The Causal Effects of Wages on Labour Supply for Married Women– Evidence from American Couples

Bob Wen

June 16, 2019

Abstract

Using individual-level panel data from PSID, I consistently estimate the causal effects of own wages on interior labour supply (the hours-wage elasticities) for married women who were between 17 and 55 years old in 2005 and surveyed every two years till 2015.

First, I consider a representative married woman's utility maximisation choice subject to her budget constraint that connects her husband's wages and non-labour income to her labour supply decisions through the factor couple relationship.

Second, suggested by the optimal hours of work equation and comparative statics, I carry out the empirical analysis with the starting-point model - pooled OLS.

Third, I take into account the endogeneity problem due to sample selection and attrition. These issues could be alleviated by adding into the original models the predicted probability of being selected into the labour force and the predicted probability of staying in the surveys.

Fourth, I control for the unobserved individual heterogeneity that is time-invariant, such as the preference for work and the family tradition, by using panel fixed effects models.

Finally, the unobserved idiosyncratic characteristics that are time-varying, such as competency and skills, are taken into account by using instrumental variables (IV) and the method of two-stage least squares. This approach could be illustrated with a graph of the simultaneity of labour supply and labour demand curves.

I find that:

(1) The causal effects of wages on labour supply (the hours-wage elasticities) drop from 0.27 in the pooled OLS to 0.21 in the panel data fixed effects 2SLS model after I account for sample selection, attrition, individual heterogeneity, and unobserved idiosyncratic characteristics.

(2) Holding other factors constant, a 1% increase in married women's wages raises their hours of work by 0.21% on average.

(3) Part-time female workers are more responsive to wage changes than their full-time counterparts.

(4) There is evidence of backwards-bending labour supply curves.

1. Introduction

1.1. Motivation

To introduce the topic of the paper, let us take a look at a graph of a seemingly paradoxical relationship and trend over time.

Using the data from Current Population Survey (CPS), I find that the time series of country-level average weekly earnings of married women who are full-time workers (over 25 hours of work per week) has an upward trend over time from 2005 to 2015. However, their average hours of work do not show a distinct pattern over time. If the average weekly earnings are replaced by the relative to husband weekly earnings, the relation between hours of work and earnings is not clear either. (Fig.1) It seems that labour supply and wages are not related.

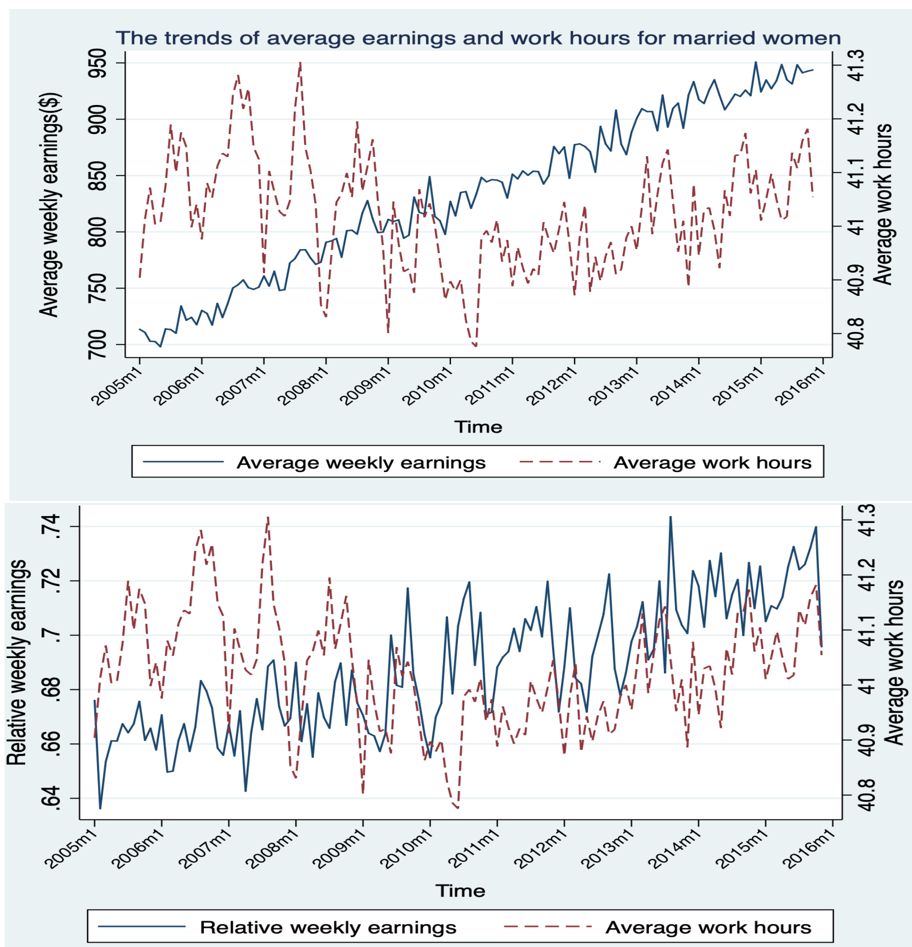

Fig. 1. The trends of the average weekly earnings and the relative weekly earnings compared to work hours for married women (from CPS)

As far as the labour force participation rate and the wages are concerned, the trends of them are opposite. (Fig.2)

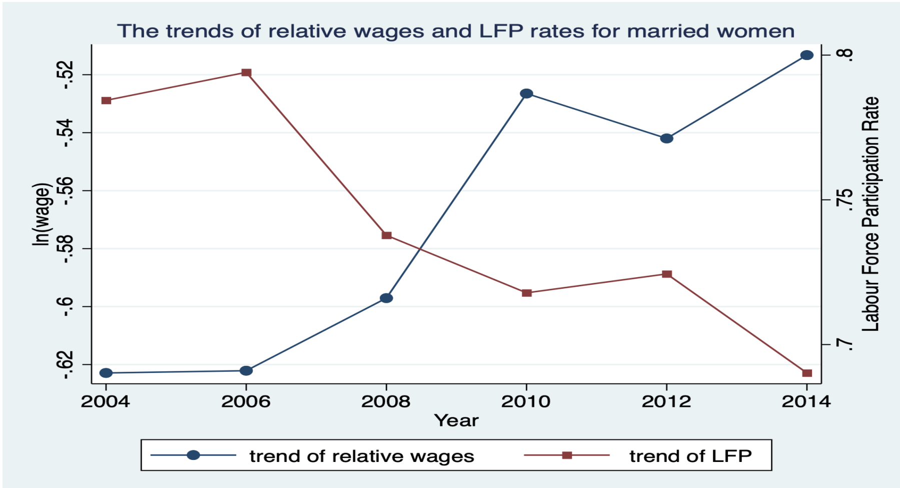

Fig. 2. The opposite trends of the relative weekly earnings compared to the labour force participation for married women (from PSID)

Does it mean that wages and labour supply are not correlated or the wages have a negative causal effect on the labour supply for married women? The answer is: ”No”. The reasons are as follows:

(1) The country-level time series data are the aggregate information that does not provide the variation within individuals over time and the variation between individuals as the panel data offer us. Besides, the interaction between couples within families could not be modelled without microdata. If we investigate at the individual level under panel structure, more detailed relation could be found.
(2) In order to obtain the causal effects of wages on labour supply, we should hold other relevant factors behind labour supply fixed and take into account bias due to endogeneity. Only in this way could we obtain the partial effects.

After modelling the interaction between couples, and alleviating self-selection, sample attrition, individual heterogeneity, and unobserved idiosyncratic characteristics bias, I found consistent, significant and positive causal effects of wages on labour supply.

1.2. Objectives of the Study

The above leads to the research questions of this paper:

(1) What are the causal effects of own wages on hours of work? In other words, how to consistently estimate the hours-wage labour supply elasticities?;

(2) Are the causal effects different between subgroups and changing over time?

Why is it important to study the relationship between labour supply and weekly labour income? The causal effects of labour income on labour supply are crucial to taxation and welfare policies making. The optimal levels of welfare depend on the estimates of the parameter of the response of labour supply to wage changes. In addition, it gives us a better understanding of the mechanism of married women's working decision that underlies the female labour supply.

1.3. Approaches and Outline

The identification challenges I am facing when estimating the causal effects of wages on labour supply include:

(1) Sample selection:
The first challenge arises when the sample used is not representative of the population of interest. The model is supposed to represent the causal effects of wages on hours of work for all married women of working age, whether or not she is working at the time of the survey. It is called the incidental truncation because hours of work and wages are missing as a result of the outcome of another variable, labour force participation. People self-selected into employment, so whether or not we observed wages and hours of work depends on an individual's labour supply decision. The OLS estimates could be biased because of this nonrandom nature of the sample.

(2) Sample attrition:
The second challenge is a unique problem for panel data. During the observation period, some respondents left the survey. If we lost them randomly, then it would not cause a problem. Otherwise, the estimates may be biased. Out of the 3618 married women under survey who were between 17 and 55 in the year 2005, 3423 took part in the survey in 2007, 3279 continued in 2009, 3076 stayed in the survey in 2011, 2,886 in 2013, and finally only 2675 respondents participated in the survey in 2015, which was about 74% of the respondents in 2005. That is to say, about 26% of the respondents dropped out of the survey during the observation period of 10 years, with an attrition rate of approximately 2.6% per year. It would not cause bias of labour supply elasticities if the attrition was random. However, if the leavers had different labour supply elasticities on average from the respondents who stayed in the survey, then the estimates would be biased based only on the stayers.

(3) Unobserved individual-specific characteristics:

For the third challenge, I divide the unobserved individual-specific characteristics that could be correlated with wages into two categories. The first category is the individual heterogeneity that is individual-specific and time-invariant, such as family tradition, preference for work, and habits. The other category is the individual idiosyncratic characteristics that are individual-specific and time-varying (different across individuals and changing over time), such as competency, skills, ability, and intelligence. For the first category, I use a fixed effects panel data model to control for individual heterogeneity; for the second category, I use instruments. Both methods are meant to alleviate the endogeneity issue.

The remainder of the paper is organized as follows. Section 2 gives a review of previous research. Section 3 presents the theoretical model, which is a guide for empirical analysis. The representative married woman chooses her consumption and dedication to her family to maximise her utility. Hypotheses about married women's labour supply are proposed. Section 4 deals with identification challenges due to endogeneity problems of sample selection, sample attrition, individual heterogeneity, and unobserved idiosyncratic characteristics. Section 5 applies the consistent estimator from section 4 to investigate how the hours-wage elasticities change over time and across different subgroups. In Section 6, the robust analysis is conducted. I compare results between the 2005 cohort and the 1995 cohort. The estimates from using PSID data are contrasted with the results using CPS data. Finally, Section 7 concludes. The brief structure of the paper is illustrated in Fig.3.

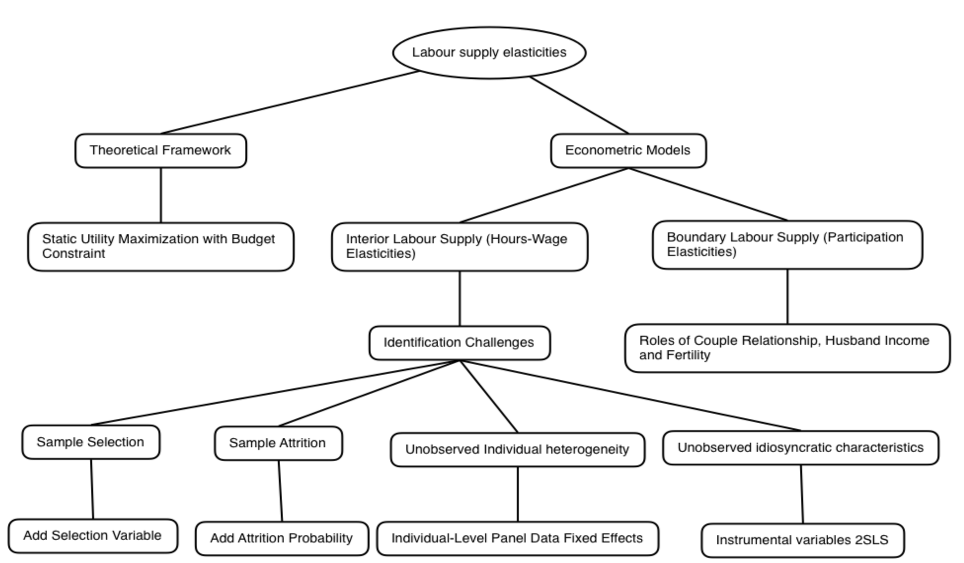

Fig. 3. The outline of the paper

2. Previous Research and Literature Review

3. Theoretical Framework

3.1. Static Constrained Utility Maximization

The representative married woman chooses her consumption and hours of dedication to her family to maximize her utility:

$$\underset{\{c,f\}}{Max}\ U = \alpha lnc + \beta\frac{f^{1-\gamma}}{1-\gamma}, 0 < \gamma < 1, \alpha > 0, \beta > 0$$
$$s.t.\ c = w(1-f) + \theta w^{H} + v$$

where

- c: composite consumption;
- f: dedication to her family; time spent on housework, child care, household production; assume h+f=1 where h is the time of work;
- α: relative importance of consumption;
- β: relative importance of dedication to family;
- w: wife's wage rate;
- w^H: husband's wage rate;
- v: non-labour income and family wealth;
- θ: the proportion of husband's wages that goes to the wife; measure of couple relationship.

I assume that the husband's labour income is in his full control at first. How much it goes to his wife depends on how close the couple's relationship is. The closer the couple's relationship is, the more proportion of the husband's wages his wife can spend. By contrast, the non-labour incomes received by the family from their relative, their friends and the government are assumed to be in the wife's complete control, and therefore it is not affected by the couple relationship. So does the family wealth, to which I assume the wife have full access.

The wife's total time is assumed to be divided into two parts: the time spent on her family and the time spent on her jobs. As a result, obtaining the optimal amount of time spent on housework, child care, and other household production is equivalent to obtaining the optimal amount of time spent on work.

As far as the functional form of the utility function is concerned, the pure Cobb-Douglas form of the utility function is fine but not as flexible as the half Cobb-Douglas and half constant relative risk aversion (half-CD-half-CRRA) form of the utility function employed here, because the half-CD-half-CRRA utility function allows the possibility of either a positive or a negative labour supply elasticity, i.e., the possibility of a backwards-bending labour supply curve.

3.2. *Interior Solution and Comparative Statics*

The FOCs give the optimal hours of work (h^*) equation:

$$\frac{\alpha}{\beta}w(1-h^*)^\gamma = wh^* + \theta w^H + v \tag{1}$$

The solution of the optimal hours of work is a function of the model parameters:
$h^* = h(w, w^H, v, \theta, \alpha, \beta, \gamma)$

Implicit differentiation of the optimal hours of work equation yields the following partial effects of each factor on married women's labour supply decision.

3.2.1. *The partial effects of own wages on labour supply (hours-wage elasticity) holding other factors constant*

$$\frac{\partial lnh^*}{\partial lnw} = [\frac{\alpha}{\beta}(1-h^*)^\gamma - h^*]/\{h^*[1+\frac{\alpha}{\beta}\gamma(1-h^*)^{\gamma-1}]\}$$

(1) The sign is uncertain. When h^* is low, the hours-wage elasticity is positive;
(2) When h^* is high, it could be negative, i.e., a backwards-bending labour supply curve is possible.

3.2.2. *The partial effects of couple relationship on labour supply holding other factors constant*

$$\frac{\partial lnh^*}{\partial \theta} = -w^H/\{wh^*[1+\frac{\alpha}{\beta}\gamma(1-h^*)^{\gamma-1}]\}$$

(1) It is negative. Closer couple relationship results in lower willingness to work or fewer hours of work;
(2) $\frac{\partial lnh^*}{\partial \theta} \to 0$ as $w >> w^H$ or $h^* \to 1$. Higher female relative wages or long work hours lead to more female economic independence and a less important role of couple relationship on female labour supply.

3.2.3. The partial effects of husband's wages on wife's labour supply (cross-wage elasticity) holding other factors constant

$$\frac{\partial lnh^*}{\partial lnw^H} = -\theta w^H/\{wh^*[1 + \frac{\alpha}{\beta}\gamma(1-h^*)^{\gamma-1}]\}$$

(1) It is negative. Higher husband's wages reduce his wife's work hours;

(2) $\frac{\partial lnh^*}{\partial lnw^H} \to 0$ as $w >> w^H$ or $h^* \to 1$ or $\theta \to 0$. The effects of husband's wages disappear when wife's wages dominate husband's labour income or when she works extremely long.

3.2.4. The partial effects of non-labour income and family wealth on married women's labour supply holding other factors constant

$$\frac{\partial lnh^*}{\partial v} = -v/\{wh^*[1 + \frac{\alpha}{\beta}\gamma(1-h^*)^{\gamma-1}]\}$$

(1) It is negative. More non-labour income and family wealth decrease the wife's willingness to work or her hours of work;

(2) $\frac{\partial lnh^*}{\partial v} \to 0$ as $w >> v$ or $h^* \to 1$. The effects of non-labour income and family wealth on labour supply disappear when wife's wages dominate non-labour income, or she works for very long hours.

3.2.5. The partial effects of the value of the wife's dedication to her family on her working decision

$$\frac{\partial lnh^*}{\partial \beta} = -\alpha(1-h^*)^{\gamma}/\{\beta^2 h^*[1 + \frac{\alpha}{\beta}\gamma(1-h^*)^{\gamma-1}]\}$$

(1) It is negative. A higher value placed on her dedication to family (e.g. more children or dependants) or higher relative importance of housework compared to her consumption lowers her probability of working.

(2) $\frac{\partial lnh^*}{\partial \beta} \to 0$ as $h^* \to 1$. The partial effects of the importance of the wife's family work on labour supply diminishes as she works extremely long.

3.3. *Boundary Solution and Reservation Wages*

The representative wife's hours of work supplied could be zero if the wages offered are lower than her reservation wages w^R.

$$h^* = 0 \; if \; w < w^R$$
$$where \; w^R = \beta(\theta w^H + v)/\alpha$$

A higher reservation wage implies a lower probability to work. The wife's reservation wage increases, or equivalently the wife's probability to work decreases when

(1) husband's wages w^H increase,
(2) non-labour income v increases,
(3) couple relation θ is getting better,
(4) the value of wife's dedication to family β is getting higher.

3.4. *Hypotheses on married women's labour supply*

Based on the theoretical model, I propose the following hypotheses on the labour supply of married women:

3.4.1. *Hypothesis I: The labour supply curve could be backwards-bending.*

Part-time female workers have a higher positive hours-wage elasticity than full-time female workers. As hours of work increase, the positive hours-wage elasticity decreases. It turns to negative for over-time workers. The corresponding labour supply curve could be backwards-bending.

3.4.2. *Hypothesis II: Couple relationship matters.*

(1) Closer couple relationship decreases the willingness to work for married women. The partial effects decrease as husband's labour income drops.
(2) The causal effects of the husband's labour income on the wife's labour supply depend on the couple relationship. The cross-wage elasticity approaches zero as couple relation gets worse and worse.

3.4.3. *Hypothesis III: Gender wage inequality plays an important role.*

(1) Obviously, a higher relative wife wage leads to stronger willingness to work for the wife, because both higher own wages and lower cross wages will increase the wife's optimal

choice of hours of work according to the model.

(2) The gender wage gap also influences the partial effects of couple relationship on female labour supply, and the partial effects of the husband's labour income on the wife labour supply. When the relative wife wages increase, the above two partial effects go towards zero.

4. Econometric Models and Estimation Methods

4.1. Data Description and Summary Statistics

4.1.1. Data source and structure

The data are taken from the Panel Study of Income Dynamics (PSID) main family data sets and contain observations on 3618 married women who were surveyed continuously every two years from 2005 to 2015.

I focus on the 2005 cohort: the family units with both wife and husband and the age of wife is between 17 and 55 in the year 2005. They were surveyed every two years until 2015 on their labour status, income information, demographic data, etc. There are six survey points, i.e., six panels of information. At the end of the observation period in the year 2015, they were between 27 and 65 years old. The choice of the age range is to ensure that the individuals in the sample were within the working age over the eleven years of observation.

The sample is unbalanced because some information is missing due to the fact that some respondents dropped out of the survey at a certain point of time and the fact that not all information of each individual was collected for each year.

The number of individuals is much larger than the number of time series or the measurements over time ($n >> T$). In the sample for the hours of work labour supply model, $n = 2996, T = 6$. In the sample for the working probability model, $n = 3618, T = 6$. The number of married women is higher in the second model due to the fact that only those married women who were working are included in the first model.

4.1.2. Summary statistics I: Within and between variation

The overall variation of the outcome variable, work hours, is derived from two sources: one is the between-individual variation, the other is the within-individual variation. The former is the variation of average work hours between individuals, and the latter is the variation of work hours over time for each individual.

The between-individual variance

$$\hat{\sigma}_b^2 = \frac{\sum_{i=1}^{n}(\overline{y}_{i.} - \overline{\overline{y}}_{..})^2}{n-1}$$

The within-individual variance

$$\hat{\sigma}_w^2 = \frac{\sum_{i=1}^{n}\sum_{t=1}^{T}(y_{it} - \overline{y}_{i.})^2}{n(T-1)}$$

The overall variance

$$\hat{\sigma}_o^2 = \frac{\sum_{i=1}^{n}\sum_{t=1}^{T}(y_{it} - \overline{\overline{y}}_{..})^2}{nT-1}$$

The variable y_{it} can be decomposed into a between component $\bar{y}_{i.}$ and a within component $(y_{it} - \bar{y}_{i.} + \bar{\bar{y}}_{..})$. The overall mean $\bar{\bar{y}}_{..}$ is added back to make results comparable, i.e., to make the means of the variable, its between component, and its within component are all identical.

Variable		Mean	Std. Dev.	Observations
Work hours	Over all	38.25	11.11	nT=11290
	Between	38.25	9.81	n=3118
	Within	38.25	6.66	
Ln(work hours)	Over all	3.58	0.42	nT=11290
	Between	3.58	0.38	n=3118
	Within	3.58	0.25	

Table 1: An example of between-individual variation and within-individual variation

For example, as Table 1 shows, the within-individual variation of wife's work hours is smaller than its between-individual variation. In other words, work hours heterogeneity across different individuals is larger than work hours heterogeneity across different years after taking into account the average level of work hours for each individual.

4.1.3. Summary Statistics II: Time-invariant and time-varying variables

(1) Time-invariant variables are different across individuals but not changing over time. Therefore, the within variation of the time-invariant variables is zero. Race is a typical time-invariant variable. For this sample, region and educational attainment are the other two candidates for the time-invariant variable. There is little variation of educational attainment and region variables, although they are not exactly time-constant.

People could move to other regions, and they could acquire an education during the observation period.

(2) The time-varying variables are variables that are not only different across individuals but also changing over time. In this model, working status, work hours, weekly labour income, spouse weekly labour income, number of children, age, couple relationship are time-varying variables.

Both of the dependent variables of the two major models are time-varying variables. In the working probability (labour force participation) model, the dependent variable is the dummy variable for work status, which varies between 0 and 1 across individuals and could be different at different points of time for each individual. In the hours of work interior labour supply model, the dependent variable is the wife's weekly work hours, which is also an individual-specific and time-varying variable.

Variable	Source	Mean	Std. Dev.	Min	Max	Observations
Wife working status (1:working;0:not working)	Overall	0.80	0.40	0	1	nT=16440
	Between	0.80	0.32	0	1	n=3618
	Within	0.80	0.26	-0.04	1.63	T-bar=4.54
Husband's labour income ($1,000/week)	Overall	1.05	1.42	0	38.36	nT=16440
	Between	1.05	1.29	0	31.91	n=3618
	Within	1.05	0.74	-21.77	26.61	T-bar=4.54
Couple relation	Overall	4.53	2.28	0	7	nT=16397
	Between	4.53	1.84	0	7	n=3618
	Within	4.53	1.44	-1.30	10.36	T-bar=4.53
Number of children	Overall	1.22	1.25	0	10	nT=16440
	Between	1.22	1.11	0	7	n=3618
	Within	1.22	0.59	-3.11	5.22	T-bar=4.54
Age of youngest child	Overall	4.27	5.26	0	17	nT=16440
	Between	4.27	4.05	0	17	N=3618
	Within	4.27	3.58	-7.53	18.43	T-bar=4.54
Wife educational attainment	Overall	13.70	2.50	0	17	nT=15714
	Between	13.70	2.36	0	17	n=3551
	Within	13.70	0.80	2.70	22.50	T-bar=4.43

Table 2: Descriptive statistics of dependent variable and explanatory variables for working probability (boundary labour supply) model (continued on next page)

Wife health	Overall	2.39	0.98	1	5	nT=16389
	Between	2.39	0.83	1	5	n=3614
	Within	2.39	0.56	-0.11	5.05	T-bar=4.53
Transfer income ($1,000/week)	Overall	0.067	0.20	0	6.60	nT=16440
	Between	0.067	0.15	0	2.42	n=3618
	Within	0.067	0.14	-2.36	4.24	T-bar=4.54
Social security income ($1,000/week)	Overall	0.023	0.089	0	1.11	nT=16440
	Between	0.023	0.064	0	0.58	n=3618
	Within	0.023	0.060	-0.53	0.94	T-bar=4.54
Family wealth ($1,000,000)	Overall	0.29	0.85	-1.86	38.09	nT=16440
	Between	0.29	0.82	-0.30	32.46	n=3618
	Within	0.29	0.48	-10.94	15.84	T-bar=4.54
Wife age	Overall	42.96	10.37	17	65	nT=16440
	Between	42.96	9.99	17	61	n=3618
	Within	42.96	3.18	36.96	48.96	T-bar=4.54
Wife housework (hours/week)	Overall	17.12	12.91	0	112	nT=16299
	Between	17.12	10.19	0	112	n=3616
	Within	17.12	8.59	-34.88	92.72	T-bar=4.51

Table 3: Descriptive statistics of dependent variable and explanatory variables for working probability (boundary labour supply) model

4.2. Primary models, Auxiliary models, and Key Variables

To estimate the causal effects of wages on hours of work, the primary models are the hours of work models. The dependent variable is the logarithm of weekly hours of work for married women, and the explanatory variable of interest is the logarithm of her weekly labour income. The hours of work model is also called the interior labour supply model.

The auxiliary models are the working probability models, which are also called the labour force participation models or the boundary labour supply models. The dependent variable is a dummy that indicates the working status of the married women. The working probability models have two functions. One is to help the hours of work model to obtain a consistent estimator of hours-wage elasticity by providing a selection variable for the hours of work model. I use the predicted probability of working from the auxiliary model as the selection variable in the primary model to control for the selection bias. The other function of the working probability model is to compare to the hours of work model and show the different roles of some key explanatory variables between boundary labour supply and interior labour supply models.

The explanatory variable and control variables:

(1) Wife's weekly labour income.
It is calculated using the wife's annual labour income divided by 52.143, the number of weeks per year.

(2) Wife's wage type.
It is a categorical variable indicating she is salaried, paid by the hour, or other in her current main job.

(3) Husband's weekly labour income.
It is calculated using the husband's annual labour income divided by 52.143, the number of weeks per year.

(4) Couple relation.
It is measured using the number of days per week the family sit down and eat the main meal of the day together. The value is from 0 to 7, indicating an increasing close couple relation. This is also a measure of family solidarity and cohesion.

(5) Number of children under 18 years old.
It represents the number of persons currently in the family who are neither "Head" nor "Wife" of the survey, from newborns through those 17 years of age, whether or not they are actually children of the "Head" or "Wife". It is the number of young dependants in the family. It reflects the importance of the wife's dedication to her family.

(6) Age of the youngest child.
It also reflects the value of the wife's dedication to her family.
(7) Hours spent on housework in an average week, including time spent cooking, cleaning, and doing other work around the house.
(8) Wife's completed education level.
The values are from 0 to 17, which represent the grade of school completed. A value of 17 indicates at least some postgraduate work.
(9) Wife's self-reported health ranking.
It is a range of values from 1 to 5, indicating excellent, very good, good, fair, and poor, respectively, self-reported by the respondents.
(10) Wife's age.
(11) Weekly family transfer income.
The components of transfer income are very diverse. It includes financial help from relatives and friends, transfers from government, and miscellaneous transfers, but excludes social security. The unit is $1000.
(12) Weekly family social security income.
The unit is $1000.
(13) Family wealth.
It is the sum of values of seven asset types (the value of farm or business, all checking and saving accounts, the real estate asset other than main home, the stocks, vehicles, private annuities or individual retirement accounts, and other assets) net of debt value (debt on business, real estate, credit cards, medical bills, legal bills, family loans, and other debts), including the value of home equity. The unit is $1000,000.
(14) Region.
The sample is limited to families in the four regions: northeast, north central, south, and west.
(15) Race of wife.
There are four categories: white, African-American, Asian, and others.

The above explanatory variables could be used in the pooled OLS models. When it comes to the fixed effects panel models, the time-invariant variables, like race, region and education, should be excluded from the model, because there is no or little within-variation of these variables. Age may be excluded from the fixed effects panel model since there is perfect collinearity within each individual between the year variable and the age variable. However, if the year dummy variable instead of the year continuous variable is used, then age could be included in the time fixed effects models.

4.3. The Estimation Strategies and Results

4.3.1. Baseline Pooled OLS Model

The starting-point model is the pooled OLS regression model without taking into account the sources of endogeneity problems.

(1) The specification:

$$\begin{aligned} ln\_wife\_weekly\_work\_hours_{it} = \beta_0 + \beta_1 ln\_wife\_wage_{it} \\ + \mathbf{Z_{1it}}\alpha + \mu_{it} \end{aligned}$$

where

- i: individual married woman; t: year.
- β_1 is the coefficient of interest the causal effects of own wages on labour supply. It is also called the Marshallian labour supply elasticity, because the non-labour income and family wealth are held constant.
- $\mathbf{Z_{1it}}$ is the row vector of exogenous control variables that are suggested by the theoretical model and help to identify the ceteris paribus effects of wages on labour supply.

$$\begin{aligned} \mathbf{Z_{1it}} = (&wife\_wage\_type_{it}, couple\_relation_{it}, hours\_on\_housework_{it} \\ &husband\_wage_{it}, family\_transfer\_income_{it}, family\_wealth_{it}, \\ &wife\_health_{it}, wife\_education_{it}, wife\_age_{it}, \\ &number\_of\_children_{it}, age\_of\_youngest\_child_{it}) \end{aligned}$$

- The error term u_{it} represents the unobserved factors that influence married women's decision of working hours.

(2) The exogeneity assumption for consistent pooled OLS estimator β_1:

$$E(\mu_{it}|ln\_wife\_wage_{it}, \mathbf{Z_{1it}}) = 0$$

This assumption states that the unobserved explanatory components of the model are not correlated with the observed explanatory components of the model, which is quite strong. The unobserved factors μ_{it} may include the individual preference for work or skills. We are assuming that they are not correlated with all the explanatory variables.

(3) Results and Interpretation.

The first column of Table.4 gives the estimates from pooled OLS.

(3.1) The pooled OLS estimate for β_1 is the average estimated hours-wage labour supply elasticities for married women holding constant their wage types, couple relationship, husband's weekly labour income, family wealth, number of children, hours spent on housework, and the demographic characteristics of women's education, health, race, age, and region.

The estimate of the hours-wage elasticity is 0.2782, which means a one percent increase in wages leads to a 0.28 percent increase in hours of work for married women. The standard deviation of $ln(wife\_wage)$ is 0.88 in the sample. One standard deviation of increase in wife's wages results in approximately 0.25 percent increase in working hours. This estimate is significantly different from zero.

(3.2) When the couple relation increases by one unit holding other factors constant, the hours of work decrease by 0.5% on average. The standard deviation of couple relation is 2.3, which means one standard deviation increase in couple relation decreases hours of work for married women by 1.2%. This is in line with the theoretical model.

(3.3) If the female workers are paid by the hour, then they work less than their salaried counterparts by 5.4%.

(3.4) The partial effects of husband's labour income on his wife's labour supply is negative, as predicted by the theoretical model. If the husband earns one thousand more per week, which is a little more than the standard deviation, then his wife's hours of work decrease by about 4.8%.

(3.5) If there is one more child under 18 years old in the family, the wife reduces her hours of work by 2.8%. When children grow up, the mother's hours of work increase.

(3.6) More time spent on housework results in less time spent on jobs for married women. One more hour spent on housework leads to 0.36% decrease in working hours, holding other factors constant. One standard deviation (13 hours per week) increase in housework hours reduces working hours by 4.7%.

(3.7) Educational attainment is also a significant factor behind married women's decision of labour supply. More educated married women work less. One more year of schooling leads to a 2.8% decrease in hours of work.

(3.9) Richer the family is, fewer hours of work are supplied by the wife. The working hours drop by 4.4% with every one million dollars more of family wealth, including the value of the home. The transfer income also contributes to less work.

(3.10) Health is another determinant of labour supply. The married women who are self-reported as "Good" or "Fair" work more than those who are "Excellent", "Very good", or "Poor".

4.3.2. *Sample Selection, Sample Attrition, and Missing Values*

(1) Sample selection bias.

(1.1) Causes:

The hours of work and wage rates are observed for the women who are employed. If we intend to estimate the elasticities only for these part of women, there will be no sample selection bias. However, if we want to estimate the labour supply elasticities of the population of all working-age married women, then exclusion of the unemployed women from the model may lead to bias because employed women and unemployed women may have a different response to the changes in wage rates. For instance, suppose the women who are out of labour market and stay at home keeping their houses are not as responsive as their counterparts already working, then OLS model without accounting for sample selection bias will overestimate the labour supply elasticities because the less elastic women are included in the sample.

(1.2) Solution:

The sample selection bias can be seen as a special type of omitted variable bias. The omitted variable is the willingness to work, or the probability of working, or the tendency towards jobs. Including the willingness to work variable into the model so that the working probability can be held fixed, which alleviates the sample selection bias and the employed women sample represents the population of all working-age married women.

The willingness to work variable can be predicted from the working probability (labour force participation) regression for all working-age married women, both employed and unemployed, including women staying at home, taking care of children and keeping their house.

(1.3) Specification.

The hours of work equation:

$$ln\_wife\_work\_hours_{it} = \begin{cases} \beta_0 + \beta_1 ln\_wife\_wage_{it} & \\ +\mathbf{Z_{1it}}\alpha + \mu_{it} & if\ working\_prob_{it} = 1 \\ & \\ unobserved & if\ working\_prob_{it} = 0 \end{cases}$$

The selection equation:

$$working\_prob_{it} = 1 \quad if \quad \mathbf{M_{it}}\gamma + \nu_{it} > 0 \quad i.e.,$$
$$Prob(working\_prob_{it} = 1|\mathbf{M_{it}}) = \Phi(\mathbf{M_{it}}\gamma)$$

where $\mathbf{M_{it}}$ include $\mathbf{Z_{1it}}$ and the excluded exogenous variables that determine the working choice but are not in the hours of work equation. I use race and region as the excluded exogenous variables.

It can be proved that

$$\begin{aligned} E(ln\_wife\_work\_hours_{it}|\mathbf{M_{it}}, working\_prob_{it} = 1) = \beta_0 + \beta_1 ln\_wife\_wage_{it} \\ + \mathbf{Z_{1it}}\alpha + \rho\lambda(\mathbf{M_{it}}\gamma) \end{aligned}$$

where $\lambda(\mathbf{M_{it}}\gamma)$ could be the inverse Mills ratio or the predicted working probability. It is the measure of willingness to work of the married women.

This equation shows that we can obtain consistent OLS estimates using only the selected sample, provided we include the willingness to work variable as an additional regressor.

(1.4) Exogeneity Assumption for consistent estimates.

With the willingness to work variable, we can rewrite the assumptions for consistent pooled OLS with sample selection issues.

$$\begin{aligned} ln\_wife\_work\_hours_{it} = \beta_0 + \beta_1 ln\_wife\_wage_{it} \\ + \beta_3 willingness\_to\_work_{it} + \mathbf{Z_{1it}}\alpha + \mu_{it} \end{aligned}$$
$$E(\mu_{it}|ln\_wife\_wage_{it}, willingness\_to\_work_{it}, \mathbf{Z_{1it}}) = 0$$

The reason we need some exogenous variables in the selection equation that are not in the hours of work equation is that the willingness to work variable (the predicted working probability) should not be a combination of the explanatory variables in the hours of work equation. If same explanatory variables are used in both equations, then the willingness to work variable, which is the predicted value of working probability in the selection equation or a function of explanatory variables in the selection equation, will cause perfect collinearity problem in the hours of work equation.

(1.5) Process:

The sample selection correction process can be done in two steps:

(1) Step 1: Run the selection equation, which is the boundary labour supply equation. I use the logit model for labour force participation choice and obtain the predicted working probability for each individual.

(2) Step 2: Run the hours of work model, interior labour supply equation, adding the predicted working probability from the selection equation. By doing so, the willingness to work has been controlled for, and the partial effects of wages on labour supply can be estimated holding the working probability constant.

(1.6) Results and Interpretation.

The second column of Table.4 gives the estimates from pooled OLS with accounting for sample selection bias.

The selection variable willingness to work is not statistically significant from zero. It implies there is no evidence that the labour supply decisions between the already employed group and the unemployed group are different. It also has no influence on the hours-wage elasticity. The reason could be that the willingness to work variable is not correlated with the wages received by the married women and therefore omitting it does not cause bias on the causal effects of wages on labour supply decision.

I will use panel data individual fixed effects model in next section. If I assume that the willingness to work is individual specific and time-invariant, then the individual fixed effects specification under panel data structure could control for the selection variable without estimating it. That is to say, using panel data gives us an alternative method to deal with sample selection bias.

(2) Sample attrition bias.

If we assume that the attrition is random, then there is no problem. If it is not random, it may lead to bias.

The sample attrition bias can also be seen as a special case of omitted variable bias in which the sample attrition probability is missing. When each individual's attrition probability is correlated with wages and has an effect on hours of work at the same time, we should control for it by including the attrition probability variable in the model. If we assume that the attrition probability is individual-specific and time-invariant, then the panel data fixed effects model in section 4.3.3 could be employed. On the other hand, if each individual's attrition probability is time-varying, then instrumental variable and the two-stage-least-square (2SLS) in section 4.3.4 could be used to solve the problem.

Among the 3616 respondents in the starting survey in 2005 who were at their wage, 943 dropped out of the survey during the study period, which amounts to 26% of the initial number of respondents.

Among the 3616 respondents in the starting survey in 2005 who were at their working age and provided employment status information (including both the employed and the unemployed), 1517 dropped out of the survey during the study period, which amounts to 42% of the initial number of respondents.

Among the 2935 respondents in the starting survey in 2005 who were at their working age and provided hours of work (only the employed), 1350 dropped out the survey, which amounts to 46% of the initial number of respondents.

Year	Respondents	Attrition	Percent Attrition
2005	3616	0	0.00
2007	3421	195	5.4
2009	3277	339	9.4
2011	3074	542	15.0
2013	2864	752	20.8
2015	2673	943	26.1

Table 4: Sample Attrition 0

Year	Respondents	Attrition	Percent Attrition
2005	3616	0	0.00
2007	3078	538	14.9
2009	2784	832	23.0
2011	2522	1094	30.1
2013	2331	1285	35.3
2015	2099	1517	42.0

Table 5: Sample Attrition 1

Year	Respondents	Attrition	Percent Attrition
2005	2935	0	0.00
2007	2488	447	15.2
2009	2189	746	25.4
2011	1901	1034	35.2
2013	1766	1169	39.8
2015	1585	1350	46.0

Table 6: Sample Attrition 2

(3) Missing values of explanatory variables In those married women who provided employment information (including both in the labour force and out of the labour force), there are missing values of some explanatory variables.

Variable	Missing	Total	Percent Missing
Wife work status	0	13037	0.00
Wife weekly hours of work	0	13037	0.00
Wife wage rate	0	13037	0.00
Couple relation	31	13037	0.24
Husband wage	0	13037	0.00
Number of children	0	13037	0.00
Age of youngest child	0	13037	0.00
Wife health ranking	36	13037	0.28
Wife age	0	13037	0.00
Wife housework hours	106	13037	0.81
Wife education	523	13037	4.01
Transfer income	0	13037	0.00
Family wealth	0	13037	0.00
Wife tenure	61	13037	0.47
Wife race	0	13037	0.00
Region	0	13037	0.00
Union contract	4668	13037	35.81
Type of labour income	0	13037	0.00

Table 7: Missing Values of Some Variables

We can see 95% of the observations in the labour force participation (working probability) model have complete information of all relevant variables. For the four variables that have missing values, multiple imputation is employed to fill in the missing values.

Missing-value patterns
(1 means complete)

Percent	Pattern 1	2	3	4
95%	1	1	1	1
4	1	1	1	0
<1	1	1	0	1
<1	0	1	1	1
<1	1	0	1	1
<1	1	0	1	0
<1	1	1	0	0
<1	0	1	0	1
<1	1	0	0	0
<1	1	0	0	1
<1	0	1	1	0
<1	0	0	0	0
<1	0	0	0	1
100%				

Table 8: Missing Value Patterns (1: couple relation; 2: wife health ranking; 3: wife housework hours; 4: wife's education level)

4.3.3. Unobserved individual heterogeneity: panel data fixed effects model

(1) Cause:

Even if the sample selection bias has been controlled for, the OLS estimates of labour supply elasticities would be still biased or inconsistent due to unobserved individual-specific characteristics and economic trends. They are the traditional examples of omitted variable bias. There are two types of unobserved individual-specific characteristics: individual heterogeneity and individual idiosyncratic characteristics. In this section, I would focus on the former. Unobserved individual heterogeneity is individual-different and time-invariant, such as family tradition, preference for work, and habits. They are different across each married woman but are constant over time for each individual. Probably the individual heterogeneity is correlated with explanatory variables in the model, like wages. They also influence how long people would like to work. For example, a married woman who has a strong preference for work is more likely to work harder and get higher paid. Besides, she would like to work longer just because she likes her job, not because of the high wage. In order words, the omitted variable, preference for work, affects the dependent variable directly or through other channels rather than through wages. Omitting the preference for work variable, the OLS estimate for the coefficient of wages would be upwards biased, because the omitted variable is positively correlated with both the explanatory variable of interest, wages, and the dependent variable, hours of work.

Another example is diligent people, they tend to work longer, and they also give a good impression to their boss and therefore are more likely to receive a higher wage. The measure of hardworking in not available. Omitting the hardworking variable results in overestimates for the effects of wages on hours of work supplied.

A third case if unobserved individual heterogeneity is the willingness to work and the tendency to drop out of the survey. The probability of self-selecting into the labour force and the probability of leaving the survey could be seen as fixed effects if they are assumed to be individual-specific and time-constant.

Panel data fixed effects model can alleviate this problem.

(2) Solution:

The advantage of using micro panel data is that the unobserved fixed heterogeneity can be controlled for by including them into the model and they can be eliminated by first differencing or demeaning methods under certain conditions. They don't need to be estimated.

We also control for the time fixed effects (the macro economic trend that is constant across individuals but changing over time) by including year dummies in the model.

(3) Specification.

$$\begin{aligned} ln\_wife\_work\_hours_{it} = \beta_0 &+ \beta_1 ln\_wife\_wage_{it} \\ &+ \beta_3 willingness\_to\_work_{it} + \mathbf{Z_{1it}}\alpha \\ &+ a_i + \varepsilon_{it} \end{aligned}$$

where a_{it} is the unobserved individual heterogeneity that is time-invariant. a_{it} could affect outcome variable directly, and it is allowed to be correlated with explanatory variables in the model.

(4) Exogeneity assumption for a consistent estimator.

$$E(a_i | ln\_wife\_wage_{it}, willingness\_to\_work_{it}, \mathbf{Z_{1it}}) \neq 0$$
$$E(\varepsilon_{it} | ln\_wife\_wage_{it}, willingness\_to\_work_{it}, \mathbf{Z_{1it}}) = 0$$

(5) Results and interpretation.

The third column of Table.4 gives the panel data fixed effects estimates. After controlling for unobserved individual heterogeneity, the estimated causal effects of wages on hours of work drops a little, which is in line with the examples mentioned. The most noticeable change is that the estimates for the partial effects of the husband labour income, the couple relation, the family wealth, and the housework hours are no longer statistically significant at any conventional significance levels.

Another change is that the selection variable becomes significant at 5% significance level.

Notice that the time-invariant variables like the race could not be estimated in the fixed effects model because it would also be removed along with the individual fixed effects when using demeaning or first differencing methods.

(6) Hausman test.

Another candidate model is panel data random effects model. The exogeneity assumption for random effects model is that the unobserved individual heterogeneity a_i is not correlated with all the explanatory variables in the model, which is not likely to be satisfied.

$$E(a_i|ln\_wife\_wage_{it}, willingness\_to\_work_{it}, \mathbf{Z_{1it}}) = 0$$
$$E(\varepsilon_{it}|ln\_wife\_wage_{it}, willingness\_to\_work_{it}, \mathbf{Z_{1it}}) = 0$$

To choose between fixed effects and random effects models, I run the Hausman test. The null hypothesis of the Hausman test is that both FE and RE are consistent, but RE is more efficient, the alternative hypothesis is that only the FE model is consistent. The null hypothesis is rejected, so the fixed effects model is more appropriate.

4.3.4. Unobserved individual idiosyncratic characteristics: Instrumental 2SLS

(1) Cause.

After holding relevant observed explanatory variables constant and taking into account sample selection bias, sample attrition bias, unobserved individual heterogeneity, we still face an endogeneity problem caused by unobserved individual idiosyncratic characteristics of married women. They are individual-specific and time-varying, such as competency, ability, and intelligence. Omitting them could lead to biased estimates because they are very likely to be correlated with the explanatory variable of interest, $ln(wage)_{it}$.

$$E(\varepsilon_{it}|ln\_wife\_wage_{it}, willingness\_to\_work_{it}, \mathbf{Z_{1it}}) \neq 0$$

(2) Solution.

The variables that only affect wage rates without influencing labour supply could be used as instrumental variables (IV) for the endogenous variable $ln(wage)_{it}$. Usually, such IVs can be found from macro economic conditions like the minimum wage and the unemployment rates. However, in the panel data fixed effects model, the macro economic trends have been captured by the year fixed effects setting, i.e., the year dummies. The IVs should be found from the microdata. One way is to consider factors from the demand side of the labour market.

The first candidate of IVs for wages that have no effects on labour supply decision is the experience the married women have in the current main job. The experience is a fundamental determinant for wages, as shown in the Mincer wage equation or the wage offering models. On the other hand, the experience is not likely to affect how long one works directly. Therefore, experience satisfies the exogeneity assumption and the relevance condition for a valid IV. Another IV is a dummy variable indicating whether the current job was covered by a union contract. If the jobs were covered by a union contract, the wage rate is expected to be higher than those without a union contract. This can be verified from the first stage of the two-stage-least-squares procedure (2SLS). It could be argued that union has nothing to do with the individual's labour supply decision, except that through the impact on wages.

(3) Specification.

$$
\begin{aligned}
ln\_wife\_work\_hours_{it} = \beta_0 &+ \beta_1 ln\_wife\_wage_{it} \\
&+ \beta_3 willingness\_to\_work_{it} + \mathbf{Z_{1it}}\alpha \\
&+ a_i + \varepsilon_{it}
\end{aligned}
$$

$$\mathbf{Z_{2it}} = (exp_{it}, union_{it})$$

where $\mathbf{Z_{2it}}$ are IVs.

(4) Exogeneity assumption and relevance condition for valid IVs.

$$Cov(\mathbf{Z_{2it}}, ln\_wage_{it}) \neq 0; Cov(\mathbf{Z_{2it}}, \varepsilon_{it}) = 0$$

The instrumental variables should be highly correlated with the endogenous explanatory variable of interest $ln(wage)_{it}$. This is the relevance condition of a valid IV. The instrumental variables should not be correlated with the unobserved idiosyncratic error term ε_{it}. This is the exogeneity assumption for a valid IV, which is also called the exclusion restriction. The relevance condition can be verified from the first stage of a 2SLS procedure. The exogeneity assumption in the panel data structure implied that the IVs could not be correlated with the error components that are individual-specific and time-varying. The IVs are allowed to be correlated with the unobserved individual-specific and time-invariant components, i.e., the individual heterogeneity a_i. The IVs could be correlated with the individual-invariant and time-varying components, captured by the year dummies. This means in a panel data structure, the IVs are less likely to violate the exogeneity assumption than in the situation where cross-sectional data are used.

(5) Illustration using simultaneity of labour supply and labour demand.
The IVs are from the labour demand shifters. The variation in the demand shifters helps to trace out the labour supply. The labour supply curve is not fixed even after controlling for fixed effects and observed relevant factors. Due to the unobserved individual idiosyncratic characteristics, it would shift. To solve this problem, I use IVs to trace out the labour supply curve. When IVs vary, the labour demand curve shifts but the labour supply curve is fixed because the change in IVs does not affect the unobserved individual idiosyncratic characteristics in the labour supply side (the exogeneity assumption). Fig.4 shows a case that the actual labour supply curve derived using IVs is steeper than the one without using IVs.

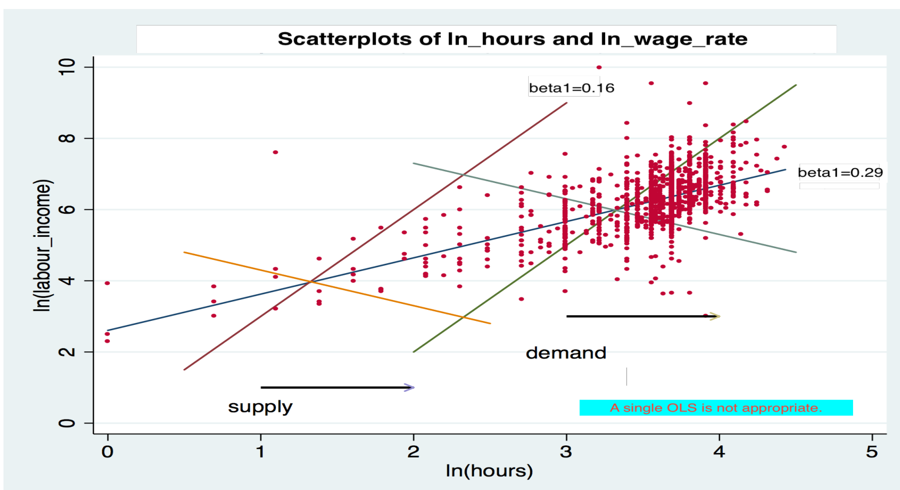

Fig. 4. The simultaneity of labour supply and labour demand

(6) The 2SLS procedure.
   - The first stage: Regress wage rate on the exogenous IVs from outside the labour supply model, i.e., the demand shifters $\mathbf{Z_{2it}}$, and on the exogenous variables from the model, i.e., the supply shifters. Then we obtain the predicted value of the wage rate.
   - The second stage: Regress hours of work on the predicted value of wage rates and all other exogenous explanatory variables in the labour supply model.

Both stages are done under the panel data structure. This panel data fixed effects 2SLS estimator is a consistent estimator for the causal effects of wages on the hours of work.

(7) Results and Interpretation.

From the fourth column of Table.4, I find that the causal effects of wages on labour supply (the hours-wage elasticities) drop from 0.27 in the pooled OLS to 0.21 in the panel data fixed effects 2SLS model. Holding other factors constant, a 1% increase in married womens wages raises their hours of work by approximately 0.21% on average.

Explanatory_Variables	Pooled_OLS	Pooled_OLS_Sele	Panel_FE	Panel_FE_2SLS
ln(wife_wage)	0.2749***	0.2749***	0.2686***	0.2123***
	(0.0087)	(0.0087)	(0.0148)	(0.0430)
Couple relation	-0.0064***	-0.0066***	0.0006	0.0005
	(0.0014)	(0.0017)	(0.0025)	(0.0025)
Wage type				
Salaried	(base)	(base)	(base)	(base)
Hourly paid	-0.0580***	-0.0578***	-0.0549***	-0.0729***
	(0.0078)	(0.0077)	(0.0123)	(0.0195)
Husband wage	-0.0534***	-0.0537***	-0.0032	-0.0036
	(0.0062)	(0.0062)	(0.0086)	(0.0086)
Number of children	-0.0271***	-0.0275***	-0.0338***	-0.0355***
	(0.0035)	(0.0040)	(0.0062)	
Age of youngest child	0.0016*	0.0017*	0.0010	0.0012
	(0.0007)	(0.0007)	(0.0009)	(0.0009)
Housework hours	-0.0038***	-0.0039***	-0.0002	-0.0002
	(0.0005)	(0.0008)	(0.0011)	(0.0011)
Education level	-0.0278***	-0.0274***	-0.0146**	-0.0133**
	(0.0019)	(0.0025)	(0.0050)	(0.0051)
Age	-0.0028***	-0.0028***	-0.0034***	-0.0026*
	(0.0004)	(0.0004)	(0.0010)	(0.0012)

Table 9: The Four Specifications (continued)

Explanatory_Variables	Pooled_OLS	Pooled_OLS_SSC	Panel_FE	Panel_FE_2SLS
Family wealth	-0.0454***	-0.0456***	-0.0056	-0.0067
	(0.0082)	(0.0081)	(0.0080)	(0.0084)
Family transfer income	-0.0616*	-0.0650*	-0.0437	-0.0587
	(0.0276)	(0.0302)	(0.0477)	(0.0505)
Health				
Excellent	(base)	(base)	(base)	(base)
Very good	0.0092	0.0093	-0.0045	-0.0056
Good	0.0325***	0.0324***	-0.0096	-0.0117
Fair	0.0562***	0.0537**	0.0390	0.0394
Poor	0.0410	0.0337	0.1341	0.1485
Sample Selection Correction	No	Yes	Yes	Yes
Working prob.		-0.0197	0.3548*	0.3968**
		(0.1106)	(0.1546)	(0.1546)
Sample Attrition Correction	No	No	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Individual fixed effects	No	No	Yes	Yes
Controlling for individual idiosyncratic characteristics	No	No	No	Yes
N=nT	9749	9749	9749	9749

Notes: Robust standard errors are in the parentheses.
*p<0.05; **p<0.01; ***p<0.001

Table 10: The Four Specifications

5. Features of Married Women's Labour Supply

5.1. *Interior Labour Supply vs Boundary Labour Supply*

We have already consistently estimated the interior labour supply elasticities. Let us look at the boundary labour supply model, with the dependent variable being a dummy that indicates whether or not the individual participates in the labour market. $wife\_LFP_{it} = 1$ means the married woman participates in the labour force.

The boundary labour supply elasticities can be estimated by a linear probability model (LPM), a probit model, or a logistic model within the individual panel data framework.

5.1.1. *Specification and Assumptions*

For the logistic model:

$$wife\_LFP_{it} = \begin{cases} 1, & if \quad \mathbf{X_{it}}\beta + a_i + \varepsilon_{it} > 0 \\ \\ 0, & otherwise \end{cases}$$

$$\varepsilon_{it}|(\mathbf{X_{it}}, a_i) \sim Logistic(0, 1)$$
$$a_i|(\mathbf{X_{it}}) \sim N(0, 1)$$

alternatively, it could be written as:

$$Prob(wife\_LFP_{it} = 1|\mathbf{X_{it}}) = G(\mathbf{X_{it}}\beta) = exp(\mathbf{X_{it}}\beta)/[1 + exp(\mathbf{X_{it}}\beta)]$$

$\mathbf{X_{it}}$ include the determinants of married women's labour supply decision suggested by the theoretical boundary labour supply model, including: hours spent on housework, her husband's weekly labour income, the couple relation measure, the number of children under 18 years old, the age of the youngest child, her education level, her self-reported health ranking, her age, her race, the family transfer income, the family wealth, and the region.

The logistic and probit specifications are preferred to the linear OLS specification for two main reasons. First, the predicted probability of working is strictly between zero and one for all values of $\mathbf{X_{it}}\beta$ for logistic or probit specifications, while the predicted probabilities could be outside of the range of zero and one using LPM. Second, the partial effects of regressors are allowed to be changing in logistic or probit models, while in LPM the partial effects are constant.

5.1.2. *Average Partial Effects (APE) and Partial Effects at the Average (PEA)*

The PEA of the j^{th} explanatory variable on the dependent variable is $g(\bar{\mathbf{x}}\hat{\beta})\hat{\beta}_j$, while the APE is $n^{-1}\sum_{i=1}^{n} g(\mathbf{x}_i\hat{\beta})\hat{\beta}_j$, where $g(z) = \frac{dG(z)}{dz}$. The APE and PEA for the linear probability model (LPM) are identical. For the logistic model, they are not the same, but close to each other, as shown in Table 6 and 7.

Dependent Variable: Working choice	Panel Random Effects			Panel Fixed Effects		Pooled	
The Explanatory Variables:	Logistic (APE)	Logistic (PEA)	LPM (APE and PEA)	LPM (APE and PEA)	Logistic (APE)	LPM (APE and PEA)	Logistic (APE)
Ln(husband wages)	-0.0178***	-0.0184***	-0.0166***	-0.0051	-0.0142	-0.0318***	-0.0348***
	(0.0047)	(0.0048)	(0.0046)	(0.0059)	(0.0107)	(0.0039)	(0.0044)
Couple relation	-0.0078***	-0.0081***	-0.0076***	-0.0041*	-0.0100*	-0.0113***	-0.0117***
	(0.0016)	(0.0016)	(0.0016)	(0.0019)	(0.0042)	(0.0014)	(0.0014)
Hours on housework	-0.0051***	-0.0053***	-0.0066***	-0.0052***	-0.0080***	-0.0088***	-0.0067***
	(0.0003)	(0.0003)	(0.0003)	(0.00039)	(0.0015)	(0.0003)	(0.0002)
Number of children	-0.0282***	-0.0291***	-0.0287***	-0.0344***	-0.0659***	-0.0236***	-0.0225***
	(0.0037)	(0.0038)	(0.0041)	(0.0058)	(0.0151)	(0.0032)	(0.0028)
Age of youngest child	0.0046***	0.0047***	0.0045***	0.0049***	0.0129***	0.0039***	0.0035***
	(0.0007)	(0.0008)	(0.0006)	(0.0008)	(0.0029)	(0.0006)	(0.0006)
Wife education level	0.0192***	0.0198***	0.0203***	0.0177***	0.0325*	0.0196***	0.0180***
	(0.0020)	(0.0021)	(0.0022)	(0.0046)	(0.0126)	(0.0015)	(0.0014)
Wife health ranking							
Excellent	base	base	base	base	base	base	base
Very good	0.0047	0.0048	0.0053	0.0062	0.0179	0.0058	0.0039
Good	-0.0099	-0.0101	-0.0087	-0.0068	-0.0103	-0.0040	-0.0070
Fair	-0.0849***	-0.0886***	-0.0894***	-0.0502**	-0.0734*	-0.1336***	-0.1304***
Poor	-0.2343***	-0.2523***	-0.2557***	-0.1360***	-0.1451*	-0.3701***	-0.3744***

Table 11: The Average Partial Effects of Explanatory Variables on Willingness to Work (to be continued)

The average partial effects (APE) are estimated using different specifications: panel linear probability model random effects, panel linear probability model fixed effects, panel logistic random effects, panel logistic fixed effects, pooled linear probability model, and

Dependent Variable: Working choice	Panel Random Effects			Panel Fixed Effects		Pooled	
The Explanatory Variables:	Logistic (APE)	Logistic (PEA)	LPM (APE and PEA)	LPM (APE and PEA)	Logistic (APE)	LPM (APE and PEA)	Logistic (APE)
Family transfer income	-0.1356***	-0.1404***	-0.1606***	-0.1436***	-0.2561***	-0.1894***	-0.1565***
	(0.0222)	(0.0231)	(0.0296)	(0.0323)	(0.0734)	(0.0275)	(0.0208)
Family wealth	-0.0129***	-0.0133***	-0.0152***	-0.009	-0.0171	-.0175***	-0.0139***
	(0.0037)	(0.0038)	(0.0050)	(0.0072)	(0.0135)	(0.0046)	(0.0034)
Number of obs	13739	13739	13739	13739	4171	13739	13739

Notes: Other control variables include region, race, age and year dummies.
Robust standard errors are in the parentheses.
*p<0.05; **p<0.01; ***p<0.001

Table 12: The Average Partial Effects of Explanatory Variables on Willingness to Work

pooled logistic model.

The difference between pooled setting and panel random effects setting is the difference in their assumptions. The panel random effects model assumes that the error term (unobserved characteristics) can be decomposed into two parts: the individual-specific and time-invariant component (individual heterogeneity) and the individual-specific and time-varying component (idiosyncratic error). The pooled setting does not take the advantage of the panel data structure and see the data as repeated cross-sectional data.

The difference between the panel random effects and the panel fixed effects is also about the assumptions. The random effects model assume that the individual heterogeneity part of the error term is not correlated with the explanatory variable, while the fixed effects model allows the correlation.

5.1.3. Results and Interpretation

(1) Choose between specifications.

The reason why the logistic model is preferred to the linear OLS model has already been discussed above. It is hard to choose between the logistic panel fixed effects model and the logistic panel random effects model. The sample size drops dramatically from 13,739 to 4,171 when fixed effects are used. The panel fixed effects model requires a variation of the dependent variable within individuals over time. In the logistic case, the dependent variable is the dummy variable working status (0 or 1). Those observations with the same working status within individuals over ten years would be dropped from the regression. There are only 875 out of 3,393 married women changed their working status during the ten years of observation time. Too much information may have been lost, and the accuracy would be affected. On the other hand, the advantage of the fixed effects model is that the unobserved individual heterogeneity could be taken into account to alleviate the endogeneity issue. It is a trade-off between consistency and accuracy.

(2) The predicted probability of labour force participation.

The following graphs in Fig 5 are based on the logistic panel random effects model (the first column of Table 6 and 7). Fig 5 shows how the predicted probabilities of working of married women change with each key explanatory variable. For instance, holding other factors at their means, the logarithm of husband's wage has a negative effect on his wife's working decision. As the husband's wage increases, his wife's willingness to work decreases. Couple relation, wife's housework hours, number of children, transfer income, and family wealth also have negative effects on married women's willingness to work, as predicted by the theoretical model. Less healthy women are less likely to work, especially for those with fair or poor self-reported health rankings. On the other hand, as the children grow up, their mothers are more likely to participate in the labour force. More educated women are more likely to work, but by contrast, once they are in the labour force, more educated women work fewer hours than less educated women on average as the interior labour supply models show in Table 4.

The estimates for the partial effects of explanatory variables showed in Table 6 and the graphs of the predicted probability of labour force participation in Fig 5 are actually two sides of a coin. Both of them are in line with the theoretical models.

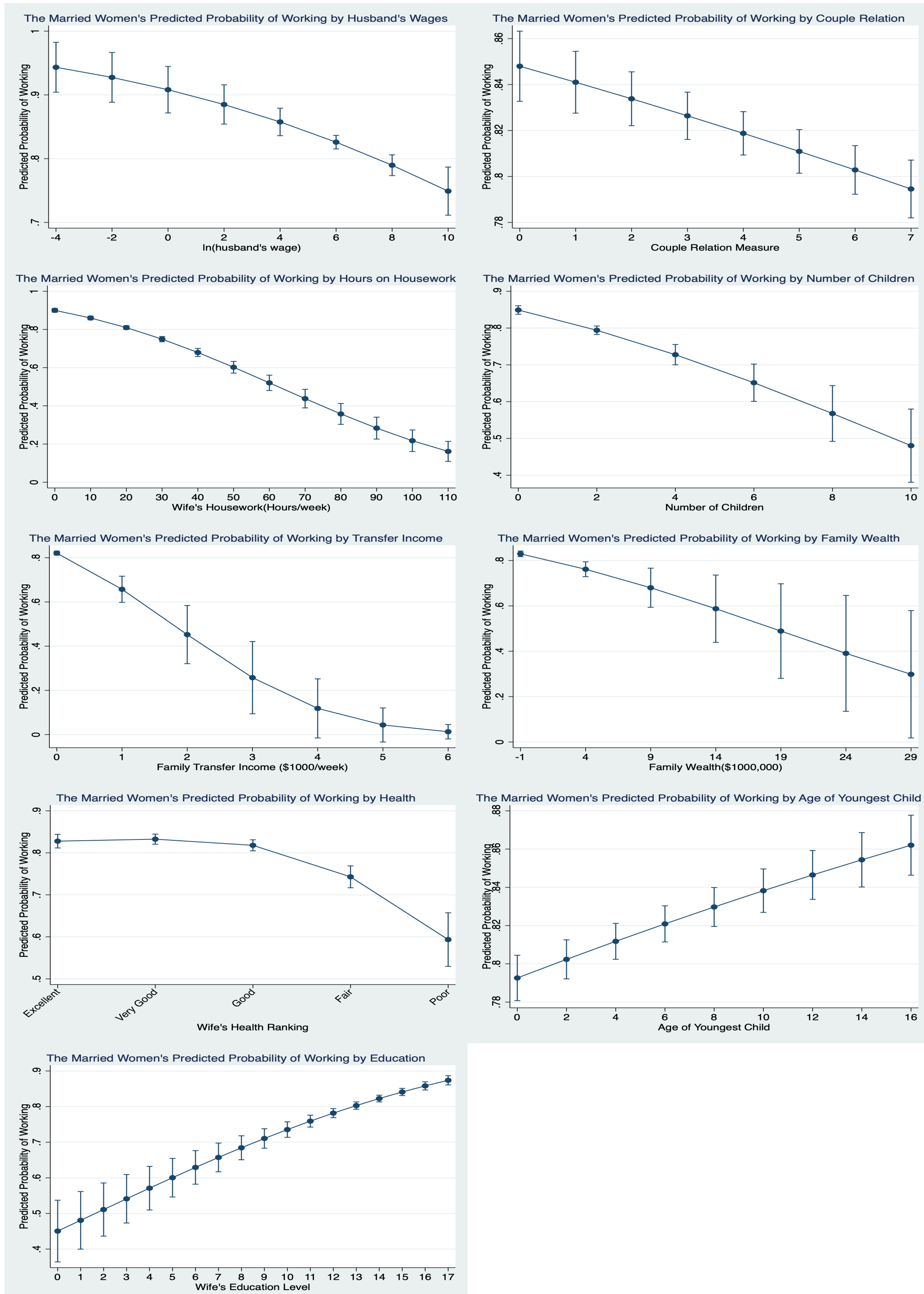

Fig. 5. The Predicted Probabilities of Working for Married Women (with 95% Confidence Intervals)

5.2. The Role of Couple Relation

5.2.1. The Partial Effects on Willingness to Work

All the specifications give negative and significant estimates for the partial effects of the couple relation on married women's willingness to work. It supports the hypothesis that a closer couple relation reduces the wife's probability of labour force participation, holding other relevant factors fixed. Just as the theoretical model implies, the partial effects of couple relation on willingness to work rely on the interaction between husband's labour income and couple relationship. As husband's wages drop, the partial effects of couple relation approach to zero. I check this hypothesis by using a sub-sample only contains families with zero husband labour income. Including all the same control variables, the estimates of the APE of couple relation become not significant in the above six specifications: pooled LPM, pooled Logistic, panel LPM random effects, panel logistic random effects, panel LPM fixed effects, and panel logistic fixed effects. This is evidence of the hypothesis that the negative causal effects of couple relation on married women's willingness to work are through its interaction with husband labour income.

Moreover, the partial effects of couple relation on hours of work is not significant in the panel data fixed effects specifications of the interior labour supply model (columns 4 and 5 in Table 4), while the panel data fixed effects specifications of the boundary labour supply model give significant estimates for couple relation (columns 4 and 5 in Table 6). This means when only the working married women are considered in the interior labour supply model, the couple relation becomes less important than when all married women are considered in the boundary labour supply model, probably because once married women earn labour income, the relative male wage and its role go down. This is also evidence of the hypothesis that couple relation's effects depend on the husband's relative wages or gender wage inequality.

5.2.2. The Influence on Cross Hours-Wage Elasticity

Including the interaction term of husband's wage and couple relation, I obtained how the partial effects of husband's wage on his wife's hours worked (the cross hours-wage elasticity) changed with the couple relationship. From Fig.6, the negative causal effects of the husband's wages on his wife's labour supply become larger as the couple has a better relationship, which is suggested by the number of meals they have together per week. When the couple eat together more frequently, they are more likely to have better relation and the wife could rely on her husband and therefore her husband's wages have a larger negative effect on her labour supply decision. Although this cross hours-wage elasticities is not significantly different from each other, it is still an evidence that support the hypothesis proposed based on the theoretical model.

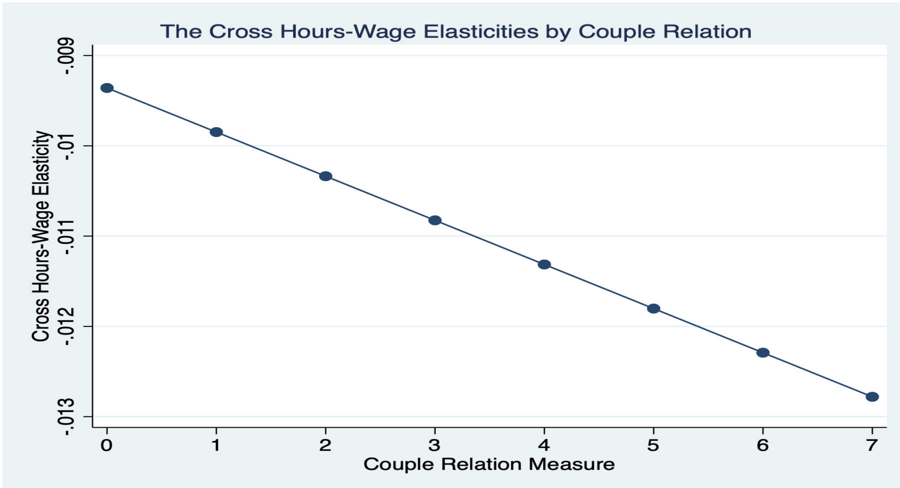

Fig. 6. The Cross Hours-Wage Elasticity by Couple Relation

5.3. Testing Backward-Bending Labour Supply Hypothesis

5.3.1. Quantile Pooled OLS Regression

The quantile pooled OLS regression gives the hours-wage elasticities for five different quantiles of married women's hours of work (Table 8). The hours-wage elasticity decreases as hours of work increase (i.e., as wages increase), holding relevant factors constant, which implies a backwards-bending labour supply curve. Since I am holding non-labour income constant in the model, the increase in the wages induces two effects: the pure substitution effect and the pure income effect. The former is due to the relative price of time spent on housework increases and therefore married women may substitute away from housework and towards working outside the home. Working hours increase as a result of this substitution effect. On the other hand, the higher wages one obtains, the richer she is. The pure income effect states that she can afford to spend more time on her dedication to family. When the wages are low, the substitution effect dominates the income effect and thus she works more. However, when the wages are very high, the income effect may be as large as the substitution effect and thus she is not so positively responsive to wages as when the wages are low because the substitution effect is offset by the income effect. When the wages are extremely high, the income effect may dominate the substitution effect and thus the married women may work less and the hours-wage elasticity is negative.

The quantile regression verifies that when married women earn higher wages, they work longer, but at a diminishing rate. In other words, they are less sensitive to wage changes than people earn low wages. This supports the backwards-bending labour supply curve hypothesis.

Another reason for a diminishing hours-wage elasticity is that when married women earn a very high wage and work for a very long time. They will reach a ceiling of time. They could no longer increase their working hours no matter how higher wages they earn. Time restriction provides a complementary explanation for a decreasing labour supply elasticity.

Quantile	0.1	0.25	0.5	0.75	0.9
Wife's weekly hours of work	20	32	40	40	50
For whole sample:					
Hours-wage Elasticity	0.4255***	0.3066***	0.1882***	0.1072***	0.0901***
For hourly paid sample:					
Hours-wage Elasticity	0.4391***	0.3142***	0.1954***	0.0841***	0.0744***

Notes: Control variables are same as in the pooled OLS model in section 4.3.1.
*p<0.05; **p<0.01; ***p<0.001

Table 13: The Hours-wage elasticities for five different quantiles

5.3.2. Panel Fixed Effects 2SLS for Subgroups

Applying the panel fixed effects 2SLS model to three subgroups, I estimate the causal effects of wages on hours of work for the part-time, full-time, and over-time married women. The three subgroups of people have different hours-wage elasticities: 0.2615 for part-time female workers who work no more than 30 hours a week; 0.0160 for full-time female workers who work between 30 and 48 hours; -0.0432 for over-time female workers who work more than 48 hours. The latter two estimates are not significantly different from zero at any conventional confidence intervals, but the diminishing trend of the labour supply elasticity can be seen clearly.

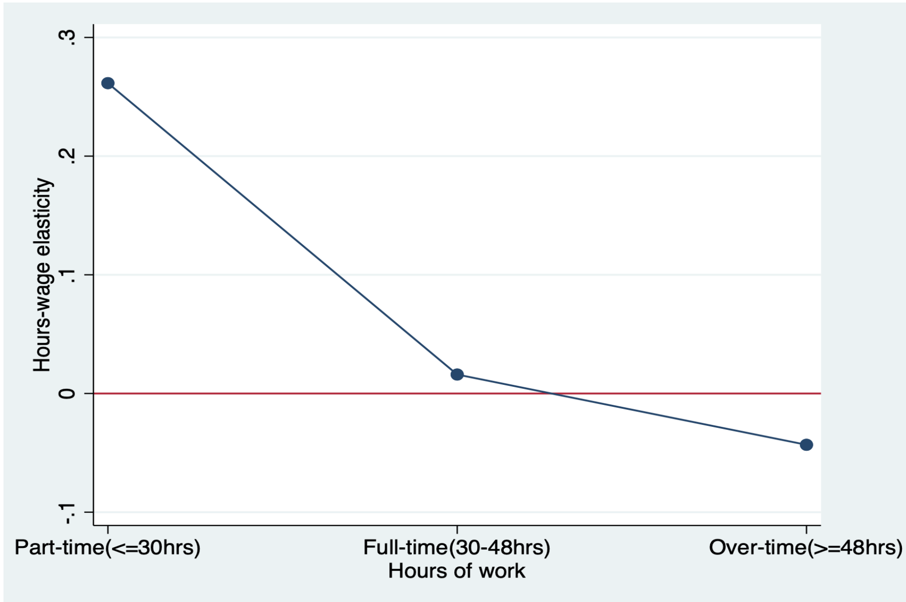

Fig. 7. The Wife's Hours-Wage Elasticities of Different Groups

5.3.3. Panel Fixed Effects 2SLS with Quadratic Term

For the whole working-age married women panel data model, I add the quadratic term of wife's wages and see how the hours-wage elasticity changes with wife's wages, holding other factors constant. There is evidence that as wages increase, the causal effects of wages on labour supply decrease towards zero, which supports the hypothesis of a backwards-bending labour supply curve.

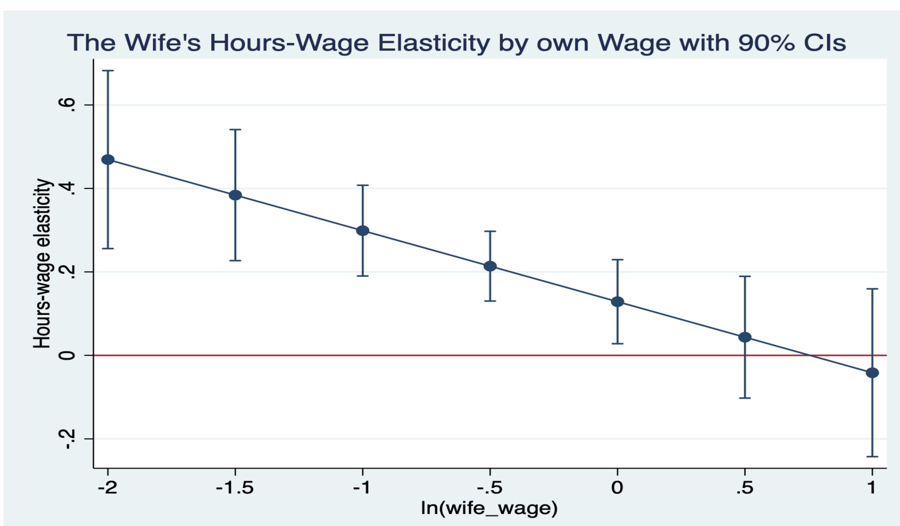

Fig. 8. The Wife's Hours-Wage Elasticities by Own Wages

5.4. The Effects of Gender Wage Inequality

The gender wage inequality is defined as the logarithm of the wife's relative wage to her husband, $ln(wife'swage/husband'swage)$.

The panel data model:

$$\begin{aligned} ln\_wife\_work\_hours_{it} = \beta_0 &+ \beta_1 ln\_relative\_wage_{it} \\ &+ \beta_3 willingness\_to\_work_{it} + \mathbf{Z_{1it}}\alpha \\ &+ a_i + \varepsilon_{it} \end{aligned}$$

As shown in the first row of Table 9, the gender wage inequality or the wife relative wage has a positive effect on married women's labour supply. After holding control variables fixed and accounting for sample selection bias, unobserved individual heterogeneity, and unobserved idiosyncratic omitted characteristics, one unit increase in the gender wage inequality leads to approximately 13% decrease in hours of work for married women on average. The value of the gender wage inequality measure in the sample is from -6.91 to 11.55, with the mean -0.44 and the standard deviation 1.12. If the gender wage inequality increases by one standard deviation, the hours of work would increase by 14.5% on average. Since the wage inequality is defined as the logarithm of the wife's relative wage, we can also have the following interpretation of the estimated coefficient: As wife's relative wages increase by 1%, her hours of work increases by 0.13% on average, holding other factors constant.

The second row of Table 9 gives the causal effects of gender wage inequality on wife's relative hours of work. The wife's relative hours of work is defined as the logarithm of the ratio of wife's hours of work to her husband's hours of work. It can be seen as the gender labour supply inequality. The value of it is from -4.38 to 3.91, with a mean of -0.18 and a standard deviation of 0.51.

The causal effects of gender wage inequality on the gender labour supply inequality can also be called the relative hours-wage elasticity. The estimates of the relative hours-wage elasticity are close to each other using different specification, as shown in the second row of Table 9. As the wife's relative wages increase by 1%, her relative hours of work supplied increase by 0.2%-0.3%.

	Pooled OLS	Pooled OLS with Sample Selection	Panel Random Effects	Panel Fixed Effects	Panel Random Effects 2SLS	Panel Fixed Effects 2SLS
The partial effects of wage inequality on hours of work	0.2148*** (0.0073)	0.2008*** (0.0093)	0.1841*** (0.0102)	0.1738*** (0.0118)	0.1256** (0.0479)	0.1329 (0.0752)
The partial effects of wage inequality on relative hours of work	0.2911*** (0.0078)	0.2946*** (0.0102)	0.2773*** (0.0113)	0.2584*** (0.0062)	0.2147*** (0.0561)	0.2726*** (0.0849)

Notes: Control variables are same as Table 4.
Robust standard errors are in the parentheses.
*p<0.05; **p<0.01; ***p<0.001

Table 14: The Causal Effects of Gender Wage Inequality on Hours of Work and Relative Hours of Work for Married Women

6. Robustness Analysis

6.1. 2005 Cohort vs 1995 Cohort

The above analysis is about the 2005 cohort, the married women who were between 17 and 55 years old in the year 2005 and were surveyed for ten years. Now I run the regressions for the 1995 cohort, the married women who were between 17 and 55 in the year 1995 and were surveyed for ten years. The comparison of the estimates for the key variables is shown in Table 10.

As far as the hours-wage elasticities are concerned, the 1995 cohort share the same pattern as the 2005 cohort. As endogeneity problems are solved using panel data fixed effects and instrumental variable two-stage-least-squares, the estimated hours-wage elasticity goes down from 0.23 in the pooled OLS to 0.16 in the panel fixed effects 2SLS model.

When it comes to the causal effect of the wife's relative wages on her relative hours of work supplied, it seems that it was smaller for the 1995 cohort.

	Pooled OLS	Pooled OLS with Sample Selection Correction	Panel Random Effects	Panel Fixed Effects	Panel Random Effects 2SLS	Panel Fixed Effects 2SLS
The 2005 Cohort						
Dependent Variable: Ln(Hours of Work)						
Ln(Wife's Wage)	.2749***	.2749***	.2734***	.2686***	.1888***	.2123***
Husband's wage	-.0534***	-.0537***	-.0375***	-0.0032	-.0429***	-0.0036
Couple Relation	-.0064***	-.0066***	-0.0035	0.0006	-.0060***	0.0005
Number of Children	-.0271***	-.0275***	-.0282***	-.0338***	-.0326***	-.0355***
The 1995 Cohort						
Dependent Variable: Ln(Hours of Work)						
Ln(Wife's Wage)	.2346***	.2345***	.2118***	.1830***	.1675***	.1553***
Husband's wage	-.0757***	-.0708***	-.0541***	-.0268*	-.0607***	-0.0317**
Couple Relation	-.0036*	-0.0024	-0.0024	-0.0019	-.0043**	-0.0032
Number of Children	-.0261***	-.0225***	-.0312***	-.0466***	-.0366***	-.0508***
The 2005 Cohort						
Dependent Variable: Ln(Relative Hours of Work)						
Ln(Wife's Relative Wage)	0.2911***	.2946***	.2772***	.2583***	/	/
Couple Relation	0.0019	0.0012	0.0033	.0061*	/	/
Number of Children	-.0295***	-.0308***	-.0358***	-.0469***	/	/
The 1995 Cohort						
Dependent Variable: Ln(Relative Hours of Work)						
Ln(Wife's Relative Wage)	.1877***	0.1559***	.1437***	.1274***	/	/
Couple Relation	0.0009	.0087***	.0063*	0.0026	/	/
Number of Children	-.0316***	-.0129*	.0063***	-.0449***	/	/

Notes: Other control variables are same as Table 4 and 5.
*p<0.05; **p<0.01; ***p<0.001

Table 15: The 1995 cohort vs 2005 cohort

6.2. *PSID vs CPS*

6.2.1. *CPS*

The PSID is a nationally representative panel household survey on roughly 5000 families that began in 1968. Information on the same families and individuals were collected annually from 1968 to 1997 and biannually since then. It returned to an annual survey from 2013.

While PSID constructs a long panel data set with a time interval of 2 years, the Current Population Survey (CPS) Outgoing Rotation Group (ORG) is basically a monthly repeated cross-sectional survey that has a special rule of survey. The respondents were interviewed for four consecutive months, ignored for 8 months, and then interviewed for 4 more months.

In this research of working-age married women (17 to 65 years old), the PSID sample size is 9,749 for 6 panels from 2005 to 2015. The sample size of CPS is 58,070 over the same period of time, which is much higher. In the CPS ORG sample, the respondents gave information about the hourly wage and the hours worked at this wage, which is are more accurate measures for the labour supply and wage rates than PSID.

The key variables used in models:

(1) Wife's hours of work: The total number of hours the wife who is paid hourly usually works per week at her main job. The dependent variable is the logarithm of it.
(2) Wife's hourly wage: This reports how much the wife earned per hour in the current job. The logarithm of it is the explanatory variable of interest. The coefficient of it is the hours-wage elasticity or the causal effects of wages on labour supply of the married women.
(3) Husband's hourly wage.
(4) Wife's educational attainment. There are four categories: high school, some college, bachelor's degree, master's degree or higher.
(5) Wife's age.
(6) Wife's race. There are four categories: White, African American, Asian, and other.
(7) Number of own children less than 18 years old.
(8) Age of the youngest own children in household.
(9) Union: A dummy variable that indicates whether, for the current job, the respondent was a member of a labour union or covered by a union contract.
(10) Year dummy variable, month dummy variable, and state dummy variables. These dummies capture the year effects, month effects, and the state effects. In other words, the macro economic policies that change over years but influence individuals the same way (individual-constant and time-varying) have been captured by the year dummies. On top of that, the month dummies represent the seasonal difference, and the state

dummies represent the geographic factors that are different between states but constant over time. So the year, month, and state heterogeneity have been controlled for in the model.

The OLS is to regress the logarithm of the wife's hours worked per week on the explanatory and control variables listed above.

The sample selection correction is performed by 2 steps. The first step is to run a selection regression of the labour force participation rate on a set of determinants of the participation, and obtain the predicted probabilities of labour force participation for each individual. The second step is to add the predicted probabilities of labour force participation into the OLS model so that the hours-wage elasticity is estimated holding the willingness to work constant. By this way can we alleviate the sample selection bias.

For the third specification, the instrumental two stage least squares (2SLS), the instrumental variables are the dummy variable that indicates whether the job is under a union contract, the state-level monthly unemployment rates, and the state-level annual minimum wage rates. The instruments of unemployment rates and minimum wage rates are not significant in the first stage of the 2SLS, which implies that they are weak instrumental variables. The union dummy is a strong IV. On the other hand, the state-level unemployment rates and minimum wage rates are more likely to meet the exogeneity requirement than the individual-level union dummy.

Dependent variable: Ln(wife_work_hours)	OLS	OLS with sample selection correction	Instrumental 2SLS
Explanatory variable:			
Ln(wife_wage)	.1934***	.1929***	.4144***
Ln(husband_wage)	-.0801***	-.0808***	-.1254***
Number of children	-.0210***	0.0056	0.0051
Education Attainment:			
High School	(base)	(base)	(base)
Some College	-.0422***	-.0418***	-.0823***
Bachelor	-.1039***	-.1035***	-.1961***
Master+	-.1603***	-.1602***	-.2939***
Number of Obs.	58,070	58,070	58,070

Notes: Other control variables include year dummies, month dummies, state dummies, the wife's age, the age of the youngest child, the wife's race.
*p<0.05; **p<0.01; ***p<0.001

Table 16: The estimates using CPS data for the hourly paid female workers

6.3. *Dynamic Model*

Dynamic models allow us to study how past values of the regressors and the outcome variable affect current outcomes.

In this dynamic model of the labour supply, I add the first lagged term of the dependent variable in the model to control for the dynamic nature of the data (the AR(1) form). The reason for doing this is due to a fairly high serial correlation of the dependent variable.

The serial correlation of ln(wife_wage)	Ln(wife_wage)	L1.Ln(wife_wage)	L2.Ln(wife_wage)
Ln(wife_wage)	1.00		
L1.Ln(wife_wage)	0.71	1.00	
L2.Ln(wife_wage)	0.61	0.73	1.00

Table 17: The serial correlation of the dependent variable

6.3.1. *Specification*

$$
\begin{aligned}
ln\_wife\_work\_hours_{it} = \beta_0 &+ \beta_1 ln\_wife\_work\_hours_{i,t-1} \\
&+ \beta_2 ln\_wife\_wage_{it} \\
&+ \beta_3 willingness\_to\_work_{it} + \mathbf{Z_{1it}}\alpha \\
&+ a_i + \varepsilon_{it}
\end{aligned}
$$

$$
\begin{aligned}
&E(a_i\varepsilon_{it}) = 0 \ for \ all \ t \\
&E(\varepsilon_{it}\varepsilon_{is}) = 0 \ for \ t \neq s \\
&E(\varepsilon_{it}) = 0 \ for \ all \ t
\end{aligned}
$$

The lagged value of the outcome variable is added into the model as an explanatory variable. We assume the unobserved time-invariant random component a_i is uncorrelated with the time-varying random error component ε_{it} and the time-varying random error components are neither serially correlated over time nor correlated with the explanatory variables.

6.3.2. *Endogeneity Problem*

(1) The unobserved individual heterogeneity – the first source of the endogeneity.

The unobserved time-invariant component must be correlated with the explanatory variables in the dynamic model because by construction it is always true that the lagged term of outcome variable is correlated with a_i. This generates an endogeneity problem.

As discussed in section 4.3.3, this endogeneity problem caused by the unobserved time-invariant factor can be solved by eliminating it using first differencing or demeaning methods. However, by doing so, we introduce a new endogeneity problem in dynamic models.

(2) The dynamic structure – the second source of the endogeneity.

For the first-differencing (FD) method,

$$\begin{aligned}\Delta ln\_wife\_work\_hours_{it} = {} & \beta_1 \Delta ln\_wife\_work\_hours_{i,t-1} \\ & + \beta_2 \Delta ln\_wife\_wage_{it} \\ & + \beta_3 \Delta willingness\_to\_work_{it} + \Delta \mathbf{Z_{1it}} \alpha \\ & + \Delta \varepsilon_{it}\end{aligned}$$

The correlation between the lagged values of the outcome variable and the time-varying error term is not zero. As a result, the estimates for the hours-wage elasticity β_2 would be biased. This is due to the dynamic structure of the model.

$$\begin{aligned}& E(\Delta ln\_wife\_work\_hours_{i,t-1}, \Delta \varepsilon_{it}) \\ = & E[(ln\_wife\_work\_hours_{i,t-1} - ln\_wife\_work\_hours_{i,t-2})(\varepsilon_{it} - \varepsilon_{i,t-1})] \\ = & E(ln\_wife\_work\_hours_{i,t-1}, \varepsilon_{i,t-1}) \\ \neq & 0\end{aligned}$$

6.3.3. Solution to Endogeneity Problem

While the first differencing method is employed to solve the endogeneity problem due to the individual heterogeneity, the instrumental variable method could be used to tackle the endogeneity problem due to the dynamic structure.

A natural candidate of IVs for the first lag of the dependent variable is the second lag of the dependent variable. The validity of the instrument comes from the fact that the time-varying random error component is uncorrelated with past values of the dependent variable. If this is the case, we have

$$E(ln\_wife\_work\_hours_{i,t-1}, \varepsilon_{it}) = 0$$

Thus,

$$E(ln\_wife\_work\_hours_{i,t-2}, \Delta\varepsilon_{it}) = 0 \; and$$
$$E(\Delta ln\_wife\_work\_hours_{i,t-2}, \Delta\varepsilon_{it}) = 0$$

Both the second lag of the dependent variable and the first difference of the second lag could be used as instrument variables for the first differencing of the first lagged value of the dependent variable because they satisfy the exogeneity requirement. Besides, the dependent variable is serially correlated over time, so the relevance condition is very likely to be met. It could be verified by the first stage of the 2SLS. Since the IVs come from the moment conditions, they are referred to as GMM-type instruments.

6.3.4. Results and Interpretation

The results from the three specifications of the dynamic panel data model are presented in Table 13. The first specification is just to use the first-difference of each variable of the dynamic panel data model. The second specification is to apply the second lagged value of the dependent variable as the instrument for the first-difference of the dependent variable, which is a regressor in the dynamic model. The third specification is to use the first-difference of the second lag as the instrument. In the second and third specifications, the standard instruments for the explanatory variable of interest, the wife's wage, are also used along with the GMM-type instruments for the dynamic term of the model, the lag of the wife's hours of work.

The hours-wage elasticities estimated using the dynamic models is similar to those using the static models.

Dependent Variable: Ln(wife_work_hours)	Panel FD	Panel FD 2SLS 1	Panel FD 2SLS 2
Explanatory Variables:			
First lag of Ln(wife_work_hours)	-0.3072***	0.1131	0.0171
Ln(wife_wage)	0.2589***	0.2120*	0.1870*
Ln(husband_wage)	-0.0197*	-0.0059	-0.0129
Couple Relation	0.0033	-0.0009	-0.0010
Number of children	-0.0429***	-0.0319**	-0.0302*
Age of youngest child	0.0016	0.0017	0.0021
Wife health ranking	-0.0189**	-0.0145*	-0.0253**
Wife housework hours	-0.0009	-0.0016*	-0.0006
Number of obs.	4798	3768	2487

Notes: Other control variables include: wife's age, wife's education, family transfer income, family wealth.
Robust standard errors are in the parentheses.
*p<0.05; **p<0.01; ***p<0.001

Table 18: The Estimates of Panel Data Dynamic Models for the 2005 Cohort

7. Conclusion