Individual and family labour supply in Thailand: A structural discrete hours approach

Abstracts

Shinawat Horayangkura[1]

Thailand is a middle-income country with several serious economic concerns, e.g., the middle-income trap, high income and wealth inequality. Understanding labour supply behaviour is important to overcome these problems. In addition, studying labour supply in the Thai context contributes to the literature on discrete hours labour supply in developing countries which is very limited.

The discrete choice approach applied in this study provides advantages over continuous models; for example, discrete hours models are more realistic since people usually have a limited number of job contracts with required hours worked. Discrete choice models also offer a simple way for estimating utility-maximising labour supply (the utility function relies on consumption-leisure preferences) in the presence of highly non-linear, due to taxes and transfers, and possibly non-convex budget constraints. Moreover, with structural techniques, the discrete choice approach permits researchers to incorporate observed and unobserved heterogeneity into empirical models.

Large cross-sectional datasets from Thailand Household Socio-Economic Survey are acquired for this study to yield precise estimates of labour supply for subgroups. Over, nine

[1] Department of Economics, University of Sheffield: shorayangkura1@sheffield.ac.uk

thousand individuals and fifteen thousand households are applied in estimation. This study estimates several models of labour supply for individuals and households. They have different degrees of flexibility based on observed heterogeneity, random heterogeneity, and alternative specific constants.

Overall, the estimated results are consistent with theory and previous empirical works. However, due to quadratic utility functions and structural equations, the estimates of labour supply models cannot be directly interpreted. Marginal utilities of arguments (consumption and leisure hours) for each agent, i.e., individual or family, are calculated. The mean values and negative proportions are then calculated in order to examine the performance of each model. The result from each model is tested to what extent the utility function is quasi-concave and monotonic. The preferred models for individual and family labour supply are selected based on these criteria and the log-likelihood test. This study finds that the model with a high degree of flexibility is in general more suitable to the data than less flexible models. In addition, the results for unmarried males are more consistent with the modelling assumptions than those for unmarried females.

JEL classification: J22; C25; C52; H31

Key words: Labour supply, Discrete choice models, Simulated maximum likelihood

1 Background

Thailand has been facing serious economic difficulties including low economic growth, the middle-income trap, high income inequality and high wealth inequality. Thai governments have been attempted to deal with these economic issues with several policies such as increasing minimum wages, revising personal income brackets, and providing welfare supports. These policies affect labour supply behaviour and subsequently lead to economic structural changes, e.g., economic growth and income inequality. Investigation of labour supply hence helps understanding how people make a decision on working and leisure; this is also beneficial for policy formation and evaluation.

Existing empirical works usually focus on labour supply in developed countries. Only Gong and van Soest (2002) studied married female labour supply in Mexico City. Most of previous studies usually investigate either individual or family labour supply. Labeaga et al. (2008) is only a paper covers labour supply at individual and household levels.

This empirical research studies both individual and household labour supply in a developing country, Thailand. It can be expected that people in countries with a different degree of economic development possibly have different labour supply behaviours. This research contributes to the existing body of literature by adding a relative comprehensive labour supply study to a limited area of studies.

Studying labour supply in a household (family) context provides a number of important extents since many tax and benefit policies are designed to have impacts on labour supply behaviour; these policies can be appropriately understood within a household labour supply framework (Blundell and Macurdy, 1999). Studying individual labour supply is also significant; the number of single individual households has increased as the demographical structure is transiting to an ageing society with a low birth rate. In fact, the percentage one-person households by the number of households have doubled in last three decades (1987-2013); and this type of family is different from multiple-member households (The United Nations Population Fund Thailand and the Office of the National Economic and Social Development Board., 2015).

This paper applies the discrete hours labour supply models, to investigate labour supply behaviour in Thailand. Several models with different degrees of flexibility for individual and household labour supply are estimated to examine the performance so that this study can find the most appropriate model for individual and household labour supply that fits the Thai dataset. Later, the additional robustness tests are also performed to see the consistency of the model.

In the next section, economic theory and econometric methodology are explained. Then, description on empirical data is presented. The results are provided in the penultimate section. The last section includes the summary of the study and future research suggestions.

2 Theory and methodology

2.1 The discrete and continuous hours labour supply models

The discrete hours choice labour supply model based on the Random Utility Maximisation (RUM) approach has been a dominating paradigm (Aaberge and Colombino, 2014). There are several advantages why it became more popular than the continuous hours models in labour supply literature.

Firstly, the discrete hours approach can be considered as more realistic, compared with the continuous hours counterpart, because, in practice, a limited number of hours worked (part-time and full-time jobs) are available. This is supported by the observed peaks in hours distributions in many countries. This coincides with the assumption of the discrete models in which agents are able to select a relatively small number of hours choices rather than being allowed to alter hours work continuously. From this perspective, the continuous approach is just an approximation to a discrete optimisation problem (Gong and van Soest, 2002).

Secondly, the discrete hours models require considerably less complex information of the budget set faced by each agent. The models simplify the information by estimating utility at a small number of hours levels, none of which is a standard tangency solution (Creedy and Kalb, 2005). By using utility maximisation with a discretised budget set, the discrete hours approach prevents the complexities arising from a non-linear budget constraint characterised by non-linear tax and benefit systems (Creedy and Duncan, 2002). On the other hand, the continue hours models must cope with the complex issues occurring from budget sets having both convex and non-convex ranges as well as multiple local equilibria when it analyses choices under piecewise-linear budget lines. Hence, it is cumbersome to evaluate the complete range of each agent's unique budget set in practice where most tax and transfer structures are complex. This issue becomes further complicated when investigating family labour supply with joint utility maximisation due to the three-dimensional budget constraint of households.

Thirdly, the discrete hours approach with a structural modelling technique early developed by van Soest (1995) and Keane and Moffitt (1998) allows researchers to incorporate random heterogeneity into the models.

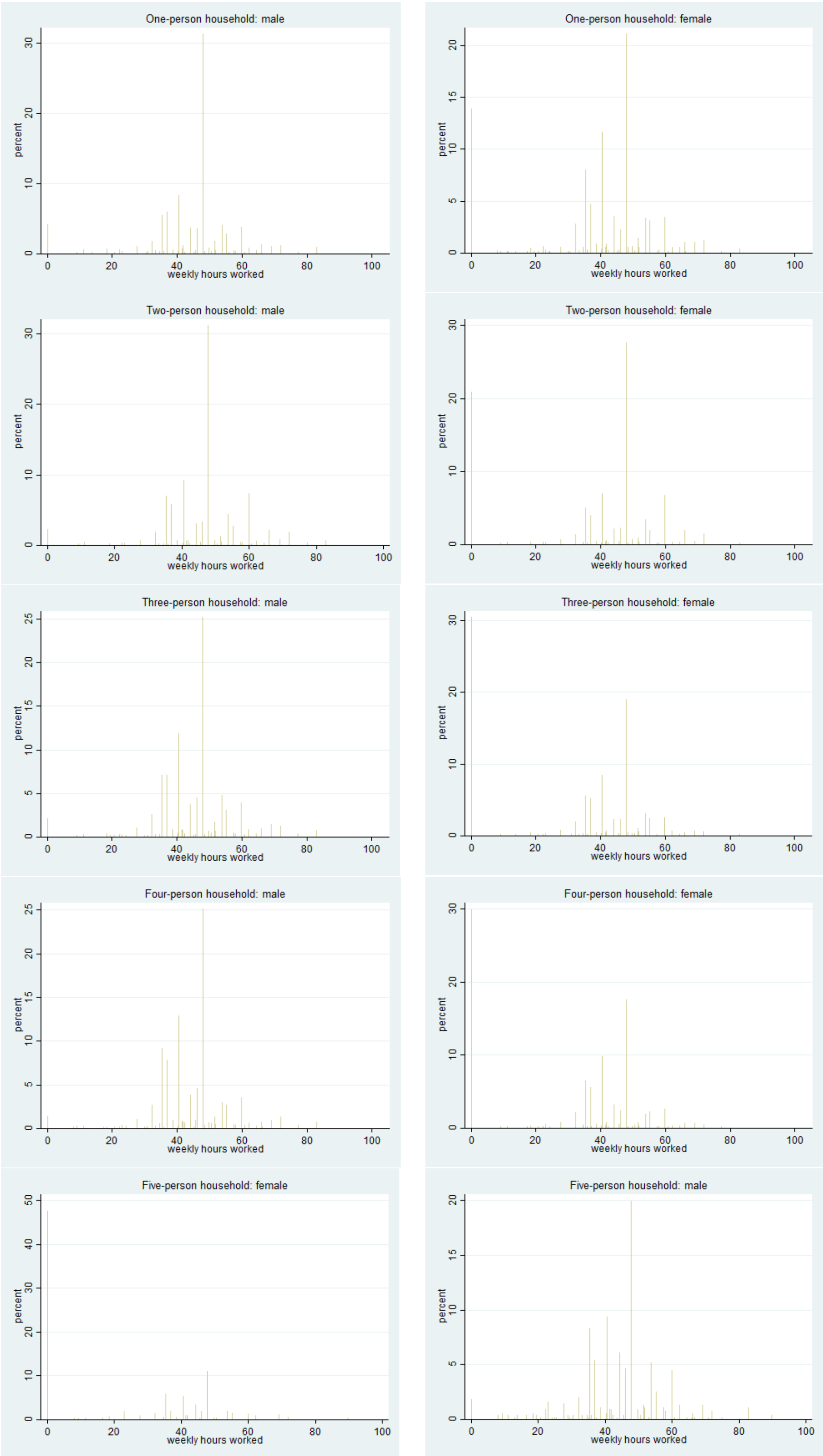
In a nutshell, the discrete hours models are less complicated than the continuous hours models in incorporating taxation, social security, and social welfare details in estimation as well as simulation.

Disadvantages of a limited number of choice opportunities are the rounding error, the difference between numbers resulting from calculated approximation and exact mathematical computation due to rounding, the incomplete utilisation of available information (van Soest, 1995).

The major distinguishing feature between the continuous labour supply models and the discrete hour choice labour supply models is that in the latter is that the choice set is discretised, i.e., each responding decision unit (individual or family) is assumed to be able to choose among a list of alternatives, $j = 1,2, \dots J$, in the choice set of combinations between income and leisure hours (van Soest, 1995). In addition, the discrete hours models assume all agents (either individuals or households) have the same available set of hours alternatives (Aaberge and Colombino, 2014).

The total number of alternatives faced by each individual are equivalent to 11 hours points which are identical to what used in Creedy et al. (2002) and Duncan and Harris (2002). This is consistent with the hours distributions in Thailand data as exhibited in FIGURE 1. Thus, in the individual labour supply setting, people are assumed to have 11 leisure hours choice opportunities; in the household labour supply, each family is allowed to select one out of 121 combinations of husband and wife leisure hours.

FIGURE 1: The distributions of hours worked for individuals and couples by gender



Hence, each individual, regardless individual and household labour supply, is assumed to maximise utility function over a set of discrete hours points $L_j \in \{L^1, L^2, \dots, L^J\}$. The observability rule for a discrete hours choice labour supply in this paper is:

$$
\begin{aligned}
L_j &= L^1 \text{ if } L \leq L_1^B \\
&= L^2 \text{ if } L_1^B < L \leq L_2^B \\
&\dots\dots\dots\dots\dots\dots \\
&= L^{J-1} \text{ if } L_{J-2}^B < L \leq L_{J-1}^B \\
&= L^J \text{ if } L > L_{J-1}^B
\end{aligned}
$$

2.2 Discrete choice model in labour supply

The basic discrete choice model applied in labour supply studies is known as the conditional logit model. Based on RUM, a unit of analysis (either individual or household level) maximises utility with respect to a budget constraint which is a function of leisure and income. The general form of the utility functions for individual and household levels is as follows:

$$U_{nj} = V_{nj}\left(Y_{nj}, L_j;\ Z_i, \beta\right) + \varepsilon_{nj} \quad (1)$$

where, U_{nj} is the utility of an individual n received from an alternative j; this study assumes the quadratic utility function in estimation; Y_{nj} represents the net income of an individual n who selects an alternative j; L_j is leisure hours an individual chooses which is assumed to be equivalent to $TE - H_j$ (TE is the total time endowment[2]; and H_j is hours worked from an alternative j); Z_n is a vector of individual characteristics; β is a vector of coefficients; and ε_{nj}

[2] In many empirical works assumed TE equal to 80 hours per week. However, this paper assumes TE to be 100 hours per week because a fair number of Thai people work longer than those in previous studies. Different TE do not affect the marginal utilities.

is an error term which reflects idiosyncrasies of an individual n on alternative j, this unsystematic component is usually assumed to be an independent and identically distributed (i.i.d) random variable with an extreme value type I distribution.

The utility of a household is more complicated than it of an individual since more arguments are included in the utility function as shown below.

$$U_{gj} = V_{gj}\left(Y_{gj}, L_{mj}, L_{fj};\ Z_g, \beta\right) + \varepsilon_{gj} \quad (2)$$

U_{gj} indicates the utility of a household g which opts an alternative j. However, it is clearly seen that two leisure arguments appearing in the utility function; they are leisure of males, L_{mj}, and females L_{fj}. Z_g is designated as a vector of household characteristics whilst β is a vector of coefficients. ε_{gj} which captures unobservable preferences of a household g which affect the utility from $j-th$ choice is an extreme value type I i.i.d. stochastic component.

On the other hand, the budget constants of individuals and households are expressed as follows:

$$Y_{nj} = w_n * H_j + y_n - T(w_n, H_j, y_n, Z_n) \quad (3)$$

$$Y_{gj} = w_m * H_m + w_f * H_f + y - \ T(w_m, H_m, w_f, H_f, y_g, Z_g) \quad (4)$$

where T represents the income tax which depends on arguments in parentheses. For an individual, the tax is calculated from individual wage, working hours, other unearned income, and individual characteristics. Considering a household agent, computation of the income tax accounts both partners' wages (male and female), working hours of both, household non-working income, and characteristics of the household.

Hence, the agent's problems for individuals and households take the following form.

$$Max\ U_{nj} \text{ subject to } Y_{nj} \leq w_n * H_j + y_n - T(w_n, H_j, y_n, Z_n) \tag{5}$$

$$Max\ U_{gj} \text{ subject to } Y_{gj} \leq w_m * H_m + w_f * H_f + y - T(w_m, H_m, w_f, H_f, y_g, Z_g) \tag{6}$$

The solution to these equations is complicated as $T(w_n, H_j, y_n, Z_n)$ and $T(w_i, H_j, y_i, Z_i)$ are non-linear, i.e., they make Y_{nj} and Y_{gj} non-linear as well. The optimisation for a given marginal tax rate is always possible and a parametric Marshallian labour supply function is always obtained; the discrete choice approach begins by utility specification and coefficient estimation instead of estimating the Marshallian labour supply parameter as done in continuous models (Labeaga et al., 2008).

Based on McFadden (1974) the probability of an alternative j being selected by an agent a is expressed as follows:

$$p_{aj} = P\left(U_{aj} \geq U_{ai}\right) = \frac{\exp\left(V_{aj}\right)}{\sum_{i=1}^{M} \exp\left(V_{ai}\right)}; i = 1, \dots, M \tag{7}$$

This is the standard conditional logit to estimate discrete hours choice labour supply models. However, by allowing random heterogeneity in the model, one needs the mixed logit model in the estimation. this paper hence applies 100 draws from Halton sequence as suggested in Train (2009). Examples of empirical works applying this estimation process are Bargain et al. (2014) and Kabátek et al. (2014).

2.3 Predicted wages for non-working individuals

Empirical studies on labour supply have a common issue as found in research on labour force participation; that is unobserved wage rates for labour market non-participants. This paper follows Kabátek et al. (2014) by applying the Heckman selection model pioneered by

Heckman (1979) to deal with the predicted wage issue. Then, the predicted hourly wages are filled for individuals with missing wages as well as those with observed wages. The sensitivity of the estimated results is tested by replacing predicted wage rates on non-workers whose wages are unobserved only.

2.4 Modelling non-participation

Creedy and Kalb (2005) state that the basic discrete hours labour supply model which does not allow for demand-sided restrictions, provides unwell-presented labour supply at some discrete hours points. In fact, labour supply models without taking account for non-participation (i.e., characteristics of hours points, hours restrictions, or fixed costs) typically yield under-predicted values for the number of non-workers and over-predicted values for the number small part-time jobs (Gong and van Soest, 2002, van Soest et al., 2002).

This paper selects includes alternative specific constants in order to enhance goodness of fit to the observed labour supply. The alternative specific constant terms for working alternatives represent pecuniary costs, e.g., search, transportation, and childcare costs, as well as non-pecuniary costs, e.g., disutility of working.

To incorporate new ad hoc terms into the discrete hours choice labour supply model, a dummy variable whether an individual is working is included in the individual direct utility function. With regard to the household labour supply model, three dummies for working males, working females, and working both partners are added into the household direct utility function.

2.5 Observed and unobserved heterogeneity

Observed heterogeneity incorporated into labour supply modes through observed socio-economic characteristics. This study allows observed taste preferences for income, leisure hours, squared income, squared leisure hours, and working alternative specific constants. Unobserved heterogeneity is allowed for income, leisure hours, and alternative specific constant for working.

This paper modifies Creedy et al. (2002) strategy in parameterisation. In fact, preference variation across agents is introduced into parameters of income and leisure hours at linear terms, squared terms, and working alternative specific constants

The functions of labor supply parameters incorporated observed and unobserved heterogeneity are shown below:

$$\alpha_{YY_a} = \alpha_{yy0_a} + \alpha'_{yy_a} Q_{yy_a} \tag{8}$$

$$\alpha_{LL_s} = \alpha_{ll0_s} + \alpha'_{ll_s} Q_{ll_s} \tag{9}$$

$$\beta_{Y_a} = \beta_{y0_a} + \beta'_{y_a} Z_a + \omega_{y_a} \tag{10}$$

$$\beta_{L_s} = \beta_{l0_s} + \beta'_{l_s} Z_s + \omega_{l_s} \tag{11}$$

where, parameters with capital letter subscripts (Y and L) designate the parameter in the utility function whilst parameters with a small letter subscripts (y and l) identify coefficients in parameterisation. The parameters with subscript 0 indicates the constant terms for observed heterogeneity while the parameters with transpose represent a vector of coefficients, Q and Z designate a vector of socio-economic characteristics. The subscript s which enters for leisure related parameters indicates each person in individual labour supply and indicates whether male or female in household labour supply. A vector Q includes dummy variables for youngest child at different age ranges (younger than three years old, between three to six years old, and between six to ten years old); hence, in individual labour supply model which covers only unmarried individuals with no children does not have parameterisation for squared terms. A vector Z includes six variables (five dummies and one variable) are included in the set of characteristics in which three of them are identical to those interacting with squared terms, the rest are a number of children, a dummy if an individual is older than 40 years old, and a dummy if an individual is holding Bachelor's degree or higher; thus, for individual labour supply a vector Z has only three characteristics. The subscript a which enters for income related parameters indicates whether the parameter is at the individual or household level. The stochastic heterogeneity terms, ω, are assumed as Callan and van Soest (1996), they follow the normal distribution with zero mean with a given standard deviation and they are not correlated with any other parameter. Hence, in practice, ω is obtained from a normal distribution with the mean equal to zero; the standard deviation of the distribution is determined by the standard deviation from estimation.

In this paper, the alternative specific constants are hence allowed for observed and unobserved preference variation across individuals through socio-economic characteristics and random heterogeneity respectively. Thus, the alternative specific constants can be expressed as following functions.

$$Par_{qj} = 0$$

$$Par_{qj} = \varphi_0 + \varphi_1 BKK + \varphi_2 PreS + \varphi_3 AS + \tau_{qj} \qquad (12)$$

where, Par_{qj} represents disutility of working; the subscript q can be replaced by n for individual labor supply and m, f, and b for family labour supply; BKK is a dummy variable if an agent lives in the capital of Thailand; $PreS$ and AS designates a number of pre-school children (younger than 6 years old) and a number of children older than pre-school age (older than 6 years old); τ_{qj} is a random heterogeneity which is assumed normally distributed with zero mean as before.

However, for the individual labour supply setting, this paper focuses only on single males and females. In other words, since they do not have any child, some socio-economic characteristics, i.e., $PreS$ and AS, are not applicable for the individual labour supply context.

2.6 Quasi-concave and monotonicity condition checks

van Soest (1995) showed two necessary conditions to check quasi-concavity after obtaining the estimation result. First, marginal utility of income must be positive, so the utility function is quasi-concave (at income and leisure for the individual level and at income, male's leisure, and female's leisure for the household level). Second, the matrix of second order derivatives of income with respect to leisure (of a respondent for the individual level and of both partners for the household level) along the indifference surface at income and leisure must be a positive definite matrix. The matrix for both individual and household levels can be expressed as follows:

$$H_n = -U_{Yn}^{-1}[y_{Ln} \quad 1]\, HU[y_{Ln} \quad 1]' \tag{12}$$

$$H_g = -U_{Yg}^{-1}\begin{bmatrix} y_{Lm} & 1 & 0 \\ y_{Lf} & 0 & 1 \end{bmatrix} HU \begin{bmatrix} y_{Lm} & 1 & 0 \\ y_{Lf} & 0 & 1 \end{bmatrix}' \tag{13}$$

where H is the matrix of second order derivatives of income with respect to leisure along the indifference surface at income and leisure arguments; U_Y represents the first order derivatives of utility, U, with respect to income, Y, while HU is designated as the Hessian matrix of second order derivatives of utility with respect to income; they are a 2*2 and 3*3 matrices for individual and household level respectively; U_Y designates the partial derivatives of utility with respect to leisure hours; y is equal to $-U_L/U_Y$, i.e., the marginal rates of substitution of leisure with income; and the subscripts n, m, and l indicates individual, male, and female in a given unit of the decision maker.

In practice, H_n is a 1x1 matrix; the positive definite matrix of H_n is satisfied if and only if an element if positive. H_g is a systematic 2x2 matrix hence determining whether H_g is a positive definite is more complicated. In fact, there are a number of tests, e.g. an eigenvalue test and a pivot test, to examine whether a systematic matrix is positive definite. This paper applies a determinants test. By giving a systematic 2x2 matrix, $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$, this matrix is defined to be a positive definite matrix if and only if $a > 0$, and $ac - b^2 > 0$.

Another important condition to be checked in a discrete hours labour supply is the monotonicity of the utility function because interior points of the budget set of the discrete model are a prior excluded since the budget set becomes discrete; the model would lose the economic meaning if the monotonicity condition in which the utility increase with income does not hold (van Soest, 1995). The utility increase in income at (income and leisure) if and only if this following expression is true.

$$2(\alpha_{YY}Y + \alpha_{YL}L) + \beta_Y > 0 \tag{14}$$

$$2\left(\alpha_{YY}Y + \alpha_{YL_m}YL_m + \alpha_{YL_f}YL_f\right) + \beta_Y > 0 \tag{15}$$

These equations (14) and (15) are for checking whether the estimation results satisfy the monotonicity condition for the individual level and the household level respectively.

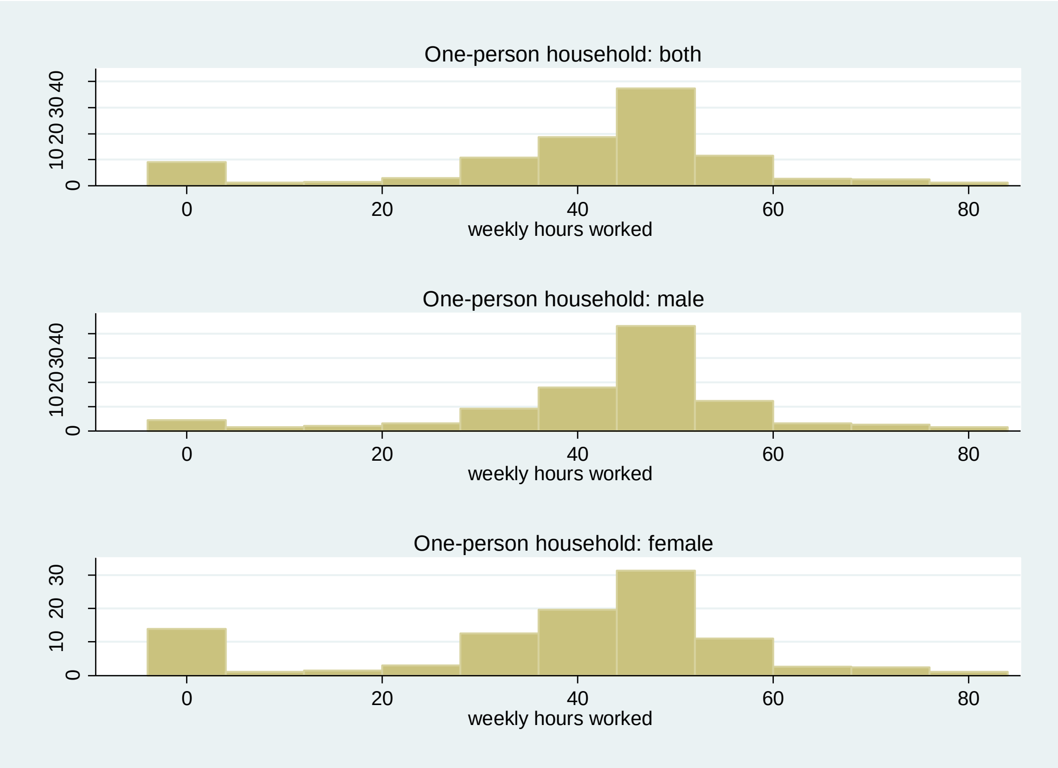
3 Empirical data

The dataset used in this paper is the same as what is in the previous one. In fact, this paper uses Household Socio-Economic Survey (HSES) which is obtained from Thailand National Statistical Office; the studied period covers 2009, 2011, 2013, and 2015.

3.1 Unmarried individuals

The total number of unmarried one-member households for analysis is 9,093 which are 4,517 males and 4,576 females. FIGURE 2: Discrete hours worked distributions for unmarried individuals exhibits hours worked of single-individual households by gender; each bar represents each hours choice as applied in the econometric model.

FIGURE 2: Discrete hours worked distributions for unmarried individuals



The top panel shows the selected hours worked of both genders; most of the people decided to work about 48 hours per week which is 37.18 percent of the whole sample. The second most sinificant proportion of people, 18.79 percent, worked 40 hours per week. The next largest group, 11.53 percent, is people who worked 56 hours per week. 9.23 percent of the whole sample selected not to work at all.

The middle panel presents hours worked of unmarried male individuals; 43.13 percent of males selected to work 48 hours per week. The second largest group, 17.85 percent, works 40 hours weekly. The next largest group which shares 12.13 percent selected to work 56 hours per week. Non-working males share 4.43 percent of the male sample.

The last panel is the histogram of weekly hours worked for females. Similar to males, the largest proportion is 48 hours per week which shares 31.30 percent. However, the second largest group of females, i.e., 19.71 percent, chose an alternative of 40 hours worked. The third largest proportion which is 13.97 percent selected to be non-participants in the labour market.

TABLE 1: Descriptive statistics of unmarried individuals by gender

	Males		Females	
Observations	4,517		4,576	
Variables	Mean	S.D.	Mean	S.D.
Predicted wage^{H}	68.00	39.96	69.50	45.72
Earned income^{Y}	151.75	93.89	140.04	107.21
Unearned income^{Y}	1.99	18.60	4.01	29.12
Disposable income^{Y}	153.14	93.75	144.22	108.91
Age	36.98	11.02	41.25	12.55
Age$\geq$ 40	0.3679	0.4823	0.5210	0.4996
Higher degree	0.2081	0.4060	0.3698	0.4828
Bangkok	0.1072	0.3093	0.0798	0.2710

^{H} Predicted hourly wage in bahts; ^{Y} yearly basis income in thousand bahts.

TABLE 1: Descriptive statistics of unmarried individuals by gender shows the descriptive statistics for important variables. The average predicted hourly wage for males and females are 68.00 and 39.50 bahts respectively. Even though females have a higher predicted wage than males on average, the yearly disposable income of males is higher (153.14 thousand baths) than females (144.22 thousand bahts). The reason is only that males, on average, spend a larger number of hours on labour supply than females do since, in general, females have a

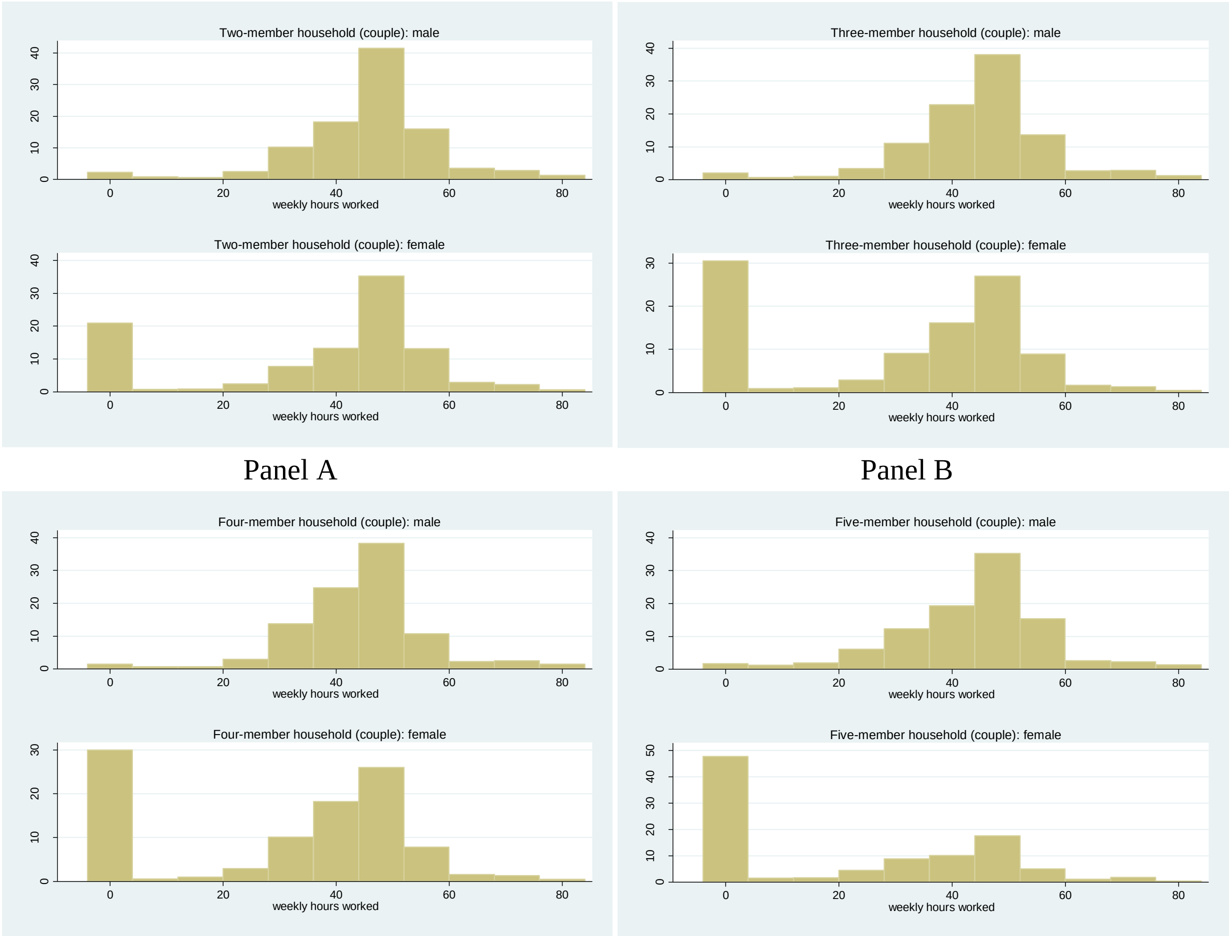
larger amount of yearly unearned income (4.01 thousand bahts) than males does (1.99 thousand bahts).

Regarding important socio-economic characteristics, the first variable is individual's age; the average age of single males is 36.98 years old whilst it of single females is 41.25 years of age. The rest of variable for characteristics are dummy variables indicating whether an individual has a considered characteristic or not. 36.79 percent of single males and 52.10 percent of single females are older than 39 years of age. In addition, 20.81 and 36.98 percent of males and females own any university degree respectively. These numbers imply why the average hourly wage of females is greater than it of males because age and education are factors determining wages. In addition, males with higher wages are more likely to get married while single females usually are those who can rely economically on themselves. 10.72 and 7.98 percent of males and females live in Bangkok, the capital city of Thailand.

3.2 Married couple households

The total number of sample for family labour supply using the discrete hours choice model is 15,578 households. There are four types of households for household labour supply estimation, namely, 7,350 of married couples without children, 4,582 households with a couple and one child, 3,089 married couples with two children, and 557 of households with three children. Other types of households, i.e., a married with more than three children are excluded because the numbers of these types of households are highly limited and the child tax-benefits in Thailand covers only three children.

FIGURE 3: Discrete hours worked distributions for couples



Panel A Panel B

Panel C Panel D

FIGURE 3 exhibits histograms of selected worked hours of four different household types by size for family labour supply analysis. Alphabetically, Panel A, B, C, D present histograms for two-member, three-member, four-member, and five-member respectively; the upper graph for each panel is the distribution for males whilst the lower one is the distribution for females. Overall, the shapes of distributions look similar across different household sizes. The most prominent spike of all males in any household type is working for 48 hours weekly.

In Panel A, 41.55 and 35.29 percent of males and females respectively select to work about 48 hours per week. The next two favourite choices for males in two-member households are 40 hours and 56 hours per week at 18.16 and 16.01 percent respectively. Only 2.24 percent of males choose not to participate in the labour market at all. On the other hand, zero worked hour per week shares 20.94 percent of females, i.e., this is the second most significant proportion. Then, 13.31 and 13.13 percent of females select to work 40 and 56 hours per week respectively.

In Panel B, The largest selected alternative for males is to work about 48 hours per week (38.04 percent) whilst the largest group of females select to be non-working (30.49 percent). With respect to males, the distribution is similar to that shown in Panel A, two-member households; two following popular choices are 40 hours (22.81 percent) and 56 hours (13.73 percent). Male non-participants are only 2.10 percent of males decide to be non-participant in the labour market. In terms of female observations, 26.95 percent decide to work about 48 hours per week; the proportion of working 48 hours weekly is just 3.55 percent below it of non-working females (zero hours worked), i.e., the percent of females working 48 hours per week is 26.95 percent. The next two most considerable proportions are those select to work 40 hours (16.11 percent) and 56 (8.93 percent). It is observable that the distribution in Panel A and B are similar.

Next, Panel C exhibits the hours worked distributions for four-member households. Similar in Panel B, while most of the males decide to work at 48 hours per week (38.26 percent), most of the females choose to provide approximately zero hours labour supply (29.98 percent). Regarding males, the next two popular for male labour supply decision makers are 40 hours (24.70 percent) and 32 hours (13.86 percent) per week respectively. A small number of males (1.49 percent) select to be inactive in the labour market. Respecting females, the second largest group of females, i.e. 26.00 percent, decides to work for 48 hours per week. Then, 18.26 and 10.13 percent of females choose to work for 40 hours and 32 hours per week respectively.

The last panel, Panel D, shows the distributions for five-member households. As same as previous panels, the most massive proportion for male observations is to work 48 hours per week (35.19 percent). The most significant group of females (47.76 percent) chooses to be inactive in the labour market, i.e., zero hours per week. In terms of males, the next two largest groups are similar to what presented in Panel A and B; they select to work 40 and 56 hours worked per week at 19.39 and 15.44 percent respectively. Only 1.80 percent of male observations are labour market non-participants. For females, the second largest proportion (17.59 percent) selects to work 48 hours per week. 10.05 and 8.80 percent of them decide to provide weekly labour supply at 40 and 32 hours respectively.

In conclusion, the distributions are similar across different household sizes, especially for male observations. Most of the males tend select to work 48 hours per week in all panels. Concerning females, when households comprise of a couple only, females tend to work. However, when households have any child, most of the females decide to be labour market non-participants. Having a child or children is presumed to have an impact on the labour supply decision. It is apparent for females that they are more likely to opt to be inactive in the labour market when they have a child or children.

TABLE 2 located below shows important descriptive statistics including pecuniary factors, such as predict wages and disposable income, couple characteristics, e.g., age and education, and household composition characteristics, e.g., the age of the youngest child and number of pre-school children.

TABLE 2: Descriptive statistics of couples by family size

	Two-member		Three-member		Four-member		Five-member	
Observation	7,350		4,582		3,089		557	
Variables	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Predicted wage^H (M)	73.64	51.59	74.75	50.00	81.91	54.58	74.13	50.21
Predicted wage^H (F)	57.41	40.36	58.74	42.17	63.99	44.15	54.60	40.72
Earned income^Y (M)	166.50	107.17	166.40	107.08	182.25	122.06	164.50	116.72
Earned income^Y (F)	108.82	95.34	98.48	103.58	108.02	109.32	70.01	98.04
Unearned income^Y	4.50	28.26	4.33	32.64	6.35	36.68	7.66	87.71
Disposable income^Y	279.86	185.64	268.04	183.59	295.35	202.53	240.94	199.87
Age (M)	40.86	11.30	40.82	9.00	41.02	7.20	40.84	6.45
Age (F)	38.68	11.31	37.66	8.97	37.82	6.78	37.11	6.27
Age$\geq$ 40 (M)	0.5146	0.4998	0.5498	0.4976	0.5801	0.4936	0.6014	0.4900
Age$\geq$ 40 (F)	0.4548	0.4980	0.4400	0.4964	0.4124	0.4924	0.3609	0.4807
Higher degree (M)	0.1921	0.3940	0.2100	0.4073	0.2671	0.4425	0.2118	0.4090
Higher degree (F)	0.2060	0.4044	0.2261	0.4184	0.2745	0.4463	0.1813	0.3856
Bangkok	0.1196	0.3245	0.1006	0.3008	0.1020	0.3026	0.1059	0.3080
Youngest child$<$ 3	-	-	0.1757	0.3806	0.1975	0.3982	0.3573	0.4796
Youngest child 3-6	-	-	0.1613	0.3678	0.1900	0.3924	0.2190	0.4140
Youngest child 6-10	-	-	0.1796	0.3839	0.2664	0.4422	0.2172	0.4127
Number of								
- children	-	-	1.00	0.00	2.00	0.00	3.00	0.00
- pre-school child	-	-	0.3370	0.4727	0.4584	0.6247	0.7307	0.7117
- school-aged child	-	-	0.6630	0.4727	1.28	0.6426	1.67	0.7018
- disabled children	-	-	0.0109	0.1039	0.0120	0.1174	0.0180	0.1329

^{H} Predicted hourly wage in bahts; ^{Y} yearly basis income in thousand bahts.

From TABLE 2, the average predicted wage of males is apperently greater than it of females. Two-member households indicate the lowest predicted wages of males at 73.64 bahts per hour whilst the four-member households have the highest average predicted wages of males at 81.91 bahts per hour. The average predicted wage of females from four-member families is also the greatest (63.99 bahts per hour) among households with different numbers of members. However, the lowest average predicted wage of females is yield from five-member households at 54.60 bahts per hour.

Earned incomes are the approximation using the predicted value of hourly wages and discretised hours worked. They are presented individually on a yearly basis. It is apparent that males earn more income than females do on average because male wages are higher than female wages. Also, hours distributions show that males tend to work longer and participate more in the market than females. Female behavioural change in labour market participation affect earned income, for example, the mean of female earned income of households without children is greater than it of household with one child although the average wage rates are similar. This is mainly because a proportion of non-working females are different between these types of households.

Unearned income, which is relatively a small amount of total income—summation between earned and unearned incomes, is however presented on a household basis. The average values of two-member and three-member households are very similar (4.50 and 4.33 thousand bahts per annum respectively). Then, when the number of children increases, the average unearned income also increases (6.35 and 7.66 thousand bahts per annum for four-member and five-member households respectively).

Disposable income, i.e., total income after tax, is the crucial variable in labour supply analysis. Disposable income is consistent with earned income since the latter shares a large proportion of the former; however, tax rules affect disposable variation as well.

In regard to socio-economic characteristics, there are several important variables relevant to this paper. The average age of males and females are similar across different types of households (ranging from 40.84 to 41.02 years old for males and from 37.11 to 38.68 years old for females). However, a dummy variable indicating if a person is 40 years old and over suggests an interesting pattern. In fact, percentages of males aged from 40 are positively correlated to a number of children whilst the correlation between the percentages of females older than 40 years and a number of children is negative.

The percentages of males who are holding any university degree are 19.21, 21.00, 26.71, and 21.18; on the other hand, the percentages of females graduated from a university are 20.60, 22.26, 27.45, and 18.13 for households with two, three, four, and five members respectively.

It is noticeable that the highest values for both genders are from four-member households; this could help explaining why average predicted wages of them for both partners are higher than others. Females from five-member families have the lowest percentage of females having any higher degree; this also implies why their average predicted wage rate is the lowest among other females. Another point is that the percentage of females holding a university degree is usually more substantial than it of males except for five-member households. The percentage of families with a couple living in Bangkok is highest (11.96) among it of other types of families (10.06, 10.20, and 10.59).

For characteristics of children, variables are not applicable for two-member households since they are comprised of a couple only. The dummy variables of the youngest child by age ranges (under three, older than 3 but younger than 6, and older than 6 but younger than 10) can be interpreted in terms of percentage. Overall, households with the larger number children have the greater probability (some of these three mean values) of having the youngest child in any of these ranges. Households with one child have the similar numbers for these variables, namely, 17.57, 16.13, and 17.96 percent for the youngest child aged under three, older than 3 but younger than 6, and older than 6 but younger than 10 respectively. However, the percentage of the youngest child is aged between 6 and 10 is highest at 26.64 for families with two children; the percentages of the youngest child in remaining groups (under three years old and between 3 and 6 years old) are 19.75 and 19.00 respectively. For the last type of families (a couple with three children), the highest percentage is the youngest child aged under three years old at 35.73; the percentages of remaining groups are 21.90 for age between 3 and 6 and 21.72 for age from 6 to 10.

The number of children is straightforward since families are differentiated by numbers of children. This is a good variable to check if all samples are categorised correctly. For three-member households, the number of pre-school children (age under six) is 0.337; the number of children older than six years old is 0.663. These numbers for this household type are, in other words, the percentages that the summation is equal to one. Regarding households with two children, the numbers of pre-school children and non-pre-school children are 0.4584 and 1.28 respectively. On average, five-member households have 0.7307 pre-school child and 1.67 non-pre-school children. The average numbers of disabled children for families with one child, two children, and three children are quite small at 0.0109, 0.0120, and 0.0180 respectively. However, having any disabled child affects the disposable income through child allowance

3.3 Calculation of the net income at each discrete point

The net incomes are not directly observed but are calculated from knowledge of the predicted hourly wage of each individual, discrete hours points, and the tax and transfer system. Obtaining net income at each discrete point for each decision maker requires five following main steps.

Firstly, after predicted hourly wages are yielded and used for all individuals in the sample, weekly earned incomes for different discrete points faced by each individual are generated by multiplication between a predicted hourly wage and weekly hours worked at each discrete hours point (11 points in total). Since income tax is levied on an annual basis, yearly earned income for different hours alternatives are calculated by multiplying weekly earned income with 52.142857 weeks, 365 days in a year divided by 7 days per week.

Secondly, the calculation for earned and unearned incomes at each hours point is taken into account. Summation between yearly earned and unearned incomes yield total yearly gross incomes (i.e., total taxable incomes). It is apparent that the total yearly gross incomes are various across hours choices due to variation in yearly earned income. Unearned incomes, on the other hand, do not vary with hours points and they enter as a household level. Therefore, they are fixed across decision makers. Because tax rules on joint tax filling changed in 2012, after than couples have options for tax filling (separate, partly-joint, and joint). Couples are assumed to be knowledgeable in these rules; they will select the option that makes them pay the least tax.

Thirdly, now the total assessable incomes after expense can be obtained. There are eight main taxable income types in the Thailand income tax system; however, they are available only four types of incomes, namely, 40(1), 40(3), 40(4), and 40(5), in Thailand HSES. Expenses for these incomes are accounted separately due to different tax rules across types of taxable incomes. This calculation is straightforward for the individual level, but it becomes more complicated for the household level due to changed tax rules of joint filling. After accounting for expenses, the final output of this stage is the total taxable incomes after expense.

Next, allowances and exemptions are deducted from the total taxable incomes yielded in the previous step. Regarding personal allowance, there is only one possibility for the individual level whilst six different scenarios of total taxable incomes after expenses and allowances due to whether they do separate, partly-joint, or joint filling. In this empirical work, disability allowance and retire-aged allowance are not applicable for the calculation as the sample does not include disabled and retire-aged individuals. For households with a child or children, a taxpayer is allowed to deduct the child allowance from the remaining taxable income; when a child is disabled, an extra amount of allowance is added. At the end of this step, the net taxable incomes at available hours points are subsequently obtained.

Lastly, tax computation for each hours points is carried out. Due to tax brackets and tax rates amendment in 2012, there are two tax calculation systems (four brackets in 2009 and 2011, and seven brackets for 2013 and 2015) in this paper. For an individual, different levels of net income are reported at different hours points; thus, taxed amounts are usually different across hours except when the net income does not exceed the exempted amount which is 150,000 baths per annum, the taxed amount is equal to zero. Regarding individual labour supply, the total amount of income tax burdens is subtracted from total gross incomes per annum for each hours point to give the net income (i.e., the disposable income). Each individual hence has 11 points of incomes for 11 hours points. For household labour supply, there are net incomes for males and females (each has 11 hours points). The summation of these two amounts yields the household net income at each combination; there are 121 possibilities of the household net income in total.

4 Results

4.1 Estimation results for individual labour supply

Appendix I and Appendix 2 show the estimation results of five various models by different degrees of flexibility for male and female labour supply respectively.

It is notable that the alternative specific constant is included in every model since it has proven that labour supply models without accounting for non-participant cannot yield well-presented labour supply. Model I1 is the simplest model among all models reported in Appendix I. It is a simple utility of individual with an alternative specific constant for working. Model I2 has two additional terms which allow random heterogeneity on income and leisure. Based on I2, observed taste shifters were introduced into Model I3 for income and leisure variables. In Model I4, the working specific constant is added to capture observed heterogeneity for income and leisure variables. The most flexible model for single individuals, I5, is I4 with an additional term, i.e., unobserved random preference for the alternative specific constant.

Interpretation of labour supply models is not straightforward due to parameterisation and random heterogeneity. The exceptions are interaction terms between leisure hours and socio-economic characteristics as these parameters indicate how these characteristics affect the marginal utility of leisure (van Soest et al., 2002). In practice, interpreting some parameters, e.g., an interaction term between income and leisure, can also provide some better understanding of utility for labour supply.

The parameters of squared income in all models show a negative sign suggesting that marginal utility of income changes at a diminishing rate. Nonetheless, a coefficient of squared income is significant (at 5% significant level) only in Model I5, it is insignificant in other remaining models. On the other hand, a negative parameter of squared leisure time indicates a diminishing rate of marginal utility of leisure hours. This parameter is statistically significant at 1% significant level in all models.

For the cross-product term between income and leisure hours, the results across model suggest that both arguments are complementary, i.e., people prefer having more consumption and leisure hours. It also shows a strong statistical significance at 1 percent level in every mode. Labour market participation, represented by an alternative specific constant, affects male utility negatively with statistical significance at 1 percent level in all five models.

Socio-economic characteristics included for income and leisure in Model I3 to I5 and an alternative specific constant of working in Model I4 and I5 appear consistent. These coefficients determine how the marginal utility changes with these characteristics. However, only those incorporated into leisure hours are statistically significant. For income variable, people gain less utility when they are older than 40 years old but higher utility if they are holding a higher degree of education. The results also suggest that, in regard to leisure, single people gain more utility of leisure when they are older than 40 years old and attain any higher degree of education. While working has a negative impact on utility with statistical significance at 1% level in all models, a characteristic of living in Bangkok, which is incorporated in Model I4 and I5, does not statistically affect the disutility of working.

For unobserved random preferences for income and leisure, the standard deviation of income and leisure are not statistically significant. This implies that the random taste variation of these arguments is not statistically different in the sample.

The results of alternative specific constants indicate that working is disutility for male individuals as expected. The parameter is statistically significant at 1 percent level. Living in the capital of Thailand does not statistically impact disutility from working. The standard deviation of the working specific constant is significant at 5% significant level suggesting the unobserved preferences across the sample are statistically different.

Appendix 2 presents estimation results of unmarried female labour supply. The parameter of squared income changed from significant negative values (at 5 percent level for Model I1 and 10 percent level for Model I2) to insignificant positive values when models are incorporated with socio-economic characteristics. The parameters of squared leisure show strongly significant negative values across models (at 1 percent significance level). This suggests a diminishing rate of marginal utility of leisure hours for single female labour supply.

The coefficients for the cross-product term between income and leisure hours in all models are statistically significant at 1 percent level. The results indicate income and leisure hours are complementary in the utility function of female labour supply. The results from all five models suggest that working decreases females significantly (it shows a statistical significance at 1 percent level in all models).

In Model I3, I4, and I5, socio-economic characteristics are incorporated with income and leisure; their coefficients are consistent across models. Single females whose age is greater than 40 prefer income and leisure hours more than other females; both coefficients are significant at 1 percent level. Females with any higher degree do not consider income and leisure hours as important factors for their utility as other females do; this effect of the characteristic is significant at 10 percent level for income but insignificant for leisure hours. The socio-economic characteristic if a person lives in Bangkok is included into Model I4 and I5. Although the coefficients in both models show positive values, they are statistically insignificant.

Regarding random heterogeneity terms, the random preferences for income indicate that the standard deviations are strongly significant at 1 percent level across models. Apart from that, other the random taste variation terms (leisure hours and labour market participation) are not statistically significant suggesting that they are not statistical different in the sample.

Similar to male labour supply, the negative value of alternative specific constants with statistical significance at 1 percent level show that working is disutility for females as theoretical expectation. However, the parameter for living in the capital of Thailand does not statistically affect disutility from working due to the insignificant result. The standard deviation of the working specific constant is also insignificant suggesting the unobserved preferences across the sample are statistically indifferent.

4.2 Marginal utilities in individual labour supply

As mentioned above, most of the coefficients in the utility function cannot be directly interpreted (van Soest et al., 2002). Marginal utilities of income and leisure hours are calculated for economic interpretation.

Calculated marginal utilities of both arguments as well as their negative proportions obtained from five models are presented in Appendix 3. Overall, all five models provide positive values of the average marginal utility of income and leisure hours for males and females as theoretical expectation.

In regarding male labour supply, the average marginal utility of income calculated from Model I1 is lowest at 0.0083136; then, it increases as the degree of model flexibility increases. The most substantial average of marginal utility of income is obtained from Model I5 at 0.0151529. The negative proportions of marginal utility of income for Model I1 and I2 are similar at 4.29 and 4.98 percent. The negative proportions drop to less than 1 percent when after introducing socio-economic characteristics into models. In other words, marginal utility of income yielded from Model I3 is 0.84 percent negative values. Model I4 and I5 yield negative proportions of marginal utility of income at 0.86 and 0.44 percent respectively.

The mean of marginal utility of leisure also increases when the model becomes more complicated. The lowest value is yield from Model I1 at 0.0145709 whilst Model I5 yields the highest value at 0.0396924. 28.89 percent of marginal utility of leisure hours are negative in Model I1, this number declines gradually when the complexity of models increases; Model I5 has 20.79 percent negative value of marginal utility of leisure hours.

For male labour supply, 9.70 and 9.32 percent of the sample do not satisfy the quasi-concave condition in Model I1 and I2 respectively. The proportion drops to 0.84 and 0.86 percent for

Model I3 and I4 respectively. 0.44 percent of the sample in Model I5 has the non-quasi-concave utility function. All models show that all samples satisfy the monotonicity condition.

Labour supply for unmarried females provides different results. Even though the lowest mean value of marginal utility of income is still from Model I1 (0.0176164), Model I2 yields the largest value, 0.0287728. Model I3, I4, and I5 obtain similar average values at 0.0219014, 0.0215334, and 0.0216825 respectively. The proportion of negative marginal utility of income is relatively low for Model I1 and I2 at 2.78 and 3.67 percent respectively. The percentage is greater than 5 percent for the rest of models.

The results of marginal utility of leisure hours have a similar pattern to those of income. Model I1 yields the lowest average value of marginal utility of leisure (0.0176141) whilst Model I2 obtains the largest (0.055557). The mean values at 0.0284791, 0.0273859, and 0.0278558 are obtained from Model I3, I4, and I5 respectively. The percentage of negative marginal of leisure hours is 28.87 for Model I1 and 22.79 for Model I2. However, the proportion raises over 30 percent at 30.62, 30.90, and 30.83 percent for Model I3, I4, and I5 respectively.

For the quasi-concave condition, the results of female labour supply indicate that all of the models (Model I1, I2, I3, I4, and I5) have some unsatisfied sample at 17.26, 15.69, 17.11, 17.55, and 17.42 respectively. However, all models provide the results that satisfy the monotonicity condition.

4.3 Estimation results for household labour supply

Appendix 4 presents the estimation results of household labour supply models by different degrees of flexibility. The total number of models for the family level is seven; similar to individual labour supply, models are constructed from the simplest to the most complex. All models include alternative specific constants for working to prevent unwell-presented labour supply.

Model H1, H indicates models for households, is the simplest model which includes only a basic utility function with alternative specific constant terms for males and females. Model H2 allows additional terms which capture random heterogeneity on income and leisure hours for both partners. Model H3 includes observed taste preferences for income, male leisure hours, and female leisure hours into Model H2. Model H4 adds observed heterogeneity terms for squared-terms as well as alternative specific constant terms. Model H5 has two additional terms to capture random preferences of labour market participation for both males and females. Last two models focuse on a new alternative specific constant. In Model H6, a new alternative constant is included to capture the effect on the household utility when both partners are working. Model H7 extends the Model H6 by allowing observed taste shifters into a new alternative specific constant; this is the most flexible model estimated in this paper.

As mentioned before, the most of parameters cannot be directly interpreted. Marginal utilities are discussed in the next section. Regarding squared-terms, the parameters of squared income in Model H1, H2, and H3 are positive which counter economic intuition. However, in Model H5, H6 and H7, the coefficients of squared income after account for characteristics are expected to be negative as the theoretical prediction because all estimates related to squared income have negative values. Only one of these estimates, an interaction between squared income and age of the youngest child is under three, is statistically significant; this suggests that diminishing rate of marginal utility of income increases significantly when a household has a very young child.

Results related to squared leisure for both partners suggest that the marginal utility is diminishing because the parameter of squared leisure hours for males and females in all models are negative with statistical significance at 1 percent level. The interaction terms in Model H4 to H7 are although positive and most of them are statistically significant; nonetheless, their effects are smaller than the effects of squared leisure hours; the total effects are negative.

The cross-product terms between income and leisure hours for males and females are positive and statistically significant at 1 percent level across models. This indicates that income complements leisure hours for both partners. The interaction terms between male leisure hours and female leisure hours in all models are also positive with statistical significance at 1 percent level. The results suggest that leisure hours of both partners are complementary. In other words, when males have longer leisure hours, females have longer leisure hours too.

Coefficients for socio-economic characteristics incorporated with income and leisure in Model H3 to H7 can be interpreted as observed taste shifters. Coefficients of interactions with income are consistent across models whilst those with leisure hours for both partners have sign swaps on characteristics of the youngest child at different age ranges. This is possibly explained by the inclusion of observed and unobserved heterogeneity into working alternative specific constants for both partners in more flexible models (H4 to H7).

Interaction terms between income and the youngest child with different age ranges imply that if a household has the youngest child which is younger than ten years old, the marginal utility of income is greater than a household without this characteristic. The degree of statistical significance decreases gradually when the age of the youngest child increases. A number of children also affect the marginal utility of income positively. In other words, a household with the larger numbers of children has the greater marginal utility of income. However, this coefficient is insignificant in all models.

Interaction terms between income and personal characteristics are interesting. The estimates reveal that two personal characteristics interacted with income have different results across genders. In fact, if male's age is greater than 40, the household marginal utility of income is lower than those who are younger than 40; this characteristic is statistically significant at 10 percent level across models. On the other hand, with statistical significance at 1 percent level in all models, if females are older than 40, the marginal family utility of income is greater than those whose age is fewer than 40.

Regarding educational characteristics for both partners, the marginal utility of income increases with statistical significance at 1 percent level in all models when males are holding any university degree but it decreases if females graduated from any university. However, the statistical significance for coefficients of this interaction between income and a dummy for female with a higher education is varied across models. They are significant at 10 percent level in Model H3, H6, and H7 but insignificant in Model H4 and H5.

Positive coefficients of interactions between leisure hours of both genders and a number of children suggest that, with the greater numbers of children, both males and females value an extra leisure hour more. People consider an additional leisure hour more valuable when they are older than 40 years old than younger people with statistical significance at 1 percent level. Their utility of leisure also depends upon age of spouse, i.e., if age of spouse is greater than 40, marginal utility of leisure increases especially for males in which the estimate are significant at 1 percent level in all models while the estimate for females is significant in Model H3 and H4 at 10 and 5 percent levels respectively. If males are holding a higher degree, their own marginal utility of leisure is significantly greater than those without a degree; they do not consider education of females in marginal utility of leisure since the coefficient is insignificant in all models. However, males with a university degree positively impact marginal utility of leisure for females with statistical significance in all models. Females holding a higher degree yield lower marginal utility for leisure than those did not graduate a university education with statistical significance for Model H3 and H4 (the rest of models provide statistical insignificant results).

Considering unobserved preferences for household income and leisure hours for males and females, the standard deviation of random heterogeneity for income is statistically significant in Model H2 and H5 and 5 and 10 percent level respectively. It for male leisure hours is always significant at 1 percent level in all models while it for female leisure hours shows statistical significance at 1 percent level until random preference for working alternative specific constants are incorporated into models.

Alternative specific constants for working alternatives suggest that working is disutility for both males and females as expected. These parameters for male working, female working, and both partner working are statistically significant at 1 percent level. Socio-economic characteristics incorporated into alternative specific constants provide various results across models; however, unobserved taste variation terms provide consistent results with high statistical significance at 1 percent level.

4.4 Marginal utilities in household labour supply

The second column of Appendix 5 presents calculated marginal utility of household income, male leisure hours, and female leisure hours. It also includes their negative proportions obtained from seven models for household labour supply.

It is apparent that marginal utility of income has positive average values for all models. The most flexible model, H7, provides the largest averaged marginal utility of income, 0.203494 for an additional thousand bahts, and the smallest proportion of negative marginal utility of income (0.37 percent of total observations) relative to other less flexible models.

With regard to marginal utility of male leisure hours, whilst Model H7 provides the highest averaged marginal utility of leisure for males (0.0666032 for an extra leisure hour), Model H1 obtains the result with the smallest percentage for negative marginal utility of male leisure hours (33.75 percent of total households).

For marginal utility of female leisure hours, the most flexible model, H7, provides the highest value of averaged marginal utility of leisure for females (-0. 0420978) as well as the smallest negative proportion, 41.80 percent of total observations. Averaged values of marginal utility of female leisure hours are not conventional. All models provide negative means of marginal utility of leisure for females in household labour supply. From the estimation results, it is possible that the unconventional results could be caused by household work responsibilities especially when couples have a very young child. The results are hence re-examined by estimating sub-groups separately; they are couples with and without children. By comparing the results in third and last columns in Appendix 5 the marginal utility of leisure hours for females in households without a child are less negative (Model H6 even yields a positive number) than it in households with a child or children. The negative numbers imply that Thai females perceive housework hours as non-leisure hours. In addition females in households without children have fewer housework responsibilities than those in households with children. Labeaga et al. (2008) give two reasons of negative marginal utility of leisure for females but it increases with age in Spain. First, due to the increasing labour market participation rate in recent years, they need to remain in employment longer to be eligible for retirement benefits. Second, women temporarily leave the labour market or work only part-time because of childbearing responsibilities; they will return to the market when their children grow up.

The proportions of negative marginal utility of leisure for both partners are largeer than those in Kabátek et al. (2014) which finds 26.62 percent of males and 34.89 percent of females have negative marginal utility of leisure respectively. However, Kabátek et al. (2014) investigates the household utility function with five main arguments, namely, income, male leisure hours, male housework hours, female leisure hours, and female housework hours; it is not possible for Thailand case due to the availability of the data. Hence, the results of these two empirical studies cannot be directly compared.

The utility functions produced by seven different models are tested for quasi-concave and monotonicity conditions. The results for the quasi-concave condition can be divided into three groups. First, Model H1 and H2 obtain the utility function which 39.06 and 41.71 of households have non quasi-concave utility. Second, Model H3 and H4 yield the result suggesting that about one-quarter of all household utility functions are not quasi-concave (27.01 and 26.04 percent respectively). The last group includes Model H5, H6, and H7; these models provide quasi-concave household utility functions at 91.88, 91.03, and 92.08 percent. The result in

Nonetheless, the results of the monotonicity condition are almost identical; most of the models produce the result that 100 percent of observations have monotonic utility functions. Only Model H2, H3, and H4 obtain the result in which 0.09, 0.03, and 0.03 percent respectively of all household utility functions fail the monotonicity condition.

4.5 Model selection

With regard to household models, Model H7 performs better than the rest of the model in almost aspects. It yields the marginal utility of income with the lowest percentage of negative values. Furthermore, the results of quasi-concave and monotonicity conditions suggest that Model H7 is more consistent with the modelling assumptions than other models. The log-likelihood ratio test confirms that Model H7 has the best goodness of fit relative to other models with statistical significance at 1 percent level.

For individual labour supply models, Model I3, I4, and I5 produce similar estimation results. By comparing the calculated marginal utility of income, Model I5 performs slightly better than Model I3 and I4. Model I5 also yields better numbers for quasi-concave and monotonicity conditions which imply it is more consistent with the modelling assumptions that the rest of models are. The log-likelihood ratio test does not indicate that Model I5 fit the data significantly better than Model I3 and I4. Nevertheless, additional variable in Model 5 is to allow observed and unobserved heterogeneity on a working alternative specific constant. An unobserved random preference term is also statistically significant for male labour supply. With these additional terms, the model is more in-line with the household labour supply model as well.

In conclusion, Model I5 is selected for the individual labour supply whilst Model H7 is the representation of family labour supply models for Thailand. In fact, both models are the most flexible model in their own kinds.

4.6 Robustness checks

The first robustness check is the estimation using 500 Halton draws. The results are shown in Appendix 6 and Appendix 7 for unmarried male and female labour supply respectively. The result is not available for the household labour supply because the sample is larger and the model is far more complicated than individual models; the estimation process consumes much longer time to obtain the result.

Overall, the labour supply results for males and females are consistent across different numbers of draws in the simulated maximum likelihood process. A concern on the result of family labour supply is the convergence of the log-likelihood function due to many random heterogeneity terms.

The second robustness check is to test the sensitivity when reported wages are applied for working people. In other words, wages of non-working people, who do not report wages in HSES, are predicted using the wage imputation in Section 2.3 whilst wages of working people are computed from the survey. The last columns in Appendix 6, 7, and 8 present the estimation results of labour supply for unmarried males, unmarried females, and married couples respectively.

The unexpected results are obtained for both individual and family levels. In fact, the average marginal utility of income turns negative for males and family labour supply, and the negative proportions of marginal utility of income increase in every case. The mean and the negative proportions of marginal utility of leisure deteriorate in all models. The proportions of failure in quasi-concave condition significantly rise in male, female, and family labour supply.

van Soest (1995) indicates that applying actual and imputed wages in the estimation could lead to inconsistent results because this approach assumes no errors in wage prediction. When predicted wages are applied only non-working previous studies usually accounted for errors by integrated the errors in the labour supply estimation process (Löffler et al., 2014). In this study, wage prediction errors are not taken into account; this could explain why the results are inconsistent.

5 Conclusion

5.1 Summary of labour supply in Thailand

This study aims to study individual and family labour supply behaviour in Thailand using the discrete choice approach. Sever models with different degrees of flexibility are estimates for individual labour supply (five models) and household labour supply (seven models).

The results of individual labour supply suggest that the most flexible model provide the best fit for the Thai dataset. Overall, the model yields the results consistent with economic theory and previous empirical works. Marginal utility of income and it of leisure hours increase with diminishing rates for both unmarried males and females. Regarding comparing quasi-concave condition, male labour supply provides the results which are more consistent with the assumption than female labour supply does. For monotonicity condition, the model for both genders performs equivalently, i.e., utility functions of all observations pass the condition.

The results of family labour supply also indicate that the most flexible model outperforms other models. The model yields the positive mean of marginal utility of income as well as the least percentage of negative marginal utility of income, i.e., it is most consistent with the theory in this regard. It also yields the positive mean of marginal utility of male leisure hours as expected. However, a negative mean of marginal utility of female leisure is not conventional. This is possibly because females in a family setting have a lot of housework responsibilities and they treat these tasks as working not leisure. The preferred model also provides diminishing rates for marginal utilities which are consistent with the theoretical prediction. The tests for quasi-concave and monotonicity conditions indicate that the most flexible model obtains the results that are most consistent with the utility function assumptions. The log-likelihood ratio test confirms that the most flexible model is the most preferable in terms of goodness of fit to the Thai dataset.

The robustness check on different numbers of Halton draws is consistent in the individual labour supply level. Nevertheless, the result of household labour supply has not been obtained due to a larger sample and a more complex model with six random heterogeneity terms. The sensitivity check on different approaches for wage prediction yields unconventional results for male, female, and family labour supply. This may be because of errors in wage prediction.

5.2 Future research on labour supply

More comprehensive datasets could make some certain improvement for labour supply research in the future. HSES has some drawbacks for labour supply studies. For example, it does not provide some certain income types at an individual level and it provides total amounts of incomes rather than sub-categories of incomes. In addition, information on allowances and exemptions is missing from HSES. If the questionnaire was more consistent with income tax rules and regulations, the calculated disposable income would be more realistic.

HSES collects data on a cross-sectional basis. Panel data provides the possibility to study dynamic labour supply which allows investigating labour supply behaviour in multiple periods because people do not make a once-and-for-all decision for labour supply.

The paper estimates several models with different degrees of flexibility. However, the models do not allow for correlation between parameters. The models will be more flexible if correlation is incorporated. Creedy et al. (2002), Duncan and Harris (2002), and Kabátek et al. (2014) are examples of empirical studies which include correlation into the models; they also find significant results on correlation and covariance matrices.

A drawback of the discrete hours labour supply models applied in this paper is that the models do not explicitly account for labour demand in the market. According to Peichl and Siegloch (2012), excluding the demand side of the labour market could lead to biased results of potential labour market outcomes. They also show that the demand effects account for approximately 25 percent of the positive labour supply impact due to policy reform.

This study mainly aims to investigate labour supply behaviour of people who are assumed to select whether to work or not and how many how they want to supply. However, in reality, people have other alternatives such as working as a farmer or doing an own-account business. The model which includes these alternatives could be restricted and reflect the whole economy in a more realistic sense.

One major current issue is the ageing society. A researcher possibly specifies a model to study labour supply decision of retire-age people. However, the model will be more complicated since these people usually live with other dependent adults. Some addition assumptions and justifications are required.

6 Bibliography

AABERGE, R. & COLOMBINO, U. 2014. Labour Supply Models. *In:* O'DONOGHUE, C. (ed.) *Handbook of Microsimulation Modelling (Contributions to Economic Analysis, Volume 293)* Emerald Group Publishing Limited.

BARGAIN, O., ORSINI, K. & PEICHL, A. 2014. Comparing labor supply elasticities in Europe and the United States: New results. *Journal of Human Resources,* 49**,** 723-838.

BLUNDELL, R. & MACURDY, T. 1999. Labor Supply: A Review of Alternative Approaches. *In:* ORLEY, C. A. & DAVID, C. (eds.) *Handbook of Labor Economics,*. Elsevier.

CALLAN, T. & VAN SOEST, A. 1996. Family Labour Supply and Taxes in Ireland. *ESRI Working Paper No. 78.*

CREEDY, J. & DUNCAN, A. 2002. Behavioural Microsimulation with Labour Supply Responses. *Journal of Economic Surveys,* 16**,** 1-39.

CREEDY, J., DUNCAN, A., HARRIS, M. & SCUTELLA, R. 2002. Microsimulation Modelling of Taxation and the Labour Market. Edward Elgar Publishing.

CREEDY, J. & KALB, G. 2005. Discrete Hours Labour Supply Modelling: Specification, Estimation and Simulation. *Journal of Economic Surveys,* 19**,** 697-734.

DUNCAN, A. & HARRIS, M. N. 2002. Simulating the Behavioural Effects of Welfare Reforms Among Sole Parents in Australia. *Economic Record,* 78**,** 264-276.

GONG, X. & VAN SOEST, A. 2002. Family structure and female labour supply in Mexico City. *Journal of Human Resources,* 37**,** 163-191.

HECKMAN, J. J. 1979. Sample Selection Bias as a Specification Error. *Econometrica,* 47**,** 153-161.

KABÁTEK, J., VAN SOEST, A. & STANCANELLI, E. 2014. Income taxation, labour supply and housework: A discrete choice model for French couples. *Labour Economics,* 27**,** 30-43.

KEANE, M. & MOFFITT, R. 1998. A Structural Model of Multiple Welfare Program Participation and Labor Supply. *International Economic Review,* 39**,** 553-589.

LABEAGA, J. M., OLIVER, X. & SPADARO, A. 2008. Discrete choice models of labour supply, behavioural microsimulation and the Spanish tax reforms. *The Journal of Economic Inequality,* 6**,** 247-273.

LÖFFLER, M., PEICHL, A. & SIEGLOCH, S. 2014. Structural labor supply models and wage exogeneity. *SOEP papers on Multidisciplinary Panel Data Research.* Berlin: eutsches Institut für Wirtschaftsforschung (DIW).

MCFADDEN, D. 1974. Conditional Logit Analysis of Qualitative Choices *In:* ZAREMBKA, P. (ed.) *Frontiers in econometrics.* New York: Academic Press.

PEICHL, A. & SIEGLOCH, S. 2012. Accounting for labor demand effects in structural labor supply models. *Labour Economics,* 19**,** 129-138.

THE UNITED NATIONS POPULATION FUND THAILAND AND THE OFFICE OF THE NATIONAL ECONOMIC AND SOCIAL DEVELOPMENT BOARD. 2015. The State of Thailand's Population 2015 "Features of Thai Families in the Era of Low Fertility and Longevity".

TRAIN, K. E. 2009. *Discrete choice methods with simulation*, Cambridge university press.

VAN SOEST, A. 1995. Structural Models of Family Labor Supply: A Discrete Choice Approach. *The Journal of Human Resources,* 30**,** 63-88.

VAN SOEST, A., DAS, M. & GONG, X. 2002. A structural labour supply model with flexible preferences. *Journal of Econometrics*, 107**,** 345-374.

7 Appendix

Appendix 1: Estimation results for unmarried male labour supply

	I1	I2	I3	I4	I5
Y^2	-0.00000400	-0.00000452	-0.0000115	-0.0000115	-0.0000117*
	(0.0000027)	(0.00000317)	(0.00000618)	(0.00000617)	(0.00000596)
L^2	-0.00256***	-0.00254***	-0.00277***	-0.00277***	-0.00276***
	(0.000121)	(0.000168)	(0.000227)	(0.000217)	(0.000213)
$Y * L$	0.000351***	0.000371***	0.000295***	0.000295***	0.000301***
	(0.0000369)	(0.0000683)	(0.0000700)	(0.0000698)	(0.0000652)
Y	-0.0102***	-0.0101***	0.000523	0.000493	0.000915
	(0.00269)	(0.00279)	(0.00420)	(0.00420)	(0.00416)
Y * age over 40			-0.00156	-0.00157	-0.000959
			(0.00170)	(0.00169)	(0.00177)
Y * higher degree			0.00692	0.00697	0.00624
			(0.00455)	(0.00455)	(0.00456)
L	0.249***	0.247***	0.284***	0.283***	0.282***
	(0.0158)	(0.0196)	(0.0264)	(0.0255)	(0.0257)
L * age over 40			0.0218***	0.0217***	0.0247***
			(0.00587)	(0.00581)	(0.00651)
L * higher degree			0.0611**	0.0611**	0.0591**
			(0.0222)	(0.0222)	(0.0219)
P	-5.993***	-5.991***	-5.980***	-5.985***	-5.200***
	(0.190)	(0.190)	(0.190)	(0.191)	(0.547)
P *BKK				0.0806	0.0692
				(0.317)	(0.351)
SD: Y		0.00260	-0.00244	-0.00244	0.00199
		(0.00385)	(0.00370)	(0.00369)	(0.00262)
SD :L		-0.00323	-0.00857	-0.00781	-0.00912
		(0.0180)	(0.0173)	(0.0174)	(0.0178)
SD :P					-1.418**
					(0.511)
Observations	49687	49687	49687	49687	49687
Log lik.	-8438.4	-8438.4	-8370.5	-8370.4	-8368.9
lrtest_chi2		0.16	135.78	0.07	3.02
lrtest_df		2	4	1	1
lrtest_p		0.9235	0.0000	0.7979	0.0823

Notes: Standard errors in parentheses; *, **, *** indicate statistical significance at 10%, 5%, and 1% levels respectively.

Appendix 2: Estimation results for unmarried female labour supply

	I1	I2	I3	I4	I5
Y^2	-0.00000695**	-0.00000855*	0.00000506	0.00000486	0.00000484
	(0.00000266)	(0.00000353)	(0.00000542)	(0.00000539)	(0.00000539)
L^2	-0.00198***	-0.00169***	-0.00203***	-0.00204***	-0.00204***
	(0.000106)	(0.000123)	(0.000134)	(0.000134)	(0.000134)
$Y * L$	0.000530***	0.000774***	0.000699***	0.000694***	0.000697***
	(0.0000333)	(0.0000674)	(0.0000677)	(0.0000674)	(0.0000678)
Y	-0.0130***	-0.0165***	-0.0221***	-0.0222***	-0.0222***
	(0.00249)	(0.00313)	(0.00377)	(0.00376)	(0.00376)
Y * age over 40			0.00570***	0.00571***	0.00572***
			(0.00150)	(0.00149)	(0.00150)
Y * higher degree			-0.00963*	-0.00929*	-0.00933*
			(0.00386)	(0.00386)	(0.00387)
L	0.185***	0.152***	0.159***	0.159***	0.159***
	(0.0137)	(0.0159)	(0.0174)	(0.0173)	(0.0173)
L * age over 40			0.0575***	0.0574***	0.0576***
			(0.00451)	(0.00450)	(0.00455)
L * higher degree			-0.0295	-0.0290	-0.0291
			(0.0163)	(0.0163)	(0.0163)
P	-6.800***	-6.838***	-6.841***	-6.861***	-6.856***
	(0.183)	(0.185)	(0.184)	(0.184)	(0.186)
P *BKK				0.301	0.302
				(0.238)	(0.239)
SD: Y		-0.0101***	-0.00671***	-0.00667***	-0.00673***
		(0.00161)	(0.00156)	(0.00155)	(0.00156)
SD :L		-0.000138	-0.000167	0.000113	-0.0000999
		(0.00722)	(0.0118)	(0.00997)	(0.00984)
SD :P					-0.163
					(0.347)
Observations	50336	50336	50336	50336	50336
Log lik.	-8540.1	-8527.3	-8280.2	-8279.4	-8279.3
lrtest_chi2		25.50	494.12	1.67	0.21
lrtest_df		2	4	1	1
lrtest_p		0.0000	0.0000	0.1957	0.6438

Notes: Standard errors in parentheses; *, **, *** indicate statistical significance at 10%, 5%, and 1% levels respectively.

Appendix 3: Mean of marginal utilities with negative percentage; and unsatisfied conditions in percentage from individual labour supply

Individual	100 draws			
	Male labour supply		Female labour supply	
	mean	neg. %	mean	neg. %
MUyI1	0.0083136	4.29	0.0176164	2.78
MUyI2	0.0092985	4.98	0.0287728	3.67
MUyI3	0.0143973	0.84	0.0219014	5.44
MUyI4	0.0143281	0.86	0.0215334	5.55
MUyI5	0.0151529	0.44	0.0216825	5.53
MUlI1	0.0145709	28.89	0.0176141	28.87
MUlI2	0.0181924	28.07	0.055557	22.79
MUlI3	0.0375659	21.61	0.0284791	30.62
MUlI4	0.0373437	21.56	0.0273859	30.90
MUlI5	0.0396924	20.79	0.0278558	30.83
	Unsatisfied (%)		Unsatisfied (%)	
Quasi-concave I1	9.70		17.26	
Quasi-concave I2	9.32		15.69	
Quasi-concave I3	0.84		17.11	
Quasi-concave I4	0.86		17.55	
Quasi-concave I5	0.44		17.42	
Monotonicity I1	0.00		0.00	
Monotonicity I2	0.00		0.00	
Monotonicity I3	0.00		0.00	
Monotonicity I4	0.00		0.00	
Monotonicity I5	0.00		0.00	
Obs.	4,517		4,576	

Appendix 4: Estimation results for household labour supply

	H1	H2	H3	H4	H5	H6	H7
Y^2	3.92e-08	1.43e-08	0.000000553	0.000000612	-0.000000993	-0.000000772	-0.000000855
	(0.000000384)	(0.000000561)	(0.000000987)	(0.00000101)	(0.00000132)	(0.00000128)	(0.00000132)
Y^2*young child 3				-0.00000966***	-0.0000148***	-0.0000147***	-0.0000154***
				(0.00000194)	(0.00000253)	(0.00000248)	(0.00000254)
Y^2*young child 6				-0.000000547	-0.00000137	-0.00000126	-0.000000865
				(0.00000166)	(0.00000217)	(0.00000209)	(0.00000217)
Y^2*young child 10				-0.000000888	-0.00000190	-0.00000177	-0.00000104
				(0.00000145)	(0.00000181)	(0.00000177)	(0.00000181)
L_m^2	-0.00355***	-0.00447***	-0.00586***	-0.00627***	-0.00944***	-0.00945***	-0.0104***
	(0.0000484)	(0.000161)	(0.000211)	(0.000226)	(0.000447)	(0.000477)	(0.000519)
L_m^2*young child < 3				0.00110***	0.00148***	0.00148***	0.00169***
				(0.000148)	(0.000199)	(0.000200)	(0.000211)
L_m^2*young child 3-6				0.000523***	0.000811***	0.000754***	0.000688**
				(0.000156)	(0.000211)	(0.000212)	(0.000227)
L_m^2*young child 6-10				0.000129	0.000400*	0.000392*	0.000255
				(0.000132)	(0.000184)	(0.000184)	(0.000201)
L_f^2	-0.00377***	-0.00411***	-0.00425***	-0.00446***	-0.00424***	-0.00422***	-0.00425***
	(0.0000542)	(0.000112)	(0.0000941)	(0.000104)	(0.0000800)	(0.0000771)	(0.0000771)
L_f^2*young child < 3				0.000767***	0.00137***	0.00130***	0.00134***
				(0.000126)	(0.000154)	(0.000151)	(0.000152)
L_f^2*young child 3-6				0.000186	0.000204	0.000208	0.000193
				(0.000114)	(0.000144)	(0.000139)	(0.000141)
L_f^2*young child 6-10				0.0000822	0.000247*	0.000215*	0.000205
				(0.0000817)	(0.000113)	(0.000108)	(0.000110)
YL_m	0.000236***	0.000301***	0.000184***	0.000180***	0.000209***	0.000207***	0.000214***
	(0.00000626)	(0.0000154)	(0.0000152)	(0.0000156)	(0.0000210)	(0.0000207)	(0.0000212)
YL_f	0.000110***	0.000147***	0.000187***	0.000184***	0.000197***	0.000196***	0.000204***
	(0.00000481)	(0.0000118)	(0.0000119)	(0.0000122)	(0.0000151)	(0.0000149)	(0.0000153)

	H1	H2	H3	H4	H5	H6	H7
$L_m L_f$	0.00181***	0.00210***	0.00202***	0.00205***	0.00285***	0.00289***	0.00306***
	(0.0000407)	(0.0000823)	(0.0000730)	(0.0000757)	(0.000129)	(0.000126)	(0.000127)
Y	-0.0127***	-0.0165***	-0.0137***	-0.0138***	-0.00979***	-0.0101***	-0.00980***
	(0.000763)	(0.00129)	(0.00175)	(0.00181)	(0.00221)	(0.00216)	(0.00225)
Y*young child $<$ 3			0.00938***	0.0215***	0.0301***	0.0300***	0.0315***
			(0.00121)	(0.00267)	(0.00360)	(0.00351)	(0.00360)
Y*young child 3-6			0.00510***	0.00656**	0.0101**	0.00959**	0.00915**
			(0.00113)	(0.00244)	(0.00333)	(0.00320)	(0.00334)
Y*young child 6-10			0.00263**	0.00415	0.00754**	0.00702*	0.00590*
			(0.000930)	(0.00218)	(0.00287)	(0.00278)	(0.00287)
Y*number of children			0.000275	0.000378	0.000824	0.000749	0.000922
			(0.000344)	(0.000363)	(0.000476)	(0.000459)	(0.000479)
Y * age40(m)			-0.00196*	-0.00202*	-0.00284*	-0.00248*	-0.00283*
			(0.000925)	(0.000971)	(0.00122)	(0.00118)	(0.00122)
Y * higher degree(m)			0.00397***	0.00474***	0.00561***	0.00582***	0.00584***
			(0.00106)	(0.00111)	(0.00143)	(0.00137)	(0.00142)
Y * age40(f)			0.00282***	0.00285**	0.00383***	0.00359**	0.00410***
			(0.000856)	(0.000900)	(0.00116)	(0.00112)	(0.00116)
Y * higher degree(f)			-0.00267*	-0.00210	-0.00321	-0.00384*	-0.00438*
			(0.00124)	(0.00128)	(0.00184)	(0.00181)	(0.00189)
L_m	0.228***	0.301***	0.445***	0.483***	0.761***	0.759***	0.844***
	(0.00629)	(0.0147)	(0.0213)	(0.0226)	(0.0418)	(0.0447)	(0.0493)
L_m*young child $<$ 3			0.0188**	-0.0814***	-0.113***	-0.114***	-0.134***
			(0.00572)	(0.0166)	(0.0225)	(0.0225)	(0.0239)
L_m*young child 3-6			0.0253***	-0.0275	-0.0441	-0.0407	-0.0352
			(0.00572)	(0.0178)	(0.0240)	(0.0241)	(0.0260)
L_m*young child 6-10			0.0201***	0.00950	-0.00422	-0.00530	0.00910
			(0.00518)	(0.0158)	(0.0217)	(0.0218)	(0.0237)
L_m*number of children			0.00295	0.00559*	0.00880**	0.00872**	0.00972**
			(0.00204)	(0.00219)	(0.00316)	(0.00312)	(0.00335)
L_m * age40(m)			0.0176***	0.0187***	0.0288***	0.0296***	0.0326***
			(0.00447)	(0.00467)	(0.00657)	(0.00651)	(0.00699)

	H1	H2	H3	H4	H5	H6	H7
L_m * higher degree(m)			0.0816***	0.0898***	0.134***	0.134***	0.143***
			(0.00781)	(0.00825)	(0.0114)	(0.0115)	(0.0122)
L_m * age40(f)			0.0384***	0.0404***	0.0600***	0.0588***	0.0657***
			(0.00456)	(0.00477)	(0.00692)	(0.00688)	(0.00740)
L_m * higher degree(f)			0.00375	0.00813	0.0195	0.0150	0.0192
			(0.00718)	(0.00746)	(0.0108)	(0.0106)	(0.0113)
L_f	0.303***	0.328***	0.315***	0.334***	0.259***	0.255***	0.249***
	(0.00677)	(0.0111)	(0.0107)	(0.0117)	(0.0110)	(0.0109)	(0.0110)
L_f*young child $<$ 3			0.0877***	-0.0102	-0.0695***	-0.0629***	-0.0643***
			(0.00529)	(0.0163)	(0.0184)	(0.0181)	(0.0183)
L_f*young child 3-6			0.0457***	0.0199	0.0245	0.0222	0.0232
			(0.00427)	(0.0150)	(0.0177)	(0.0172)	(0.0175)
L_f*young child 6-10			0.0194***	0.0133	0.00541	0.00701	0.00643
			(0.00354)	(0.0119)	(0.0143)	(0.0138)	(0.0142)
L_f*number of children			0.00202	0.00871***	0.00629**	0.00635***	0.00682***
			(0.00135)	(0.00174)	(0.00198)	(0.00193)	(0.00198)
L_f * age40(f)			0.0246***	0.0265***	0.0373***	0.0353***	0.0374***
			(0.00309)	(0.00329)	(0.00404)	(0.00391)	(0.00402)
L_f * higher degree(f)			-0.0229***	-0.0202**	-0.000356	-0.00550	-0.00540
			(0.00595)	(0.00618)	(0.00883)	(0.00850)	(0.00884)
L_f* age40(m)			0.00730*	0.00852**	0.00629	0.00691	0.00577
			(0.00296)	(0.00314)	(0.00392)	(0.00378)	(0.00390)
L_f * higher degree(m)			0.0113*	0.0157**	0.0181**	0.0185**	0.0173*
			(0.00503)	(0.00534)	(0.00700)	(0.00668)	(0.00693)
P_m	-6.010***	-6.201***	-6.491***	-7.123***	-7.282***	-5.982***	-6.689***
	(0.116)	(0.127)	(0.136)	(0.174)	(0.380)	(0.556)	(0.440)
P_m*BKK				-0.523	-1.254*	-1.169*	-2.428***
				(0.300)	(0.494)	(0.494)	(0.677)
P_m* $PreS$				1.265***	1.516***	1.287***	2.536***
				(0.272)	(0.353)	(0.360)	(0.501)
P_m* AS				0.570***	0.728***	0.724***	1.004**
				(0.139)	(0.202)	(0.206)	(0.307)

	H1	H2	H3	H4	H5	H6	H7
P_f	-8.414***	-8.465***	-8.555***	-8.822***	-6.860***	-5.680***	-5.855***
	(0.111)	(0.114)	(0.116)	(0.133)	(0.291)	(0.319)	(0.362)
P_f*BKK				-0.317**	-0.341	-0.324	-1.686*
				(0.0979)	(0.190)	(0.171)	(0.753)
P_f* $PreS$				0.0972	-1.224***	-0.966***	0.565
				(0.167)	(0.307)	(0.272)	(0.535)
P_f* AS				0.474***	0.335**	0.366***	0.511
				(0.0680)	(0.112)	(0.103)	(0.326)
P_{hh}						-1.620***	-1.278***
						(0.216)	(0.282)
P_{hh}*BKK							1.375
							(0.743)
P_{hh}* $PreS$							-1.756***
							(0.475)
P_{hh}* AS							-0.154
							(0.315)
SD:Y		-0.00290**	-0.000558	-0.000419	0.00187*	0.00124	0.000496
		(0.000895)	(0.00105)	(0.00109)	(0.000939)	(0.000954)	(0.000848)
SD:L_m		0.0537***	0.0785***	0.0840***	0.163***	0.162***	0.186***
		(0.00553)	(0.00565)	(0.00577)	(0.0108)	(0.0114)	(0.0125)
SD:L_f		0.0332***	0.0389***	0.0439***	0.00345	0.00369	-0.00332
		(0.00568)	(0.00367)	(0.00375)	(0.00284)	(0.00292)	(0.00263)
SD:P_m					2.655***	2.989***	-3.133***
					(0.548)	(0.715)	(0.593)
SD:P_f					-4.878***	-4.279***	4.630***
					(0.431)	(0.357)	(0.365)
Observations	1884938	1884938	1884938	1884938	1884938	1884938	1884938
Log lik.	-55968.5	-55956.5	-54803.7	-54692.8	-54603.4	-54571.4	-54534.7
lrtest_chi2		24.12	2305.49	221.78	178.83	63.99	73.47
lrtest_df		3	24	15	2	1	3
lrtest_p		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Notes: Standard errors in parentheses; *, **, *** indicate statistical significance at 10%, 5%, and 1% levels respectively.

Appendix 5: Mean of marginal utilities with negative percentage; and unsatisfied conditions in percentage from household labour supply

Household	100 draws					
	Whole sample		Couples without child		Couples with children	
	mean	neg. %	mean	neg. %	mean	neg. %
MUyH1	0.0076722	4.09	0.0071059	6.34	0.0078508	1.35
MUyH2	0.009882	4.94	0.0072434	7.13	0.008256	1.40
MUyH3	0.0116173	2.10	0.0128455	1.39	0.0100903	2.73
MUyH4	0.0124662	2.53	0.0132897	0.93	0.0114399	4.38
MUyH5	0.0196232	0.51	0.048266	15.89	-	-
MUyH6	0.0192221	0.51	0.0492227	14.46	0.0124494	4.23
MUyH7	0.0203494	0.37	0.0215967	0.18	0.0123765	3.25
MUmH1	0.0236368	33.75	0.0218322	38.91	0.0241949	30.87
MUmH2	0.0319568	35.88	0.0225007	37.43	0.0253316	36.57
MUmH3	0.0390586	36.91	0.0442799	35.61	0.0333768	37.23
MUmH4	0.0418657	36.67	0.0459061	35.36	0.0379877	36.84
MUmH5	0.064857	37.70	0.1593477	13.01	-	-
MUmH6	0.0631281	37.94	0.1627688	12.53	0.0412771	36.05
MUmH7	0.0666032	38.42	0.0718767	38.30	0.0405449	37.12
MUfH1	-0.0710917	47.61	-0.0576752	41.22	-0.0840294	53.45
MUfH2	-0.0647838	46.40	-0.0572354	41.14	-0.0829449	53.04
MUfH3	-0.060665	47.66	-0.0440603	42.38	-0.0765721	52.30
MUfH4	-0.0572101	47.32	-0.0429986	42.18	-0.0707873	51.23
MUfH5	-0.0429309	42.36	-0.0003381	30.31	-	-
MUfH6	-0.0441392	42.46	0.00401	30.07	-0.0679959	50.44
MUfH7	-0.0420978	41.80	-0.0336678	35.02	-0.0686935	50.70

Household	100 draws		
	Whole sample	Couples without child	Couples with children
Conditions	Unsatisfied %	Unsatisfied %	Unsatisfied %
Quasi-concave H1	39.06	41.02	33.91
Quasi-concave H2	41.71	44.29	25.95
Quasi-concave H3	27.01	24.12	26.12
Quasi-concave H4	26.04	23.35	26.42
Quasi-concave H5	8.12	25.09	-
Quasi-concave H6	8.97	20.90	20.27
Quasi-concave H7	7.92	10.61	24.91
Monotonicity H1	0.00	0.22	0.00
Monotonicity H2	0.09	0.15	0.00
Monotonicity H3	0.03	0.00	0.07
Monotonicity H4	0.03	0.00	0.07
Monotonicity H5	0.00	1.09	-
Monotonicity H6	0.00	1.05	0.09
Monotonicity H7	0.00	0.00	0.06
Obs.	15578	7,350	8,228

Appendix 6: Robustness checks for estimation results for unmarried male labour supply

	100 draws	500 draws	Report wages
Y^2	-0.0000117*	-0.0000113	0.00000661
	(0.00000596)	(0.00000594)	(0.00000379)
L^2	-0.00276***	-0.00274***	-0.00542***
	(0.000213)	(0.000156)	(0.000905)
$Y * L$	0.000301***	0.000290***	0.000751***
	(0.0000652)	(0.0000635)	(0.000197)
Y	0.000915	0.000566	-0.0747***
	(0.00416)	(0.00411)	(0.0148)
Y * age over 40	-0.000959	-0.00118	0.0164***
	(0.00177)	(0.00172)	(0.00382)
Y*higher degree	0.00624	0.00643	0.0210***
	(0.00456)	(0.00449)	(0.00508)
L	0.282***	0.280***	0.408***
	(0.0257)	(0.0202)	(0.0678)
L * age over 40	0.0247***	0.0230***	0.0882***
	(0.00651)	(0.00590)	(0.0182)
L*higher degree	0.0591**	0.0587**	0.0368*
	(0.0219)	(0.0217)	(0.0160)
P	-5.200***	-5.475***	-0.507
	(0.547)	(0.618)	(2.301)
P *BKK	0.0692	0.0869	1.771
	(0.351)	(0.328)	(1.140)
SD: Y	0.00199	0.00127	0.0170***
	(0.00262)	(0.00467)	(0.00507)
SD :L	-0.00912	0.00165	0.0807**
	(0.0178)	(0.0200)	(0.0305)
SD :P	-1.418**	-1.113	-8.032***
	(0.511)	(0.678)	(1.319)
Observations	49687	49687	49687
Log lik.	-8368.9	-8370.2	-8015.4

	100 draws	500 draws	Report wages
MUyI5	0.0151529	0.0142514	-0.0208388
neg. %	0.44	0.42	80.10
MUlIH5	0.0396924	0.0366561	-0.0629028
neg. %	20.79	21.10	63.12
Quasi-concave I5	0.44	0.42	83.09
Monotonicity I5	0.00	0.00	0.00
Observations	4,517	4,517	4,517

Notes: Standard errors in parentheses; *, **, *** indicate statistical significance at 10%, 5%, and 1% levels respectively.

Appendix 7: Robustness checks for estimation results for single female labour supply

	100 draws	500 draws	Report wages
Y^2	0.00000484	0.00000505	0.00000193
	(0.00000539)	(0.00000538)	(0.00000139)
L^2	-0.00204***	-0.00204***	-0.00269***
	(0.000134)	(0.000134)	(0.000104)
$Y * L$	0.000697***	0.000689***	0.000824***
	(0.0000678)	(0.0000671)	(0.0000505)
Y	-0.0222***	-0.0222***	-0.0604***
	(0.00376)	(0.00375)	(0.00310)
Y * age over 40	0.00572***	0.00572***	0.0149***
	(0.00150)	(0.00148)	(0.00144)
Y*higher degree	-0.00933*	-0.00930*	0.0118***
	(0.00387)	(0.00385)	(0.00199)
L	0.159***	0.160***	0.158***
	(0.0173)	(0.0174)	(0.0132)
L * age over 40	0.0576***	0.0572***	0.0824***
	(0.00455)	(0.00448)	(0.00462)
L*higher degree	-0.0291	-0.0291	-0.00684
	(0.0163)	(0.0162)	(0.00676)
P	-6.856***	-6.859***	-8.209***
	(0.186)	(0.184)	(0.210)
P *BKK	0.302	0.302	0.961***
	(0.239)	(0.237)	(0.253)
SD: Y	-0.00673***	0.00642***	-0.00906***
	(0.00156)	(0.00155)	(0.00114)
SD :L	-0.0000999	0.000370	0.0000131
	(0.00984)	(0.00985)	(0.00552)
SD :P	-0.163	0.0256	-0.115
	(0.347)	(0.380)	(0.313)
Observations	50336	50336	50336
Log lik.	-8279.3	-8279.8	-7951.8
	100 draws	500 draws	Report wages
MUyI5	0.0216825	0.0212242	0.0030985
neg. %	5.53	5.53	47.71
MUlIH5	0.0278558	0.0263832	-.0196929
neg. %	30.83	31.16	55.22
Quasi-concave I5	17.42	17.61	62.19
Monotonicity I5	0.00	0.00	0.00
Observations	4,576	4,576	4,576

Notes: Standard errors in parentheses; *, **, *** indicate statistical significance at 10%, 5%, and 1% levels respectively.

Appendix 8: Robustness checks of estimation results for family labour supply

	100 draws	500 draws	Wage prediction
Y^2	-0.000000855		0.000000558
	(0.00000132)		(0.000000524)
Y^2*young child 3	-0.0000154***		0.00000207
	(0.00000254)		(0.00000183)
Y^2*young child 6	-0.000000865		0.00000114
	(0.00000217)		(0.00000153)
Y^2*young child 10	-0.00000104		-0.00000149**
	(0.00000181)		(0.000000526)
L_m^2	-0.0104***		-0.00652***
	(0.000519)		(0.000183)
L_m^2*young child < 3	0.00169***		0.00162***
	(0.000211)		(0.000174)
L_m^2*young child 3-6	0.000688**		0.00100***
	(0.000227)		(0.000191)
L_m^2*young child 6-10	0.000255		0.000566**
	(0.000201)		(0.000178)
L_f^2	-0.00425***		-0.00548***
	(0.0000771)		(0.000125)
L_f^2*young child < 3	0.00134***		0.00168***
	(0.000152)		(0.000172)
L_f^2*young child 3-6	0.000193		0.000260
	(0.000141)		(0.000170)
L_f^2*young child 6-10	0.000205		0.000376**
	(0.000110)		(0.000135)
YL_m	0.000214***		0.000385***
	(0.0000212)		(0.0000153)
YL_f	0.000204***		0.000309***
	(0.0000153)		(0.0000117)
L_mL_f	0.00306***		0.00216***
	(0.000127)		(0.0000772)
Y	-0.00980***		-0.0718***
	(0.00225)		(0.00230)
Y*young child < 3	0.0315***		0.00578*
	(0.00360)		(0.00286)
Y*young child 3-6	0.00915**		0.00370
	(0.00334)		(0.00268)
Y*young child 6-10	0.00590*		0.00559***
	(0.00287)		(0.00170)
Y*number of children	0.000922		0.00118*
	(0.000479)		(0.000514)
Y * age40(m)	-0.00283*		0.00892***
	(0.00122)		(0.00129)
Y * higher degree(m)	0.00584***		0.00829***
	(0.00142)		(0.00122)
Y * age40(f)	0.00410***		0.00864***
	(0.00116)		(0.00120)

	100 draws	500 draws	Wage prediction
Y * higher degree(f)	-0.00438*		0.00851***
	(0.00189)		(0.00115)
L_m	0.844***		0.383***
	(0.0493)		(0.0142)
L_m*young child $<$ 3	-0.134***		-0.156***
	(0.0239)		(0.0181)
L_m*young child 3-6	-0.0352		-0.0773***
	(0.0260)		(0.0202)
L_m*young child 6-10	0.00910		-0.0304
	(0.0237)		(0.0194)
L_m*number of children	0.00972**		0.00474*
	(0.00335)		(0.00235)
L_m * age40(m)	0.0326***		0.0255***
	(0.00699)		(0.00470)
L_m * higher degree(m)	0.143***		0.0129
	(0.0122)		(0.00719)
L_m * age40(f)	0.0657***		0.0484***
	(0.00740)		(0.00486)
L_m * higher degree(f)	0.0192		-0.00718
	(0.0113)		(0.00632)
L_f	0.249***		0.324***
	(0.0110)		(0.0108)
L_f*young child $<$ 3	-0.0643***		-0.122***
	(0.0183)		(0.0199)
L_f*young child 3-6	0.0232		0.0200
	(0.0175)		(0.0200)
L_f*young child 6-10	0.00643		-0.00395
	(0.0142)		(0.0164)
L_f*number of children	0.00682***		0.00418*
	(0.00198)		(0.00201)
L_f * age40(f)	0.0374***		0.0515***
	(0.00402)		(0.00396)
L_f * higher degree(f)	-0.00540		-0.0544***
	(0.00884)		(0.00550)
L_f* age40(m)	0.00577		0.0303***
	(0.00390)		(0.00386)
L_f * higher degree(m)	0.0173*		0.0105
	(0.00693)		(0.00560)
P_m	-6.689***		7.437***
	(0.440)		(2.224)
P_m*BKK	-2.428***		0.129
	(0.677)		(0.892)
P_m* $PreS$	2.536***		6.888***
	(0.501)		(0.850)
P_m* AS	1.004**		1.636***
	(0.307)		(0.490)

	100 draws	500 draws	Wage prediction
P_f	-5.855***		-8.630***
	(0.362)		(0.452)
P_f*BKK	-1.686*		-1.723
	(0.753)		(0.922)
P_f* $PreS$	0.565		2.231**
	(0.535)		(0.767)
P_f* AS	0.511		1.043*
	(0.326)		(0.454)
P_{hh}	-1.278***		0.886*
	(0.282)		(0.409)
P_{hh}*BKK	1.375		2.357*
	(0.743)		(0.924)
P_{hh}* $PreS$	-1.756***		-4.280***
	(0.475)		(0.730)
P_{hh}* AS	-0.154		-0.513
	(0.315)		(0.445)
SD:Y	0.000496		-0.0151***
	(0.000848)		(0.000698)
SD:L_m	0.186***		0.0549***
	(0.0125)		(0.00518)
SD:L_f	-0.00332		0.0178*
	(0.00263)		(0.00720)
SD:P_m	-3.133***		12.93***
	(0.593)		(1.341)
SD:P_f	4.630***		-7.036***
	(0.365)		(0.383)
Observations	1884938		1884938
Log lik.	-54534.7		-52817.7
	100 draws	500 draws	Report wages
MUyH7	0.0203494		-0.0147495
neg. %	0.37		74.26
MUmH7	0.0666032		-0.0333784
neg. %	38.42		59.09
MUfH7	-0.0420978		-0.1443651
neg. %	41.80		70.16
Quasi-concave H7	7.92		9.55
Monotonicity H7	0.00		18.28
Observations	15578		15578

Notes: Standard errors in parentheses; *, **, *** indicate statistical significance at 10%, 5%, and 1% levels respectively.