Does housework affect personal employment participation with different education levels? ——Mediation effect test of econometric models based on Stata

Yang Shi

Email:fighting8306@126.com

Department of Accounting, Shandong Youth University of Political Science, Jinan, 250103, China

Abstract: This paper selects relevant data from the CGSS database, uses Stata statistical software and micro econometric models, and completes data analysis through logistic regression and multiple regression analysis, while controlling for variables such as education levels. It tests the impact of housework on personal employment participation through the mediating variable of personal income. The results show that, housework has a significantly negative impact on personal income; housework has a negative impact on personal employment participation; the mediation effect of personal income exists. Although the sample size of this paper has some limitations, the establishment of the mediation effect is of great practical significance for recognizing the value of housework and promoting a more reasonable family division of labor.

Keywords: Housework; Employment participation; Computer software； Econometric model; Mediation effect test; Big data regression analysis

1 Introduction

Family economics believes that the family is a small production unit, and the division of housework also needs to follow the principle of maximizing overall interests, that is, individual rational choice is based on the theory of comparative advantage[1]. Therefore, from the perspective of gender comparative advantage, the division of housework in most families is dominated by women, supplemented by men, that is, “men take the lead outside, women take the lead inside”. This is also in line with the inherent physiological and personality traits of both sexes[2]. People obtain income by putting in working time, namely employment participation, to meet their physical and psychological needs. But people’s life is limited, so we need to allocate time between various activities, including work and housework, that is, housework will crowd out work time to some extent[3]. Then, does housework affect personal employment participation by affecting personal income?

Considering that personal income may be affected by the original accumulation of family wealth, we choose an appropriate database (such as CGSS, CFPS and other large social survey databases), use personal income as a mediating variable, use Stata and other computer software, establish econometric models to study the impact of housework on personal employment participation through Logistic regression and multiple regression analysis, and verify the existence

ICEKIM 2023, May 26-28, Nanjing, People's Republic of China

DOI 10.4108/eai.26-5-2023.2337256

of mediating effect of personal income, and provide reference for individual employment decision-making.

2 Variable selection and model building

2.1 Data sources

This paper selects the data of Chinese General Social Survey (CGSS) in 2015, and the total number of samples of CGSS2015 is 10968.

2.2 Variable selection

2.2.1 Housework

This is the main explanatory variable, which is related to Part C of the questionnaire. Then eliminate samples with missing data and calculated values of zero, and according to the questionnaire, the time period of housework is selected for division, the final statistics are as follows(Table 1).

Table 1. Descriptive statistics of housework

Time Groups	Number of Valid Samples	Gender	Number of Valid Samples by Gender	Proportion	Average Housework (AHW，hour)
Weekdays	1390	Male(M)	553	39.78%	1.69
		Female(F)	837	60.22%	2.86
Rest Days	1439	Male(M)	598	41.56%	2.00
		Female(F)	841	58.44%	3.31

Data source: Obtained from CGSS2015 database C241, C242, C261 and C262.

2.2.2 Personal income

It's the mediating variable. In order to ensure that the relationship between housework and personal income is not affected by the original wealth stock of the household, the survey results corresponding to the question of "your personal occupational/labor income for the whole year last year" are selected, as shown in Table 2.

Table 2. Descriptive statistics of personal income

Time Groups	Number of Valid Samples	Minimum(¥)	Maximum(¥)	Mean(¥)	Standard Error
Weekdays	861	160	1000000	35657.88	74604.272
Rest days	895	160	1000000	36391.32	77743.493

Data source: Obtained according to CGSS2015 database.

2.2.3 Employment participation

This is the explained variable. In order to simplify the research question, the samples of item 2 and item 3 in the corresponding question are eliminated, and item 1 and item 4 are retained as the basis for the assignment of employment participation. If item 1 "not engaged in any work

for the purpose of obtaining economic income" is selected, the personal employment participation is assigned a value of 0; If item 4 "engaged in work for the purpose of obtaining economic income (including joining the army)" is selected, the personal employment participation is assigned a value of 1, that is, the explained variable is a binary variable, as shown in the following Table 3.

Table 3. Descriptive statistics of employment participation

Time Groups	Number of Valid Samples	Minimum	Maximum	Mean
Weekdays	527	0	1	0.47
Rest Days	547	0	1	0.52

Data source: Obtained according to CGSS2015 database.

2.2.4 Control variables

Age, gender, nationality, political status, marital status, ownership and educational background are selected as control variables. Due to space constraints, descriptive statistics of control variables are no longer described in detail.

2.3 Model building

Based on the research path set earlier and the characteristics of the selected variables, the econometric models used in this paper involve binary logistic regression and multiple linear regression. The above two types of regression analysis are commonly used models in econometrics. According to the research purpose, explanatory and explained variables are selected, and some other variables are controlled to construct models. With the help of commonly used computer software such as Stata or SPSS, the significance of model coefficients are tested to study the relationship between explanatory and explained variables, which have been widely applied in various application studies. Therefore, this paper selects the personal income of CGSS database as the mediating variable, constructs the above econometric models, and uses Stata software to test the mechanism of housework's impact on employment participation through personal income by using the mediation effect analysis method. Based on the approach of Wen Zhonglin and Ye Baojuan (2014) [4], the following econometric models were constructed and validated through multiple regression analysis:

$$\boldsymbol{Work}=\boldsymbol{\alpha_0}+\boldsymbol{\beta_0 Housework}+\boldsymbol{\gamma_0 X_0}+\boldsymbol{\mu_0} \quad (1)$$

$$\boldsymbol{Y}=\boldsymbol{\alpha_1}+\boldsymbol{\beta_1 Housework}+\boldsymbol{\gamma_1 X_1}+\boldsymbol{\mu_1} \quad (2)$$

$$\boldsymbol{Work}=\boldsymbol{\alpha_2}+\boldsymbol{\beta_2 Housework}+\boldsymbol{\delta Y}+\boldsymbol{\gamma_2 X_2}+\boldsymbol{\mu_2} \quad (3)$$

Among them, $\boldsymbol{Housework}$ shows housework time, which is the main explanatory variable; The logarithmic form of personal income is used as the mediating variable $\boldsymbol{Y}$ after smoothing; $\boldsymbol{Work}$ represents the explained variable of employment participation; $\boldsymbol{X_i}$ represents the control variables,including gender, nationality, etc. $\boldsymbol{\beta_i, \gamma_i, \delta}$ are parameters to be estimated, $\boldsymbol{\mu_i}$ is random error term, and $\boldsymbol{\alpha_i}$ is a constant term.

According to the relevant research by Wen Zhonglin and Ye Baojuan (2014) [4], the standards for judging the mediation effect of the above model are: the explained variable $\boldsymbol{Work}$ has a regression on the explanatory variable $\boldsymbol{Housework}$, the coefficient $\boldsymbol{\beta_0}$significant; the medi-

ating variable $\boldsymbol{Y}$ has a regression on the explanatory variable $\boldsymbol{Housework}$, the explained variable $\boldsymbol{Work}$ has a regression on the mediating variable $\boldsymbol{Y}$ and the explanatory variable $\boldsymbol{Housework}$, if the coefficients $\boldsymbol{\beta_1}$ and $\boldsymbol{\delta}$ are both significant, it means that there is a mediating effect.

3 Methods: econometric test of the mediation effect

This paper uses Stata to test the mediation effect.Stata is a set of computer statistics software that relies on big data to analyze data, manage data, and draw professional charts for users. Stata is powerful and a very convenient computer software.

3.1 Test of the effect of housework on employment participation

With the help of Stata computer software, this paper applies binary logistic regression to analyze the selected samples and tests the significance of the regression coefficients in model (1). The regression results are as follows Table 4.

Table 4. Logistic regression results of model(1)

Housework on Weekdays(N=527)		Housework on Rest Days(N=547)		Total Sample(N=566)	
Variable	Coefficient	Variable	Coefficient	Variable	Coefficient
Constant	1.531* (2.739)	Constant	1.311 (2.076)	Constant	1.477 (2.569)
Housework	-0.247*** (12.366)	Housework	-0.167*** (7.111)	Housework	-0.115*** (10.473)
Age	0.009(0.292)	Age	0.010 (0.365)	Age	0.010 (0.382)
Gender (Control Women)	0.615** (4.425)	Gender (Control Women)	0.643** (4.837)	Gender (Control Women)	0.602** (4.206)
Nationality (Control Non Han)	-0.223 (0.190)	Nationality (Control Non Han)	-0.132 (0.070)	Nationality (Control Non Han)	-0.198 (0.151)
Political Status(Control Non Party Members)	-0.259 (0.170)	Political Status(Control Non Party Members)	-0.342 (0.306)	Political Status(Control Non Party Members)	-0.293 (0.220)
		Education level(Control Primary)			
Junior High School	0.620* (3.766)	Junior High School	0.630** (3.967)	Junior High School	0.631** (3.919)
High School	0.992** (5.256)	High School	1.176*** (7.428)	High School	1.091** (6.408)
Colledge	0.358 (0.651)	Colledge	0.514 (1.379)	Colledge	0.439 (0.994)
Undergraduate	18.731 (0.000)	Undergraduate	19.015 (0.000)	Undergraduate	18.869 (0.000)
Ownership (Control Non-State Owned Enterprises)	2.691*** (6.781)	Ownership (Control Non-State Owned Enterprises)	2.768*** (7.196)	Ownership (Control Non-State Owned Enterprises)	2.719*** (6.936)
Marital Status	-0.314	Marital Status	-0.408	Marital Status	-0.363

(Control Unmarried)	(0.565)	(Control Unmarried)	(0.947)	(Control Unmarried)	(0.748)
Hosmer and Lemeshow Test	0.764	Hosmer and Lemeshow Test	0.543	Hosmer and Lemeshow Test	0.271

Note1: The data in parentheses are wals values, ***, **, * indicate significance at the 1%, 5%, and 10% confidence levels, respectively.

The overall sample is divided into the sample of housework on weekdays, housework on rest days and the total sample. In the three samples, the impact of housework on employment participation is significant at the level of 1%, and the coefficient is negative, indicating that housework has a negative effect on personal employment participation. For the control variables, the regression results of the three samples all show that males have a higher probability of employment participation; If you work in the traditional sense of "system", the probability of employment participation is higher; For different stages of schooling, the probability of employment participation is higher for the junior high school and high school groups. However, the remaining control variables are not significant.

3.2 Test of the effect of housework on income

Apply the multiple OLS regression analysis in Stata to test the significance of the regression coefficients in model (2). The regression results are as follows Table 5.

Table 5. Multiple OLS regression results of model(2)

Housework on Weekdays(N=527)		Housework on Rest Days (N=547)		Total Sample(N=566)	
Variable	Coefficient	Variable	Coefficient	Variable	Coefficient
Constant	8.898*** (34.109)	Constant	8.870*** (34.035)	Constant	8.906*** (34.064)
Housework	-0.069*** (-2.732)	Housework	-0.052** (-2.374)	Housework	-0.035*** (-2.752)
Age	-0.007 (-1.532)	Age	-0.007 (-1.506)	Age	-0.006 (-1.468)
Gender (Control Women)	0.464*** (5.844)	Gender (Control Women)	0.466*** (5.835)	Gender (Control Women)	0.456*** (5.706)
Nationality (Control Non Han)	0.512*** (3.257)	Nationality (Control Non Han)	0.520** (3.308)	Nationality (Control Non Han)	0.509*** (3.235)
Political Status(Control Non Party Members)	0.067 (0.462)	Political Status(Control Non Party Members)	0.083 (0.565)	Political Status(Control Non Party Members)	0.075 (0.515)
Education Level(Control Primary School)					
Junior High School	0.551*** (5.147)	Junior High School	0.564*** (5.248)	Junior High School	0.559*** (5.220)
High School	0.811*** (6.695)	High School	0.849*** (7.038)	High School	0.830*** (6.888)
Colledge	1.351*** (9.994)	Colledge	1.389*** (10.325)	Colledge	1.369*** (10.172)
Undergraduate	1.596*** (5.425)	Undergraduate	1.653*** (5.628)	Undergraduate	1.622*** (5.527)

Ownership(Control Non-State Owned Enterprises)	0.068 (0.639)	Ownership(Control Non-State Owned Enterprises)	0.056 (0.530)	Ownership(Control Non-State Owned Enterprises)	0.064 (0.608)
Marital Status (Control Unmarried)	0.296*** (2.843)	Marital Status (Control Unmarried)	0.282*** (2.710)	Marital Status (Control Unmarried)	0.288*** (2.774)
R^2	0.340	R^2	0.338	R^2	0.340
Adjust R^2	0.327	Adjust R^2	0.324	Adjust R^2	0.327

Note: The data in parentheses are t values, ***, **, * indicate significance at the 1%, 5%, and 10% confidence levels, respectively.

The regression results show that, no matter for the sample of housework on weekdays, the sample of housework on rest days or the total sample, under the influence of controlling other variables, housework has a significant impact on personal income (that is, $\boldsymbol{\beta_0}$ is significant), and the regression coefficients are all negative, that is, the impact of housework on personal income is negative, and housework can significantly negatively affect personal income.

For the control variables, the results show higher personal income for male and Han samples. However, the impact of educational background on personal income is basically significantly positive. In addition, the regression coefficient of marital status of the sample is significantly positive. The influence of gender, political status and ownership on personal income in the selected sample is not significant.

3.3 The test of the effect of housework and personal income on personal employment participation

Similarly, using Stata binary logistic regression to test the significance of the regression coefficients in model (3), the results are shown in Table 6.

Table 6. Logistic regression results of model(3)

Housework on Weekdays(N=527)		Housework on Rest Days(N=547)		Total Sample(N=566)	
Variable	Coefficient	Variable	Coefficient	Variable	Coefficient
Constant	-2.137 (1.904)	Constant	-2.488 (2.648)	Constant	-2.236 (2.098)
Housework	-0.227*** (10.314)	Housework	-0.152** (5.635)	Housework	-0.106***(8.602)
Personal Income	0.417*** (8.805)	Personal Income	0.434*** (9.698)	Personal Income	0.423*** (9.086)
Age	0.013 (0.671)	Age	0.014 (0.804)	Age	0.014 (0.803)
Gender (Control Women)	0.403 (1.770)	Gender (Control Women)	0.420 (1.906)	Gender (Control Women)	0.388 (1.623)
Nationality (Control Non Han)	-0.508 (0.924)	Nationality (Control Non Han)	-0.444 (0.731)	Nationality (Control Non Han)	-0.493 (0.880)

Political Status(Control Non Party Members)	-0.201 (0.102)	Political Status(Control Non Party Members)	-0.267 (0.185)	Political Status (Control Non Party Members)	-0.227 (0.131)
Education Level(Control Primary School)					
Junior High School	0.421 (1.618)	Junior High School	0.417 (1.611)	Junior High School	0.426 (1.662)
High School	0.684 (2.334)	High School	0.837* (3.485)	High School	0.771* (2.975)
Colledge	-0.208 (0.183)	Colledge	-0.091 (0.035)	Colledge	-0.142 (0.086)
Undergraduate	18.031 (0.000)	Undergraduate	18.263 (0.000)	Undergraduate	18.153 (0.000)
Ownership(Control Non-State Owned Enterprises)	2.680*** (6.729)	Ownership(Control Non-State Owned Enterprises)	2.756*** (7.133)	Ownership(Control Non-State Owned Enterprises)	2.707*** (6.878)
Marital Status (Control Unmarried)	-0.440 (1.069)	Marital Status (Control Unmarried)	-0.538 (1.594)	Marital Status (Control Unmarried)	-0.490 (1.322)
Hosmer and Lemeshow Test	0.680	Hosmer and Lemeshow Test	0.008	Hosmer and Lemeshow Test	0.073

Note: The data in parentheses are wals values, ***, **, * indicate significance at the 1%, 5%, and 10% confidence levels, respectively.

The regression results in Table 6 show that no matter for the above samples, the effect of housework on employment participation is significantly negative, and the effect of personal income on employment participation is significantly positive.

For other control variables, age, gender, nationality, political status, educational background and marital status have no significant impact on employment participation.

To sum up, With the help of Stata computer software, the regression results of the above three models are as follows: the explained variable $\boldsymbol{Work}$ has a regression on the explanatory variable $\boldsymbol{Housework}$, the coefficient $\boldsymbol{\beta_0}$ is significant; the mediating variable $\boldsymbol{Y}$ has a regression on the explanatory variable $\boldsymbol{Housework}$, the coefficient $\boldsymbol{\beta_1}$ is significant; the explained variable $\boldsymbol{Work}$ is regressed on the mediating variable $\boldsymbol{Y}$ and the explanatory variable $\boldsymbol{Housework}$, and $\boldsymbol{\delta}$ is significant, so according to the sequential test method of the mediating effect, it can be seen that the mediating effect of the personal income exists on different education levels.

4 Conclusions and further discussion

Relying on big data and using Stata computer statistical software to complete the mediation effect test, the results of the above studies show that:

The mediation effect of personal income in the impact of housework on personal employment participation with different education levels exists;

Housework can affect personal employment participation by affecting income.

However, there are still some theoretical limitations in this conclusion:

Firstly, in terms of data screening of personal employment participation, Item 2 and Item 3 “paid leave, study, temporary work stoppage or seasonal closure, etc.” are eliminated; As a result, the total number of samples meeting the requirements is only 566, which is relatively small and will have a certain impact on the conclusion. There is some subjectivity in the screening of the above two items of employment, so the method of this paper is to eliminate these two items and only keep the option of clear employment or non-employment, which will inevitably lead to too small sample size and affect the persuasion of the conclusion to a certain extent.

Secondly, this paper selects personal income from database, which is the income that can only be obtained through work. However, this choice will also delete some samples, resulting in a small sample size.

Finally, in the regression results of model (3), some control variables are not significant, and there is a certain deviation from the reality. For example, marital status, relevant studies have shown that the traditional family division of labor and gender awareness will affect and distort marriage and labor performance in the family [5], resulting in a greater impact of housework on married women. Therefore, the marital status has an impact on the personal employment participation decision, and gender should also have an impact. The reason for this conclusion may be related to the small sample size.

To sum up, only for the current screening sample, although this paper has verified the mediating effect of personal income in the impact of housework on employment participation, there are still many shortcomings that need to be further improved based on the actual situation, which is undoubtedly of great practical significance for recognizing the value of housework and promoting a more reasonable family division of labor.

This research was supported by a project of social science research of Shandong Youth University of Political Science (Grant No.: SJYBXM202212)

Referance

[1] Weiling Wang. Research on the sexual division of labor [J]. Journal of Chongqing University(Social Science Edition),2016(05):135-143.

[2] Jie Xiao.The Influence of Housework on Gender Pay Gaps Based on the 3rd National Survey of the Status of Chinese Women [J]. Journal of Chinese Women’s Studies,2017(11):72-84.

[3] Jiamei Zheng,Shisong Qing. A Review of Literature on the Division of Household Labour and Gender-based Earning Disparities [J]. Journal of Chinese Women’s Studies,2014(01):107-114.

[4] Zhonglin Wen,Baojuan Ye. Analyses of Mediating Effects: The Development of Methods and Models[J]. Advances in Psychological Science,2014(05):731-745.

[5] Ji Xu, Yana Huang. Gender Identity，Marriage and Labor Behavior within Households [J]. Economic Research Journal,2018(04):136-150.