The Impact of Network Application on Income Gap in Digital Economy Era —— Based on Gender Differences

Tongyang Wang

School of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu 233000, China.

Abstract: Based on the data of China Household Tracking Survey in 2020, this paper studied the change of female salary and the decomposition of salary difference caused by Internet application, as well as the heterogeneity and internal influence mechanism of population Internet application on the change of female salary gap at different stages. The study found that Internet use has a significant positive impact on the overall wage level and significantly reduces the gender wage gap. With the increase of income quantile, the average gap of per capita income between men and women has gradually narrowed, and the problem of income imbalance between men and women in low-income groups has become more and more serious.

Keywords: Network Application; Wage Differentials; RIF Regression Decomposition; Digital Economy

1. Introduction

With the rapid development of new generation information technologies such as the Internet of Things, cloud computing, big data and artificial intelligence, digital economy has become an important engine driving China's economic growth[1]. According to the 50th Statistical Report on the Development of Internet in China released by CNNIC[2], the proportion of female Internet users in China increased from 12.3% in October 1997 to 48.3% in June 2022. The Internet is profoundly changing the way of social participation of both sexes, providing more opportunities for the self-development and value realization of individuals of different genders. As the proportion of female Internet users in China continues to rise, how will online apps affect the gender wage gap?

Based on this, this paper uses the data of China Household Tracking Survey in 2020 and uses unconditional quantile regression to study the effect of Internet application on male and female salary and decompose it. Compared with the existing research, the possible contributions of this paper mainly include the following aspects: (1) In terms of theoretical analysis, the existing data mainly study the impact of Internet application on urban-rural income gap, but this paper, starting from the perspective of gender difference, puts forward a new point of view for understanding the internal mechanism of income imbalance in the digital economy society. (2) In terms of empirical research, most of the current data decompose the gender income gap from the average level. In this paper, unconditional quantile regression decomposition is adopted to reflect the distribution of the income gap at each quantile.

2. Empirical research design

2.1 Model setting and research methods

Based on the classical Minser income equation, this paper introduces virtual variables such as network application and region, and builds the benchmark regression model as follows:

$$\ln(Wage_i) = \beta_0 + \beta_1 Internet_i + \sum_j \beta_i X_{ij} + \varepsilon_i$$

Where Wagei represents the hourly wage of the worker, Interneti represents whether the worker uses the Internet, and X is the control variable, including the worker's gender, household registration, years of education, health status, work experience, regional characteristics, etc.

On this basis, in order to study the impact of distribution changes of explanatory variables on outcome variables, this paper adopts Recentered Influence Function(RIF) to analyze income.This method was proposed by Firpo et al.[3]. RIF uses the influence function in robust estimation to construct an unconditional quantile regression function, and its prominent role is to decompose the total effect of income gap into each feature vector.

First, the income distribution is recorded as distribution statistic g_t=g(F_t), and g is the relevant statistic to describe distribution F. When the statistic is quantile, RIF regression is unconditional quantile regression. t represents the group that compares the income gap, $F_t(\cdot)$ represents the distribution function defined by the eigenvector on the basis group, and g(F_t) is the statistic defined on F_t .After that, construct the RIF:

$$RIF(\ln y_t; g) = g(F_t) + IF(\ln y_t; g)$$

Where $IF(\ln y_t; g)$ is the influence function of $\ln y_t$:

$$IF(\ln y_t; g) = \frac{\tau - (\ln y_t \le g)}{f_{\ln y_t(g)}}$$

Therefore, the income gap can be broken down into the following forms:

$$RIF(\ln y_{t_1}; g) - RIF(\ln y_{t_2}; g) = (\overline{X}_{t_1} - \overline{X}_{t_2})\beta_{t_1}^g + (\beta_{t_1}^g - \beta_{t_2}^g)\overline{X}_{t_2}$$

In the above formula, $(\overline{X}_{t_1} - \overline{X}_{t_2})\beta_{t_1}^g$ represents the characteristic effect of income gap at the G-quintile, and $(\beta_{t_1}{}^g - \beta_{t_2}{}^g)\overline{X}_{t_2}$ represents the coefficient effect. Moreover, the total characteristic effect can be further decomposed into the effects of several specific explanatory variables:

$$(\overline{X}_{t_1} - \overline{X}_{t_2})\beta_{t_1}^g = \sum_k (\overline{X}_{kt_1} - \overline{X}_{kt_2})\beta_{kt_1}^g$$

Similarly, the coefficient effect of income difference can also be decomposed into each specific characteristic variable.

The decomposition method of unconditional quantile regression based on RIF can reflect the influence of each quantile on the income difference. It can divide the income difference into the explainable part caused by the characteristic effect and the unexplainable part caused by the characteristic return difference, so as to estimate the effect of each explanatory factor on the coefficient of the characteristic effect. Thus, we can analyze the influence of various explanatory factors on the income difference more deeply.

2.2 Variable description

The difference in working hours is an important factor affecting the gender wage difference[4]. In order to exclude the influence of working hours, the logarithm of hourly income was used as the explanatory variable in this paper. Whether the core explanatory variable uses the Internet as a dummy variable, according to the question in the questionnaire: "Do you use mobile devices, such as mobile phones, tablets, to access the Internet?" If the answer to one of the questions is "Yes", the value is assigned to 1, and if the answer to both questions is "no", the value is assigned to zero. At the same time, considering other factors that affect the income level of workers, the control variables selected in this paper mainly include: (1) gender. Male is set as the reference group, and the value is 0, while female is 1. (2) Years of education, according to the length of the academic year assigned value, illiterate/semi-illiterate assigned value 0, primary school assigned value 6, junior high school assigned value 9, high school/secondary school/vocational high assigned value 12, junior college assigned value 15, university undergraduate assigned value 16, master's assigned value 19, doctor assigned value 22. (3) Work experience is represented by the year of "which month and year was the last time I left school" in the questionnaire, using 2020 minus the year of leaving school. (4) Urban and rural classification, this paper assigns a value of 1 to urban workers, and 0 to others. (5) Health status, there are five answers to the question "How do you think your health status is?" : unhealthy, average, relatively healthy, very healthy and very healthy, which are assigned integers 1 to 5 in order. (6) Regional characteristics: According to the regional classification of national cities by the National Bureau of Statistics of China, Northeast China is classified into the eastern region, which includes three regions: east, middle and west. At the same time, the western region is taken as the reference group, and two regional dummy variables are introduced in the eastern region and the central region.

2.3 Data sources and descriptive statistics

The data on the individual level of the labor force comes from the data of the 2020 China Household Tracking Survey, and other macro data are from the China Statistical Yearbook. After removing the samples with missing data, the descriptive statistical results are shown in Table 1.

Table 1 Descriptive statistical analysis of variables

Variables	Overall		Male		Female	
	Mean	St.d	Mean	St.d	Mean	St.d
Logarithmic hourly income	2.592	1.084	2.723	1.024	2.405	1.139
Network application	0.816	0.387	0.801	0.399	0.838	0.369
Gender	0.412	0.492	——	——	——	——
Educational level	10.309	4.450	10.160	4.219	10.522	4.753
Work experience	29.492	14.526	30.770	14.622	27.668	14.191
Urban	0.602	0.489	0.577	0.494	0.638	0.481
Health	3.274	1.064	3.323	1.078	3.203	1.040
East	0.501	0.500	0.491	0.500	0.515	0.500
Middle	0.240	0.427	0.238	0.426	0.242	0.428
West	0.260	0.438	0.271	0.445	0.242	0.429

3. Results and analysis

3.1 Quantile effect of Internet application on gender wage

Control variables such as years of education, working experience, urban and rural classification, and health status were combined and placed on the three typical wage subpoints of 0.25, 0.50, and 0.75, which were mainly considered. The regression results are shown in Table 2. The results show that network application has a more obvious effect on the income improvement of low-income groups, and from the perspective of gender, it also has a more obvious effect on reducing the gender income gap of this group. It can be seen that the vigorous development of the digital economy has produced many new forms of employment, most of which have a low threshold, while the work forms and work terms of online services are more flexible and diverse, creating more job opportunities for some people with low and medium education, and postpartum female workers.

Table 2 RIF regression results of the impact of Internet application on gender wage

Variables	Overall			Male			Female		
	0.25	0.5	0.75	0.25	0.5	0.75	0.25	0.5	0.75
Network application	0.295***	0.178***	0.156***	0.189***	0.110***	0.115***	0.356**	0.226*	0.148**
Control variable					Controlled				
Observed value		9409			5532			3877	

3.2 Quantile breakdown of gender wage gap

Wage differences are further decomposed, as shown in Table 6. The results show that: (1) With the increase of quantile, the total difference decreases gradually, indicating that the gender wage difference of low-income groups is more obvious than that of high-income groups. (2) The total difference values of characteristics at each sub-site were significantly negative, indicating that the endowment characteristics of individuals significantly reduced the total gender wage difference. (3) Among the characteristic effects, network application is significantly negative on each quantile, indicating that it is conducive to narrowing the gender income gap, especially for low-income groups. Hypothesis H2 is verified. (4) In the characteristic effect, the years of schooling are significantly negative on each quantile, indicating that it is conducive to narrowing the gender income gap. Different from the network application, the improvement of education level is more effective in narrowing the gender income gap of high-income groups. (5) Among the characteristic effects, the coefficient in the eastern region is significantly negative, while the coefficient in the central region is negative but not significant, indicating that the gender income gap in the eastern region is smaller than that in the western region.

Table 3 Decomposition results based on RIF unconditional quantile regression

Variables	0.25		0.5		0.75	
	Decomposition result	Specific gravity	Decomposition result	Specific gravity	Decomposition result	Specific gravity
Total variance	0.3696***	100.00%	0.3297***	100.00%	0.1964***	100.00%
Characteristic effect	-0.0466***	-12.61%	-0.0412***	-12.50%	-0.0316***	-16.09%
Coefficient effect	0.4162***	112.61%	0.3709***	112.50%	0.2281***	116.14%
（Characteristic effect）						
Network application	-0.0129***	-3.49%	-0.0082***	-2.49%	-0.0053***	-2.70%
Gender	-0.0226***	-6.11%	-0.0234***	-7.10%	-0.0336***	-17.11%
Educational level	0.0033	0.89%	0.0037	1.12%	0.0214***	10.90%
Work experience	-0.0119***	-3.22%	-0.0102***	-3.09%	-0.0104***	-5.30%
Urban	0.0048**	-1.30%	0.0027*	0.82%	0.0018	0.92%
Health	-0.006**	-1.62%	-0.0052**	-1.58%	-0.0051**	-2.60%
East	-0.0009	0.24%	-0.0006	-0.18%	-0.0003	-0.15%
constant	0.7787***	210.69%	0.6939***	210.46%	0.4322**	220.06%

4. Conclusion and suggestion

The findings of this study are as follows: (1) Network application has a significant positive impact on labor income level, and its characteristic effect sign is negative, indicating that network application significantly reduces the gender income gap. (2) With the increase of quantile, the total gender income difference gradually decreases, and the gender income inequality of low-income groups becomes more serious. In order to give full play to the dividends of the Internet, further narrow the gender income gap, and realize the development goal of the Internet to help common prosperity, this paper puts forward the following suggestions: (1) Encourage the use of employment opportunities generated by the development of the digital economy, improve network service skills, and promote network office, so as to reduce the employment discrimination of women in the labor market. (2) Encourage various new forms of employment under the digital economy, such as we-media creation, network broadcast and other new employment channels that are more in line with women's "soft skills", so as to provide opportunities for low-income female workers to realize job transformation.

References

[1] Xu XC, Zhang MH. Research on the scale measurement of China's digital economy: from the perspective of international comparison [J]. China Industrial Economy, 2020, (05): 23-41.

[2] China Internet Network Information Center: The 50th Statistical Report on Internet Development in China.

[3] Firpo S, Fortin NM, Lemieux T. Unconditional Quantile Regressions[J]. Econometrica, 2009, 77(3): 953-973.

[4] Mandel H., and Semyonov M. Gender Pay Gap and Employment Sector of Earnings Disparities in the United States, 1970-2010. Demography, 51(5):1597-1618.