.105*		0.146**		0.100
		(0.063)		(0.063)		(0.061)		(0.069)
Anderson-Rubin test	0.06	0.16	0.03	0.10	0.00	0.02	0.08	0.15
First stage	0.033 (0.002)	0.126 (0.006)	0.033 (0.002)	0.126 (0.006)	0.033 (0.002)	0.127 (0.006)	0.031 (0.002)	0.115 (0.006)
K-P F statistics	474	517	474	516	476	514	395	413
Observations	32789	43832	32785	43825	32558	43042	29225	39775

Notes: This table reports OLS and IV estimates of e-commerce's impact on job-switching. Observations are at the individual-by-year level. The sample is restricted to respondents who are 18-55 years of age and not full-time students. Two e-commerce measures are adopted: the percentage of the population adopting online shopping, serving as a measure of extensive margins, and the online retail transaction volumes per capita (10,000 CNY, 2009 price), serving as a measure of intensive margins. Both measurements vary by province and year. For IV estimates, the e-commerce measures are instrumented with a stringency measure of an Internet policy. The p-values from the weak-IV-robust Anderson-Rubin tests of the explanatory variable, first-stage estimates of the policy stringency measure, and K-P F statistics are reported. All the regressions control individual and year fixed effects. Standard errors in parentheses are clustered at the individual level.

Significant levels: $***\, p < 0.01,\ **\, p < 0.05,\ *\, p < 0.1$

Data sources: CFPS 2014 – 2018 are used when using the percentage measure, and CFPS 2010 – 2018 are used when using the volume measure.

3.2. The impact of e-commerce on income

There are mixed findings on how new technology affects incomes and wages depending on the context and the technology examined (Hjort and Poulsen 2019; Acemoglu and Restrepo 2020; Chava et al. 2022; Albanesi et al. 2023; Braxton and Taska 2023). To complete the picture of e-commerce's impacts on labor market outcomes,we examine its impacts on household incomes and personal wages. Table 7 reports the results. All the monetary measurements are expressed in the 2009 price before taking the logarithm form. Household and year fixed effects are controlled for regressions of total household incomes. Individual and year fixed effects are controlled for regressions of wages.

The IV estimates imply that if the share of the population adopting online shopping increases by ten percentage points, total household incomes will increase by 43 percentage points, and individual wages will increase by 13 percentage points. All these estimates are significant at the 5% level or above. On the other hand, if the online retail volumes per capita increase by 1000 CNY (2009 price), total household incomes will increase by 13 percentage points and individual wages will increase by 3.5 percentage points. However, only the estimate for total household incomes is significant at the 5% level. These results suggest that, on average, e-commerce can increase total household incomes and individual wages.

Table 7 The impact of e-commerce on income

	(1)	(2)	(3)	(4)
	Ln(Total household income)		Ln(Wages)	
Panel A: OLS estimates				
Percentage of population adopting online shopping	-0.124		0.170	
	(0.193)		(0.179)	
Online retail transaction volumes per capita		0.070**		0.060*
		(0.032)		(0.034)
R-squared	0.52	0.51	0.54	0.53
Panel B: IV estimates				
Percentage of population adopting online shopping	4.305***		1.262**	
	(1.586)		(0.558)	
Online retail transaction volumes per capita		1.312**		0.348
		(0.526)		(0.235)
Anderson-Rubin test	0.01	0.01	0.02	0.14
First stage	0.018 (0.001)	0.051 (0.002)	0.050 (0.001)	0.113 (0.006)
K-P F statistics	1101	459	1604	384
Observations	102632	110571	58863	62283
Household fixed effects	Yes	Yes		
Individual fixed effects			Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes

Notes: This table reports OLS and IV estimates of e-commerce's impact on incomes and wages. Observations are at the household-by-year level in columns 1 – 2 and are at the individual-by-year level in columns 3 – 4. For the analysis at the individual level, the sample is restricted to respondents who are 18-55 years of age and not full-time students. Two e-commerce measures are adopted: the percentage of the population adopting online shopping, serving as a measure of extensive margins, and the online retail transaction volumes per capita (10,000 CNY, 2009 price), serving as a measure of intensive margins. Both measurements vary by province and year. For IV estimates, the e-commerce measures are instrumented with a stringency measure of an Internet policy. The p-values from the weak-IV-robust Anderson-Rubin tests of the explanatory variable, first-stage estimates of the policy stringency measure, and K-P F statistics are reported. Standard errors in parentheses are clustered at the household level in columns 1 – 2 and individual level in columns 3 – 4.

Significant levels: $*** p < 0.01$, $** p < 0.05$, $* p < 0.1$
Data sources: CHFS 2013 – 2019 are used when using the percentage measure, and CHFS 2011 – 2019 are used when using the volume measure.

3.3. Validity and robustness checks

3.3.1. Evidence on the exclusion restriction of the IV

To provide supportive evidence on the exclusion requirement of the IV, we draw on the idea of the zero-first-stage (ZFS) test first suggested by Bound and Jaeger (2000), but the logic is slightly different. In the context of this paper, we restrict the sample to respondents who are locals (i.e., those who are not migrant workers), never shop online, and live in rural and economically underdeveloped areas. These people should be exempt from the impact of e-commerce. Suppose the SUCI policy affects occupational choices only through the development of e-commerce. In that case, reduced-form regressions on this subsample, i.e., regressing employment outcomes on the policy stringency measurement, should yield insignificant estimates.

In practice, we first restricted the respondents who are locals, never shop online, and live in rural areas. We then divided the sample into three groups based on the percentile of province GDP year by year. Panels A, B, and C of Table 8 show the results from reduced-form regressions over the respondents residing in provinces with low, middle, and high levels of GDP, respectively. Table 9 presents the same exercise for the outcomes from the CFPS.

In line with the assumption, estimates based on the sample from economically underdeveloped regions are insignificant (Panel A of Table 8 and Table 9). Significant estimates in the same direction as benchmark analyses appear as the GDP level rises. These results make sense as people residing in more developed regions will be indirectly affected by e-commerce as it transforms the market structure. Moreover, the significant results in Panel B and C imply that the insignificant results in Panel A are not merely due to the small sample size. In summary, the results in Table 8 and Table 9 give us more confidence that the exclusion restriction is satisfied under the specification of Equations (1) and (2).

Table 8 Test for the exclusion restriction of the IV: outcomes from the CHFS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GDP level	Have a job	Employer or self-employed worker	Employee	Farmer	Public institution	Enterprise	Individual businesses and others	Ln(Wages)
Panel A: Low								
L1.Ln(Accumulative Number of articles in CPC newspapers)	-0.030	0.014	-0.015	0.007	0.016	-0.030	-0.001	-0.147
	(0.028)	(0.014)	(0.025)	(0.030)	(0.011)	(0.020)	(0.023)	(0.236)
R-squared	0.56	0.54	0.62	0.63	0.57	0.56	0.47	0.55
Observations	11380	11380	11380	11380	11380	11380	11380	1201
Panel B: Middle								
L1.Ln(Accumulative Number of articles in CPC newspapers)	-0.027	0.008	0.026	-0.065***	0.010	0.019	0.003	-0.081
	(0.017)	(0.010)	(0.018)	(0.021)	(0.008)	(0.014)	(0.017)	(0.180)
R-squared	0.58	0.57	0.65	0.67	0.61	0.57	0.51	0.56
Observations	10374	10374	10374	10374	10374	10374	10374	1629
Panel C: High								
L1.Ln(Accumulative Number of articles in CPC newspapers)	0.031*	0.022**	0.003	0.023	-0.010	0.029*	-0.008	-0.046
	(0.017)	(0.011)	(0.018)	(0.019)	(0.007)	(0.015)	(0.017)	(0.125)
R-squared	0.58	0.60	0.64	0.68	0.61	0.56	0.49	0.50
Observations	12235	12235	12235	12235	12235	12235	12235	2276

Notes: This table reports reduced-form estimates of the policy stringency's impact on labor market outcomes by subsample. Observations are at the individual-by-year level. The sample is restricted to respondents who are 18-55 years of age, not full-time students, locals (i.e., those who are not migrant workers), and never shop online. Based on the percentile of the GDP each year, the sample is divided into three groups. Panels A, B, and C show the estimates for the respondents residing in provinces with low, middle, and high levels of GDP, respectively. The explanatory variable is the one-year-lag accumulative weighted sum of news articles mentioning the "Speed up and Cheapen the Internet" policy in the CPC newspapers (in logarithm form), which measures the policy's stringency. All the regressions control individual and year fixed effects. Standard errors in parentheses are clustered at the individual level.

Significant levels: $*** p < 0.01$, $** p < 0.05$, $* p < 0.1$

Data sources: CHFS 2011 – 2019

Table 9 Test for the exclusion restriction of the IV: outcomes from the CFPS

	(1)	(2)	(3)	(4)
GDP level	Change in occupations (1-digit)	Change in occupations (3-digit)	Change in occupations (5-digit)	Change in industries
Panel A: Low				
L1.Ln(Accumulative number of articles in CPC newspapers)	-0.037	-0.081	-0.064	-0.027
	(0.055)	(0.052)	(0.054)	(0.056)
R-squared	0.58	0.66	0.68	0.56
Observations	3230	3230	3228	3041
Panel B: Middle				
L1.Ln(Accumulative number of articles in CPC newspapers)	0.056	0.066*	0.074**	0.046
	(0.034)	(0.034)	(0.033)	(0.029)
R-squared	0.58	0.64	0.68	0.59
Observations	3560	3560	3558	3252
Panel C: High				
L1.Ln(Accumulative number of articles in CPC newspapers)	0.062**	0.058**	0.061**	0.060**
	(0.027)	(0.027)	(0.025)	(0.027)
R-squared	0.55	0.61	0.67	0.55
Observations	3971	3971	3968	3501

Notes: This table reports reduced-form estimates of the policy stringency's impact on labor market outcomes by subsample. Observations are at the individual-by-year level. The sample is restricted to respondents who are 18-55 years of age, not full-time students, locals (i.e., those who are not migrant workers), and never shop online. Based on the percentile of the GDP each year, the sample is divided into three groups. Panels A, B, and C show the estimates for respondents residing in provinces with low, middle, and high levels of GDP, respectively. The explanatory variable is the one-year-lag accumulative weighted sum of news articles mentioning the "Speed up and Cheapen the Internet" policy in the CPC newspapers (in logarithm form), which measures the policy's stringency. All the regressions control for individual and year fixed effects. Standard errors in parentheses are clustered at the individual level.

Significant levels: $***\,p < 0.01$, $**\,p < 0.05$, $*\,p < 0.1$

Data sources: CFPS 2010 – 2018

3.3.2. Robustness checks

We conduct several robustness checks for the benchmark analyses. Figure A2 illustrates the checks for employment status and positions, Figure A3 shows that for employment providers and incomes, and Figure A4 presents that for the job-switching outcomes. For comparison, we draw the benchmark estimates of the percentage and volume measures at the top two rows in each figure.

Some may concern that the main results are sensitive to the measure of e-commerce. We adopt two additional measurements to check the robustness. First, rather than using per capita online retail transaction volumes,we use the logarithm form of online retail transaction volumes as an alternative measure. Second,we use the logarithm form of monthly online expenditure per household (by province), estimated from the CHFS data, as another measure (see the details in Section 2.2.1).

[31] These exercises are denoted as “Log(vol.)” and “Log(exp.)” in the left axis of each figure, respectively. Except for the explanatory variables, all the regression settings are the same as their counterparts in Section 3.

Some may challenge that people from the same community are not independent as they face the same labor market conditions. Therefore, clustering the standard errors at the individual level may be insufficient.We apply a two-way clustering approach to address this concern: we re-estimate the equation by clustering the standard errors at the individual and community-by-year levels. These checks are denoted as “SE (pct.)” and “SE (vol.)” in the figures for the percentage and volume measures of e-commerce, respectively.

Finally, the SUCI policy may directly affect occupational choices due to Internet accessibility (Hjort and Poulsen 2019; Zuo 2021; Chiplunkar and Goldberg 2022; Bhuller et al. 2023). To address such a concern, based on the benchmark specification, we additionally control three province-level variables (lagged by one year, the same period as the IV) to capture Internet development, including the length of long-distance optical cables per capita, the length of optical cables per capita, and the number of Internet users (in logarithm form). Moreover, based on the data availability, we define two dummy variables indicating whether a household has computers or phones in the CHFS and two dummy variables indicating whether an individual is an Internet user or phone user in the CFPS, including them as control variables in the relative analysis. These exercises are denoted as “Ctrl (pct.)” and “Ctrl (vol.)” in each figure for the percentage and volume measures of e-commerce, respectively.

As shown in the figures, although the magnitude of the estimates differs, the signs and significant levels remain comparable across different robustness checks.

3.4. Heterogeneous effects of e-commerce

E-commerce’s benefits may be shared disproportionately by younger and wealthier people who live in remote areas (Sinai and Waldfogel 2004; Fan et al. 2018; Couture et al. 2021; Relihan 2022). Existing studies also point out that workers with

[31] We take these two measures as robustness checks for the following reasons. Compared to online retail transaction volumes per capita used in the benchmark analysis, the logarithm form of online retail transaction volumes, which makes the coefficient have implications for the percentage change, is less intuitive. Compared to statistics provided by the National Bureau of Statistics, the online expenditure documented in the CHFS is recalled by the respondents, which may suffer more measurement errors, such as recall bias.

higher education are winners in the era of technological progress (Forman et al. 2012; Michaels et al. 2014; Akerman et al. 2015; Dillender and Forsythe 2022). In this subsection, we conduct subsample analyses by respondents' socio-demographic characteristics to investigate the heterogeneous effect of e-commerce.

Figure 5 reports age heterogeneity by dividing the sample into three age groups: 18-30, 31-40, and 41-55 (inclusive). Figure 6 shows the subsample estimations by gender. For educational levels, we denote junior high education or below as the low level, senior high or vocational education as the middle level, and undergraduate or above as the high level. Figure 7 shows the results by educational level. Finally, we examine the heterogeneity between migrant and local workers. Figure 8 reports the results.

[32] Panel A of each figure shows the outcomes of employment status, positions, employment providers, and wages, whereas Panel B shows indicators of job-switching. The left axis of each graph indicates the dependent variables and corresponding subsamples. Estimates are plotted with a line representing the 95% confidence interval and capped spikes representing the 90% one. The *p*-value from the AR test is attached to each estimate. The left panel in each figure shows estimates of the percentage of the population adopting online shopping, and the right one shows estimates of the online retail transaction volumes per capita.

Figure 5 shows interesting patterns across age groups. E-commerce raises employment rates only for the young and middle-aged labor force (i.e., 40 years of age or below). In contrast, the estimates for employment rates among the oldest group are negative though statistically insignificant, suggesting that older workers could be losers under the shock of e-commerce. There is no discernible disparity across age groups in e-commerce's impact on getting employed or working for enterprises, with estimates across subsamples all significantly positive. On the other hand, e-commerce significantly reduces the probability of being a farmer for the younger (18-30) and older (41-55) age groups, and it significantly lowers the likelihood of working in individual businesses or others among the older labor force (aged 41-55). For wages, while estimates for the younger workers tend to be larger, none of them are significant. When investigating the job-switch outcomes, we find that people of older ages show a clear pattern: conditioning on having jobs in successive survey years, people aged 41-55 are significantly more likely to switch occupations and change working industries as e-

[32] "Migrant workers" are defined as those whose registered prefecture in the household registration system (*Hukou*) is different from their resident prefecture. In the analysis sample, around 18% of respondents aged 18-55 are defined as migrants in the CHFS. The number is 14% in the CFPS.

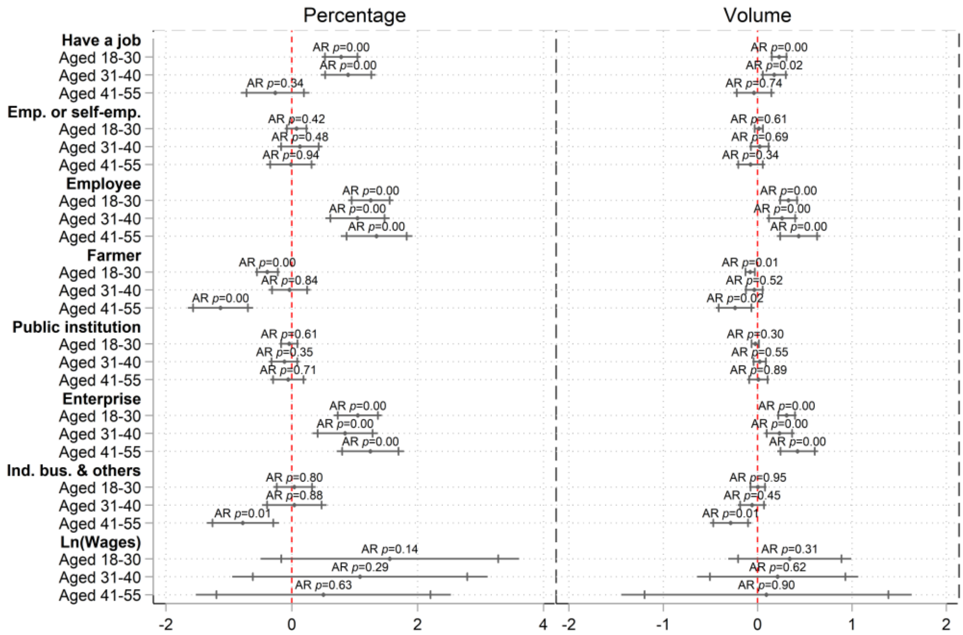
commerce develops. More importantly, the results show that the older workers who succeed in getting employed successively face more dramatic changes in job tasks, as the estimate is significant for the dummy variable indicating changing occupations at the one-digit level.

For gender heterogeneity, Figure 6 shows no apparent disparity in most occupational outcomes. One exception is that females are likelier than males to work as employers or self-employed workers under the extensive-margin shock of e-commerce (i.e., the percentage measure), which is in line with the findings that e-commerce facilitates more entrepreneurship for females than males (Huang et al. 2021). However, conditioning on having jobs successively, males tend to change their jobs more frequently than females with e-commerce development, although the switches are generally restricted to similar tasks (significant estimates for lower-level changes).

Educational attainment is another salient dimension by which e-commerce exerts its heterogeneous effects. Figure 7 suggests that e-commerce development is not a skill-biased technological change: It increases employment rates mainly for the less-educated workforce by relocating them to jobs in enterprises. Consistently, there is a significant decline in the less-educated labor force working as farmers. As a result, as e-commerce develops, less-educated workers tend to earn higher wages. Similarly, for workers who keep employed in successive waves, less-educated workers are more likely to switch occupations under the impact of e-commerce: estimates of the percentage measure are all significantly positive, whereas that of the volume measure is only significant for changing occupations at the five-digit level. These results are consistent with some industry observations that e-commerce and the standardization of programming allow less-educated workers to work as website designers or applet developers (see a case study in Appendix II).

Lastly, Figure 8 shows the heterogeneity analysis by migrant status. As location-based household surveys, the CHFS and CFPS may not collect a representative sample of migrant workers as their precise information is hard to obtain. Nonetheless, excluding migrant workers to check whether there are new patterns is still a meaningful investigation. Once excluding migrant observations, the coefficients of e-commerce's impact on employment rates turn negative, although the estimates of both e-commerce measures are insignificant. Such results imply that labor mobility is essential for e-commerce to benefit the workforce. It is worth noting that local people exposed to advanced e-commerce markets are less likely to work in individual businesses or others. The phenomenon may closely relate to the previous discussion about e-commerce crowding out small offline retailers (see a case study in Appendix II). In addition, estimates suggest that local workers who keep jobs in successive waves are more likely to switch jobs under the shock of e-commerce.

Panel A: Occupations and wages



Panel B: Job-switching

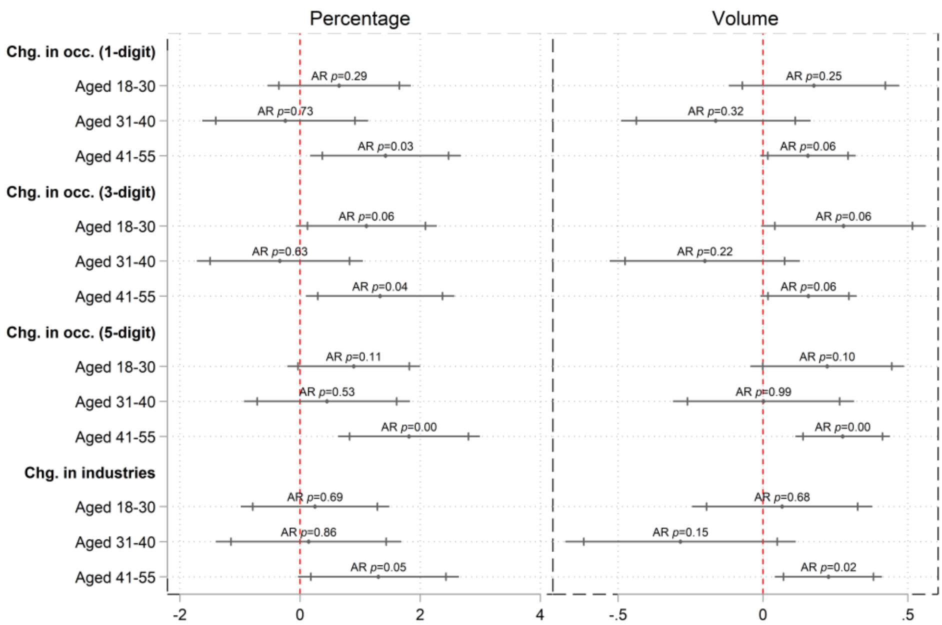
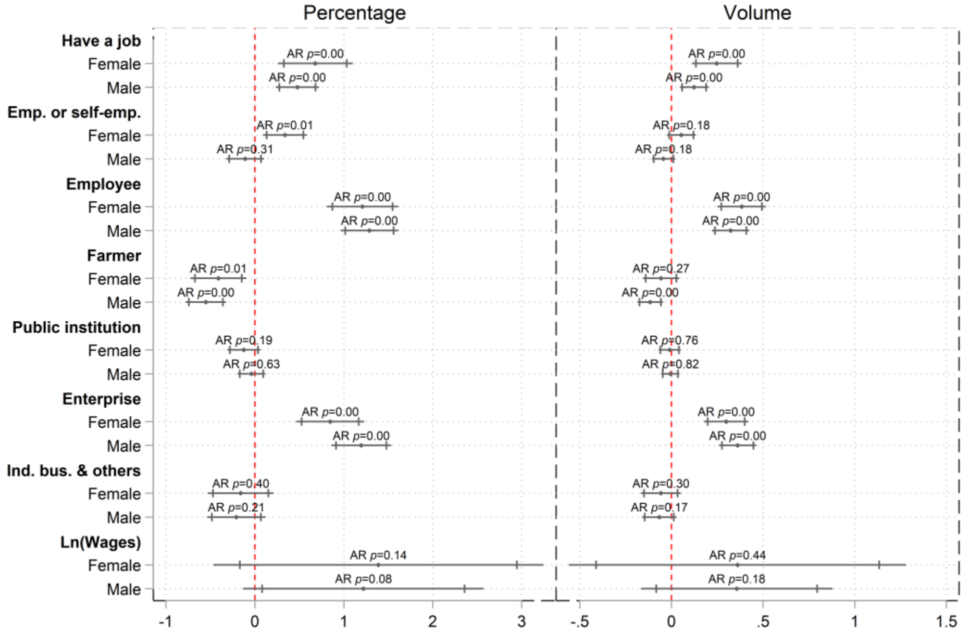


Figure 5 Heterogeneity by age group

Notes: This figure reports the IV estimates by age group. The e-commerce measures are instrumented with a stringency measure of an Internet policy. The left-handed axis indicates the dependent variables and corresponding subsamples. Estimates are plotted with a line representing the 95% confidence interval and capped spikes representing the 90% one. The *p*-value from the Anderson-Rubin test is attached to each estimate. The left panel shows estimates of the percentage of the population adopting online shopping, and the right panel shows estimates of the online retail transaction volumes per capita (10,000 CNY, 2009 price). Both e-commerce measurements vary by province and year.

Data sources: CHFS 2011 – 2019, CFPS 2010 – 2018

Panel A: Occupations and wages



Panel B: Job-switching

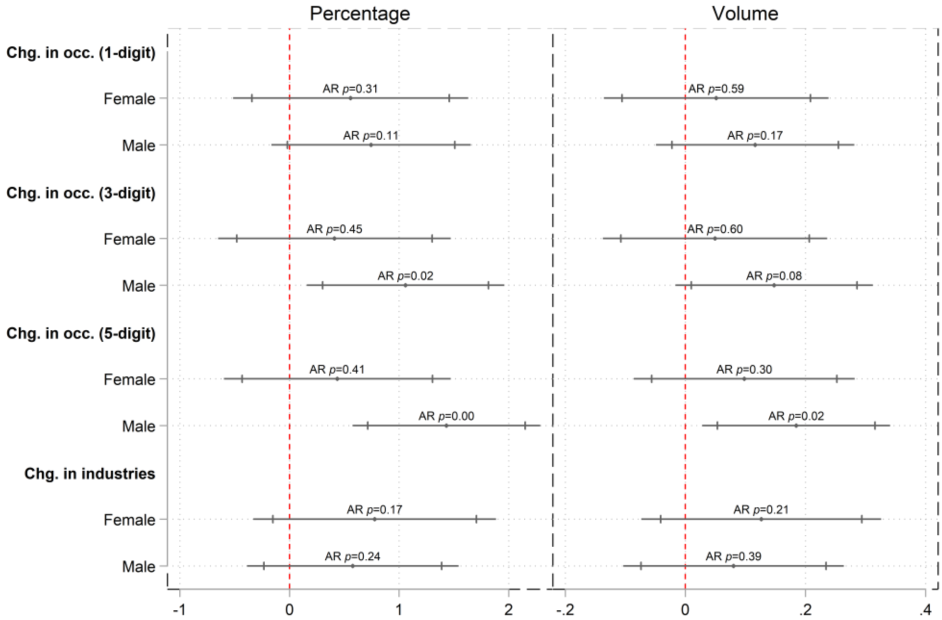
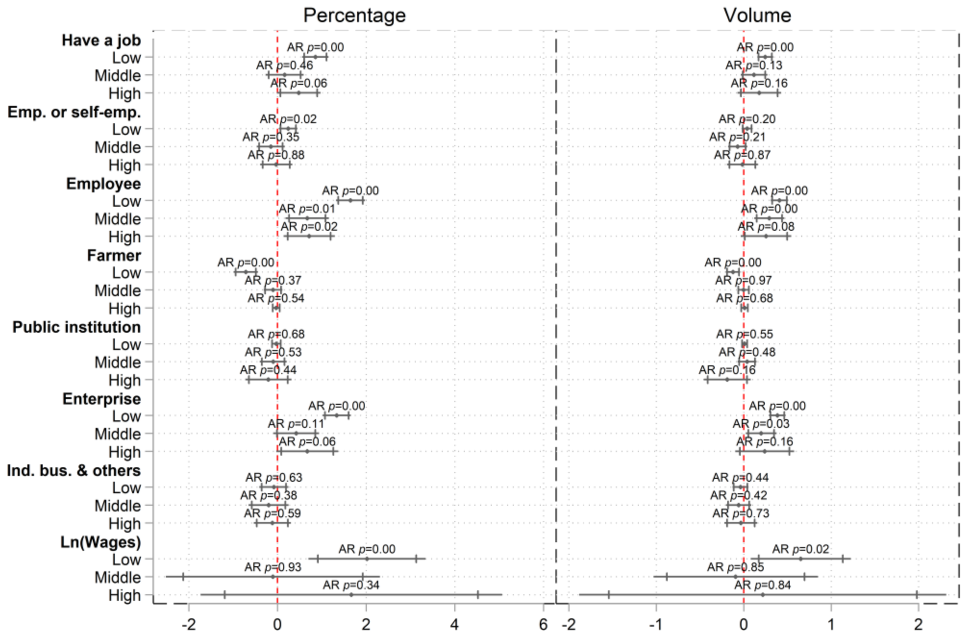


Figure 6 Heterogeneity by gender

Notes: This figure reports the IV estimates by gender. The e-commerce measures are instrumented with a stringency measure of an Internet policy. The left-handed axis indicates the dependent variables and corresponding subsamples. Estimates are plotted with a line representing the 95% confidence interval and capped spikes representing the 90% one. The *p*-value from the Anderson-Rubin test is attached to each estimate. The left panel shows estimates of the percentage of the population adopting online shopping, and the right panel shows estimates of the online retail transaction volumes per capita (10,000 CNY, 2009 price). Both e-commerce measurements vary by province and year.

Data sources: CHFS 2011 – 2019, CFPS 2010 – 2018

Panel A: Occupations and wages



Panel B: Job-switching

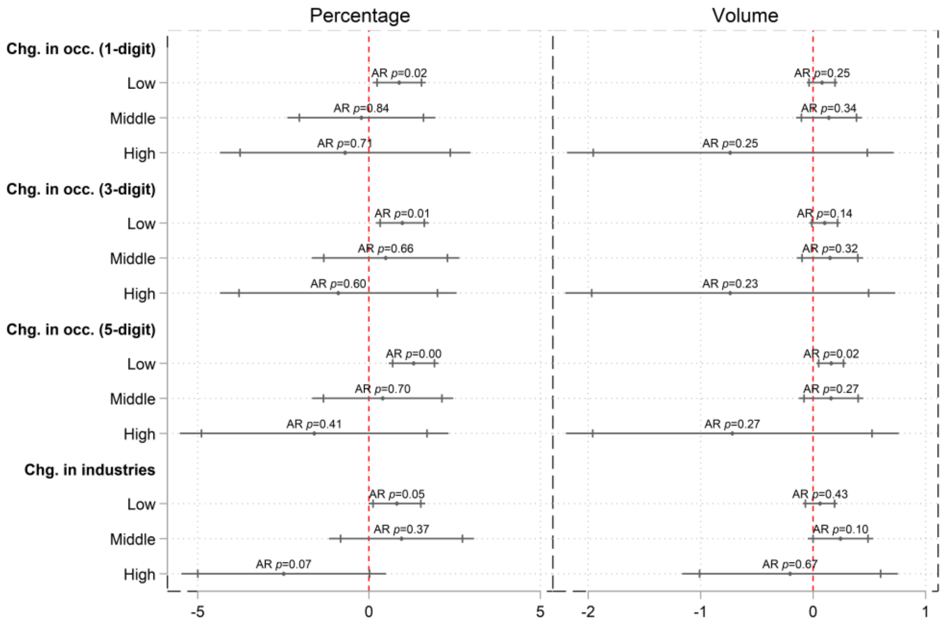
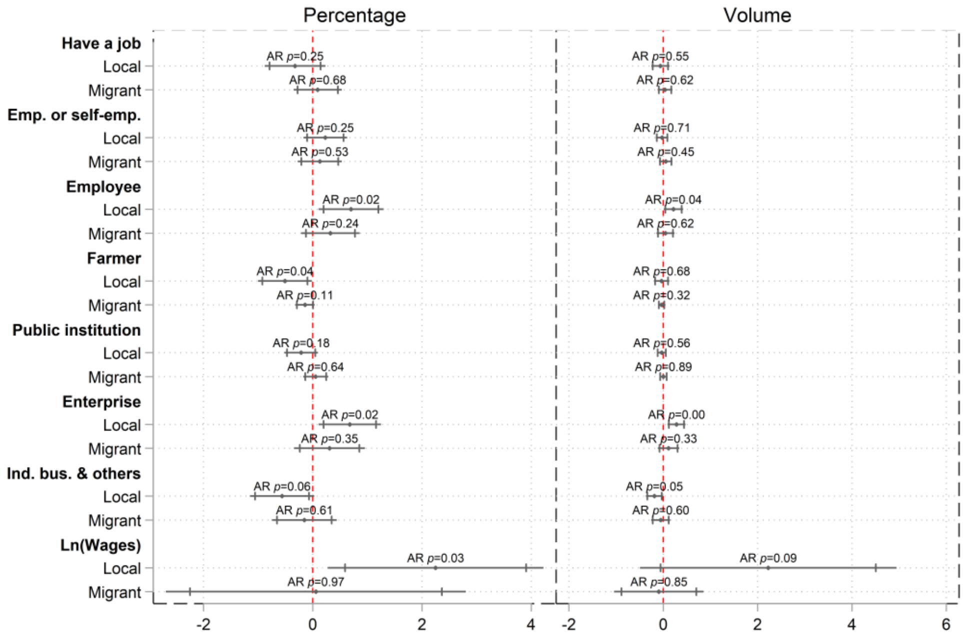


Figure 7 Heterogeneity by educational level

Notes: This figure reports the IV estimates by educational level. The e-commerce measures are instrumented with a stringency measure of an Internet policy. The left-handed axis indicates the dependent variables and corresponding subsamples. Estimates are plotted with a line representing the 95% confidence interval and capped spikes representing the 90% one. The *p*-value from the Anderson-Rubin test is attached to each estimate. The left panel shows estimates of the percentage of the population adopting online shopping, and the right panel shows estimates of the online retail transaction volumes per capita (10,000 CNY, 2009 price). Both e-commerce measurements vary by province and year.

Data sources: CHFS 2011 – 2019, CFPS 2010 – 2018

Panel A: Occupations and wages



Panel B: Job-switching

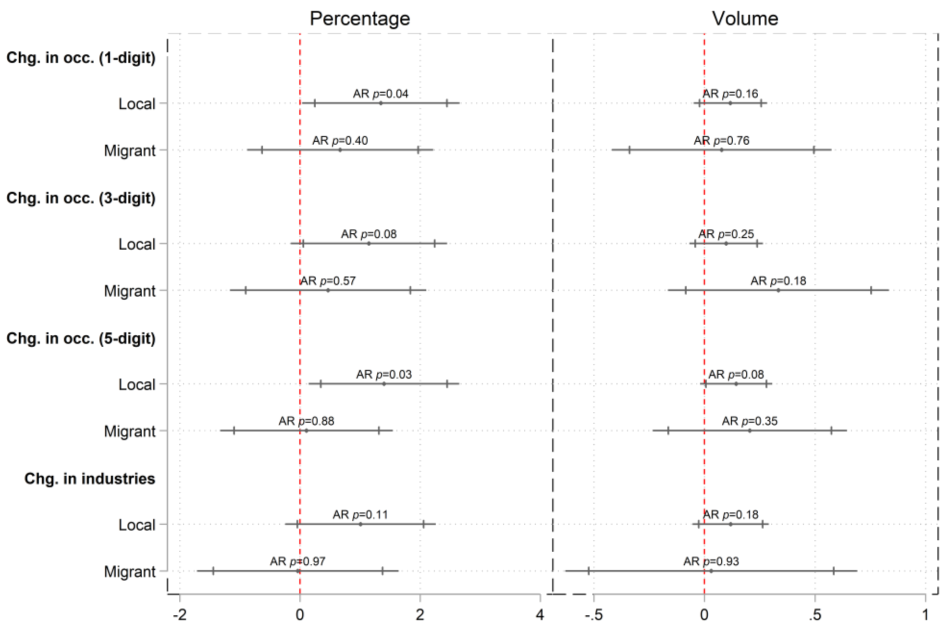


Figure 8 Heterogeneity by migrant status

Notes: This figure reports the IV estimates by migrant status. The e-commerce measures are instrumented with a stringency measure of an Internet policy. The left-handed axis indicates the dependent variables and corresponding subsamples. Estimates are plotted with a line representing the 95% confidence interval and capped spikes representing the 90% one. The *p*-value from the Anderson-Rubin test is attached to each estimate. The left panel shows estimates of the percentage of the population adopting online shopping, and the right panel shows estimates of the online retail transaction volumes per capita (10,000 CNY, 2009 price). Both e-commerce measurements vary by province and year.

Data sources: CHFS 2011 – 2019, CFPS 2010 – 2018

4. Mechanisms: Firm-level Evidence

In previous sections, we have established that e-commerce affects occupational choices based on individual-level data. The workforce can relocate only when new employment opportunities are created. Existing literature has shed light on e-commerce's impact on job creation. Using data from 1998 to 2007 in the United States, Lieber and Syverson (2012) find that a ten percentage point increase in the share of consumers shopping online in a market will induce a 2.2 opening of online businesses in local areas. Some industry observers also point out that millions of online sellers link to millions of factories in China (Lowrey 2016), creating labor demand on both online sellers' and factories' sides. Moreover, e-commerce can change the market structure, dramatically altering the allocation and demand of labor. In this section, we try to provide evidence of e-commerce's impact on firm creation and firms' employment decisions.

4.1. Firm creation

One ideal way of linking e-commerce to job creation is to use job vacancy data. As there is no publicly available data set on job vacancies in China, we draw on the Firm Registration Database from the State Administration for Market Regulation to gauge job creation indirectly. In addition, the Firm Registration Database contains information on the businesses that firms involve. Leveraging this information, we define a firm running e-commerce-related businesses if the business description or firm name hits one of a list of e-commerce-related keywords (see the details in Section 2.2.4 and the list of keywords in Table A2). The broad definition of "e-commerce-related" will encompass businesses ranging from online selling services to derivative activities such as e-commerce training.

I start by presenting descriptive evidence. Figure 9 shows the trend of firms running e-commerce-related businesses and the annual growth rate of such firms. Since 2013, many related firms have been established, with annual growth rates maintaining more than 50% in the following several years. Recalling that consumers' online shopping behavior has experienced a surge since 2012, as shown in Figure 1, the trend suggests that the growth of online shopping behavior is indeed associated with a rise in firms running related businesses.

Figure 9 The trend of new firms in e-commerce-related businesses

Data sources: Firm registration database

E-commerce penetration can vary significantly across sectors (Lieber and Syverson 2012). Figure 10 illustrates the top ten industries with the highest e-commerce penetration. The upper panel shows the percentage of newly registered firms running e-commerce-related businesses over the total newly registered firms, and the lower panel reports the number of such firms. For each industry, the three bars stand for statistics in 2010, 2015, and 2019, respectively.

Most top ten industries with the highest e-commerce penetration are in the service and manufacturing sectors. In terms of the absolute number, "wholesale and retail trade," "information and communication," "leasing and business services," "manufacturing," and "professional, science, and technology activities" are the top five industries with the most firms running e-commerce-related businesses. Notably, the number of firms in the wholesale and retail trade sector significantly outstrips the number in the second rank, with 1,257,946 compared to 110,393 in 2019. Such a fact echoes the results in Figure 4, which shows that e-commerce significantly relocates more workforces to manufacturing and wholesale and retail trade.

Finally, by investigating the time dimension, we observe that the rankings are mostly robust across years. It may be due to the nature of the industry and path dependence on adopting new technologies.

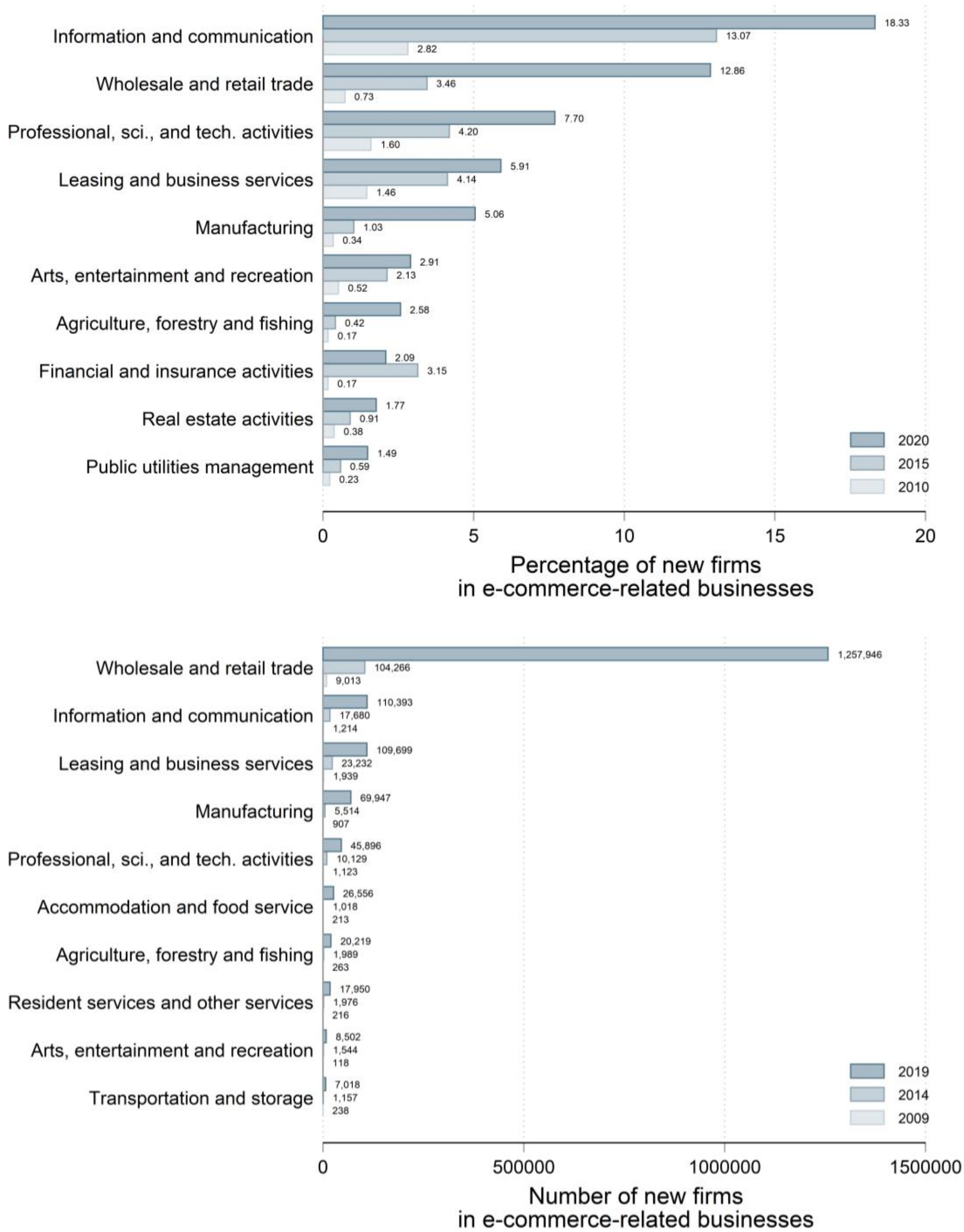


Figure 10 E-commerce penetration by year and industry

Notes: This figure illustrates the distribution of newly registered firms in e-commerce-related businesses by year and industry. For aesthetic reasons, several names of industries adopt abbreviations, with “sci” and “tech” standing for science and technology, respectively. The upper panel shows the top ten industries with the highest e-commerce penetration in terms of the percentage of new firms in e-commerce-related businesses over the total newly registered firms. The lower panel shows the top ten such industries in terms of the absolute number of new firms in e-commerce-related businesses. Over each industry, the three bars stand for statistics in 2010, 2015, and 2019, respectively.
Data sources: Firm registration database

Next, we perform the IV identification to estimate the causal effect of e-commerce on firm registration outcomes. The second-stage equation is specified as follows:

$$y_{ct} = \beta_0' + \beta_1' EC_{pt} + \gamma_c' + \delta_t' + \varepsilon_{ct}, \quad (3)$$

where y_{ct} is the outcome of interest by prefecture c and year t, including the number of new firms, the number of new firms in e-commerce-related businesses, and the percentage of new firms in e-commerce-related businesses. EC_{pt} is the measure of e-commerce development over province p. As elaborated in Section 2.2.1, we use two e-commerce measures: the percentage of the population adopting online shopping and the online retail transaction volumes per capita (10,000 CNY, 2009 price). Based on data availability, the period examined for the percentage measure is 2012, 2015, 2016, and 2018, and the period for the volume measure is 2010 – 2019. γ_c' and δ_t' are the prefecture and year fixed effects, respectively. The standard errors, ε_c, are clustered at the prefecture level.

Table 10 reports the results, with Panel A showing the OLS estimates and Panel B showing the IV estimates. Neither the percentage nor volume measure significantly affects the total number of newly registered firms. In contrast, the estimate of the volume measure for the number of new firms in e-commerce-related businesses is significant at the 5% level. If the online retail volumes per capita increase by 1000 CNY (2009 price), the number of new firms in e-commerce-related businesses will increase by 5.6 percentage points. The estimate of the percentage measure for the same outcome is positive but statistically insignificant.

As for the percentage of new firms in e-commerce-related businesses, estimates of both the percentage and volume measures are significantly positive. If the percentage of the population adopting online shopping increases by ten percentage points, the percentage of such new firms will increase by 5.4 percentage points. If the online retail volumes per capita increase by 1000 CNY (2009 price), the percentage will increase by 0.6 percentage points.

In summary, Table 10 establishes that, with the development of the online consumption market, more firms running e-commerce-related businesses will be established, and, presumably, more related jobs will be created.

Table 10 The impact of e-commerce on firm creation

	(1)	(2)	(3)	(4)	(5)	(6)
	Ln(Number of new firms)		Ln(Number of new firms conducting e-commerce -related businesses)		Percentage of new firms conducting e-commerce	
Panel A: OLS estimates						
Percentage of population adopting online shopping	0.025		-0.302		-0.003	
	(0.292)		(0.408)		(0.027)	
Online retail transaction volumes per capita		0.017		0.114		0.035***
		(0.080)		(0.142)		(0.007)
R-squared	0.98	0.98	0.96	0.95	0.68	0.48
Panel B: IV estimates						
Percentage of population adopting online shopping	-0.546		2.325		0.539***	
	(1.697)		(2.412)		(0.166)	
Online retail transaction volumes per capita		-0.272		0.561**		0.061***
		(0.202)		(0.257)		(0.014)
Number of clusters	327	357	327	357	327	357
Anderson-Rubin test	0.74	0.13	0.33	0.03	0.00	0.00
First stage	0.016 (0.004)	0.130 (0.023)	0.016 (0.004)	0.130 (0.023)	0.016 (0.004)	0.130 (0.023)
K-P F statistics	18	33	18	33	18	33
Observations	1302	3550	1302	3550	1302	3550

Notes: This table reports OLS and IV estimates of e-commerce's impact on firm registration. Observations are at the prefecture-by-year level. Two e-commerce measures are adopted: the percentage of the population adopting online shopping, serving as a measure of extensive margins, and the online retail transaction volumes per capita (10,000 CNY, 2009 price), serving as a measure of intensive margins. Both measurements vary by province and year. For IV estimates, the e-commerce measures are instrumented with a stringency measure of an Internet policy. The p-values from the weak-IV-robust Anderson-Rubin tests of the explanatory variable, first-stage estimates of the policy stringency measure, and K-P F statistics are reported. All regressions control prefecture and year fixed effects. Standard errors in parentheses are clustered at the prefecture level.

Significant levels: $***p < 0.01,\ **p < 0.05,\ *p < 0.1$

Data sources: Firm Registration Database

4.2. Firms' employment decisions

Section 4.1 provides evidence based on the annual flow statistics of firm registration. However, the stock of firms could be vital as they may offer more job vacancies. To provide related evidence, we draw on two firm surveys maintained by the National Bureau of Statistics: the ASIE from 2002 to 2012 and the 2008 Economic Census.

By merging the retail transaction volumes per capita with firm survey data based on located provinces, we can investigate the association between firms' employment and the development of the province-level e-commerce market. Furthermore, we merge the firm registration information with the survey data to identify whether a firm runs e-commerce-related businesses. Hence, we can directly examine the association between employment size and a firm's business choices. However, as the period covered by these two firm data sets is before the launch of the SUCI policy, we can only provide suggestive OLS evidence.

In column 1 of Table 11, conditioning on prefecture and year fixed effects, we first regress the logarithm form of employment size on the online retail transaction volumes per capita. Based on column 1, we further control the firm age and its squared term; estimates are shown in column 2. Next, columns 3 – 4 present the estimates, without and with firm age control variables, by replacing the province-level e-commerce measure with the firm-level dummy variables indicating a firm runs e-commerce-related businesses. Column 5 shows the results by including all the explanatory variables, and column 6 presents that by additionally including an industry (four-digit) fixed effect. Standard errors are clustered at the industry level.

As shown, all estimates of both the province-level and firm-level e-commerce measures are significantly positive at the 1% level. These results suggest that both province-level e-commerce market development and running e-commerce-related businesses on their own are associated with more employment in the manufacturing sector, echoing Figure 4 showing that more labor force relocated to the manufacturing sector under the shock of e-commerce.

Table 11 The association between e-commerce and manufacturing firms' employment

	(1)	(2)	(3)	(4)	(5)	(6)
	Ln(Number of employees)					
Online retail transaction volumes per capita	0.518***	0.324**			0.323**	0.272**
	(0.143)	(0.140)			(0.140)	(0.126)
Firms running e-commerce-related businesses			0.085***	0.131***	0.129***	0.115***
			(0.027)	(0.026)	(0.026)	(0.022)
Firm age		0.022***		0.022***	0.022***	0.023***
		(0.001)		(0.001)	(0.001)	(0.001)
Firm age squared		-0.000***		-0.000***	-0.000***	-0.000***
		(0.000)		(0.000)	(0.000)	(0.000)
R-squared	0.77	0.77	0.77	0.77	0.77	0.79
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effect						Yes
City fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2982543	2982543	2982543	2982543	2982543	2982543

Notes: This table reports OLS estimates of e-commerce's impact on manufacturing firms' employment decisions. Observations are at the firm-by-year level. Standard errors in parentheses are clustered at the industry level.

Significant levels: $*** p < 0.01,\ ** p < 0.05,\ * p < 0.1$
Data sources: The Annual Survey of Industrial Enterprises

However, the AISE does not provide information on the background of their workers. To overcome such a shortcoming, we draw on data from the 2008 Economic Census, which includes detailed information on employment size by workers' gender and educational background. Table 12 presents the results. For each outcome, we first show the estimate of the dummy variable indicating a firm runs e-commerce-related businesses, conditioning on industry (three-digit) and prefecture fixed effects. Standard errors are clustered at the industry level. Next, we additionally control the firm age and its squared term.[33]

Most estimates of the e-commerce-related-business dummy variable are significantly positive. These results imply that running e-commerce-related businesses is associated with a larger employment size, more female employees, and more employees with senior high education, junior college education, and college education. For outcomes regarding employment composition by education, the estimates for junior college education are the largest. While the evidence of the heterogeneous effects of e-commerce by educational level (Figure 7) shows that e-commerce benefits workers with junior high education or below, the firm-level data show no strong evidence of favoring workers with too little education. The reasons may be due to the different periods covered by the data. As some industry observers have noted, with the development of e-commerce and digital technology, more and more work is standardized and becomes routine jobs, allowing workers with less education to handle it (see a case study in Appendix II). Nonetheless, the data as early as 2008 have indicated a tendency to employ less-educated workers in e-commerce-related businesses.

[33] The variations of province-level measurements are too limited in cross-sectional data as there are only 31 provinces in mainland China. Hence, we do not include the province-level e-commerce measure in this analysis.

Table 12 The association between e-commerce and firms' employment decisions

	(1)	(2)	(3)	(4)	(5)	(6)
	Ln(Number of employees)		Ln(Number of female employees)		Ln(Number of employees with a junior high education or below)	
Firms in e-commerce-related businesses	0.283***	0.306***	0.295***	0.314***	0.038	0.057*
	(0.049)	(0.051)	(0.050)	(0.051)	(0.032)	(0.033)
Firm age		0.042***		0.036***		0.031***
		(0.005)		(0.004)		(0.005)
Firm age squared/100		-0.043***		-0.040***		-0.025***
		(0.006)		(0.005)		(0.004)
R-squared	0.25	0.28	0.16	0.18	0.30	0.31
	Ln(Number of employees with senior high education)		Ln(Number of employees with junior college education)		Ln(Number of employees with bachelor degrees or above)	
Firms in e-commerce-related businesses	0.174***	0.194***	0.295***	0.308***	0.260***	0.270***
	(0.048)	(0.050)	(0.044)	(0.045)	(0.036)	(0.036)
Firm age		0.036***		0.024***		0.019***
		(0.004)		(0.004)		(0.003)
Firm age squared/100		-0.034***		-0.026***		-0.023***
		(0.005)		(0.006)		(0.005)
R-squared	0.22	0.24	0.13	0.14	0.17	0.18
Observations	2596199	2596199	2596199	2596199	2596199	2596199

Notes: This table reports OLS estimates of e-commerce's impact on firms' employment decisions. Observations are at the firm-by-year level. All regressions control industry and prefecture fixed effects. Standard errors in parentheses are clustered at the industry level.

Significant levels: $*** p < 0.01,\ ** p < 0.05,\ * p < 0.1$

Data sources: The 2008 Economic Census

5. Conclusions

In this paper, we use large, nationally representative Chinese data to identify the causal impact of e-commerce on labor market outcomes. Utilizing a stringency measure based on media contents in CPC newspapers, we extract exogenous variations in the share and volumes of online shopping on the consumer side and establish that e-commerce increases employment rates, individual wages, and total household incomes on average. Moreover, people exposed to more developed e-commerce markets tend to quit the agricultural sector and get employed in enterprises of the manufacturing and service sectors. In addition, evidence suggests that e-commerce involves creative destruction, and people who remain employed tend to switch occupations and handle new tasks. Moreover, older workers may be the losers in the e-commerce era as they tend to lose their jobs and face more dramatic changes in job tasks. Interestingly, e-commerce is not a technological-biased change as it benefits less-educated workers. Firm-level evidence confirms that e-commerce facilitates more related firms and jobs to be created. These results further our understanding of e-commerce's socioeconomic impacts.

It is worth noting some limitations in the current research. For example, we can only measure occupation switches at the category level. Whether there are regularities regarding job-hopping within the same occupation category remains unanswered. In addition, due to the data availability, mechanism investigations at the firm level only draw on data during the earlier period, with job creation being measured indirectly. We may need novel data, such as job vacancies, to complete the picture of e-commerce's socioeconomic impacts on the labor market.

Reference

Acemoglu, Daron, and David Autor. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings," *Handbook of Labor Economics.* Elsevier, 1043-171.

Acemoglu, Daron, and Pascual Restrepo. 2019. "Automation and New Tasks: How Technology Displaces and Reinstates Labor." *Journal of Economic Perspectives*, 33(2), 3-30.

____. 2020. "Robots and Jobs: Evidence from Us Labor Markets." *Journal of political economy*, 128(6), 2188-244.

Agrawal, Ajay, John Horton, Nicola Lacetera, and Elizabeth Lyons. 2015. "Digitization and the Contract Labor Market: A Research Agenda." *Economic Analysis of the Digital Economy*, 219-50.

Akerman, Anders, Ingvil Gaarder, and Magne Mogstad. 2015. "The Skill Complementarity of Broadband Internet." *The Quarterly Journal of Economics*, 130(4), 1781-824.

Albanesi, Stefania, Antonio Dias da Silva, Juan F Jimeno, Ana Lamo, and Alena Wabitsch. 2023. "New Technologies and Jobs in Europe." *National Bureau of Economic Research Working Paper Series*.

Angrist, Joshua D, and Jörn-Steffen Pischke. 2008. "Instrumental Variables in Action: Sometimes You Get What You Need," *Mostly Harmless Econometrics.* Princeton university press,

Autor, David H. 2022. "The Labor Market Impacts of Technological Change: From Unbridled Enthusiasm to Qualified Optimism to Vast Uncertainty." *National Bureau of Economic Research Working Paper Series*.

____. 2001. "Wiring the Labor Market." *Journal of Economic Perspectives*, 15(1), 25-40.

Autor, David, Anna Salomons, and Bryan Seegmiller. 2021. "New Frontiers: The Origins and Content of New Work, 1940–2018." *MIT Working Paper*.

Baker, Scott R, Nicholas Bloom, and Steven J Davis. 2016. "Measuring Economic Policy Uncertainty." *The quarterly journal of economics*, 131(4), 1593-636.

Bakos, Yannis. 2001. "The Emerging Landscape for Retail E-Commerce." *Journal of economic perspectives*, 15(1), 69-80.

Barwick, Panle Jia, Shengmao Cao, and Shanjun Li. 2021. "Local Protectionism, Market Structure, and Social Welfare: China's Automobile Market." *American Economic Journal: Economic Policy*, 13(4), 112-51.

Bhuller, Manudeep, Domenico Ferraro, Andreas R. Kostøl, and Trond C. Vigtel. 2023. "The Internet, Search Frictions and Aggregate Unemployment." *National Bureau of Economic Research Working Paper Series*.

Bound, John, and David A Jaeger. 2000. "Do Compulsory School Attendance Laws Alone Explain the Association between Quarter of Birth and Earnings?," *Research in Labor Economics.* Emerald Group Publishing Limited, 83-108.

Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang. 2014. "Challenges of

Working with the Chinese Nbs Firm-Level Data." *China Economic Review*, 30, 339-52.
Braxton, J Carter, and Bledi Taska. 2023. "Technological Change and the Consequences of Job Loss." *American Economic Review*, 113(2), 279-316.
Brynjolfsson, Erik, Long Chen, and Xijie Gao. 2022. "Gains from Product Variety: Evidence from a Large Digital Platform." *National Bureau of Economic Research Working Paper Series*.
Brynjolfsson, Erik, Yu Hu, and Duncan Simester. 2011. "Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales." *Management Science*, 57(8), 1373-86.
Brynjolfsson, Erik, Yu Hu, and Michael D Smith. 2003. "Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers." *Management science*, 49(11), 1580-96.
Bustos, Paula, Bruno Caprettini, and Jacopo Ponticelli. 2016. "Agricultural Productivity and Structural Transformation: Evidence from Brazil." *American Economic Review*, 106(6), 1320-65.
Chava, Sudheer, Alexander Oettl, Manpreet Singh, and Linghang Zeng. 2022. "Creative Destruction? Impact of E-Commerce on the Retail Sector." *National Bureau of Economic Research Working Paper Series*.
Chiplunkar, Gaurav, and Pinelopi Koujianou Goldberg. 2022. "The Employment Effects of Mobile Internet in Developing Countries." *National Bureau of Economic Research Working Paper Series*.
Couture, Victor, Benjamin Faber, Yizhen Gu, and Lizhi Liu. 2021. "Connecting the Countryside Via E-Commerce: Evidence from China." *American Economic Review: Insights*, 3(1), 35-50.
Cravino, Javier, and Sebastian Sotelo. 2019. "Trade-Induced Structural Change and the Skill Premium." *American Economic Journal: Macroeconomics*, 11(3), 289-326.
Dillender, Marcus, and Eliza Forsythe. 2022. "Computerization of White Collar Jobs." *National Bureau of Economic Research Working Paper Series*.
Dolfen, Paul, Liran Einav, Peter J Klenow, Benjamin Klopack, Jonathan D Levin, Larry Levin, and Wayne Best. 2023. "Assessing the Gains from E-Commerce." *American Economic Journal: Macroeconomics*, 15(1), 342-70.
Duch-Brown, Néstor, Lukasz Grzybowski, André Romahn, and Frank Verboven. 2017. "The Impact of Online Sales on Consumers and Firms. Evidence from Consumer Electronics." *International Journal of Industrial Organization*, 52, 30-62.
Fan, Jingting, Lixin Tang, Weiming Zhu, and Ben Zou. 2018. "The Alibaba Effect: Spatial Consumption Inequality and the Welfare Gains from E-Commerce." *Journal of International Economics*, 114, 203-20.
Fang, Hanming, Ming Li, and Zenan Wu. 2022. "Tournament-Style Political Competition and Local Protectionism: Theory and Evidence from China." *National Bureau of Economic Research Working Paper Series*.
Forman, Chris, Avi Goldfarb, and Shane Greenstein. 2012. "The Internet and Local Wages: A Puzzle." *American Economic Review*, 102(1), 556-75.
Goldfarb, Avi, and Catherine Tucker. 2019. "Digital Economics." *Journal of*

Economic Literature, 57(1), 3-43.

Goldin, Claudia, and Lawrence F. Katz. 2007. "Long-Run Changes in the Wage Structure: Narrowing, Widening, Polarizing." *Brookings Papers on Economic Activity*, 38(2), 135–65.

Goldmanis, Maris, Ali Hortaçsu, Chad Syverson, and Önsel Emre. 2010. "E-Commerce and the Market Structure of Retail Industries." *The Economic Journal*, 120(545), 651-82.

Goos, Maarten, Alan Manning, and Anna Salomons. 2014. "Explaining Job Polarization: Routine-Biased Technological Change and Offshoring." *American Economic Review*, 104(8), 2509-26.

Herrendorf, Berthold, Richard Rogerson, and Akos Valentinyi. 2014. "Growth and Structural Transformation," *Handbook of Economic Growth.* Elsevier, 855-941.

Hjort, Jonas, and Jonas Poulsen. 2019. "The Arrival of Fast Internet and Employment in Africa." *American Economic Review*, 109(3), 1032-79.

Huang, Yi, Chen Lin, Sibo Liu, and Gustavo Manso. 2021. "The Salience of Entrepreneurship: Evidence from Online Business." *Working Paper*.

Huang, Yun, and Paul Luk. 2020. "Measuring Economic Policy Uncertainty in China." *China Economic Review*, 59, 101367.

Jin, Yizhou, and Zhengyun Sun. 2021. "Lifting Growth Barriers for New Firms: Evidence from an Entrepreneurship Training Experiment with Two Million Online Businesses." *Working Paper*.

Jo, Yoon J, Misaki Matsumura, and David E Weinstein. 2019. "The Impact of E-Commerce on Relative Prices and Consumer Welfare." *National Bureau of Economic Research Working Paper Series*.

Kolko, Jed. 2012. "Broadband and Local Growth." *Journal of Urban Economics*, 71(1), 100-13.

Kuhn, Peter, and Hani Mansour. 2014. "Is Internet Job Search Still Ineffective?" *The Economic Journal*, 124(581), 1213-33.

Lee, Donghoon, and Kenneth I Wolpin. 2006. "Intersectoral Labor Mobility and the Growth of the Service Sector." *Econometrica*, 74(1), 1-46.

Lendle, Andreas, Marcelo Olarreaga, Simon Schropp, and Pierre-Louis Vézina. 2016. "There Goes Gravity: Ebay and the Death of Distance." *The Economic Journal*, 126(591), 406-41.

Levine, Ross, and Yona Rubinstein. 2018. "Selection into Entrepreneurship and Self-Employment." *National Bureau of Economic Research Working Paper Series*.

Li, Hongbin, and Li-An Zhou. 2005. "Political Turnover and Economic Performance: The Incentive Role of Personnel Control in China." *Journal of public economics*, 89(9-10), 1743-62.

Lieber, Ethan, and Chad Syverson. 2012. "Online Versus Offline Competition." *The Oxford handbook of the digital economy*, 189.

Lowrey, Ying. 2016. *The Alibaba Way: Unleashing Grass-Roots Entrepreneurship to Build the World's Most Innovative Internet Company*. McGraw Hill Professional.
Michaels, Guy, Ashwini Natraj, and John Van Reenen. 2014. "Has Ict Polarized Skill Demand? Evidence from Eleven Countries over Twenty-Five Years." *Review of Economics and Statistics*, 96(1), 60-77.
Ministry of Commerce. 2021. "E-Commerce in China 2020."
Porzio, Tommaso, Federico Rossi, and Gabriella Santangelo. 2022. "The Human Side of Structural Transformation." *American Economic Review*, 112(8), 2774-814.

Pozzi, Andrea. 2013. "The Effect of Internet Distribution on Brick-and-Mortar Sales."

The RAND Journal of Economics, 44(3), 569-83.
Qin, Bei, David Strömberg, and Yanhui Wu. 2018. "Media Bias in China." *American Economic Review*, 108(9), 2442-76.
Randall, Taylor, Serguei Netessine, and Nils Rudi. 2006. "An Empirical Examination of the Decision to Invest in Fulfillment Capabilities: A Study of Internet Retailers." *Management Science*, 52(4), 567-80.
Relihan, Lindsay. 2022. "Is Online Retail Killing Coffee Shops? Estimating the Winners and Losers of Online Retail Using Customer Transaction Microdata." *London School of Economics CEP Discussion Paper*, (1836).
Sinai, Todd, and Joel Waldfogel. 2004. "Geography and the Internet: Is the Internet a Substitute or a Complement for Cities?" *Journal of Urban Economics*, 56(1), 1-24.
Song, Yang. 2014. "What Should Economists Know About the Current Chinese Hukou System?" *China Economic Review*, 29, 200-12.
Stockmann, Daniela. 2013. *Media Commercialization and Authoritarian Rule in China*. Cambridge University Press.
Xia, Jun. 2017. "China's Telecommunications Evolution, Institutions, and Policy Issues on the Eve of 5g: A Two-Decade Retrospect and Prospect." *Telecommunications Policy*, 41(10), 931-47.
Xu, Chenggang. 2011. "The Fundamental Institutions of China's Reforms and Development." *Journal of economic literature*, 49(4), 1076-151.
Zhang, Jipeng, Jin Huang, Junhui Wang, and Liang Guo. 2020. "Return Migration and Hukou Registration Constraints in Chinese Cities." *China Economic Review*, 63, 101498.
Zhao, Yuezhi. 1998. *Media, Market, and Democracy in China: Between the Party Line and the Bottom Line*. University of Illinois Press.
Zhou, Li-an. 2007. "Research on the Promotion Tournament of Local Cadres in China." *Economic Research (in Chinese)*, 15, 36-50.
Zuo, George W. 2021. "Wired and Hired: Employment Effects of Subsidized Broadband Internet for Low-Income Americans." *American Economic Journal: Economic Policy*, 13(3), 447-82.

Appendix I. Supplementary Description of the Data

China Household Finance Survey (CHFS)

The China Household Finance Survey (CHFS) is a biennially longitudinal survey executed by the Survey and Research Center for China Household Finance of the Southwestern University of Finance and Economics. The data are collected via face-to-face interviews using the technology of computer-assisted personal interviewing (CAPI). The survey topics cover various facets of households and individuals, such as demographics and employment, asset and liability, income and consumption, social insurance, and subjective attitudes.

The baseline survey of the CHFS was conducted in 2011. By 2019, the CHFS had completed five waves of surveys, covering over 70,000 households from 29 provinces (Xinjiang and Tibet are excluded). The baseline survey collects a nationally representative sample. Since 2013, the CHFS has extended its sample size to ensure that the data has representativeness at the nation and province levels. For surveys in 2015 and 2017, the samples also have representativeness for several big cities, including Shenyang, Dalian, Changchun, Harbin, Nanjing, Hangzhou, Jinan, Qingdao, Ningbo, Xiamen, Wuhan, Guangzhou, Shenzhen, Chengdu, and Xi’an.

China Family Panel Studies (CFPS)

China Family Panel Studies (CFPS) is a biennially longitudinal survey executed by the Institute of Social Science Survey of Peking University. Typically, the data are collected via face-to-face interviews using the CAPI technology. The survey team will try phone interviews for those who cannot be reached in person. The CFPS gleans data at the community, family, and individual levels, including information such as economic activities, education, family structure, migration, and health.

The baseline survey of the CFPS was conducted in 2010. By 2020, the CFPS had completed five waves of surveys. Its nationally representative sample has covered over 20,000 households from 31 provinces. [34] In addition, the CFPS sample has representativeness for several big provinces, including Shanghai, Liaoning, Henan, Gansu, and Guangdong.

The occupational information in the CFPS should be of high quality. To avoid measurement errors, interviewers of the CFPS only recorded the contents of the respondents' work. Typically, an interviewer will ask, "*What do you do for a living? Where do you work?*" After the survey, the CFPS will appoint skilled staff to code the occupational categories and industries based on the text. At least two experienced workers will code the text independently into occupational and industrial categories. If the results are the same, the result will be adopted; otherwise, a professional will code the text independently again. If the third result can match any of the two previous results, the third result will be adopted; otherwise, a senior professional will examine the three outcomes, deciding which one should be adopted or assigning new codes. As of writing this paper, the detailed categories of occupations are only available in CFPS 2010 – 2018.

[34] Initially, the survey design restricts the sampling regions to 25 provinces. As the sample households migrate and the CFPS follows them up, the observations have covered all provinces in Mainland China.

The time series of the online retail transaction volumes

The time series of the online retail transaction volumes by province is constructed from two sources. *The China Trade and External Economic Statistical Yearbook*, edited by the Department of Statistics on Trade and External Economic Relations, the National Bureau of Statistics of China, published the statistic by province from 2015 to 2020. Before 2015, a national-level statistic was published annually in the report *E-commerce in China*, edited by the Ministry of Commerce, which can be traced back to 2002. Assuming the share of the online retail transaction volumes among the provinces remained constant as in 2015, we extended the province-by-year series to 2002.

WiseNews Database

The WiseNews Database is maintained by Wiser, a private data company established in 1998. As one of the largest Chinese media databases in the world, the WiseNews Database covers over 1200 print newspapers from the Greater China Region (Mainland China, Hong Kong, Macao, and Taiwan). In addition, the database contains content from blogs, websites, and social media. It also gleans media data from countries in the rest of the world, such as the United States and Europe.

This paper uses the contents published in the province and prefecture CPC newspapers. The WiseNews Database includes all province CPC newspapers. However, its coverage of prefecture CPC papers concentrates on those circulating in relatively large cities. Table A1 reports statistics on WiseNews' coverage at the prefecture level. The first two columns of Table A1 present the percentage of prefectures without or with their CPC papers recorded in the WiseNews Database. Columns 3 – 4 show the population percentage in the two types of prefectures. We calculate the statistics by region (i.e., eastern, middle, and western areas) in the first three rows and report the number over the whole sample in the last row.

The WiseNews covers a higher proportion of prefectures in the eastern area (66%) than those in the middle (50%) or western (38%) regions. Half of the prefectures have their CPC papers recorded in the WiseNews Database. Looking at the population in these prefectures, we can observe a significant increase in the coverage rate compared to the rate in terms of the prefecture, with 73% for the eastern, 59% for the middle, and 55% for the western areas. Such an increase indicates that the WiseNews targets more newspapers circulating in relatively large cities. The total coverage rate in terms of the population is 64%. These results suggest that the WiseNews Database maintains a representative sample of prefecture CPC newspapers in relatively large cities and has broader coverage in the eastern area.

These features should be enough to construct an accurate measure at the province level in the context of this paper. Articles published in smaller cities will be assigned a smaller weight when generating the provincial stringency measure (see the details in Section 2.2.3). Therefore, the omission of newspapers circulating in small cities should not change the measurement too much.

Table A1 Coverage of prefecture CPC newspapers in the WiseNews Database

	(1)	(2)	(3)	(4)
	Not cover	Cover	Not cover	Cover
	Prefecture		Population	
Eastern area	34.0%	66.0%	27.3%	72.7%
Middle area	50.0%	50.0%	41.2%	58.8%
Western area	62.3%	37.7%	44.7%	55.3%
Total	50.2%	49.8%	36.5%	63.5%

Notes: This table shows the coverage of the WiseNews Database in terms of the number of prefectures and population. Columns 1 – 2 show the percentage of prefectures without or with their local CPC papers recorded in the WiseNews Database. Columns 3 – 4 present the percentage of the population in the two types of prefectures. The first three rows show statistics by region, and the last row reports the situation in the whole sample. The demarcation of eastern, middle, and western areas is defined by the National Bureau of Statistics.

Data sources: The WiseNews Database, 2010 population census

The Annual Survey of Industrial Enterprises

The Annual Survey of Industrial Enterprises (ASIE) is an economic census for "big industrial firms" conducted by the National Bureau of Statistics (NBS). It is also dubbed various names in academic works, such as "the China Industrial Census" and "the Annual Survey of Manufacturing Enterprises."

Every year, the NBS requires firms in industrial sectors, including mining, manufacturing, and public utilities, to compile statistical documents and submit them to the local bureau. The survey unit of the AISE is a "legal person." According to the NBS, a legal person unit is an organization that possesses assets, assumes liabilities, and performs economic activities independently, which resembles the definition of "establishments" in the United States.

The target firms of the AISE change several times. From 1998 to 2006, the AISE covers all state-owned industrial firms and non-stated-owned firms with principal business revenue of five million CNY or above. From 2007 to 2010, the ASIE stopped gleaning information on "small" state-owned industrial firms, only covering those whose principal business revenue is above five million CNY. Starting in 2011, the definition of "big firms" changed to those whose principal business revenue exceeds 20 million CNY. Although the ASIE does not include smaller firms, it accounts for the majority of total manufacturing output and employment.

The AISE records firms' basic information on location, industry, establishment year, and ownership type. It also provides statistics on firms' economic activities, including inputs, outputs, revenue, assets, and employment.

We restrict the period of examination from 2002 to 2012. While the earliest year we can access is 1998, we exclude data series before 2002, as business-to-consumer and consumer-to-consumer e-commerce activities have only been significantly observed in China since 2003.[35] In addition, we restrict observations to the manufacturing firms as a correspondence table of the industrial category, provided by Brandt et al. (2014), is only available for manufacturing firms in and before 2012.[36]

[35] Meemi O. 2022. "A Timeline of China's E-Commerce Sector and the Evolution of Alibaba, JD, and Pinduoduo." https://perma.cc/G8HP-ZZMA.

[36] The data after 2013 are not available for academic use.

Appendix II. Case Studies

Digital economy, programming, and education

E-commerce lowers the entry cost for doing business, and its long-tail effect allows various small firms to survive and cater to niche markets. As a result, many online sellers need to develop customized applets, applications, or websites to promote and sell their products.

As the technologies and programming procedures become mature and standardized, software development (e.g., applets) has been dramatically simplified, opening up opportunities for less-educated workers to handle these tasks. Numerous teams with no more than ten programmers are formed in megacities such as Guangzhou and Shenzhen. Most members of these teams only have a junior college or vocational education.

One practitioner likened the applet development to the assembly line. Most difficult programming tasks have been “automated” as they are compiled into various modules or packages. However, the design and development of customized applets to fulfill special needs cannot be automated. Therefore, the primary work of programmers is to “assemble” the modules or packages to build applets based on customers’ demands. While the tasks may not necessarily be hard, the working environment of these programmers is usually better than that of factory workers. Such assembly-line programming jobs are attractive to those young migrant workers with limited education.

Source: Junittt, “Everyone can be ‘wild’ programmers (in Chinese),” *Neweekly*, https://perma.cc/3S8G-25LW (accessed on 28 June 2023).

E-commerce, offline store shutdown, and workforce relocation

The coming of e-commerce renders many offline sellers shut down, especially those intermediary businesses of low-value-added products.

This case study focuses on a wholesale market of bags in *Wuhu*, a prefecture in Anhui Province, China. The golden age of this wholesale market was before 2009, when Taobao, a large e-commerce platform in China, entered people’s daily lives. After that, people found they could obtain similar products at lower prices for a product sold in the offline market. Gradually, Taobao and other e-commerce platforms undermined offline wholesale businesses.

After 2013, when the double-eleven event hit a historical transaction volume, stores run by those older owners who had difficulties using the online tools had to shut down. However, sellers who adopted e-commerce selling channels still maintained running.

At the end of the story, the daughter of a wholesale store owner chose to work in an Internet company. The case may be typical: the next generation of traditional wholesale store owners tends to become employees in companies rather than take over family businesses.

Source: Aoli, "The dead of offline wholesale stores with the coming of *Taobao* (in Chinese)," *Work in China*, https://perma.cc/DHS6-J4S4 (accessed on 28 June 2023).

Appendix III. Supplementary Figures

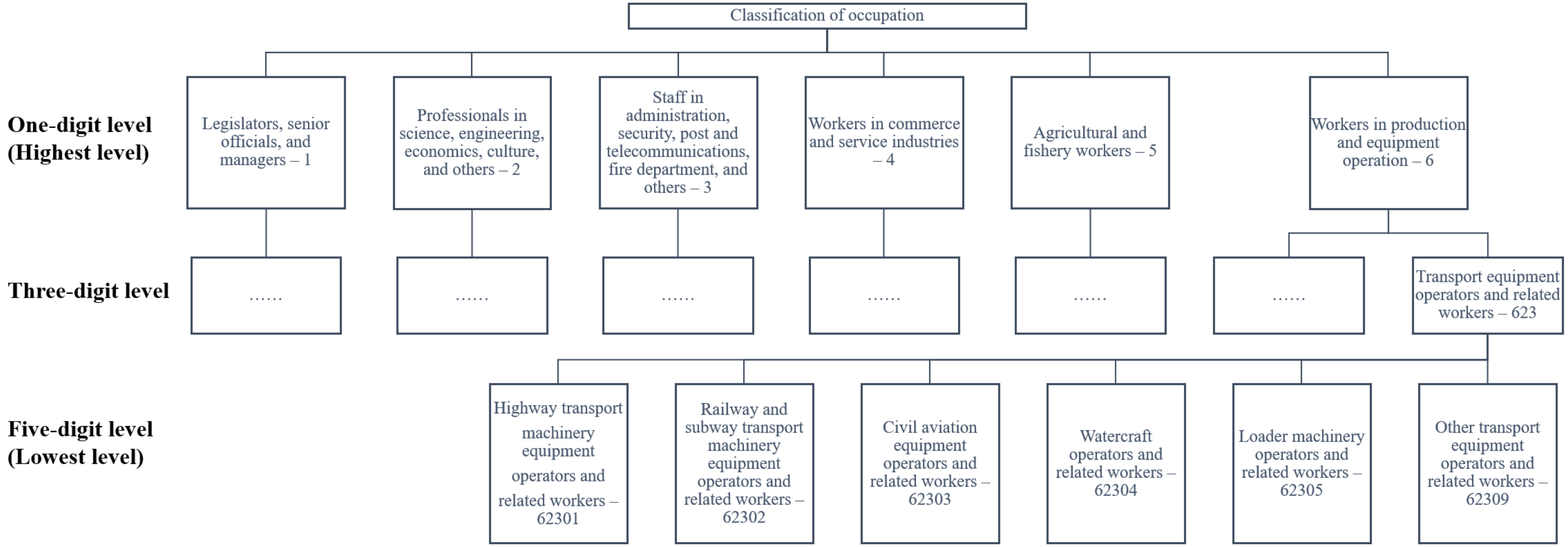


Figure A1 Illustration of Chinese Standard Classification of Occupations

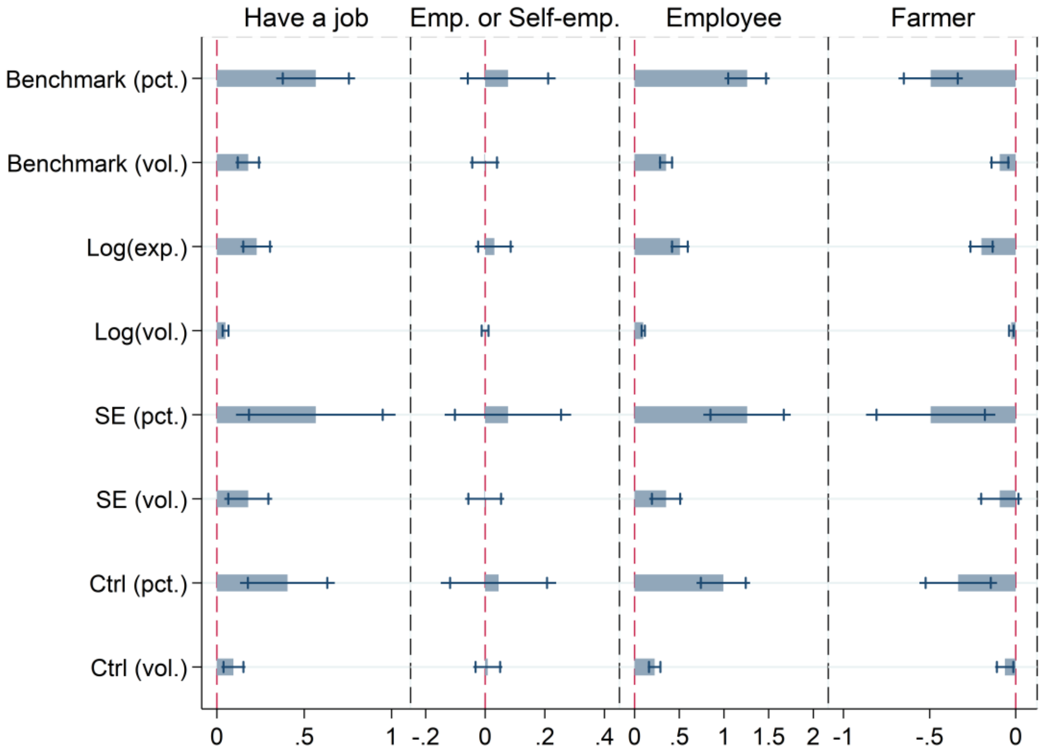


Figure A2 Robustness checks for employment status and positions

Notes: This figure reports the IV estimates for outcomes regarding employment status and positions, including dummy variables indicating having a job, being an employer or self-employed worker, being an employee, and being a farmer. The e-commerce measures are instrumented with a stringency measure of an Internet policy. Each column shows estimates for one dependent variable. Each bar in the graph represents the estimate, with a line representing the 95% confidence interval and capped spikes representing the 90% one. The left-handed axis indicates the types of robustness checks. For comparison, the figure presents the benchmark estimates of the percentage and volume measures at the top two rows, denoted as "Benchmark (pct.)" and "Benchmark (vol.)." "Log(vol.)" and "Log(exp.)" denote the checks by replacing the benchmark explanatory variables with the logarithm form of online retail transaction volumes and the logarithm form of monthly online expenditure per household (estimated from the CHFS data), respectively. "SE(pct.)" and "SE(vol.)" denote the checks by clustering the standard errors at the individual and community-by-year levels for the percentage and volume measures of e-commerce, respectively. "Ctrl (pct.)" and "Ctrl (vol.)" denote the checks by adding province-level and individual-level control variables to capture potential confounders regarding Internet development and usage for the percentage and volume measures of e-commerce, respectively.

Data sources: CHFS 2011 – 2019

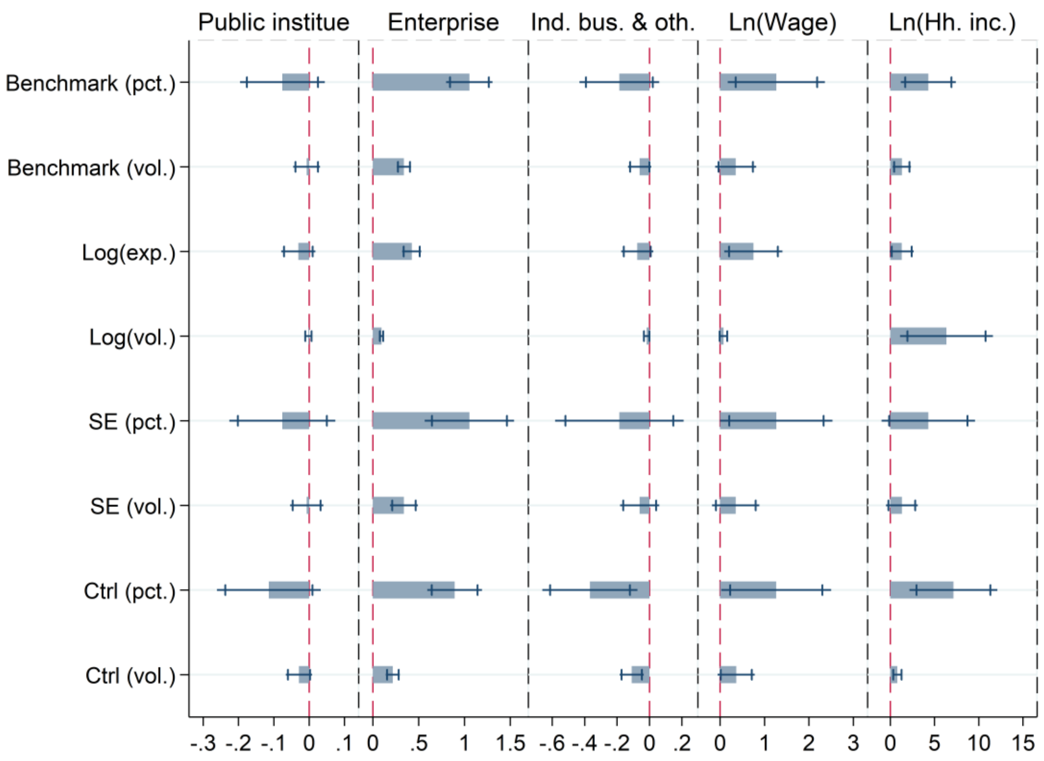


Figure A3 Robustness checks for employment providers and incomes

Notes: This figure reports the IV estimates for outcomes regarding employment providers and incomes, including dummy variables indicating working for public institutes, enterprises, and individual businesses or others, and the logarithm form of individual wages and household incomes. The e-commerce measures are instrumented with a stringency measure of an Internet policy. Each column shows estimates for one dependent variable. Each bar in the graph represents the estimate, with a line representing the 95% confidence interval and capped spikes representing the 90% one. The left-handed axis indicates the types of robustness checks. For comparison, the figure presents the benchmark estimates of the percentage and volume measures at the top two rows, denoted as “Benchmark (pct.)” and “Benchmark (vol.).” “Log(vol.)” and “Log(exp.)” denote the checks by replacing the benchmark explanatory variables with the logarithm form of online retail transaction volumes and the logarithm form of monthly online expenditure per household (estimated from the CHFS data), respectively. “SE (pct.)” and “SE (vol.)” denote the checks by clustering the standard errors at the individual and community-by-year levels for the percentage and volume measures of e-commerce, respectively. “Ctrl (pct.)” and “Ctrl (vol.)” denote the checks by adding province-level and individual-level control variables to capture potential confounders regarding Internet development and usage for the percentage and volume measures of e-commerce, respectively.
Data sources: CHFS 2011 – 2019

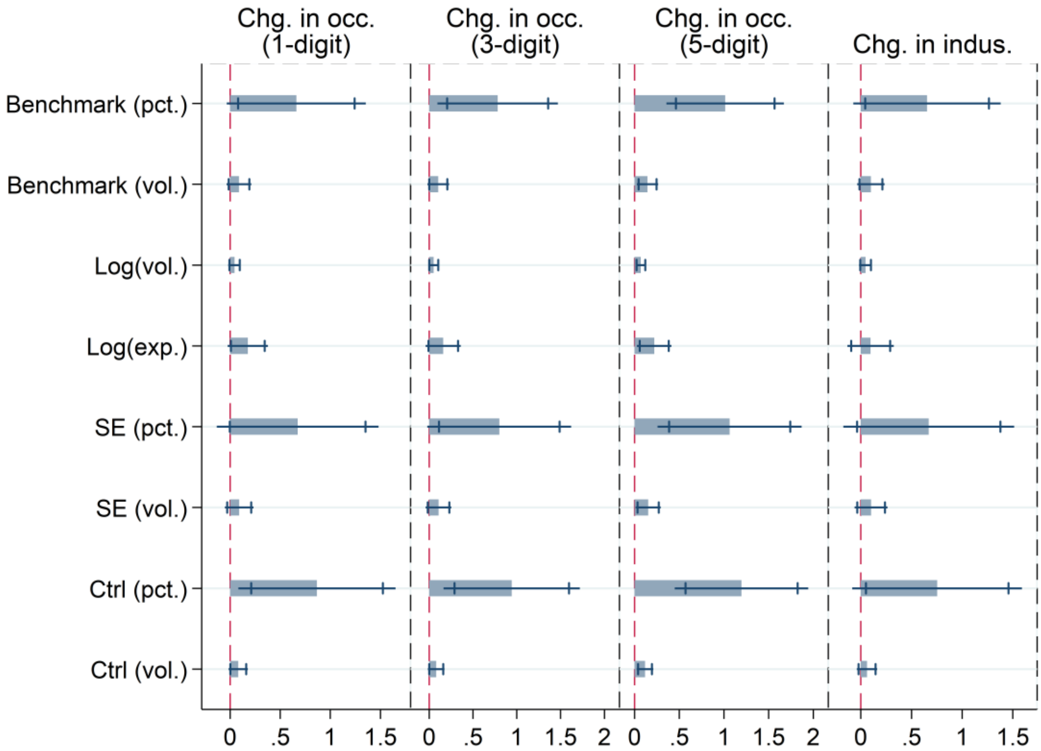


Figure A4 Robustness checks for job-switching

Notes: This figure reports the IV estimates for outcomes regarding job-switching, including dummy variables indicating a change in occupations at the one-, three-, and five-digit levels and a change in working industries. The e-commerce measures are instrumented with a stringency measure of an Internet policy. Each column shows estimates for one dependent variable. Each bar in the graph represents the estimate, with a line representing the 95% confidence interval and capped spikes representing the 90% one. The left-handed axis indicates the types of robustness checks. For comparison, the figure presents the benchmark estimates of the percentage and volume measures at the top two rows, denoted as "Benchmark (pct.)" and "Benchmark (vol.)." "Log(vol.)" and "Log(exp.)" denote the checks by replacing the benchmark explanatory variables with the logarithm form of online retail transaction volumes and the logarithm form of monthly online expenditure per household (estimated from the CHFS data), respectively. "SE (pct.)" and "SE (vol.)" denote the checks by clustering the standard errors at the individual and community-by-year levels for the percentage and volume measures of e-commerce, respectively. "Ctrl (pct.)" and "Ctrl (vol.)" denote the checks by adding province-level and individual-level control variables to capture potential confounders regarding Internet development and usage for the percentage and volume measures of e-commerce, respectively.

Data sources: CFPS 2010 – 2018

Appendix IV. Supplementary Table

Table A2 List of e-commerce-related keywords

Chinese Characters	Chinese phonetic alphabet	Meaning
电子商务	*dianzi shangwu*	e-commerce
电商	*dianshang*	e-commerce (abbreviation form)
网络购物	*wangluo gouwu*	Synonyms or abbreviations of online shopping
网上购物	*wangshang gouwu*	
网购	*wanggou*	
互联网购物	*hulianwang gouwu*	
电子平台购物	*dianzipingtai gouwu*	
网上销售	*wangshang xiaoshou*	Synonyms or abbreviations of online selling
网络销售	*wangluo xiaoshou*	
互联网销售	*hulianwang xiaoshou*	
电子平台销售	*dianzipingtai xiaoshou*	

Table A3 The impact of the SUCI policy on Internet costs and usage

	(1)	(2)	(3)	(4)	(5)
	Ln(Communication costs)		Ln(Internet users)		
Ln(Accumulative number of articles in CPC newspapers)	-0.093***	-0.015	0.177***	0.130**	0.130**
	(0.029)	(0.029)	(0.063)	(0.052)	(0.059)
Household workforce	0.345***	0.356***			
	(0.008)	(0.007)			
Rural areas	-0.801***	-0.633***			
	(0.018)	(0.019)			
Ln(Population)				0.079	0.079
				(0.553)	(0.932)
Ln(GDP in 2009 price)				0.886***	0.886**
				(0.213)	(0.331)
p-value (wild bootstrap)					0.07
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes		Yes	Yes	Yes
City fixed effects		Yes			
R-squared	0.22	0.26	0.99	0.99	0.99
Observations	148400	148400	217	217	217

Notes: This table shows the impact of the stringency of the"Speed up and Cheapen the Internet" policy measured by news articles on household communications costs and the number of Internet users. Observations in columns 1 – 2 are at the household-by-year level, whereas those in columns 3 – 5 are at the province-by-year level. For household-level regressions, standard errors are clustered at the household level, and sampling weights are used. For province-level regressions, robust standard errors are used in columns 3 – 4, and clustered standard errors at the province level are used in column 5.

Significant levels: $*** p < 0.01$, $** p < 0.05$, $* p < 0.1$

Data sources: CHFS 2011-2019, the National Bureau of Statistics 2014 – 2020

Table A4 The predictive power of socioeconomic factors on the policy stringency

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Ln(Accumulative number of articles in CPC newspapers)									
Ln(Population)	1.682*									2.120
	(0.864)									(1.390)
Percentage of the population aged 0-14		2.835								2.346
		(1.957)								(4.185)
Percentage of the population aged 15-64			0.687							1.187
			(1.790)							(2.907)
Ln(GDP)				0.204						-0.552
				(0.290)						(1.486)
Ln(GDP in the primary sector)					0.074					0.087
					(0.180)					(0.298)
Ln(GDP in the secondary sector)						0.084				-0.227
						(0.144)				(0.725)
Ln(GDP in the tertiary sector)							0.311			0.216
							(0.348)			(0.793)
Ln(Consumption per capita)								0.259		-0.223
								(0.464)		(0.883)
Ln(Disposable income per capita)									0.715	1.569
									(0.754)	(1.593)
p-value (wild bootstrap)	0.08	0.17	0.72	0.49	0.72	0.57	0.41	0.59	0.35	
R-squared	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
Observations	372	372	372	372	372	372	372	372	372	372

Notes: This table shows the predictive power of province socioeconomic factors to the stringency of the "Speed up and Cheapen the Internet" policy measured by news articles. Observations are at the province-by-year level. Monetary measurements are in the 2009 price. All regressions control province and year fixed effects. Standard errors in the parentheses are clustered at the province level. The province-clustered p-values adjusted by the wild bootstrap are reported.

Significant levels: $*** p < 0.01$, $** p < 0.05$, $* p < 0.1$

Data sources: The National Bureau of Statistics 2009 – 2020

Table A5 The impact of e-commerce on employers and self-employed workers

	(1)	(2)	(3)	(4)
	Employer		Self-employed worker	
Panel A: OLS estimates				
Percentage of population adopting online shopping	0.083***		-0.022	
	(0.029)		(0.040)	
Online retail transaction volumes per capita		0.007		0.012*
		(0.006)		(0.007)
R-squared	0.59	0.59	0.64	0.64
Panel B: IV estimates				
Percentage of population adopting online shopping	-0.093		-0.128	
	(0.107)		(0.139)	
Online retail transaction volumes per capita		-0.016		-0.018
		(0.017)		(0.022)
Anderson-Rubin test	0.38	0.36	0.36	0.41
First stage	0.074	0.467	0.074	0.467
	(0.004)	(0.023)	(0.004)	(0.023)
K-P *F* statistics	419	425	419	425
Observations	39228	39302	39228	39302

Notes: This table reports OLS and IV estimates of e-commerce's impact on the probability of being an employer or self-employed worker. Observations are at the individual-by-year level. The sample is restricted to respondents who are 18-55 years of age and not full-time students. Two e-commerce measures are adopted: the percentage of the population adopting online shopping, serving as a measure of extensive margins, and the online retail transaction volumes per capita (10,000 CNY, 2009 price), serving as a measure of intensive margins. Both measurements vary by province and year. For IV estimates, the e-commerce measures are instrumented with a stringency measure of an Internet policy. The p-values from the weak-IV-robust Anderson-Rubin tests of the explanatory variable, first-stage estimates of the policy stringency measure, and K-P F statistics are reported. All the regressions control individual and year fixed effects. Standard errors in parentheses are clustered at the individual level.

Significant levels: $*** p < 0.01$, $** p < 0.05$, $* p < 0.1$

Data sources: The CHFS 2017 and 2019