

# Digital Countryside, Rural Tourism, and (Female) Labor Supply -Taking CFPS and Taobao Village Data as Examples

Yuhan Yuan

College of Economics and Management, Beijing Forestry University, Beijing, China

yh18811268768@163.com

**Abstract.** In recent years, how to use digital technology to accelerate rural development and farmers' well-being has become a topic of the times that has brought widespread concern at this stage. Based on the database of China family tracing survey from 2010 to 2022 and the digital rural index of Taobao Village, this paper takes rural tourism as a regulating variable to investigate the influence of digital technology on rural women's labor supply. The results show that digital technology has significantly improved rural women's working hours and labor supply. Further research shows that rural tourism plays a regulatory role, and its development can strengthen the positive impact of digital villages on rural women's labor supply. Expansive analysis shows that the labor supply of middle-aged and elderly people and low-education groups has a greater positive impact.

**Keywords:** digital countryside, rural tourism, female labor supply.

## 1. Introduction

With the improvement of medical and health care and the decline in birth rate, the population structure changes, the working-age population decreases, the aging population and the labor participation rate decrease, and the problem of insufficient labor supply in China becomes increasingly apparent. In the *Outline of Women's Development in China (2021-2030)* issued by the State Council, China in 2021, it was pointed out that in the new era, the issue of women's livelihood security in rural areas, especially in underdeveloped areas, is prominent. In addition, a large number of young rural laborers flow to cities, which harms production efficiency and brings about rising labor costs. In addition, according to the instruction of the Chinese government in 2024, the problem of rural women's marginalization needs to be improved, and the *No.1 Document of the Central Committee* particularly emphasizes caring for women and promoting their employment. The above phenomenon shows that women's employment is still facing various difficulties, getting low attention and great potential for labor development. Therefore, this issue has further exploration value.

Since the 18<sup>th</sup> National Congress of the Communist Party of China, the Central Committee of the Communist Party of China has attached great importance to the digital construction of agriculture. At present, the development level of digital infrastructure in county rural areas is relatively high, but the level of the rural economy, governance, and digitalization of life is low and there are obvious differences among regions.

At the same time, the rise of rural tourism has created a large number of jobs, which has spawned the "sideline occupations during the agricultural off-season" model and realized the seasonal complementarity between agriculture and tourism. Therefore, with the rapid development of information and the constant adjustment of industrial structure in the Internet age, which is both an opportunity and a challenge for rural school-age women who are facing the dual roles of market and family. It can be seen that it is an urgent proposition of the times to explore how to use digital technology to boost women's employment and realize rural revitalization.

Under the background of digital development of county villages, this paper uses the microscopic data of China Family Panel Studies (CFPS) in 2018 and 2020, and the *Index of County Digital Villages* to investigate the influence of digital villages on rural women's labor supply. In particular, this paper adds rural tourism as a regulating variable. In the research, the following questions will be discussed: Does the construction of digital villages have an impact on the labor supply of rural women? What is the transmission mechanism behind it? Can the development of rural tourism enhance this

influence? Is there heterogeneity among groups of different ages, different education levels, and different job categories?

The contribution of this paper is mainly reflected in the following two aspects: Firstly, from the perspective of existing literature, domestic and foreign research mainly focuses on the mechanism and implementation strategy of digital villages empowering rural revitalization, with less micro and macro data, and failing to concentrate on rural women. This paper is helpful to explore the influencing factors of gender in the labor market, enriching the theory in the field of gender and labor research. Secondly, starting from the method dimension, this paper controls exogenous variables while adjusting variables in the process of inquiry, which provides a new perspective and ideas for other studies in terms of methods. Finally, from the policy point of view, it can rationally allocate and optimize the labor resources of rural middle-aged women, attract more rural female workers to join the labor market, alleviate the shortage of rural labor supply, and thus promote rural economic development and social progress.

## 2. Literature Review

In recent years, with the gradual development of digital technology, all fields have begun to develop digitally. As an important national development strategy, digital villages have gradually become a hot topic. Most of the previous literature focuses on the macro field, discussing the realistic representation, implementation strategy [1], how to promote rural revitalization [2], the problems currently faced[3], and so on. Some studies focus on a certain field. Sun Yujie et al. found that rural education under digital technology promoted educational equity[5]. In addition, society and villages can also realize the modernization of the rural governance system and governance capacity using digital space and improving villagers' digital literacy [7]. From the microscopic point of view, many studies have verified that the digital economy not only promotes the development of agriculture and the rural economy, but also brings long-term income increases to farmers, and the income gap among rural residents is also significantly reduced [8-9]. Some studies have also found that when the rural economy and farmers improve, the consumption level of rural residents will also upgrade accordingly [10]. Generally, most of the literature about the development of digital villages solved problems from a macro perspective, and few did theoretical research based on microdata. At the same time, from the topic of this paper, the existing research has not paid attention to the impact of digital construction on rural residents, especially female labor supply.

Although the previous literature did not pay enough attention to this, we can also learn from some experiences of the influence of digital technology on labor supply. Most studies showed that the development of digital technology can open up new jobs and improve workers' working hours. Lin Longfei et al. showed that the development of the digital economy has improved the digital literacy of workers, thus alleviating the shortage of labor supply, and its impact on female workers was more significant [11]. There are many channels for its action. Qi Yudong believed that the popularity of the Internet has increased work preferences and income levels, thus prolonging working hours [12]; Shi Yupeng et al. found that the deep integration of digital technology and education improved the labor participation rate and working time of urban women [13]. However, some scholars believe that while technological progress creates a large number of jobs, low-tech jobs will be replaced in large numbers, and the overall labor supply level of workers employed on digital platforms will not be high [14].

Therefore, based on the existing literature, this paper, based on the county digital village index in 2018 and 2020, combined with the individual micro-data of CFPS, investigates the influence of digital village construction on women's labor supply, explores its transmission mechanism and the regulatory effect of rural tourism in it, and makes a series of heterogeneity analysis, which effectively expands the topic of related literature and provides new ideas for related research in methods.

### 3. Data Source and Model Setting

This section explains the data and variables, builds a regression model, and uses the panel data of 100 digital villages in 2018 and 2020 to investigate the impact of digital village construction on rural women’s labor supply.

#### 3.1 Data source

This paper uses both individual and macroeconomic data. First, at the individual level, the data source is the China Family Panel Studies (CFPS) database. This survey was conducted by the China Social Science Investigation Center of Peking University, and the sample covered 25 provinces, municipalities, and autonomous regions in China. Since 2010, it has been tracked for six rounds, with a total of seven issues. CFPS database contains rich social, economic, demographic, and other information, which is more in line with this study and provides a guarantee for our in-depth analysis.

Second, the macro-economic data. The county digital village measurement index is the county digital village index jointly published by Peking University New Rural Development Research Institute and Ali Research Institute, and the time is 2018-2020. For the first time, this data takes the county as the basic unit, comprehensively combs the digital content and specific representation of rural infrastructure, rural economy, rural life, and rural governance, systematically constructs the county digital village index system, and comprehensively evaluates the actual level of digital village development in 1880 counties in China. The county digital rural index is obtained by adopting the standardized processing method and uniformly dimensionless processing of the above four indicators with different properties and measurement units.

#### 3.2 Variable structure

The explained variable labor supply in this paper is measured by labor time, which comes from the “working hours per week” in the CFPS questionnaire. To facilitate observation and calculation, this paper uniformly averages the working hours per week (divided by 7) to every day. There are two questions in CFPS about working hours per week, main work and general work. In the benchmark regression, we calculate the explained variables by adding all working hours and then averaging them. In the heterogeneity analysis, we distinguish agricultural work from non-agricultural work according to the “nature of work” in the main work, and then calculate the corresponding working hours.

The core explanatory variable of this paper is the digital village index to capture the macro-digital impact on individuals. In the benchmark regression, the development level of county digital villages, which is formed by the above four indicators and released by Peking University, is adopted. In the analysis of heterogeneity, four specific indicators are adopted: rural digital infrastructure, rural economy digitalization, rural governance digitalization, and rural life digitalization. This paper also controls other factors that affect labor supply, including individual age, marital status, education level, etc. And social and economic state variables, such as the total population at the end of the year. Meanwhile, this paper also controls the fixed effect of city and time. Table 3.1 is descriptive statistics of each variable.

Table 3.1 Descriptive Statistics

Variable name	Obs	Mean	SD	Median	Min	Max
Working hours	9488	1.5181	0.9373	1.7918	-4.2485	3.1781
Digital rural index	9488	54.8148	11.7560	54.8302	22.5776	106.5505
Year-end total population	9279	3.7639	0.6651	3.7842	1.6094	5.0434
Years of education	9185	6.0351	4.7982	6.0000	0.0000	19.0000
age	9488	47.1487	14.0156	48.0000	16.0000	87.0000
Marital status	9488	4.5963	1.1145	5.0000	1.0000	5.0000

### 3.3 Model setting

#### 3.3.1 Benchmark regression model

In this paper, the following benchmark OLS regression model is constructed to examine the impact of digital villages on women’s labor supply:

$$Labor - supply_{it} = \alpha_0 + \alpha_1 DVI_{ct} + \alpha_2 X_{cit} + \lambda_{it} + \delta_{ct} + \varepsilon_{cit} \quad (1)$$

Among them,  $Labor - supply_{it}$  is the labor supply of the individual  $i$  in the year  $t$ ,  $DVI_{ct}$  is the digital rural index of the county  $c$  where the individual is located in the year  $t$ , and  $X_{cit}$  is the control variables at the individual and district level, including age, marriage, education level, total population at the end of the year, etc.  $\lambda_{it}$  and  $\delta_{ct}$  are fixed effects of city and time respectively, and  $\varepsilon_{cit}$  are random error terms.

#### 3.3.2 Regulatory effect model

The theoretical analysis of this paper shows that rural tourism can strengthen the influence of digital construction on rural women’s labor supply, so the following econometric model is set:

$$Labor - supply_{it} = \beta_0 + \beta_1 DVI_{ct} + \beta_2 M_{cit} + \beta_3 DLTI_{ct} + \beta_4 X_{cit} + \delta_{cit} + \varepsilon_{cit} \quad (2)$$

Among them,  $M_{cit}$  is the intersection term representing the digital rural index of the county and rural tourism,  $DLTI_{ct}$  represents the rural cultural tourism index and other variables have the same meaning as formula (1).

## 4. Estimation Results and Analysis

### 4.1 Benchmark regression results

Table 4.1 reports the results of the baseline estimation. Column (1) and column (2) have significantly positive variable coefficients before and after adding control variables. It shows that the digital countryside can promote the improvement of rural women’s labor supply. According to the estimation coefficient, every time the digital countryside index increases by 1 point, the rural women’s labor time increases by 0.005%. This is because, on the one hand, digital technology can create some new jobs and provide employment opportunities for some rural women. On the other hand, the digital countryside has introduced new technology, and the digitalization of life and infrastructure has eased the burden of rural women to take care of the elderly and children, so they have more time to work. However, it cannot be ruled out that the digital industry, with increasing work content and a threshold for the low-skilled labor force, requires the labor force to spend extra time learning to adapt to new technologies. Among the control variables, the influence of years of education on rural women’s working time is significantly positive, which is also in line with our cognition. Women with high skills and high education can adapt to the digital industry well, so they can undertake more work and increase their labor supply. In addition, age has a significant negative impact on labor supply. This is because older women have a weak ability to accept and learn new technologies. The development of digital technology has crowded out and replaced low-skilled laborers through mechanization, intelligence, and industrial restructuring, thus reducing their labor supply. But on the whole, the development of digital villages has had a positive impact on the labor supply of rural women.

Table 4.1 Benchmark Regression Results

Explanatory variable	(1)	(2)
	Working hours	Working hours
Digital rural index	0.005*	0.005*
	(0.003)	(0.003)

Marital status		0.012 (0.009)
Years of education		0.007** (0.003)
Age		-0.012*** (0.001)
Year-end total population		0.031 (0.041)
Constant	1.208*** (0.156)	1.598*** (0.241)
Time-fixed effect	YES	YES
<i>N</i>	9488	8977

Note: The standard error in brackets is robust; \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

## 4.2 Robustness test

In this section, we test the robustness by changing the explained variables, adjusting samples, changing models, and adding new control variables. According to the results, the results obtained by benchmark regression are still valid.

### 4.2.1 Replacing the explained variables

Here, we replace the explained variable with the digital index of rural life, which is more comprehensive and closer to farmers' lives. The results are shown in Table 4.2, column (1).

### 4.2.2 Eliminating those who may have no labor force

Considering that women younger than 18 and older than 60 may have fewer working hours due to their insufficient physical conditions, to test the robustness of the results in this paper, we only keep the samples of rural women aged 18-60 for regression. Column (3) removes 1% of the data before and after to avoid the influence of extreme values on the results. The results are shown in Table 4.2.

### 4.2.3 Replacing the model

In this paper, the high-dimensional fixed model is used to replace the original model for testing, and the results are shown in Table 4.2 column (4).

### 4.2.4 Adding new control variables

In this paper, new individual and regional control variables are added to the original model for regression. The results are shown in Table 4.2, column (5).

Table 4.2 Robustness Test

Explanatory variable	Replacing the explained variable	Adjusting sample		Replacing the model	Adding a new control variable
	(1)	(2)	(3)	(4)	(5)
	Working hours	Working hours	Working hours	Working hours	Working hours
Digital rural index		0.006* (0.003)	0.006* (0.003)	0.005* (0.003)	0.005* (0.003)

Digital index of rural life	0.006***				
	(0.002)				
Square of age					-0.000***
					(0.000)
political profile					-0.079
					(0.057)
Constant	1.563***	1.594***	1.549***	1.598***	1.282***
	(0.200)	(0.260)	(0.234)	(0.241)	(0.258)
Control variable	YES	YES	YES	YES	YES
Time-fixed effect	YES	YES	YES	YES	YES
N	8977	7225	8977	8976	8977

Note: The standard error in brackets is robust; \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

## 5. Expansive Analysis

### 5.1 Regulatory effect

To test the adjustment function of rural tourism, this paper selects the digital cultural tourism index in the county digital rural database published by Peking University New Rural Development Research Institute and observes the changes in the intensity of digital rural influence on labor supply with the development of tourism. Based on the original model, we add the cross-product term (M) of county digital rural index and rural tourism to make a regression, and the estimated results are shown in Table 5.1. The regression coefficient of column (1) is significantly positive at the level of 1%, which shows that with the continuous development of rural tourism, the influence of digitalization on rural women's labor supply has increased. The possible reasons are as follows: firstly, tourism has increased the number of jobs and enhanced the flexibility of work. Digital rural construction has lowered the employment threshold of rural tourism through digital tools such as e-commerce platforms and live delivery and enabled the labor force to gain more freedom in working place and time. Secondly, with the coordinated development of rural tourism and digital villages, some people began to work part-time and developed tourism during the slack season, which extended the working hours. At the same time, the intelligent digital platform reduces labor costs, and these technologies replace some repetitive and inefficient labor, optimize labor allocation, and improve the efficiency of time utilization.

Table 5.1 Regulatory Effect

Explanatory variable	(1) Working hours
Digital rural index	-0.002**
	(0.001)
Digital village index × digital cultural tourism index	0.015***
	(0.004)

Digital travel index	-0.190** (0.086)
Constant	2.234*** (0.074)
Control variable	YES
Time-fixed effect	YES
N	8840

Note: The standard error in brackets is robust; \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

## 5.2 Heterogeneity analysis

The influence of digitalization on individuals may vary with age, educational background, region, and other factors. Therefore, in this section, we make group regression according to these characteristics to investigate the heterogeneity of the influence of digital villages on rural women's labor supply.

### 5.2.1 Heterogeneity at the age level

According to the World Health Organization's age classification standard: young people are under 44 years old, and elderly people are over 60 years old. Because both agricultural work and non-agricultural work in rural areas need certain physical conditions, we divide the samples into four groups: under 18 years old, 18-35 years old, 35-59 years old, 60 years old and above, and the corresponding estimation results are shown in Table 5.2.1. The results of columns (1), (3) and (4) are not significant, indicating that teenagers and the elderly are not so sensitive to digitalization, are in a passive acceptance state, and may not be able to increase working hours because of their poor physical condition compared with middle-aged and elderly people. The middle-aged and elderly people are active in thinking and can support longer working hours. At the same time, employment pressure drives them to learn and iterate technology, and they also assume the role of education and maintenance of home digital devices. Some studies show that the frequency of middle-aged and middle-aged people using digital collaborative tools (such as nailing/flying books) is 2.4 times that of teenagers, so this group is an active digital practitioner, which is greatly affected by the advancement of digitalization.

Table 5.2.1 Heterogeneity Analysis of Age Level

Explanatory variable	Working hours			
	Under 18 years old (1)	18-35 years old (2)	35-59 years old (3)	60 years old and above (4)
Digital rural index	0.004 (0.011) (.)	-0.002 (0.005) (0.069)	0.008* (0.004) (0.054)	0.006 (0.008) (0.118)
Constant	-13.585* (2.078)	2.330*** (0.428)	1.311*** (0.338)	1.808** (0.728)
Control variable	YES	YES	YES	YES
Time-fixed effect	YES	YES	YES	YES
N	2268	4957	1728	24

Note: The standard error in brackets is robust; \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

### 5.2.2 Heterogeneity of education level

Will groups with low academic qualifications be more impacted by the process of digitalization? This paper analyzes the heterogeneity according to the years of education, and the estimated results are shown in Table 5.2.2. The results show that only the regression coefficient in column (1) is

significant, that is, the deepening of digitalization will increase the working hours of people with low education. This is because digitalization has spawned “low-skill-friendly” jobs, and the platform economy has risen: jobs such as take-out and online car rental have absorbed a large number of low-educated laborers, and they belong to the “piece-by-piece system”. They have actively extended working hours in pursuit of maximizing income and forced labor supply, hoping to make up for the income gap by increasing labor hours. In addition, people with low academic qualifications do not know enough about digital technology, so they need to spend more time learning and adapting. In the early stage, there may be cases of low efficiency and redundant operation leading to extended working hours, and they have become the group most affected by technology in the digital age.

Table 5.2.2 Heterogeneity Analysis of Age Level

Explanatory variable	Working hours			
	Primary school and below (1)	Junior high school (2)	Senior high school (3)	University or above (4)
Digital rural index	0.012*** (0.004)	0.000 (0.005)	0.007 (0.007)	-0.011 (0.009)
Constant	0.886*** (0.331)	2.575*** (0.582)	1.609** (0.768)	2.687** (1.084)
Control variable	YES	YES	YES	YES
Time-fixed effect	YES	YES	YES	YES
N	4819	2803	585	770

Note: The standard error in brackets is robust; \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

## 6. Conclusions and Suggestions

Based on the data of CFPS from 2018 to 2020, this paper studies the influence of rural digitalization on rural women’s labor supply, constructs the OLS model to verify it, and considers whether to develop rural tourism as a regulating variable. This paper draws the following conclusions: firstly, there is a positive correlation between rural digital development and female working hours. According to the estimation coefficient, the development level of county digital villages will increase by 1% and the working hours of rural women will increase by 0.005% when other factors remain unchanged. The result has passed the robustness test, and the result is consistent with the benchmark estimation result. Secondly, the expansive analysis further introduces the development of rural tourism as a regulating variable. We find that in the rural areas where rural tourism has been developed, the degree of digitalization has expanded the number of jobs, and the working time has become fragmented, resulting in longer working hours for women. Finally, the influence of digital empowerment on rural women’s labor supply is heterogeneous. Middle-aged and elderly women with low academic qualifications have relatively low labor skills and greater pressure, and more jobs and skills demands make their labor time increase.

Combined with the above research conclusions, this paper puts forward three inspirations and suggestions. First, digital technology has promoted the increase of rural women’s labor supply. Therefore, we should continue to improve the degree of rural digitalization and further tap the rural female labor market, to alleviate the problems caused by aging, such as labor shortage, marginalization of women, and employment difficulties. Specifically, on the one hand, we should promote the development of digital technology, introduce smart homes, and make the workplace more flexible, so that women can take care of family and work at the same time. On the other hand, for low-skilled rural female workers, it is necessary to increase technical training and adopt dialect teaching to lower the learning threshold. The second is to actively develop tourism. Through rural tourism, more service jobs will be created for local women, such as tour guides, catering, and accommodation management, which will solve the employment difficulties of rural women, increase the income of local people, and promote rural revitalization. Through the Internet, platformization,

increasing the compensation for the return of talents, and other methods, we can solve the increased work content due to digital empowerment, such as the need to deal with online booking and publicity through social media. Third, the results of heterogeneity analysis mean that we should pay special attention to the development of digital technology in the central region, and at the same time vigorously carry out infrastructure construction in the western region. Through cultural reconstruction, we should transform intangible skills such as embroidery and batik into digital collections, develop tourism under the premise of protecting the ecology, and solve the multiple constraints in the western region, thus improving the labor supply of women in that region.

## References

- [1] Sheng Wang, Yu Na, Fu Rui. Digital Rural Construction: Mechanism, Realistic Challenges, and Implementation Strategies [J]. *Reform*, 2021, (04):45-59.
- [2] Zhang Yunping, Luan Jing. Digital Economy Empowering Rural Revitalization: Theoretical Mechanism, Constraints and Advancing Path [J]. *Reform*, 2022, (05):79-89.
- [3] Zeng Yiwu, Song Yixiang, Lin Xiazhen, et al. On some issues of digital rural construction in China [J]. *China Rural Economy*, 2021, (04): 21-35.
- [4] Sun Yujie, Yang Wei 'an. The logical context, predicament, and path relief of digital technology enable the high-quality development of rural education [J]. *Educational Theory and Practice*, 2025,45(07):28-33.
- [5] Ding Bo. Digital governance: a new model of village governance in digital villages [J]. *Journal of Northwest A&F University (Social Science Edition)*, 2022,22 (02): 9-15.
- [6] Deng Yue, Xu Hongkai. Can the digital economy promote farmers' long-term income increase? -Based on the elaboration of economic resilience [J]. *Rural Economy*, 2025, (01):131-142.
- [7] Wang Shangao. Research on the influence of digital rural construction on the common prosperity of rural residents [J]. *Statistics and Decision*, 2025,41 (02): 48-53.
- [8] Yang Huahua, Lu Qi, Zhang Jieyu. Digital rural construction empowers rural residents to upgrade their consumption: A mechanism test based on income and circulation efficiency [J]. *Commercial Economic Research*, 2025, (03):39-44.
- [9] Lin Longfei, Zhu Zhongkun. Can the digital economy alleviate the tight labor supply in China? [J]. *Economic System Reform*, 2024, (01):15-23.
- [10] Qi Yudong, Ding Shulei, Liu Cuihua. Internet use and labor supply of flexible employees under the background of the digital economy: theory and demonstration [J]. *Contemporary Finance and Economics*, 2021, (05): 3-16.
- [11] Shi Yupeng, Wang Yuyue, Li Xinrong. Liberated women: online education and women's labor participation [J]. *Sankei Review*, 2024,15 (06): 130-146.
- [12] Yang Weiguo, Wang Qi. Participation groups in digital platform work: labor supply and influencing factors-evidence of car drivers based on U platform network [J]. *Population Research*, 2018, 42(04):78-90.