The Impact of Internet Use on Household Per Capita Income - An Empirical Analysis Based on 2022CFPS Data

Wendi Liu, Gongjing Gao*

School of Political Science and Law, University of Jinan,
Jinan City, Shandong Province, China

lwd_determined@163.com, *sl_gaogj@ujn.edu.cn

Abstract. Based on 2022 CFPS data, this paper uses OLS and quantile regression models to systematically analyze the impact of Internet use on household per capita income. The study found that Internet use can significantly increase the per capita income of households, and it is ' pro-poor '. The marginal effect on low-income groups is 50.3 % higher than that of rich groups. In addition, there is structural differentiation in educational returns, and the top premium of higher education coexists with the bottom pulling effect of middle and high education. The research provides a precise empirical target for optimizing digital inclusion policies and breaking down institutional barriers.

Keywords: Internet use; family per capita income; quantile regression

1 Introduction

With the rapid development of information technology, the Internet has become an important infrastructure for modern economic activities. As of December 2024, China 's Internet penetration rate has reached 78.6 %, and the number of Internet users has reached 1.108 billion. The popularity of the Internet has not only changed people 's way of life, but also profoundly affected the efficiency and mode of economic operation. In recent years, the impact of Internet use on economic growth, social equity and family economic conditions has become a research hotspot. The per capita income of a family is the core indicator to measure the economic status of a family and an important symbol to reflect the level of social and economic development. Internet use may affect household income in many ways, and the impact of Internet use on household income is not a single linear relationship. Its role may be regulated by many factors, such as family characteristics, regional differences and the depth of Internet use.

Existing studies generally recognize the role of Internet technology in promoting macro income growth, but there are still theoretical controversies about its net effect on micro household income. On the one hand, digital technology may increase household

B. Majoul et al. (eds.), *Proceedings of the 2025 11th International Conference on Humanities and Social Science Research (ICHSSR 2025)*, Advances in Social Science, Education and Humanities Research 945,
https://doi.org/10.2991/978-2-38476-440-2_73

income by reducing information search costs[1]and expanding non-agricultural employment channels[2]. On the other hand, the heterogeneity of digital skills acquisition may lead to the ' digital divide ' effect[3], which may lead to the widening income gap.

With regard to the impact of Internet use on household per capita income, existing research has formed two theoretical explanations. The first is the theoretical explanation of technology empowerment mechanism. Supportive research emphasizes that the Internet empowers the family economy through two paths: First, the expansion of information channels reduces market friction. In China 's rural areas, the construction of digital villages has brought new growth points to the rural economy and played an important role in promoting farmers ' income. Second, the value-added effect of human capital. A study conducted in Niger found that technology-assisted adult education programme (such as mobile phone modules) can sustain learning outcomes in the medium to long term[4].The second is the theoretical explanation of the technology gap effect. Opposing research points out that digital technology may exacerbate income differentiation. The theory of skill-biased technological progress holds that high-skilled groups are more likely to benefit from Internet access[5].China 's research shows that the ' Broadband China ' demonstration cities have significantly alleviated the relative income deprivation of individuals in their groups[6], reflecting the synergistic constraints and effects of infrastructure and human capital. This differentiation stems from two barriers : first, the digital skills gap, the lack of information screening and transformation capabilities of vulnerable groups[7]; second, the lack of complementary assets, poor families often face equipment shortages and credit constraints.

Although existing studies have extensively explored the relationship between Internet use and economic growth, there are still deficiencies in the existing literature at the household level. On the one hand, most of the existing research focuses on the economic growth effect at the macro level, while there are relatively few studies on the impact mechanism of household income at the micro level. On the other hand, existing studies have not reached a consensus on the causal relationship between Internet use and household income. Therefore, in-depth analysis of the impact mechanism of Internet use on household per capita income not only helps to understand the changes in household economic behavior in the digital economy era, but also provides a theoretical basis for formulating more accurate Internet popularization policies. The innovation of this study is reflected in the method of breaking through the average effect analysis, and using quantile regression to reveal the heterogeneous impact of Internet use in different intervals of income distribution.

2 Data, Variables and Models

The data in this paper are from China Family Panel Studies (CFPS). This paper selects the 2022 CFPS data, mainly using the variables in the adult questionnaire and the family questionnaire. After eliminating missing values and outliers, 18,622 valid observations are finally obtained. The samples are widely representative and meet the needs of empirical analysis.

The explained variable is the logarithm of household per capita income. The variable is derived from the 2022 CFPS household economy questionnaire section, which is obtained by logarithming the household per capita income data. The explanatory variable is Internet use, which is obtained according to the question of " whether the mobile device is online " in the questionnaire. Among them, the sample assigned to the Internet is 1, and the individual assigned to the Internet is 0, so as to generate a dummy variable to measure whether the respondents use the Internet.

To mitigate omitted variable bias, control variables include: gender (dummy, male=1); age (continuous) and squared term (nonlinear effects); household registration (dummy, non-agricultural=1); education level (primary school/below, junior high school, Senior high school/technical school/vocational high school, college, bachelor+); marital status (dummy, married=1); employment status (dummy, employed=1); region (1=eastern, 2=central, 3=western, 4=northeastern China); self-rated social status (1-10); and health status (self-rated, 1="very healthy" to 5="very unhealthy").

In order to explore the impact of Internet use on household per capita income, this paper establishes the following multiple linear regression model :

$$\ln(\mathrm{PCI})_i = \alpha + \beta_1 \mathrm{Internet}_i + \beta_2 X_i + \varepsilon_i \quad (1)$$

Among them : $\ln(\mathrm{PCI})_i$ represents the logarithm of household per capita income ; $\mathrm{Internet}_i$ is a dummy variable indicating whether the individual uses the internet (1=user, 0=non-user); X_i represents control variables including: gender, age and its quadratic term, household registration (hukou), education level, marital status, employment status, region, self-rated social status, and health status; α is a constant term ; ε_i is a random error term ; β_1 and β_2 are the coefficients to be estimated.

3 Empirical Analysis Results

Based on the CFPS data in 2022, this paper systematically examines the impact of Internet use on the logarithm of household per capita income by using ordinary least squares (OLS) and quantile regression. Table 1 reports the full sample OLS regression (Model 1) and the quantile regression results of 0.25, 0.5, and 0.75 quantiles (Model 2-4). All models passed the variance inflation factor test (VIF < 5), and the R2 value indicated that the model had good explanatory power (OLS model R2 = 0.227). The following analysis is carried out from the two dimensions of core variables and key control variables:

From the perspective of core explanatory variables, Internet use (internet) showed a significant positive impact at the 1 % statistical level in all four models, but its marginal effect showed significant quantile heterogeneity. Specifically, the regression coefficient of the 0.25 quantile is 0.275, which is higher than the 0.183 of the 0.75 quantile, indicating that the income-increasing effect of Internet use on low-income groups is more prominent. This result may be due to the expansion of employment opportunities for low-income families by digital technology, such as obtaining income through non-traditional channels such as e-commerce platforms and gig economy, while high-

income families have stable career paths and multiple sources of income. The marginal contribution of the Internet is relatively limited.

The level of education shows a significant gradient return characteristic. Based on the primary school and below groups, the income effect of junior high school, senior high school, college and undergraduate education is gradually enhanced. The return rate of undergraduate and above education in the OLS model is as high as 0.929, and it still maintains a significant effect of 0.864 at the 0.75 high income quantile. The quantile regression further reveals that the coefficient of middle and high education levels decreases with the increase of income quantile, indicating that education has a more prominent marginal improvement effect on low-income groups.

The linear term (OLS coefficient 0.024) and the quadratic term (-0.000) of age verified the inverted U-shaped curve of income over the life cycle, and the positive effect of age on low-income groups was stronger (0.25 quantile 0.033), or reflected the early return of experience accumulation in the manual labor industry. The gender income gap is concentrated in the low-income group (0.25 quantile male dominance coefficient 0.038).

The income effect of non-agricultural household registration is stable and significant (OLS coefficient is 0.438, quantile fluctuation is only ± 7.3 %), revealing the rigid constraint of urban-rural dual structure. The regional difference shows that the coefficients of the central and western regions and the northeast region are significantly negative, especially the western region has the strongest inhibitory effect in the high-income quantile (0.75 quantile-0.490). Married status is negatively correlated with income (OLS coefficient − 0.158), and its negative effect increases with income (0.75 quantile − 0.226), which may be related to the restriction of family responsibility on the career development of high-income groups. The positive effect of employment is more significant in the bottom group (0.25 quantile 0.045 vs 0.75 quantile 0.051). The role of health capital is verified. Self-rated health deterioration significantly reduces income (OLS coefficient-0.031), while the impact of subjective social status does not pass the significance test, suggesting that the theoretical correlation between subjective and objective indicators needs to be further explored.

The empirical analysis reveals that the promotion effect of Internet use on household income is statistically robust, but its intensity has significant heterogeneity with income distribution. Specifically, the marginal promotion effect on low-income families is the most prominent, and with the increase of household income quantile, the effect shows a systematic decreasing feature. This gradient change shows that the income-increasing effect of digital technology has a clear pro-poor tendency, and its marginal contribution at the bottom of the income distribution is 50.3 % higher than that at the top.

Table 1. Regression Results of Internet Usage on Household Per Capita Income

Variables	(1) OLS	(2) Q25	(3) Q50	(4) Q75
Internet Usage	0.257***	0.275***	0.206***	0.183***
Gender	0.030**	0.038**	0.018	0.022
Age	0.024***	0.033***	0.018***	0.016***
Age Squared	0.000***	0.000***	0.000***	-0.000***

Household Registration	0.438***	0.427***	0.412***	0.374***
(Ref: Primary) Junior High School	0.253***	0.247***	0.206***	0.162***
Senior high/Vocational school	0.370***	0.362***	0.296***	0.253***
College	0.612***	0.608***	0.572***	0.548***
Bachelor degree or higher	0.929***	0.867***	0.885***	0.864***
Married	0.158***	0.110***	0.172***	-0.226***
Employed	0.095***	0.045**	0.045**	0.051***
Self-rated Health	0.031***	-0.015**	0.020***	-0.024***
Self-rated Social Status	-0.003	0.003	-0.003	-0.002
(Ref: East) Central China	0.332***	0.285***	0.275***	-0.349***
Western China	0.474***	0.454***	0.439***	-0.490***
Northeastern China	0.250***	0.149***	0.226***	-0.345***
Constant	9.221***	8.485***	9.449***	10.130***
Observations	18,622	18,622	18,622	18,622
R-squared	0.227	-	-	-

Note:* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4 Conclusion

Based on the data of China Family Panel Studies (CFPS) in 2022, this paper systematically examines the impact of Internet use on household per capita income and its heterogeneity through OLS and quantile regression models. The study found that : First, Internet use has a significant role in promoting household income growth, and shows obvious ' pro-poor ' characteristics. The quantile regression results show that its marginal effect on low-income groups is 50.3 % higher than that of rich groups. Second, the return to education presents nonlinear heterogeneity. The ' premium effect ' of bachelor 's degree and above at the top of income distribution is significantly higher than that of basic education level, but the marginal improvement effect of secondary and higher education on low-income groups is more prominent.

This study has three implications for policy formulation : First, strengthen digital inclusive development. We will carry out targeted digital skills training for low-income groups, expand informal employment channels through the ' Internet + employment ' model, and increase investment in rural network infrastructure to reduce access costs. Second, optimize the structure of human capital investment. While promoting the improvement of the quality of higher education, we should strengthen the digital transformation of vocational education, and cultivate high-end skilled talents to adapt to the digital economy through the integration of production and education. Third, deepen institutional reform. Accelerate the linkage reform of the household registration system and public service supply, and focus on supporting the digital industry layout and human capital accumulation in the western region.

Acknowledgment

This work was supported by the Research Project on the Improvement of the Life long Vocational Skills Training Policy System in the Digital Era. (Number: 24ASH016).

References

1. Pham V K, Pham Thi T D, Duong N T. A study on information search behavior using AI-Powered engines: evidence from chatbots on online shopping platforms[J]. Sage Open, 2024, 14(4): 21582440241300007.
2. Zhan Y, Gao D, Feng M, et al. Digital finance, non-agricultural employment, and the income-increasing effect on rural households[J]. International Review of Financial Analysis, 2025, 98: 103897.
3. Lythreatis S, Singh S K, El-Kassar A N. The digital divide: A review and future research agenda[J]. Technological Forecasting and Social Change, 2022, 175: 121359.
4. Aker J C, Ksoll C. Can ABC lead to sustained 123? The medium-term effects of a technology-enhanced adult education program[J]. Economic Development and Cultural Change, 2020, 68(3): 1081-1102.
5. Atasoy H, Banker R D, Pavlou P A. Information technology skills and labor market outcomes for workers[J]. Information Systems Research, 2021, 32(2): 437-461.
6. Tang Lizhi, Chen Mingcong, Ma Qingshan. Network infrastructure construction and residents ' income gap - empirical evidence from the ' Broadband China ' strategy [J].Economic Review, 2025, (01) : 117-132.
7. Van Deursen A J A M, Van Dijk J A G M. The first-level digital divide shifts from inequalities in physical access to inequalities in material access[J]. New media & society, 2019, 21(2): 354-375.