



Research on the Internet Access and Usage Gap in Rural Areas and Its Influencing Factors —Empirical Analysis Based on Heckman’s Two-Stage Model

Mingbo Ji¹, Rui Zhou², and Yaqin Fang³(✉)

¹ School of Economics, Sichuan University, Chengdu 610065, People’s Republic of China

² School of Marxism, Sichuan University, Chengdu 610065, People’s Republic of China

³ School of Public Administration, Sichuan University, Chengdu 610065, People’s Republic of China

294854375@qq.com

Abstract. With the implementation of the digital village development strategy, it is of great significance to explore the Internet access and usage divide and its influencing factors in rural areas of China to promote the rural revitalization strategy with high quality. Based on the data from the 2018 China Household Tracking Survey, this paper analyzes the Internet access and usage divide and its influencing factors in rural areas using the Heckman two-stage model. The empirical results show that people with higher education and better economic conditions are more likely to access the Internet, and men are more likely to access it through mobile Internet. Young people with higher levels of education and digital skills are more likely to use the Internet. Patterns of Internet use vary by gender, age, education level, occupation and geographic location. Young people are more likely to engage in e-commerce online activities, while older people are less likely to participate in online activities. The findings of this paper provide empirical evidence to analyze the digital divide in rural China and provide a basis for the government to formulate targeted telecommunications policies.

Keywords: rural · digital divide · digital access divide · digital usage divide · Heckman two-stage modeling

1 Introduction

Accessing and using the Internet can help residents gain digital advantages, such as obtaining valuable information, finding a job, consulting on health issues, or accessing public services. However, access to and use of the Internet presupposes access to technology and infrastructure, as well as the acquisition of skills to cope with innovations in the digital world [1], this difference in technology, facilities, and skills creates a digital divide, which primarily describes the differences between different countries in terms of technology, facilities, and skills. The digital divide primarily describes differences in Internet access and use by gender, income, and urban and rural residents [2]. Although China’s Internet access facilities have achieved widespread coverage [3], but China’s

rural areas still face a serious digital divide problem. Therefore, this paper utilizes CFPS 2018 data to conduct a detailed analysis of the Internet diffusion in rural areas of China. The reason for choosing this database is that the CFPS database focuses on the economic and non-economic well-being of Chinese residents, as well as a number of research topics including economic activity, educational outcomes, health, and internet use, etc. It is a nationwide, large-scale, multidisciplinary social tracking survey project, with a sample covering 25 provinces/municipalities/autonomous regions. Data such as Internet use in the database allow us to determine the profile of Internet users and define various patterns of Internet use.

This paper contributes to the study of the digital divide in rural areas of developing countries in three ways. First, there are more existing studies on the urban digital divide, but relatively few on the rural digital divide. Secondly, by studying the rural digital divide through microeconomic data, this study can further clarify the status quo and reasons for the relative lagging behind of rural development under China's urban-rural dualistic system, especially in the presentation of "digital space". Third, this study provides solid empirical evidence for the Chinese government to formulate public policies, such as rural revitalization and common prosperity, aimed at increasing the benefits of adequate and effective handling of information technology (IT), including the Internet, and thus making it easier for a growing number of people living in rural areas of developing countries to access and use IT.

The structure of this article is as follows: the first part describes the significance of the study and leads to the research question. The literature review in Part II identifies the new characteristics of rural China in the new era, the digital divide, and the determinants of Internet access and use. The third part presents the CFPS survey data and the variables used in the econometric model. The fourth section includes regression results on Internet access, usage, and usage patterns in rural China. Finally, the conclusions, policy recommendations, and research deficiencies of this article are given.

2 Theoretical Analysis and Literature Review

2.1 New Characteristics of China's Rural Areas in the New Era

Since the reform and opening up for more than 40 years, China's rural society has undergone great changes, presenting the following main features: First, with the increasing urbanization rate, the number of villages in China shows an overall downward trend, and the rural space "gives way" to the urban space [4]; Secondly, rural population mobility has increased and heterogeneity has strengthened, especially the outflow of young and strong people, and the rural areas are mainly left with the elderly people, children, and women [5]; Thirdly, as population mobility reshapes the rural mechanism, there are also significant differences within rural China [6]. The new features presented in China's rural areas in the new era have led to non-agricultural activities in the countryside being strengthened, with more and more linkages between rural areas and towns, highlighting the multifunctional role of rural space [7].

In the context of the rural revitalization strategy, the Chinese government has put forward a digital rural development strategy, which attempts to modernize agriculture, rural areas and farmers through rural digital empowerment, and to crack the problem

of unbalanced and insufficient development of agriculture, rural areas and farmers, in order to narrow the development gap between urban and rural areas, and to find the endogenous power for rural development as well as to provide a new momentum for the modernization of the rural society [8]. Information and communication technologies are naturally incorporated into the countryside as a key element of changes in the life of the country [9], which in turn makes rural and urban areas more connected.

Influenced by factors such as relative geographic isolation and low population density, digital access in rural areas falls behind the urban areas [8]. Although the Chinese government has dramatically improved the coverage of Internet access in rural areas through telecommunication policies such as “Village to Village”, the lack of access devices [10], educational limitations [11] and other factors, resulting in the fact that there is still a large portion of rural residents without access to ICT products and services. Although mobile devices such as cell phones have improved rural residents’ Internet access, most villagers have not yet been able to enjoy the digital dividends brought by the Internet due to the limitations of cell phone products and the generally low digital skills of the residents [12].

2.2 Digital Divide

The term “digital divide” originated in the United States in the 1990s and quickly became a topic of concern with the widespread use and penetration of ICT [13]. Existing research on the digital divide consists mainly of the Level 1 digital divide, where studies have focused on differences in Internet use and access [14]. The second-level digital divide, began to study the digital skills needed to use the Internet and differences in carrying out various types of activities online [15]. The third-level digital divide, which focuses on the types of inequalities that result from Internet use outcomes [16], suggesting that differences created through Internet use may exacerbate the social stratification in reality [17]. In addition, the current studies provide further details about differences in Internet use. For example, Internet use provides more opportunities and resources to improve their education, employment, professional life and social status [18]. Connected usage provides more insight into differences in devices, uses, skills and purposes of Internet use [19].

2.3 Literature Review: From Internet Access to Internet Use

Studies on Internet access and use have focused on both developed and developing countries. As far as China is concerned, some scholars point out that the digital divide exists, mainly manifested as the urban-rural divide [20] and the intergenerational divide [21], the essence of which is that the Matthew effect in the process of synergistic interaction between information and the social economy makes the digital society “replicate” the inequality of the real society, and even “create” new inequality. Research has shown that in the elderly people [22], low-income [23] and other groups in the digital divide, the choice of online activities is determined by the specific characteristics of each group, and the government needs to help these groups to Internet access and use through public decision-making.

The study found that the main factors affecting internet penetration between countries include income, education level, infrastructure [24]. Some scholars have found that income, telephone costs, and years of education are major determinants of online access [25]. Existing studies generally agree that the availability of Internet infrastructure in rural areas is in its infancy, which is largely determined by low population density and high network costs [26].

From a demand perspective, income disparities within countries hinder Internet diffusion [27]. Dohse and Cheng consider the important impact of geographic location on the digital divide [24]. This finding can be interpreted as a consequence of the lack of telecommunications infrastructure in rural and remote areas. Some studies have found that internet connections made by households are influenced by income, education level and number of children [28]. Some scholars have focused on the factors influencing the choice of Internet modes, arguing that the choice of online activities, such as communication, entertainment, social networking and e-commerce, is largely dependent on digital skills [29].

In contrast, the literature on the digital divide in China is less extensive and focuses mainly on comparative studies of urban-rural differences [30]. Regarding Internet diffusion, some scholars believe that in addition to traditional socio-economic factors determining Internet access, the factor of children in the family also plays an active role [31]. Scholars generally believe that education is still the biggest limitation for Chinese residents to use the Internet [32]. For the quality of Internet usage, it was found that economic income [33], education level, age [34], type of work [35] and digital skill [36] can affect Chinese residents' Internet use.

In the case of China, few studies have addressed the digital divide in the rural context by using microeconomic data. Some scholars have analyzed the disadvantages of rural areas in the digital divide through comparison between urban and rural areas [37], partly studying the impact of digital divide on income in rural areas [38], consumption structure [39], etc. There are also some studies focus on the impact of digital divide on youth [40], the Elderly people [41] and other special groups. However, existing studies have not comprehensively analyzed the situation of Internet access, use, and usage patterns (entertainment, social networking, e-commerce, study and work) in rural areas, as well as the factors influencing them. Therefore, this paper analyzes Internet access, usage and usage patterns in rural areas and their influencing factors based on differentiating Internet access (including mobile access and regular access) and usage patterns (entertainment, social networking, e-commerce, study and work).

3 Data Sources and Research Methodology

3.1 Data Sources

The data used in this paper come from the 2018 China Family Tracking Survey (CFPS) organized and implemented by the China Center for Social Science Research at Peking University. The CFPS observes and records China's characteristics and changes in social, economic, demographic, educational, and health aspects in an all-round way by tracking the data at the three levels of the individual, the family, and the community over a long period of time. A sample of 15,605 rural people aged 12 years and above in CFPS was

selected for this study. The sample data contains socio-demographic information such as gender, age, education level and occupation. The main topic of the survey was Internet access (both in the form of regular Internet access, such as computers, and mobile Internet access, such as cell phones) and usage, and the respondents were asked about Internet access, frequency of use, online activities, digital skills, and so on.

The data shows the descriptive statistical information from the sample. In terms of gender, 51.84% of the respondents were male, which is basically in line with the percentage of male population (51.13%) in the China Statistical Yearbook 2019; 22.33% of the respondents were 32 years old or younger, which reflects the young population in rural areas. In terms of education, 45.11% were in elementary school and below, 32.40% in middle school, 12.41% in high school, and only 10.08% in college. This data reflects the reality of the low proportion of people with higher education typical of rural China, coupled with the fact that education is an important factor influencing the digital divide [42], which deepens the possibility of an internal digital divide in rural areas, exacerbating digital exclusion and educational backwardness of the rural population [8]. In terms of work, 59.14% of the respondents were mainly engaged in their own agricultural production and management, accounting for 35.71% of the employed. Interestingly, the western region has the highest percentage of farmers realizing Internet access at 39.99%, followed by the eastern and central regions respectively. As for digital connectivity in rural China, 35.53% of respondents realized Internet connectivity (either regular access or mobile access). Among them, the proportion of mobile Internet access (42.08%) has exceeded the proportion of regular Internet access (11.48%). A higher proportion of Internet users have accessed the Internet via mobile Internet. The proportion of rural users using the Internet is 34.27%, and the proportion of Internet users who fulfill the conditions for Internet use set in this paper is 96.45%.

3.2 Research Methodology

In order to determine the factors influencing Internet usage patterns in rural China, this paper identifies five usage patterns: entertainment, social networking, e-commerce, study and work. Overall, rural users in the sample used e-commerce activities and study and work activities relatively less. This also shows that rural residents' Internet applications are less used for value-creating activities such as work, study, and transactions, which is similar to the findings of the China Academy of Information and Communication Research Institute's "Research Report on China's Urban and Rural Digital Inclusive Development - Digitalization Helps Revitalize the Rural Areas and Shared Prosperity," which concluded that China's The proportion of rural residents using the Internet for learning and business activities is low [43]. First, we use a logistic regression model to estimate the determinants of Internet access. The dependent variable consists of two variables, regular access (regular) and mobile access (mobile), both of which are binary (1 indicates an Internet connection and 0 indicates no Internet connection). The independent variables are a set of socioeconomic and demographic characteristics at the household and individual level. Second, two equations were created to model Internet use decisions and usage patterns in rural areas. It is worth noting that the choice of usage pattern depends on the Internet usage decision. It is important to note that given the different demographic characteristics in the two models, Internet usage patterns are

influenced by Internet usage decisions, meaning that the second equation may have a sample selection problem. Considering that the dependent variable of Internet use decision is a binary variable (1 means use, 0 means not applicable), according to the research of Shen Hongbo et al. [44], splitting the Heckman two-stage analysis so as to solve the sample selection bias problem and provide consistent and asymptotically valid estimates for the parameters.

3.2.1 Using Probabilistic Models to Estimate Whether Residents Use the Internet or Not

This view is based on the utility maximization model, which argues that the decision to use the Internet depends on a range of individual and household characteristics, reflecting differences in education, Internet use skills, financial status, social capital, and age [45]. For these reasons, the decision to use the Internet depends on the maximization of the utility of its use by each individual, which can be expressed as:

$$y_{io}^* = X_{io}\beta_0 + \varepsilon_{i0} \quad (1)$$

where X_{io} represents a matrix of independent variables (such as socio-demographic characteristics, social capital and digital skills), β_0 represents a vector of coefficients, and ε_{i0} is a normally distributed random error term. Total utility is unobservable, but the decision to access the Internet is observable. Therefore, ε_{i0} is the result of a decision-making process that is influenced by the explanatory variables. Thus, $y_{io} = 1$ indicates an individual's decision to access the Internet (regular access or mobile access), $y_{io} = 0$ indicates not doing so.

3.2.2 Internet Usage Pattern Selection Model

After determining whether the residents use the Internet or not, the Internet usage pattern selection model is determined as shown in (2):

$$y_{ij} = X_{ij}\beta_j + \varepsilon_{ij} \quad (2)$$

Among them, j (where $j = 1, \dots, J$) represents the Internet usage pattern. y_{ij} measurement of Internet usage patterns, X_{ij} represents the matrix of independent variables and ε_{ij} denotes the normally distributed random error term. This section uses the Heckman two-stage method for regression analysis and assumes a binary normal distribution with zero mean and correlation. In applying this method, the presence of sample bias is indicated if the estimated Lambda coefficient is significant [46], and is suitable for use with sample selection models.

3.2.3 Variable Selection

This paper chooses five modes of use: entertainment, study, work, socialization, and e-commerce. In order to describe the use of rural residents on the Internet, this paper sets that as long as the frequency of use of any of the above five modes reaches at least once a month and above is deemed to have used the Internet. The independent variables

in Internet use (Eq. 1) and Internet use pattern (Eq. 2) are divided into three categories, including the socioeconomic and demographic characteristics of the respondents, digital skills, and children.

The socio-economic variables include gender, age, education level, household economic status, and personal occupation. Regarding age, this paper refers to the study of Marlen et al., who identified three age ranges of 12–32, 33–64, and 65 and above [47]. Education level is measured using four categories: primary school and below (reference), middle school, high school/technical school/technical school/vocational high school, and university and above. Similarly, When dividing the work types of respondents, they are divided into own agricultural production and operation (reference), private enterprises/individual industrial and commercial households/other self-employment, agriculture work, employment, and non-agricultural casual workers. This paper chooses “net household income per capita” from the CFPS2018 database as a measure to reflect the living conditions of households. And we uses the questions “Do you send and receive emails?”, “Online shopping expenses (yuan/year)?”, “How many hours do you spend online in your spare time every week?” to measure. The number of children in the family are used to measure the impact of children in the family on Internet use. Finally, this article divides China into eastern, central and western to consider the Internet usage in different areas.

4 Empirical Results

4.1 Determinants of Internet Access

Table 1 shows the regression results of Internet access in rural areas of China. Our research results confirm that education level and family income are core influencing factors of Internet access, which means that residents with higher education and higher household income are more likely to have access to the Internet. This result is consistent with Chaudhuri et al. [48]. Possible reasons for this are that the level of education enhances a person’s ability to receive information and learn, while mobile Internet access requires a certain level of learning ability, and the devices that enable mobile Internet access, etc., require a certain level of ability to pay for them. The significance of age and its squared term in relation to mobile Internet access suggests that older people are more likely to adopt mobile Internet access, but with decreasing marginal effects. However, it can be seen from the numerical value of the age coefficient that the impact of age on the mobile Internet is not very strong. Gender is another core factor affecting mobile access, with men (relative to women) more likely to have mobile Internet access, which may be related to the fact that men in rural areas are more likely to be valued in education. Residents who are ‘private enterprise/self-employed/enterprise self-employed’ and ‘employed’ are more likely to have both types of Internet access than those who have their own agricultural business. A possible reason why employed residents are more likely to have access to the Internet is that the organizations they work for have certain requirements for Internet use, which increases their opportunities and likelihood of joining the Internet. Geographically, residents in Western China (compared to Eastern China) are less likely to have access to the Internet (both regular and mobile), and residents in Central China (compared to Eastern China) are less likely to have mobile

Internet access. This reveals that Internet access in rural China may be closely linked to economic development across regions, suggesting that the digital divide in rural areas in the digital era is homogeneous with the development divide in the industrialized era Qiu Zeqi. Existing data do not confirm the relationship between household size, number of children and Internet access. In general, the group of farmers accessing the Internet has the following characteristics: better educated in the eastern part of the country, better family income, and men employed by an organization.

Table 1. Regression results of factors influencing rural Internet access

variant	Effects on Mobile (Std.Err)	Effects on Regula (Std.Err)	variant	Effects on Mobile (Std.Err)	Defects on Regula (Std.Err)
Gender (Female = 1)	0.113 (0.0701)	0.627*** (0.0830)	Non-farm labor	0.251 (0.1711)	0.214 (0.2577)
Age	0.047 (0.0244)	0.101* (0.0425)	Household income	0.279*** (0.0456)	0.384*** (0.0578)
Age squared	-0.002*** (0.0003)	-0.002*** (0.0001)	Family size	-0.013 (0.0117)	0.011 (0.0193)
Education	0.562*** (0.0538)	0.951*** (0.0457)	Number of children	-0.029 (0.0268)	-0.009 (0.0337)
Own Agricultural	reference	reference	Eastern	reference	reference
Private or self-employed	0.722*** (0.11355)	0.826*** (0.1560)	Central	0.015 (0.0942)	-0.206* (0.1009)
Working in agriculture	-0.468 (0.3706)	-1.287 (1.0945)	Western	-0.200** (0.6556)	-0.583*** (0.0977)
Employed	0.382*** (0.0835)	0.587*** (0.1241)	Observations	5970	5970

4.2 Determinants of Internet Use (First Stage)

Table 2 shows the results of the first stage regression of the Heckman two-stage estimation model. It shows that age, education level, digital skills, type of job, and household economic status are determinants of Internet adoption. Overall, the decision to use the Internet depends on the expected benefits of going online and the associated costs. The sample shows that young people in rural areas with a university degree or higher, higher digital skills, and better household economic circumstances are more likely to use the Internet. There is a clear disparity in terms of age: the younger the user, the more likely they are to use the Internet, suggesting that younger people are more willing to experiment with and adapt to new technologies, while older people are less comfortable with them. Residents with a university degree or higher are more likely to use the

Internet, which is consistent with existing research and confirms that the more educated users are more likely to use the Internet. This positive effect is still significant in rural areas. Residents in the “private enterprise/self-employed/enterprise self-employed” and “employed” categories are more likely to use the Internet than residents in their own agricultural operations, probably because they have a greater need to use the Internet at work.

Table 2. Regression results of factors influencing rural Internet use

Variables	Marginal Effects (Std.Err)	Variables	Marginal Effects (Std.Err)
Gender (Female = 1)	-0.077 (0.0975)	Employed	0.389** (0.1121)
Age (12-32 years)	0.460*** (0.1111)	Non-farm labor	0.074 (0.2212)
Age (13-64 years)	reference	Household economic status	0.230*** (0.0538)
Age (65 years and older)	-1.277 (0.3601)	Numerical Skills	4.577*** (0.1427)
Primary and below	reference	Number of children	0.038 (0.0378)
Secondary school	0.155 (0.1046)	Eastern	reference
High School	0.303 (0.1905)	Central	0.233 (0.1379)
University and above	1.250*** (0.2814)	Western	0.233* (0.1114)
Own Agricultural	reference	Log likelihood	-392.2732
Private or self-employed	0.601*** (0.1480)	Wald chi2	1204.82
Working in agriculture	-0.366 (0.5178)	Number of observations	5970

Access to wealth is crucial when using the Internet, as people of higher economic status are more likely to use the Internet, which is consistent with the findings of Kilenthong et al. [49]. As for digital skills, the likelihood of using the Internet is higher when a person has adequate work-study skills or business skills or recreational skills, this result is consistent with the findings of Hu Ying’s study on Chinese rural residents [50], suggesting that digital skills are the main reason for promoting residents’ Internet use.

Likewise, geographical location is crucial to Internet usage. Data shows that people living in rural areas of western China are more likely to use the Internet, which is an interesting phenomenon. Scholars generally believe that since the reform and opening up, China’s economy has been doing better in the east and weaker in the central and

western parts of the country. Information and communication technology (ICT) has been an important support for less developed regions to achieve “catch-up” development, but weak infrastructure and other factors have made Internet access and use in the west significantly less favorable than in the east [51]. However, our study found that residents of rural areas in the west are more likely to use the Internet, which may be related to China’s western development and the “village-to-village” telecommunication policy, which encourages and supports more residents of rural areas in the west to use the Internet.

4.3 Determinants of Internet Use (Second Stage)

The Lambda coefficient of the treatment variable (Internet use) can be calculated based on the regression results of the first stage. After adding the Lambda coefficient, the regression results of the second stage of the two-stage estimation model can be calculated, as shown in Table 3. The Lambda coefficients of the five Internet usage patterns are all significant at the 0.001 level, Indicates that it is appropriate to use a sample selection model. The results show the regression results of the Heckman two-stage model on factors influencing Internet usage patterns (entertainment, social networking, e-commerce, learning, and work) in rural areas of China. Internet usage patterns vary by gender, age, education level, occupation and geographic location. In gender, women are more likely to use the Internet for social networking and electronic trading activities, while men are more likely to use the Internet to learn. It shows that Chinese rural women mainly use the Internet to maintain contact with family and friends, and these activities are related to the traditional role of Chinese rural women in the family. In terms of age, young people are more likely to use the Internet for entertainment, social networking, e-commerce, study and work, etc.; Except for the elderly people who did not show any significant relationship in learning, they are unlikely to carry out related activities through the Internet in other activities.

In education, those most likely to study and work via the Internet are those with university education and above, while those most likely to engage in e-commerce are those with high school education and above. It implies that rural users with higher levels of education are more likely to engage in value-added activities via the Internet. The conclusion is interesting. Residents with different education levels are relatively similar in their likelihood of using online entertainment. It is consistent with Hu Ying’s (2022) study, currently, most residents use the Internet mainly for entertainment activities. This Internet usage pattern structure needs to be improved and adjusted urgently [52].

In terms of occupation, those residents who are “employed” and work in the private sector, for example, are more likely to participate in value-added online activities such as learning and working, while “agricultural wage earners” or “non-agricultural casual” users are less likely to participate in value-added online activities such as learning and working. “Working agricultural” or “non-farm casual” users are less likely to participate in such value-added activities, shows that a digital divide at the level of Internet use has formed in rural areas due to differences in occupational types.

Digital skills significantly increase the likelihood of engaging in entertainment, social networking, e-commerce and learning via the Internet, except for a non-significant effect

Table 3. Regression results of factors influencing rural Internet use

variant	Entertainment	Social networks	E-commerce	Learning	Working
Marginal Effects (Std.Err)					
Gender (Female = 1)	0.009 (0.0085)	-0.016* (0.0070)	-0.056*** (0.0137)	0.038* (0.0147)	0.023 (0.0146)
Age (12–32 years)	0.060*** (0.0092)	0.040*** (0.0079)	0.161*** (0.0159)	0.055** (0.0162)	0.067*** (0.0162)
Age (33–64 years)	reference	reference	reference	reference	reference
Age (65 years and older)	-0.539*** (0.1112)	-0.329*** (0.0802)	-0.306*** (0.0552)	-0.109 (0.0670)	-0.133*** (0.0363)
Primary and below	reference	reference	reference	reference	reference
Secondary school	0.079*** (0.0127)	0.005 (0.0106)	0.154*** (0.0197)	0.172*** (0.0190)	0.128*** (0.0185)
High School	0.083*** (0.0138)	0.021 (0.0121)	0.264*** (0.0235)	0.338*** (0.0250)	0.300*** (0.0259)
University and above	0.064*** (0.0149)	0.037** (0.0111)	0.031*** (0.0231)	0.535*** (0.0246)	0.548*** (0.0242)
Own Agricultural	reference	reference	reference	reference	reference
Private or self-employed	0.050** (0.0158)	0.064*** (0.0134)	0.226*** (0.0255)	0.103*** (0.0270)	0.171*** (0.0271)
Working in agriculture	0.072*** (0.0173)	0.060*** (0.0130)	0.145 (0.1133)	-0.070 (0.1043)	-0.032 (0.1043)
Employed	0.054*** (0.0122)	0.050*** (0.0111)	0.175*** (0.0201)	0.080*** (0.0198)	0.185*** (0.0198)
Non-farm labor	0.069*** (0.0228)	0.027 (0.0235)	0.113* (0.0454)	-0.008 (0.0404)	-0.008 (0.0370)
Household Economic Status	0.044*** (0.0057)	0.020*** (0.0045)	0.051*** (0.0086)	0.023* (0.0095)	0.034*** (0.0094)
Number of children	-0.000 (0.0036)	0.001 (0.0027)	0.003 (0.0055)	0.012* (0.0060)	0.003 (0.0058)
digital skill	0.819*** (0.0748)	1.069*** (0.1239)	0.575*** (0.1497)	0.496*** (0.1055)	0.197 (0.1055)
Eastern	reference	reference	reference	reference	reference
Central	0.016 (0.0105)	0.019* (0.0097)	0.003 (0.0176)	0.021 (0.0191)	-0.006 (0.0194)

(continued)

Table 3. (continued)

variant	Entertainment	Social networks	E-commerce	Learning	Working
Marginal Effects (Std.Err)					
Western	0.011 (0.0101)	0.042*** (0.0085)	0.008 (0.0164)	0.0861*** (0.0176)	0.008 (0.0178)
IMR	0.367*** (0.0063)	0.397*** (0.0078)	0.189*** (0.0091)	0.123*** (0.0079)	0.010*** (0.0095)
Cons	-0.479*** (0.0892)	-0.415** (0.1298)	-0.783* (0.1649)	-0.634*** (0.1330)	-0.464** (0.1448)
N	3860	3862	3859	3864	3471
R-sq	0.2427	0.296	0.2326	0.1878	0.274

on work patterns. It is broadly consistent with existing research, which generally agrees that people with higher digital skills are more likely to use different Internet modes [53].

Geographic location is crucial in explaining Internet usage patterns, and Table 5 illustrates the heterogeneity in Internet usage patterns across different rural areas in China. In rural China, users in the central region are more likely to use the Internet for social activities, and users in the western region are more likely to use the Internet for social and learning activities. It is inconsistent with existing research, as some scholars believe that residents in rural areas of western China are drowning in entertainment and other recreational “junk information”. However, we feel that as China’s western region continues to develop, the attitudes of rural residents towards the Internet have changed dramatically, and that the Internet may become an effective means of catching up by leaps and bounds in the western region. The higher the economic income, the higher the likelihood of using the Internet for entertainment, socializing, e-commerce, learning and work. Finally, users with more children in the household are more likely to use the Internet for learning, which may be related to the current promotion of online education and the importance that families place on their children’s education.

5 Conclusions and Recommendations

This paper uses the Heckman two-stage model based on the China Family Tracking Survey (CFPS) 2018 data to study the influencing factors of Internet access, usage and usage patterns in rural China, which provides empirical evidence to prove the primary and secondary digital divide in rural China. This paper explains the influencing factors of rural residents’ Internet access and usage behaviors at the micro level. It will help further enrich theoretical research on the digital divide. (1) It is found that the factors influencing Internet access, usage and usage patterns in rural China are similar to those observed in other developing countries at the early stage of Internet diffusion, mainly as follows: (1) We can infer the profile of Internet access user groups in rural areas of China: young and well-educated people with good economic income and digital skills; and there is a

big gap in access methods in terms of Internet access. Mobile Internet access is significantly higher than conventional Internet access; (2) Age, education level, digital skills and family economic status are important factors affecting Internet use; (3) Internet use patterns (entertainment, social networking, e-commerce, learning, and geographic location and work) have significant differences in gender, age, education level, occupation and geographic location; and (4) rural users' Internet use is mainly dominated by non-value-creating activities such as entertainment and socializing. Accordingly, this paper puts forward the following policy recommendations:

Based on the conclusions drawn in this paper, the following countermeasures are proposed for Internet popularization and use in rural areas. First, it is necessary to comprehensively understand the evolution and trend of the digital divide in rural areas in practice, establish a sound digital policy in light of the fact that rural areas mainly rely on smartphones for Internet access, and increase the promotion of online services and applications related to employment, education, and public health, so that more people in rural areas of developing countries can enjoy the digital dividend. Secondly, continue to improve rural Internet access infrastructure to help more rural residents access the Internet through broadband; encourage communication companies to continuously strengthen smart phone products and related services so that mobile Internet access can obtain the same functions as conventional Internet access; establish a sound training system for rural users' Internet use skills, and realize the use of the Internet by improving the digital skills of rural users; and enhance the advantages brought by Internet use. Enhance the advantages and benefits brought by the use of the Internet. Once again, continuously improve e-commerce-related policies, provide policy support for rural young people to engage in e-commerce. And take e-commerce as an big way to solve the problem of young people's employment in rural areas; continuously improve young people's digital skills, and provide support to promote the sustainable and healthy development of e-commerce in rural areas. Finally, the government must pay attention to the problem of digital divide among rural elderly groups. On the one hand, we should use the community as a unit to improve the digital skills of the elderly people and minimize the digital divide among the elderly people; on the other hand, in various specific scenarios of grassroots governance (especially governance scenarios for the purpose of "intelligent"), focus on Taking into account the fact that the elderly people cannot use the Internet well reflects the temperature of grassroots governance.

It should be pointed out that this study only deals with the two levels of digital divide, Internet access and Internet use, and does not discuss the three levels of digital divide, such as the inequality brought by Internet use. This is also a breakthrough for future research, i.e., discussing the current situation of the three-level digital divide in rural areas and its influencing factors, so as to continue to enrich the research on the impact of Internet access and use on the life satisfaction of residents in rural areas in China and other developing countries.

References

1. Prieger, J.E.: The broadband digital divide and the economic benefits of mobile broadband for rural areas. *Telecommun. Policy* **37**(2013), 483–502 (2013)
2. Cooper, M., Kimmelman, G.: *The Digital Divide Confronts the Telecommunications Act of 1996: Economic Reality Versus Public Policy*. Consumer Union, Washington, DC (1999)
3. Qiu, Z., Zhang, S., Liu, S., Xu, Y.: From digital divide to dividend difference - a perspective of internet capital. *China Soc. Sci.* (10), 93–115+203–204 (2016)
4. Gübli, G.A., Xue, D., Song, Y., Zhang, X., Yang, B.: Regional differences and factors influencing the evolution of the number of villages in China. *China Agric. Resour. Zoning* **42**(11), 85–96 (2021)
5. Zhou, Z.: Hollowing out of China's rural population and its challenges. *Popul. Stud.* (02), 45–52 (2008)s
6. Li, Y., Wang, H.: Spatial distribution of hollow and solid villages of China's population-evidence from the administrative village sample of the third agricultural census. *China Rural Econ.* (04), 124–144 (2020)
7. Blumberg, R.: Alternative food networks and farmer livelihoods. A spatializing livelihoods perspective. *Geoforum* **88**, 161–173 (2018)
8. Yang, R., Cao, Y.: On the tension between digital empowerment and digital divide in the countryside and its dissolution. *J. Nanjing Agric. Univ. (Soc. Sci. Ed.)* **21**(05), 31–40 (2021)
9. De Grammont, H.: Hacia una ruralidad fragmentada. *Nueva Soc.* **262**, 51–63 (2016). http://nuso.org/media/articles/downloads/2.TC_de_Grammont_262.pdf
10. Wang, D., Liu, Z.: Technological empowerment in rural areas: connotation, dynamics and boundaries. *J. Huazhong Agric. Univ. (Soc. Sci. Ed.)* (3), 138–148 (2020)
11. Zhang, J., Cheng, M., Gong, X.: Research on the characteristics and influencing factors of digital divide between urban and rural areas in China. *Stat. Inf. Forum* **36**(12), 92–102 (2021)
12. Qiu, Z., Li, Y., Xu, W.: Digitalization and rural governance structure change. *J. Xi'an Jiaotong Univ. (Soc. Sci. Ed.)* **42**(02), 74–84 (2022)
13. Erdiaw-Kwasie, M., Khorshed, A.: Towards understanding digital divide in rural partnerships and development: a framework and evidence from rural Australia. *J. Rural. Stud.* **43**, 214–224 (2016)
14. Compaine, B.M. (ed.): *The Digital Divide: Facing a Crisis or Creating a Myth?* MIT Press, Cambridge (2001)
15. Büchi, M., Just, N., Latzer, M.: Modeling the second-level digital divide: a five-country study of social differences in Internet use. *New Media Soc.* **18**(11), 2703–2722 (2016)
16. Van Deursen, A.J., Helsper, E.J.: The third-level digital divide: who benefits most from being online? *Communication and information technologies annual*, pp. 29–52. Emerald Group Publishing Limited (2015)
17. Van Dijk, J.A.G.M.: *The Digital Divide*. Polity Press, Cambridge (2019)
18. Van Deursen, A., Van Dijk, J.: The digital divide shifts to differences in usage. *New Media Soc.* **16**(3), 507–526 (2014)
19. DiMaggio, P., Hargittai, E., Celeste, C., Shafer, S.: From unequal access to differentiated use: a literature review and agenda for research on digital inequality. In: Kathryn, N. (ed.) *Social Inequality*. Russell Sage Foundation, New York (2004)
20. Zhou, L., Feng, D., Yi, X.: Digital inclusive finance and urban-rural income gap: “digital dividend” or “digital divide”. *Economist* (05), 99–108 (2020)
21. Huang, Y.: Causes and empowering governance of China's digital divide. *Learn. Practic.* (09), 23–33 (2022)
22. Liu, Y., Li, X.: How to cross the digital divide of the elderly people in the digital age? *Southeast Acad.* **05**, 105–115 (2022)

23. Liu, W.: The effect of digital financial inclusion on residents' relative poverty. *J. South China Agric. Univ. (Soc. Sci. Ed.)* **20**(06), 65–77 (2021)
24. Dohse, D., Cheng, L.: Bad neighborhood and internet adoption in poor countries: what is behind the persistent digital gap? *Growth Chang.* **49**(1), 241–262 (2018)
25. Kiiski, S., Pohjola, M.: Cross-country diffusion of the internet. *Inf. Econ. Policy* **14**(2), 297–310 (2002)
26. Gwaka, L., May, J., Tucker, W.: Toward low-cost community networks in rural communities: the impact of context using the case study of Beitbridge, Zimbabwe. *Electron. J. Inf. Syst. Dev. Ctries.* **84**(3), 1–11 (2018)
27. Srinuan, C., Bohlin, E.: Analysis of fixed broadband access and use in Thailand: drivers and barriers. *Telecommun. Policy* **37**(8), 615–625 (2013)
28. Michailidis, A., Partalidou, M., Nastis, A., Papadaki-Klavdianou, A., Charatsari, C.: Who goes online? Evidence of internet use patterns from rural Greece. *Telecommun. Policy* **35**(4), 333–343 (2011)
29. Garín-Muñoz, T., López, R., Pérez-Amaral, T., Herguera, I., Valarezo, A.: Models for individual adoption of eCommerce, eBanking and eGovernment in Spain. *Telecommun. Policy* **43**(1), 110–111 (2019)
30. Cao, J., Mei, W.: The gap between urban and rural starting lines: an analysis of the digital divide in transitional China. *Contemp. Commun. (02)*, 17–21+78 (2017)
31. Hong, J., Li, X.: WeChat's "feedback" diffusion in rural families: an examination based on Chen District village in Shanxi Province. *Int. J.* **41**(10), 50–74 (2019)
32. Wang, C., Li, J., Huang, Y.: Classification, impact and response of digital divide. *Financ. Sci. (04)*, 75–81 (2022)
33. Fan, G., Zhang, X.: *China's Information Economy in Global Perspective: Development and Challenges*. Renmin University of China Press, Beijing (2003)
34. Cheng, M., Zhang, Z.: Internet development and consumption gap between urban and rural residents in the context of the new era. *Res. Quant. Econ. Tech. Econ. (7)*, 22–41 (2019)
35. Song, H.: Digital divide or information empowerment? – a study on mobile phone usage of migrant workers in the Yangtze River Delta. *Modern Commun. (J. Commun. Univ. China)* **38**(06), 132–137 (2016)
36. Lu, J.H., Wei, X.D.: Analytical framework, concepts and path selection of digital divide governance for the elderly - based on the perspectives of digital divide and knowledge divide theory. *Popul. Stud.* **45**(03), 17–30 (2021)
37. Su, H.: Digital urban-rural construction: the road to urban-rural integration and common wealth. *E-Government (10)*, 88–98 (2022)
38. Li, W., Zhou, D., Li, X.: Research on the impact of digital divide on rural household income and mechanisms - an empirical analysis based on China family tracking survey (CFPS) data. *J. Southwest Univ. Natl. (Hum. Soc. Sci. Ed.)* **43**(10), 116–127 (2022)
39. Huang, M., Dou, X.: Will the urban-rural digital divide hinder the upgrading of rural residents' consumption structure? – an analysis based on data from the China Family Tracking Survey (CFPS). *Econ. Issues Explor.* **09**, 47–64 (2022)
40. Wu, X., Wang, M.: A study of digital inequality among rural youth based on ethnographic futures research method. *Libr. Constr. (03)*, 35–38+42 (2016)
41. Wang, M., Yan, H.: The value of social capital in the process of rural residents crossing the episodic digital divide - a report of a field survey in Jinghai, Tianjin. *Chin. Libr. J.* **39**(05), 39–49 (2013)
42. Huang, Y.: Causes and empowering governance of China's digital divide. *Learn. Practic. (09)*, 23–33 (2022)

43. Research Report on China's Urban and Rural Digital Inclusive Development-Digitalization Boosts Rural Revitalization and Common Wealth. China Academy of Information and Communication Research Website. http://www.caict.ac.cn/kxyj/qwfb/ztbg/202112/t20211230_394812.htm
44. Shen, H., Yang, Y., Pan, F.: Political affiliation, securities violations and surplus quality of privately listed companies. *Financ. Res.* (01), 194–206 (2014)
45. Alderete, M.V.: Examining the drivers of Internet use among the poor: the case of Bahía Blanca city in Argentina. *Technol. Soc.* **59**, 42–58 (2019)
46. Wooldridge, J.M.: *Econometric Analysis of Cross Section and Panel Data*, 2nd edn. MIT Press, Cambridge (2010)
47. Martínez-Domínguez, M., Mora-Rivera, J.: Internet adoption and usage patterns in rural Mexico. *Technol. Soc.* (60), 101226 (2020)
48. Chaudhuri, A., Flamm, S., Horrigan, J.: An analysis of the determinants of internet access. *Telecommun. Policy* **29**(9–10), 731–755 (2005)
49. Kilenthong, T., Odton, P.: Access to ICT in rural and urban Thailand. *Telecommun. Policy* **38**(11), 1146–1159 (2014)
50. Hu, Y.: A review of the urban-rural digital divide in the context of rural revitalization. *Res. Soc. Chin. Charact.* (04), 60–69 (2022)
51. Wang, C.: Research on the innovation model generated in the process of digital economy development. *Sci. Manag. Res.* **41**(01), 74–79 (2023)
52. Hu, Y., Zhong, Y.: Digitalization helps rural revitalization and common wealth – the case of Maoming City, Guangdong Province. *New Econ.* (02), 47–53 (2023)
53. Ding, S., Liu, C.: Internet use, income improvement and skill premium of youth groups in the digital economy era - a perspective based on rural-urban differences. *Contemp. Econ. Manag.* **44**(08), 64–72 (2022)