Family Structure and Female Labour Supply in Mexico City

Xiaodong Gong and Arthur van Soest[1]

Tilburg University

December 1997

Correspondence:

Xiaodong Gong, Arthur van Soest
Dept. of Econometrics, Tilburg University
P.O. Box 90153,
5000 LE, Tilburg, the Netherlands
E-mail: X.Gong@kub.nl, Avas@kub.nl

Abstract: This paper investigates labour supply of the wives of the heads of households in Mexico City, with a focus on the impact of family structure. A static neoclassical structural model is used. We assume that each woman chooses her labour supply and corresponding income so that her utility is maximized, conditional upon her husband's labour supply and earnings. We use a direct translog specification, and include family composition variables as taste shifters. Also taken into account are fixed costs of working, nonlinear taxes, unobserved preference variation, prediction errors in wages of nonworkers, and potential endogeneity of wages. The models are estimated by smooth simulated maximum likelihood using data from Mexico's Urban Employment Survey drawn in 1992. We find income elasticities of labour supply of about -0.35, and wage elasticities of about 0.5. The latter is substantially overestimated if wage rate endogeneity is not taken into account. The results are robust with respect to other specification choices. We find that the impact of family structure variables on participation is different from that on hours worked, so that their total effect is ambiguous.

JEL CODES: C31, C35, J16, J22, O54.

KEYWORDS: Female labour supply, family structure, discrete regression.

---

[1] Elizabeth Villagomez is gratefully acknowledged for her help and for providing the data. We thank Rob Euwals and Bernard Conlon for useful comments. Research of the second author has been made possible by a fellowship of the Netherlands Royal Academy of Arts and Sciences (KNAW).

I. Introduction

We analyse labour supply and labour force participation of married women in Mexico City. Labour force participation of women in Mexico is low. For example, in 1987, 45% of females (age 15 to 64) in EC countries and 55% of females in the seven major industrialized countries were employed (CBS, 1993), while according to our 1992 survey, only 38% of women (age 15 to 64) were employed in Mexico City. For the urban areas of Mexico as a whole, the women's participation rate was 36% in 1993 (Fleck & Sorrentino, 1994). We aim at explaining this phenomenon using a structural model of female labour supply. Such models have been estimated for many countries, but this is, to our knowledge, the first such study for Mexico. This seems particularly interesting since Mexico shares characteristics of industrialized and developing countries. It has gone through a period of serious economic transition and as of 1992, its per capita income (measured at purchasing power parity) is about that of lower income OECD countries such as Greece and Turkey.

As usual in this type of analysis, we estimate wage and other income elasticities. Moreover, we focus on the role of family structure, which is a potentially important determinant of labour supply. The average family in Mexico City had 4.8 persons in 1992. About 20% of the families had more than five persons and about 26% of the households were extended families. In many families, the presence of other females offers a potential substitute for child care services. In developing countries, large size and complicated composition of families are prevalent and surely influence people's behaviour. This has received little emphasis in the empirical structural labour supply literature. An exception is Newman & Gertler (1994), who examined rural Peru. They find that family structure closely relates to the household production and so to labour supply; the family values the leisure of various members differently, and the own marginal return to farm work is affected by the amount of work performed by other family members.

Wong & Levine (1992) study the effect of household structure on labour force participation of recent mothers in urban Mexico. They formulate reduced-form equations of female labour force participation and fertility. They find that the presence of a "mother substitute" significantly increases the labour force participation of females who have newly given birth to a child. Such findings are not confined to developing countries. Tienda & Glass (1985), using a similar model, find that in the US, the presence of other adults increased the probability of labour force participation of mothers who were heads of households. Neither Wong & Levine (1992) nor Tienda & Glass (1985) use a structural labour supply model.

Our data define a household as “the set of individuals living in the same house sharing a common income (provided by one or more of the individuals making up the household) destined to cover food, rent and housing utilities for all the members. This includes individuals living on their own” (see also Villagomez, 1996). This means that households not only include nuclear families but also extended families of more complicated structure. The number of household members in our sample varies from 2 to 19. In some households, the servants and their families are also included. We confine ourselves to the female spouse (including cohabiting partners) of the head of the household, whom we call the wife or the mother (if she has children).

Our paper differs from previous work by examining one structural model, in which the impact of wages, other income, and family structure on labour force participation and hours worked is analysed simultaneously. We follow the framework of van Soest (1995) and its extensions by Callan & van Soest (1995) and Euwals & van Soest (1996). We analyse female labour supply and take the husbands behaviour as given, following, for example, Hausman (1985). This is simpler than the family labour supply model with joint utility maximization used by, for example, Hausman & Ruud (1984). The simplification can be justified by the empirical finding in the latter type of models that cross elasticities of male labour supply with respect to the wife’s wage tend to be small (see Van Soest, 1995, for example).

The paper is organised as follows. In section II, we describe the model. Section III describes the data, which stem from the Urban Employment Survey conducted in the second quarter of 1992 by Mexico Statistical Institute. In section IV the estimation results are presented. In section V we discuss the results of some simulations and sensitivity and misspecification analysis. Section VI concludes.

II. Model and Estimation Method

We follow the discrete choice approach of Van Soest (1995). He assumes that the agent maximizes utility over a finite choice set. This approach has several advantages compared to the traditional (continuous) models.[2] First, it does not require convexity of budget set or preferences. Second, the approach makes it computationally feasible to incorporate nonstandard budget restrictions (fixed costs, hours constraints, nonlinear taxation, unemployment benefits, etc.), which enlarges the scope for policy analysis. Third, flexible functional forms of the direct utility function can be used, without the need for analytic expressions of the labour supply or the expenditure function. Fourth, the stochastic specification can be allowed to be quite rich, for example allowing for prediction errors of unobserved wage rates of nonworkers and wage rate endogeneity, random preferences, and random errors in fixed costs.

We assume that the woman decides on her leisure, *l*, and after-tax income, *y*, composed of her own labour income, her husband's and her children's earnings.[3] Leisure is set to be equal to *TE-h*, where *h* is working hours per week and *TE* is the time endowment, which we set equal to 80 hours per week.[4] We also assume that each woman maximises utility given by the direct translog specification:

$$U(v) = v'Av + b'v \quad (1)$$

where $v = (\log y, \log l)'$. A is a symmetric 2×2 matrix with entries A_{ij} (i, j = 1,2), and $b = (b_1, b_2)'$. Preference variation across individuals through observed and unobserved characteristics is incorporated through one of the parameters:

$$b_2 = \Sigma_k \beta_{2k} x_k + \epsilon^r, \quad (2)$$

---

[2] The traditional models is, for example, described in Hausman (1985) or Moffitt (1986).

[3] Due to lack of information , we are not able to consider the nonlabour income of the family such as asset income.

[4] Setting *TE*=80 hours per week is *ad hoc*, but results of Van Soest (1995) and Euwals & van Soest (1996) suggest that results are insensitive to this.

where $\boldsymbol{x}=(x_1,...,x_K)'$ is a vector of exogenous characteristics, such as age and family composition. The error term ϵ^r is interpreted as random preferences due to the unobserved characteristics. It is assumed to be normally distributed with mean zero, independent of $\boldsymbol{x}$.

We assume that U is increasing in y, implying that each woman will choose a point on the frontier of her budget set. The woman's before tax wage rate w is assumed not to depend on hours worked. Thus, once l is chosen, after tax income y is determined by w: $y=y(l,w)$.

In the traditional standard continuous model, the individual solves the problem:

$$\text{Max } U(y, l) \text{ s.t. } y \leq y(l, w), \; l \leq TE. \tag{3}$$

This can be solved using Lagrange techniques, but the shape of the budget set determines the complexity of the solution. Following Van Soest (1995), we discretize the budget set, replacing the budget frontier by some of its points. The optimization problem then becomes:

$$\begin{aligned} &\text{Max } U(y, l) \\ &\quad \text{s.t. } (y, l) \in CS(w) \end{aligned} \tag{4}$$

where the choice set is given by

$$CS(w) = \{(y, TE\text{-}h);\; h \in \{0, IL, ..., (m\text{-}1)IL\},\; y=y(TE\text{-}h, w)\} \tag{5}$$

Here IL is a fixed interval length for the working hours. These are rounded to a multiple of IL and censored at $(m\text{-}1)IL$. The choice set with m points is denoted by $\{(y_0, l_0), ..., (y_{m-}, l_{m-1})\}$. In our base case empirical specification, we use IL = 10 and m = 8. In sensitivity analysis, we also try IL = 4 and m = 19.

To the utilities of all the alternatives in the choice set, random disturbances are added as in the multinomial logit model (Maddala, 1983):

$$U_j = U(y_j, l_j) + \epsilon_j \qquad (j = 0, ..., m\text{-}1) \tag{6}$$

where the ϵ_j is i.i.d. with a type I extreme value distribution, and are independent of $\boldsymbol{x}$ and of other error terms in the model. ϵ_j can be interpreted as an optimization error, but not as random preferences. The latter are already represented by ϵ^r in (2). The individual chooses j if U_j is the larger than the other U_i . Thus, conditional on ϵ^r, $\boldsymbol{x}$, and w, the probability that j is chosen is

$$Pr[U_j \geq U_i \ \ for \ \ all \ \ i] = \frac{\exp(U(y_j,\ l_j))}{\sum_{i=0}^{m-1} \exp(U(y_i,\ l_i))} \tag{7}$$

Wage rates of the nonworkers are not observed. We need a wage equation to predict them. The wage equation is also needed to investigate endogeneity of wages in the labour supply model. It is defined as:

$$\text{Log}\ w = \pi\text{'z} + \mu \tag{8}$$

where z is the vector of individual characteristics (education level, for example), π is the vector of parameters, and μ is the error term which we assume to be normally distributed with mean zero, independent of z. Initially, we assume that μ is uncorrelated with the random preference term in (2). We relax this assumption below to allow for endogeneity of wage rates in the labour supply model. In the latter case, we assume that ϵ^r is correlated to μ with an arbitrary covariance structure and that the correlation efficient is ρ.

As in van Soest (1995), the model described so far appears to underpredict the number of nonworkers and overpredicts the number of part-time jobs involving a few hours a week. Unobserved fixed costs of working, such as commuting costs or child care costs, might be responsible for this. Because we use a direct utility function, we can incorporate fixed costs in a natural way: fixed costs are subtracted from income y if h>0. Equivalently and computationally more conveniently, we add fixed revenues of not working (FR) to the income at zero hours of work. So $U(y_0, l_0)$ is replaced by $U(y_0+FR, l_0)$. FR is specified as follows:

$$FR = \delta\text{'}t + \epsilon^f \tag{9}$$

t is a vector of exogenous variables, δ a vector of parameters. We assume that the error term ϵ^f is normally distributed with mean zero and is independent of explanatory variables and other error terms in the model. Positive fixed revenues increase the probability of nonworking by increasing the utility of nonparticipation (since utility increases with income). The fixed revenues are fully incorporated in the structural model. For example, an increase in wages will increase $U(y_j, l_j)$ for $j>0$, but does not change the utility of not working, and so increase the participation rate. Therefore, the effects of wage (or tax, benefits, etc.) changes on participation can be easily analysed in the simulations. Compared with the model conditional on participation (for example, Blundell, 1987), this is an important advantage for policy analysis.

Estimation

For computational convenience, we first estimate equation (8) separately. To account for selection bias, we add a reduced form participation equation, and estimate the two-equation system by maximum likelihood, following Heckman (1979).

For estimating the structural labour supply model, we use a simulated maximum likelihood method approach, as in Van Soest (1995). The standard model, without random preferences or fixed revenues (or fixed costs) and with observed wage rates only, could be estimated by maximum likelihood. The likelihood contribution would be given by (7). For unobserved wages, the wage prediction errors have to be integrated out using (8). When random preferences and fixed revenues are considered, two additional error terms have to be integrated out as well. This requires multidimensional numerical integration. We denote the probabilities of working hours conditional on ϵ^r and ϵ^f as a function of the wage w (and for given earnings of the husband and children, and explanatory variables $\boldsymbol{x}$, z and t) by

$$Pr[\, h = h_j | \epsilon^r, \epsilon^f] = F_j(w | \epsilon^r, \epsilon^f) \; (j=1, ..., m), \tag{10}$$

where F_j is given by (7). The exact likelihood contribution is then given by

$$L=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} F_j(w^0,\ x|\epsilon^r,\epsilon^f)p_1(\epsilon^r|w^0)p_2(\epsilon^f)d\epsilon^r d\epsilon^f, \tag{11}$$

if w^0 is observed, or,

$$L=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}F_j(w,\ x|\epsilon^r,\epsilon^f)p(\epsilon^r,w)p_2(\epsilon^f)d\epsilon^r d\epsilon^f dw, \tag{12}$$

if w is not observed. p_1, p_2 are density functions of ϵ^r *and* ϵ^f, respectively. $p(\epsilon^r,w)$ is the joint density of ϵ^r and w. To avoid the multidimensional integration, we approximate the integral by a simulated mean: for each individual, we take R drawings from the distribution of the error terms, and calculate the average of the R likelihood values conditional on the drawn errors. Then the integral (12) can be approximated by

$$L_R=\frac{1}{R}\sum_{q=1}^{R}F_j(w_q,\ x|\epsilon^r{}_q,\ \epsilon^f{}_q) \tag{13}$$

where $(w_q,\ \epsilon^r{}_q,\ \epsilon^f{}_q)$, $q= 1, ..., R$, are independent draws from the conditional distribution of $(w,\ \epsilon^r,\ \epsilon^f)$ on Z. The integral (11) can be replaced by

$$L_R=\frac{1}{R}\sum_{q=1}^{R}F_j(w^0,\ x|\epsilon^r{}_q,\ \epsilon^f{}_q) \tag{14}$$

The resulting estimator is inconsistent for fixed R, but will be consistent if R tends to infinity with the number of observations (n). If $n^{1/2}/R \to 0$, the method is asymptotically equivalent to maximum likelihood, see Gourieroux and Monfort (1993).

III. Data

The data we use are drawn from Mexico's Urban Employment Survey (Encuesta Nacional de Empleo Urbano, second quarter of 1992), conducted by *Instituto Nacional de Estadistica, Geografia e Informatica* (INEGI, i.e. Mexican Statistical Institute) in 32 Mexican cities. The survey is the only quarterly household survey in Mexico and the source of official open unemployment rates. It provides detailed information on the economic activities of all the household members older than twelve years of age, such as employment status, employment conditions, working hours, labour income, characteristics of the workplace, etc., but no information on nonlabour income. It has been used for unemployment analysis in urban Mexico, by, for example, Fleck & Sorrentino (1994). It has been used by Villagomez (1996) for an analysis of the labour market in Mexico City, with the focus on the impact of segmentation on the individual's labour supply. In this paper, we use the sub-sample for married couples in Mexico City in which each partner is less than 65 years old and the husband is employed. This gives observations on 3008 households. Some observations, however, are incomplete. 302 observations have no information on the husband's income, 173 observations do not have information on the wife's income, in 3 observations both "husband" and wife are females, in 20 observations the wife is retired or a full-time student. After eliminating these observations, we get a sample of 2510 families. The means and standard deviations of the variables used in the analysis are presented in Table 1.

In 74 percent of the households, the wife does not work. Figure 1 gives the distribution of working hours for the working wives. 31 percent of them work less than 30 hours per week (compared to only 8.6 percent of their husbands). About 48 percent of the wives have 6-year basic education or less.

Figure 2 gives the distribution of the working wives' after-tax wage rates. The mean of (*log*) wage rate is 1.61, which is almost the same as their husbands'. The mean of the latter is 1.63. So, wage differential is not the reason of the low female participation. To get an idea about the pattern of the participation rate against age, we plot the nonparametric estimate of the participation probability as a function of age. Figure 3 concerns the complete sample. The pattern is inversely U-shaped. Women of about 30 years of age work most. Figure 4 shows the pattern for different subsamples. The figure as a whole does not reveal a clear relation between participation and the presence of children or other adult females in the household if age is controlled for. A model with more structure seems needed for this.

We use number of adults, number of young children (of the head of the household), presence of other adult female(s), and joint presence of other female and young children as indicators of family structure.[5] The average family has 3.4 members who are older than 12 years. 48 percent of the families have young children (age at most 12), 26 percent of these have adult females (Age > 12, including sisters of young children) other than the mother. Adult females

[5] We also replaced joint presence of other female and young children with joint presence of elder sister and young children in our model, and find that the results are very similar.

other than the mother are present in about 35 percent of all families.

Table 1: Sample statistics

Variables	*Description*	*Mean*	*Std. Dev.*
age	age	35.6	10.0
agesq	age square	1368	776
hour	working hours	8.86	16.7
dnosch	dummy no schooling	.049	.217
dbasic	dummy 6-year basic education	.428	.495
djunior	dummy 9-year education	.187	.390
dsenior	dummy 12-year education	.048	.213
dcolleg	dummy university education	.078	.270
dtechb	dummy vocational education plus 6-year-schooling	.041	.197
dtechj	dummy vocational education plus 9-year-schooling	.160	.365
dtechs	dummy vocational education plus 12-year-schooling	.010	.097
chidern	total weekly earnings of all unmarried children of head of household (pesos)	52.3	154
husern	weekly earnings of husband (pesos)	312	442
othern	total weekly earnings of other household members (pesos)	24.0	95.6
child12	number of children of head of household aged 12 or less	.876	1.09
chdofem	dummy presence of both other adult female and young children	.127	.333
adult	number of adults (older than 12)	3.40	1.67
ofemale	dummy presence of other female adult	.353	.478
eldis	number of elderly and disabled	.029	.181
lnwage	log hourly wage(after-tax, pesos)	1.63	.785

Table 2 gives the means of some variables for several sub-samples. The low educated group refers to females with at most six years of schooling, and the high educated group are all the others. Compared to those with lower education level, highly educated individuals tend to be younger, participate more often, and, given participation, have higher wage rates and work more hours. Their husbands' incomes are higher because the husband also tends to have high education level. Wives with high education level have more young children but fewer adult females in the household.

Table 2. Means by subsample

Variable	***Education***		***Family with children***			***Family without children***	***Family size (No. of Adults)***	
	Low	***High***	***Other female***	***No other female***	***As a whole***		***>4***	***≤4***
obs.	1199	1311	319	888	1207	1303	572	1938
lnwage	1.20	1.87	1.51	1.74	1.67	1.60	1.55	1.66
age	38.3	33.1	36.8	30.0	31.8	39.1	44.5	33.0
huern	203	411	300	314	310	314	324	309
chidern	77.4	29.4	33.1	8.2	14.8	87.1	178.2	15.2
child12	.78	.96	1.66	1.88	1.82	0	0.37	1.03
Hours for workers	33.65	35.02	32.48	34.59	33.90	35.18	35.68	34.23
adults	3.91	2.94	4.29	2.36	2.87	3.90	5.94	2.65
participation rate	18.7	32.0	32.9	24.1	26.4	24.9	24.7	27.3

In families with children, the wives participate more often if there is another adult female. Given participation however, they work fewer hours than the wives in other families. This might be because in families with young children and another adult female, the wife is usually younger (cf. figure 2) and can earn more per hour. Comparing all wives with children with those without children, we find that the former participate more often. Conditional on participation, however, wives without children work more hours. The wives in large families (with more than four adult members) are older, participate less often, but, conditional on participation, they work more hours.

Many of these results suggest that the effects of family composition on participation and hours worked are different. We allow for this in the structural model by incorporating fixed costs of working.

In the empirical literature on structural labour supply models, the role of progressive income taxes is often emphasized (see Hausman, 1985, for example). For the current study on Mexico, however, the relevance of explicitly including the income tax in the analysis is doubtful. Mexico's tax system is described in the appendix. The relation between before and after tax incomes is almost linear. Moreover, nearly 36% of the workers in the sample work in the informal sector, where income taxes are hardly collected. Finally, although the situation has improved, tax evasion is still common due to poor tax collection, long collection lags, and high inflation rates. In 1991, only 17.1 million individuals registered tax payers in a population of about 86 million (Armella, 1992), and total tax income was 15.6% of GDP. Individual income tax is just a small part of this, amounting to 2.0% of GDP in 1990 (OECD Economic Surveys, 1992). We will use only after-tax income in the benchmark model, ignoring the nonlinearity in the income tax, and discuss a model which incorporates the features of the income tax system in our sensitivity analysis. The survey contains information on after-tax earnings per week. Log after-tax hourly wages are computed from this and hours worked per week. In the model with income tax, before-tax wage rates are recovered from the after-tax income using the details of the tax brackets. It is assumed that before tax wage rates do not vary with hours worked. In the benchmark model, it is assumed that after-tax wage rates do not vary with hours worked.

The husband's earnings, together with earnings of unmarried children, are considered as the wife's nonlabour income. In the models we present, we do not include income of other family members. We included this as a separate explanatory variable, but found it was insignificant. This is plausible because the individual's labour supply decision is independent of those incomes which she cannot control.

IV. Estimation Results

Table 3 shows the estimates of the Heckman (1979) model. They are in line with the common findings in the literature. For example, the higher the education level, the higher the woman's wage offer. The more husband and children earn, the less likely the wife works. On the other hand, income of other family members does not affect the wife's participation decision. The correlation coefficient between error terms in participation and wage equation (ρ^w), is significantly positive, implying that the self-selection induces an upwards bias on wages.

Table 3. Estimation Results Heckman (1979) Model

	wage equation	*t-value*	*participation equation*	*t-value*
age	0.142	5.50	0.130	6.06
agesq	-0.002	-5.04	-0.002	-5.72
nosch	-0.111	-0.65	-0.123	-0.84
djunior	0.351	3.84	0.171	2.08
dsenior	0.784	5.47	0.420	3.10
dtechb	0.403	2.54	0.109	0.75
dtechj	1.062	11.7	0.663	8.18
dtechs	1.362	5.57	0.897	3.46
dcolleg	1.692	15.4	1.069	9.85
ofemale	_		0.024	0.34
chdofem	_		0.178	1.92
child12	_		-0.060	-2.25
huern/100	_		-0.029	-4.46
chidern/100	_		-0.030	-1.97
othern/100	_		-0.001	-0.03
constant	-2.64	-4.93	-3.272	-8.23
....	....	....	....	....
ρ^w	.808	11.9		
σ	.884	16.5		
λ	.714			

We present the estimation results for the labour supply models with fixed costs[6] in Table 4. We present the results for two models, the first with wage rates assumed to be exogenous, and the second allowing for endogenous wage rates. The SML estimates of both models are based

[6] The estimated standard deviation of the error term in fixed revenues (σ^f) appeared to be close to zero with the standard error larger than the estimate in all specifications. We simply ignore it in estimation.

upon R = 20 draws per household. We used IL=10 and m=8. The estimated utility function is increasing in income at almost all sample points (except for 6 observations in the model with endogenous wage rates).

Table 4. Estimation Results Structural Models

	model (exogenous wage)	***model (endogenous wage)***
A_{11}	1.10(4.46*)	1.18(5.23*)
A_{12}	0.67(5.34*)	1.35(9.75*)
A_{22}	-2.68(-8.22*)	-3.30(-8.79*)
b_1	-3.31(-0.91)	-13.88(-4.43*)
logl	32.61(6.27*)	33.62(6.48*)
logl*ofemale	-0.14(-0.23)	0.10(0.16)
logl*chdofem	0.44(0.63)	0.69(0.94)
logl*age	-0.37(-2.65*)	-0.69(-4.53*)
logl*agesq	0.005(2.76*)	.009(4.65*)
logl*child12	.0.36(1.50)	0.46(1.92**)
logl*adult	-0.14(-0.80)	0.003(0.02)
σ^r	7.66(10.8*)	7.56(9.37*)
ρ	_	-0.30(-6.73*)
fixed costs		
constant	61.59(5.50*)	80.5(4.90*)
ofemale	-10.72(-1.70**)	-13.3(-1.41)
chdofem	-13.66(-1.90**)	-22.7(-2.10*)
age	-0.29(-1.12)	-0.29(-0.76)
child12	-0.26(-0.12)	-0.25(-0.08)
adult	3.94(1.91**)	5.28(1.76**)
eldis	7.48(0.61)	11.1(0.64)

t-values are in parentheses. * significant at 5% level. ** significant at 10% level.

The signs of the parameters determine the way in which characteristics affect preferences.

A positive β_{2k} implies a positive effect on the marginal utility of leisure, and a negative effect on labour supply. A positive parameter of a variable in fixed costs indicates that the variable relates negatively to the probability of participation. Age significantly affects labour supply in both models, implying maximum labour supply at age about 37 *(ceteris paribus)*. Age is insignificant in the fixed costs equation. A Wald test shows that the number of adults in the family is jointly insignificant in labour supply and fixed costs equations. The number of young children has a significantly negative effect on labour supply in the second model. It plays no role for the fixed costs. The same holds for the number of elderly and disabled people.

The joint presence of young children and another adult female decreases the fixed costs of working, thus increasing labour force participation of mothers. Nevertheless, it negatively affects labour supply. Thus, the overall effect is also ambiguous. To show the effects of these variables, we have drawn some labour supply curves in figures 5-7. These are based on simulations using the complete model, and take account of fixed costs and error terms.[7] The wife's age is set to 35 years. Figure 5 presents labour supply curves for families including and excluding other adult females. In both of the families there are 1 child and 4 adults. At low wages, the joint presence of child and other females increases labour supply. Beyond a certain wage, the effect has the reverse sign. Most wages in the sample are between 2 and 12 pesos per hour, where the effect of fixed costs dominates and the total effect is positive. Figure 6 shows that the more children the wife has, the fewer hours she will work. Figure 7 shows that the number of adults in the family does not have much influence. Overall, the three figures suggest that the impact of family composition be quite limited, though statistically significant.

The significant estimate of ρ indicates that exogeneity of wage rates is rejected. Its negative sign implies a positive correlation of the error in the wage equation with unobserved preferences for labour supply.

Due to the complex structure of the model, the parameters A_{ij} is hard to interpret. In Figure 8, we draw some indifference curves in the *(y,h)* plane for families with one child, four adults including one other female. For other type of families, the figures are similar. We use the model with fixed costs and endogenous wages. Error terms and fixed costs are not taken into account. Age is set to 35. Utility levels increase from solid to dashed lines. The curves have the

[7] The curves thus reflect the expected value of labour supply, given wages and exogenous characteristics, but for unknown values of the error terms.

expected convex shape. At low hours, many curves are almost flat, indicating that nobody (ignoring error terms) would want to work part-time. This can explain why relatively few part-time jobs are observed in the data. Together with fixed revenues, random preferences, etc., the shape of the indifference curves determines the sensitivity of labour supply for changes in wages, other income, etc. This will be discussed in next section.

V. Simulations

Using simulations, we first examine the goodness of fit of the models, and then analyse the sensitivity of average labour supply and participation with respect to wage rate and other income. The third purpose of the simulation is to compare the different specifications of the model. In particular, we study the relevance of incorporating the tax system. The simulated hours per individual are the "expected hours," computed as a weighted sum of hours levels. From these, we compute average values for the whole sample and for several subsamples. Simulated hours given participation are the ratios of hours and the probability of participation. Wage and income elasticities are derived by increasing all wage rates or other incomes by 1% and calculating the percentage change of average hours.

Table 5 presents means of observed and simulated hours (given participation) and participation rates for different models. We split the sample according to the women's education level, family size and family type. The models fit the data reasonably well, though they somewhat overestimate the participation rates. For example, the model predicts the difference of working hours between those with high education and those with low education quite well, though there are no education variables in the structural model. Participation rates reflect the same pattern. The two models perform quite similarly; the assumption of wage rate exogeneity does not matter much for goodness of fit.

Simulated elasticities for the two benchmark models are presented in Tables 6 for the full sample and in Table 7 for some sub-samples. These numbers are point estimates. By repeating the simulations for a large number of draws from the asymptotic distribution of the parameter estimates, we also calculated 95% confident intervals. According to the model with exogenous wage rates, a rise of 1% in all wage rates would lead to a rise of average working hours by 0.72%, while participation would increase by 0.70%. If other income (i.e. husband's and unmarried children's income) increases by 1%, the average working hours would fall by 0.36%, and participation would decrease by 0.19%. When we drop the exogeneity assumption, the wage elasticities are smaller. This corresponds to the findings of Mroz (1987) for the US that imposing exogeneity of wage rates leads to overstated wage effects.

Table 5. Goodness of Fit: Hours worked (given participation) and Participation

	sample distribution	**model (exogenous wage)**	**model (endogenous wage)**
Whole sample	34.6(26.5%)	33.9(30.8%)	34.0(30.6%)
Low educated	33.7(18.7%)	32.9(22.2%)	32.5(24.1%)
High educated	35.0(32.0%)	35.0(38.6%)	35.6(36.5%)
Family with 4 or fewer adults	34.2(27.3%)	34.3(31.2%)	34.4(30.9%)
Family with more than 4 adults	35.7(24.7%)	32.5(29.4%)	32.6(29.3%)
Family with kids:			
All	33.9(26.4%)	33.7(32.1%)	33.6(31.5%)
No other female	34.6(24.1%)	34.4(29.6%)	34.4(29.0%)
With other females	32.5(32.9%)	31.8(39.0%)	31.5(38.7%)
Family without kids	35.2(24.9%)	34.1(29.5%)	34.4(29.7%)

Note: See Table 2 for number of observations. Percentages in parentheses are participation rates.

Table 6 Elasticities for the whole sample

	elasticities (exogenous wage)	***95% confidence intervals***	***elasticities (endogenous wage)***	***95% confidence intervals***
ϵ_{hw}	0.768	[0.721, 0.803]	0.556	[0.485, 0.615]
ϵ_{hi}	-0.380	[-0.425, -0.335]	-0.354	[-0.396, -.0.314]
ϵ_{pw}	0.747	[0.712, 0.774]	0.580	[0.528, 0.622]
ϵ_{pi}	-0.189	[-0.234, -0.148]	-0.166	[-0.210, -0.128]

Table 7 shows that other income elasticities are stable across different groups. The wives who are higher educated, are in a smaller family, or have kids, are more sensitive to wage changes than their counterparts.

Table 7 Elasticities for some sub-samples

	ϵ_{hw}	ϵ_{hi}	ϵ_{pw}	ϵ_{pi}
Model with endogenous wage				
(1)	0.519(.029)	-0.303(.024)	0.559(.031)	-0.090(.026)
(2)	0.581(.039)	-0.389(.020)	0.593(.021)	-0.212(.019)
(3)	0.571(.034)	-0.357(.021)	0.599(.025)	-0.176(.021)
(4)	0.502(.033)	-0.345(.024)	0.513(.027)	-0.132(.024)
(5)	0.587(.032)	-0.374(.023)	0.608(.024)	-0.190(.022)
(6)	0.527(.034)	-0.384(.026)	0.536(.031)	-0.199(.025)
(7)	0.612(.037)	-0.369(.023)	0.642(.030)	-0.185(.022)
(8)	0.528(.035)	-0.336(.021)	0.553(.028)	-0.143(.022)
Model with exogenous wage				
(1)	0.805(.032)	-0.316(.028)	0.814(.026)	-0.097(.029)
(2)	0.747(.017)	-0.416(.022)	0.711(.015)	-0.237(.019)
(3)	0.787(.021)	-0.388(.024)	0.770(.016)	-0.205(.022)
(4)	0.696(.029)	-0.350(.025)	0.664(.026)	-0.134(.025)
(5)	0.792(.022)	-0.405(.025)	0.768(.018)	-0.217(.022)
(6)	0.727(.036)	-0.417(.028)	0.688(.037)	-0.223(.024)
(7)	0.822(.024)	-0.400(.026)	0.806(.021)	-0.214(.023)
(8)	0.743(.024)	-0.355(.024)	0.725(.020)	-0.161(.022)

Standard errors of the elasticities are in parentheses.

(1) low educated; (2) high educated; (3) family with 4 or fewer adults; (4) family with more than 4 adults; (5) all families with young children; (6) family with young children and other females; (7) family with children without other females; (8) family without young children.

We present the simulation results of three alternative specifications with endogenous wages in Table 8. The first row repeats the results of the benchmark model. The second row uses a model with more points in the choice set (multiples of 4 hours per week instead of 10). In the third row, SML is based upon 30 instead of 20 draws per observation. The fourth row combines the two decisions. The table shows that in terms of predicted hours and participation, the specifications give similar results. In terms of elasticities, the difference between those R=20 and R=30 draws is small. But the wage elasticities do vary with the number of points in the choice set, with smaller values in the model with a finer hours grid.

The final row in Table 8 shows the results when we incorporate the progressive income tax system. The estimated elasticities are somewhat smaller than in the benchmark model.

Table 8. Simulation results of different models with endogenous wage rates

	hours	***participation rates***	ϵ_{hi}	ϵ_{pi}	ϵ_{hw}	ϵ_{pw}
IL=10, R=20	34.0	30.6%	-0.354	-0.166	0.556	0.580
IL=4, R=20	33.8	29.6%	-0.299	-0.060	0.430	0.430
IL=10, R=30	35.0	30.4%	-0.377	-0.219	0.554	0.601
IL=4, R=30	35.0	29.7%	-0.318	-0.107	0.430	0.463
IL=10, R=20, including tax	33.5	30.5%	-0.290	-0.075	0.412	0.440

To test formally for model misspecification, we use the chi-square goodness of fit tests introduced by Andrews (1988). These tests are generalizations of the traditional Pearson chi-square tests for the multinomial model. They are based upon partitioning the sample space into a given number of cells and comparing sample probabilities with predicted probabilities, given the parameter estimates and the covariates. In our case the test statistic is based upon the ML estimates, and the test statistic can be computed straightforwardly. Let A be the $n \times J$ matrix with the differences between observed and predicted cell probabilities per observation.[8] Let B be the $n \times K$ matrix of scores. And let C be the matrix $[A|B]$. The test statistic is then equal to the

[8] The observed probability is one if the individual belongs to the cell and zero otherwise. The predicted probability is the probability that the individual belongs to the cell, given the covariates and the parameter estimates.

explained sum of squares of the regression of a vector of ones the columns of *C*. If the cells are constructed as products of a partition of the covariates space (*X*) and the space of endogenous variables (working hours, *H*), then under the null of a correct specification, the test statistic follows a chi-square distribution with *J-G* degrees of freedom, where *G* is the number of cells in the partition of *X*.

We computed eight test statistics based on different partitions of $X \times H$. For *X*, we use four partitions: a) no partitioning, b) two education categories (low level vs. high level), c) two family size categories (small vs. large families), d) three types of families (with children and other females, with children but no other females, without children). For *H* we use two partitions: a) a partition into eight hours categories, and b) a partion into the three categories not working, working part-time, and working full-time. The results are presented in Table 9. All the tests lead to rejecting the null-hypothesis of no missspecification. This reflects the finding that participation rates in the data are not reproduced too well by the model. The fact that the model specification if formally rejected by the data is not uncommon in the literature. As far as we know, in the few studies that explicitly test for this, the result is the same (Magnac, 1991, Pradhan and van Soest, 1997).

Table 9. Chi-square tests

Part. of X	***Part. of H***	***Test statistic***	***J-G***	***Critical value $\alpha = 0.01$***
a)	a)	336	7	18.5
a)	b)	238	2	9.2
b)	a)	405	14	29.1
b)	b)	272	4	13.3
c)	a)	364	14	29.1
c)	b)	242	4	13.3
d)	a)	395	22	40.3
d)	b)	244	6	16.8

Partitions of *X*: a) no partition
b) two education categories
c) two family size categories
d) three family type categories

Partitions of *H*: a) eight cells: non-workers, seven cells of workers partitioned according to working hours; b) three cells: 0 hours, between 0 and 40 hours, 40 hours of work or more.

VI. Conclusions

We have analysed labour supply of married women in Mexico City, emphasizing the influence of family structures. We used a static structural neoclassical model, extending the model of Van Soest (1995). We have taken account of fixed costs of working, and found that they are positive for all the observations. We have incorporated random preferences, we have corrected for prediction errors in imputed wages of nonworkers, and we have allowed for wage rate endogeneity.

We find that the overall effects of family structure on labour supply are ambiguous, with opposite effects through fixed costs of working and preferences. Nevertheless, in the range where most observed wages are found, the presence of another female increases labour supply of mothers with young children. Similar to Mroz (1987), we find that ignoring endogeneity of wage rates leads to overestimated wage elasticities. The elasticities we find are in line with those in the literature: in the benchmark model with endogenous wage rates, the uncompensated wage elasticity of average hours worked is 0.56, the other income elasticity is -0.35.[9] Test results show that the model is still misspecified, a common finding with this type of structural models.

References

Andrews, D. (1988), "Chi-square Diagnostic Tests for Econometric Models - introduction and applications," *Journal of Econometrics 37, 135-156.*

Armella, P. (1992), "Tax Policy in Mexico," *Bulletin for international Fiscal Documentation: Official Organ of the International Fiscal Association* 46, 414-416.

Blundell, R. (1987), "Econometric approaches to the specification of life-cycle labour supply and commodity demand behaviour," *Econometric reviews* 6, 103-165.

Callan, T. & van Soest, A. (1995), "Family Labour Supply and Taxes in Ireland," working paper, Tilburg University.

[9] It should be noted that many different elasticities are used in the empirical literature, limiting the value of comparisons. For example, many studies consider the elasticity for the average household, while we study the elasticity of average hours worked. Also, the other income elasticity depends on what is included in other income.

CBS, (1993), "Arbeidsdeelname in Nederland vanaf 1961 in Intenationaal Perspectief," *Supplement Sociaal-economische Maandstatistiek* 1993(2): 13-19.

Euwals, R. & van Soest, A. (1996), "Desired and Actual Labour Supply of Unmarried Men and women in the Netherlands," Discussion paper 9623, CentER of Economic Research, Tilburg University.

Fleck, S. & Sorrentino, C. (1994), "Employment and Unemployment in Mexico's labour force," *Monthly Labour Review* 117(11), 3-31.

Gourieroux, C. & Monfort, A., (1993) "Simulation based inference: A Survey with Special Reference to Panel Data," *Journal of Econometrics* 59(1-2), 5-34.

Hausman, J. & Ruud, P. (1984), "Family Labour Supply with Taxes," *American Economic Review* 74, 242-248.

Hausman, J. (1985), "The Econometrics of nonlinear Budget Sets," *Econometrica* 42, 679-694.

Heckman, J. (1979), "Sample Selection Bias as a Specification Error," *Econometrica* 47(1), 153-161.

Maddala, G. (1983), *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.

Magnac, Th. (1991), "Segmented or Competitive Labour Markets," *Econometrica* 59, 165-187.

Moffitt, R. (1986), "The Econometrics of Piecewise-Linear Budget Constraints," *Journal of Business and Economic Statistics* 4, 317-328.

Morz, T. (1987), "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions," *Econometrica* 55, 765-99.

Newman, J. & Gertler, P. (1994), "Family Productivity, Labour Supply, and Welfare in a Low Income Country," *The Journal of Human Resources* 29, 989-1026.

OECD Economic Surveys -Mexico, 1992, OECD publications, Paris.

Pradhan, M. & van Soest, A. (1997), "Household Labour Supply in Urban Areas of Bolivia," *The Review of Economics and Statistics* 74, 300-310.

Tienda, M. & Glass, J. (1985), "Household Structure and Labour Force Participation of Black, Hispanic, and White Mothers," *Demography* 22, 381-394.

Van Soest, A. (1995), "Structural Models of Family Labour Supply -A Discrete Choice Approach," *The Journal of Human Resources* 30, 63-88.

Villagomez, E (1996), "Informal Markets: A Model of Household Labour Supply," working paper, Universidad de Alcala de Henares.

Wong, R. & Levine, R. (1992), "The Effect of Household Structure on Women's Economic Activity and Fertility: Evidence from Recent Mothers in Urban Mexico," *Economic Development and Cultural Change* 41, 89-102.

Appendix: Mexico's income tax system

Mexico has an individual based income tax system. It is progressive with 8 brackets, which are indexed to inflation. To compensate for the high inflation rate, subsidies are used to reduce the effective marginal rate. The subsidy rates are also progressive. The marginal tax rate ranges from 3% to 35%, but if subsidies are taken into account, they vary only from 3% to 17%. Many components of income are tax-exempt, for example, fringe benefits, overtime pay, and social insurance benefits (OECD Economic Surveys, 1992).

In the sample, 545 out of 644 workers fall in the brackets with a marginal rate of 17% or less. Taking account of the subsidy, their effective marginal rate is at most 8.5%. In Figure A1, the wives' yearly after-tax earnings are plotted against yearly before-tax earnings. The relation is close to linear. Regressing after-tax earnings on before-tax earnings for workers, gives a slope coefficient of about 0.82 (with standard error 0.002), and an R^2 of 0.99.

To incorporate the tax system in the structural model, we also need to predict before-tax wages for nonworkers instead of after tax wages. Table A1 gives the results of Heckman (1979) selection model. They are very similar to those for after-tax wages in Table 3.

Table A1. Estimation Results Heckman (1979) Model for Before-tax Wages

	wage equation	*t-val.*	*participation equation*	*t-val.*
age	.148	5.60	.130	6.06
agesq	-.002	-5.13	-.002	-5.72
nosch	-.122	-0.69	-.123	-.841
djunior	.363	3.85	.171	2.08
dsenior	.806	5.44	.417	3.08
dtechb	.415	2.52	.110	.751
dtechj	1.10	11.9	.662	8.18
dtechs	1.42	5.61	.901	3.49
dcolleg	1.76	15.8	1.07	9.89
ofemale	_		-.023	-.332
chdofem	_		.175	1.91
child12	_		-.059	-2.25
huern/100	_		-.029	-4.44
chidern/100	_		-.030	-1.60
othern/100	_		-.001	-.052
constant	-2.78	-5.12	-3.27	-8.30
…………………	………………………	………	……………………………	………………
ρ^w	.824	11.7		
σ	.924	18.1		
λ	.762			

Figure 1. Distribution of hours of workers

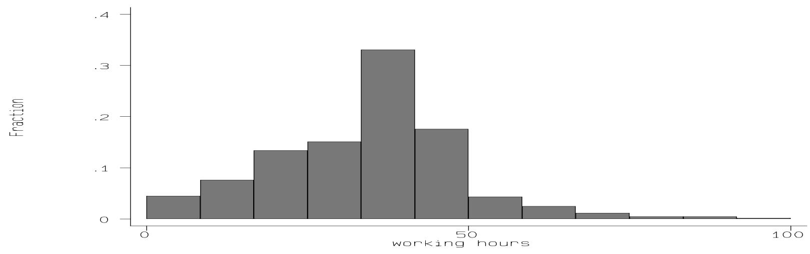

Figure 2 Distribution of after-tax wage rates

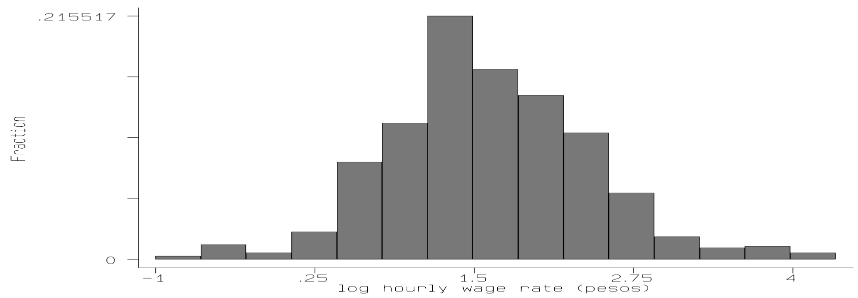

Figure 3. Nonparametric estimate of participation rates on age with 95% uniform confidence interval (bandwidth = 8 hours)

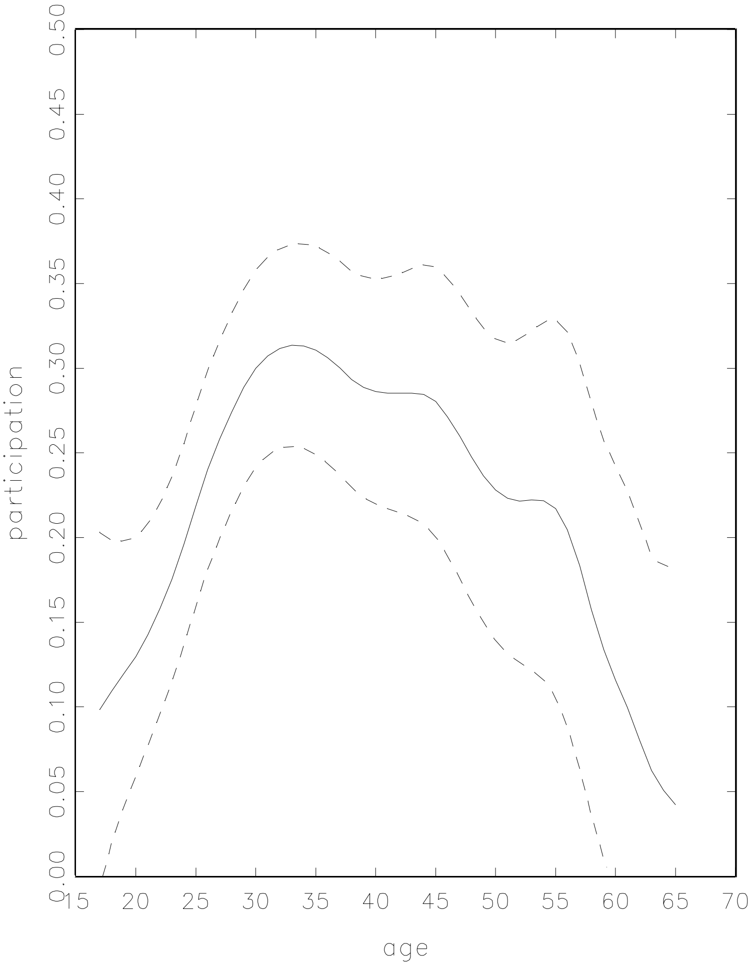

Figure 4. Nonparametric estimate of participation rates on age for subsamples with 95% uniform confidence intervals

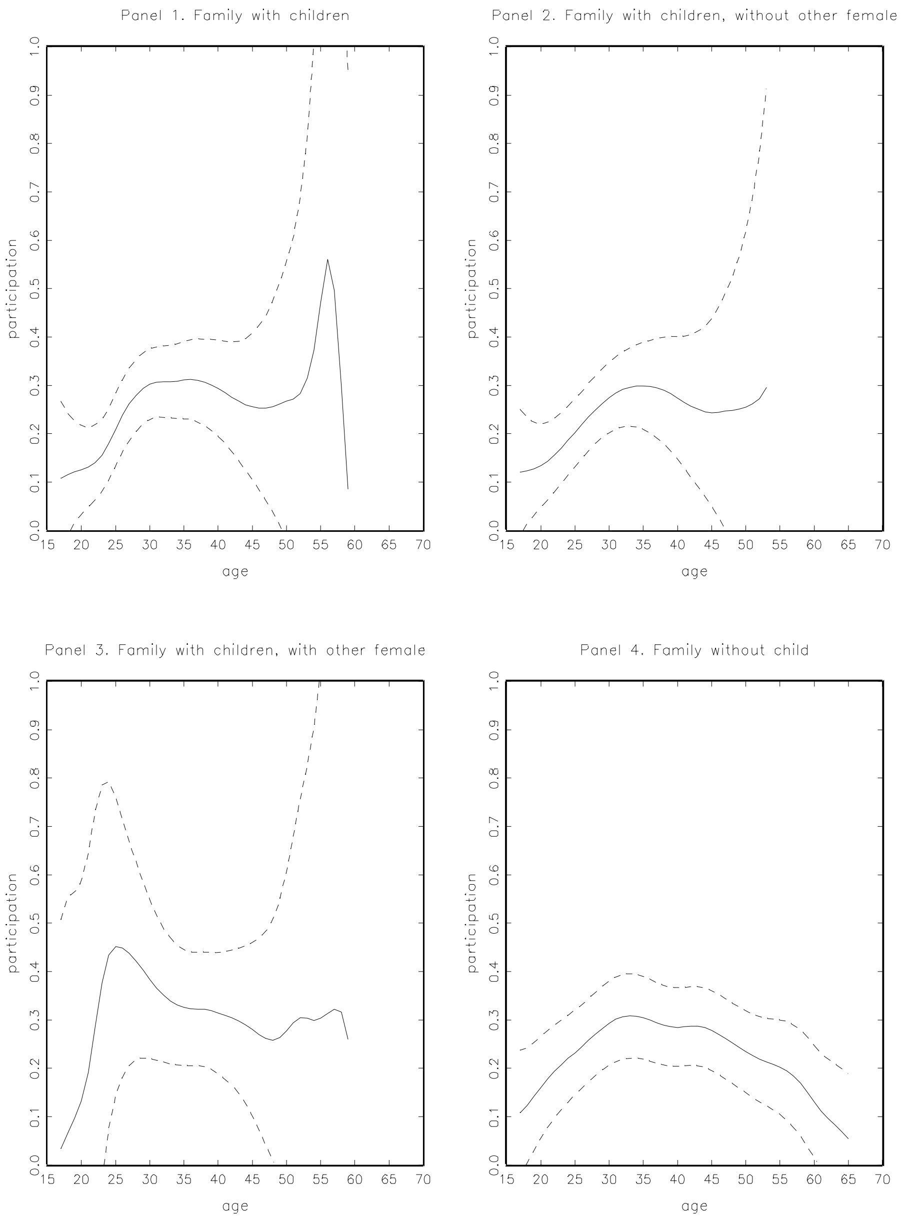

Figure 5. Simulated Labour Supply Curve: Family with one child, four adults, with or without other female.

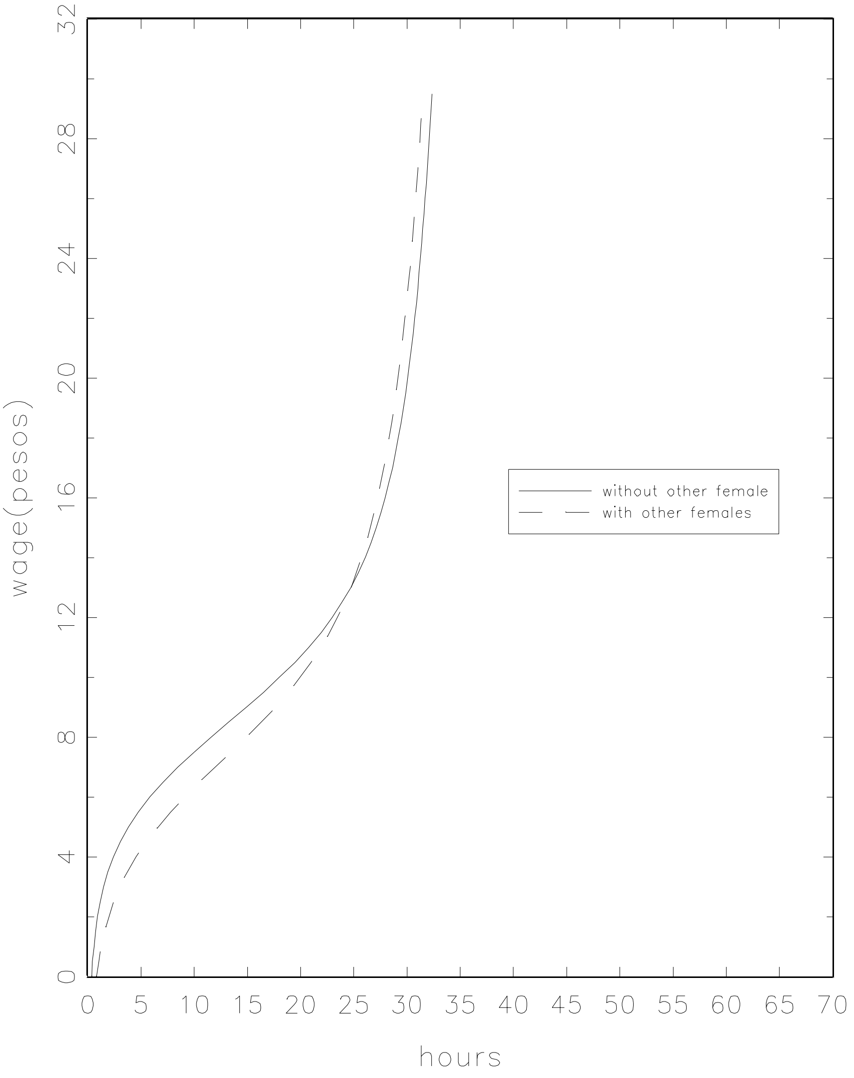

Figure 6. Simulated Labour Supply Curve: Family with zero or two children, four adults, and without other female.

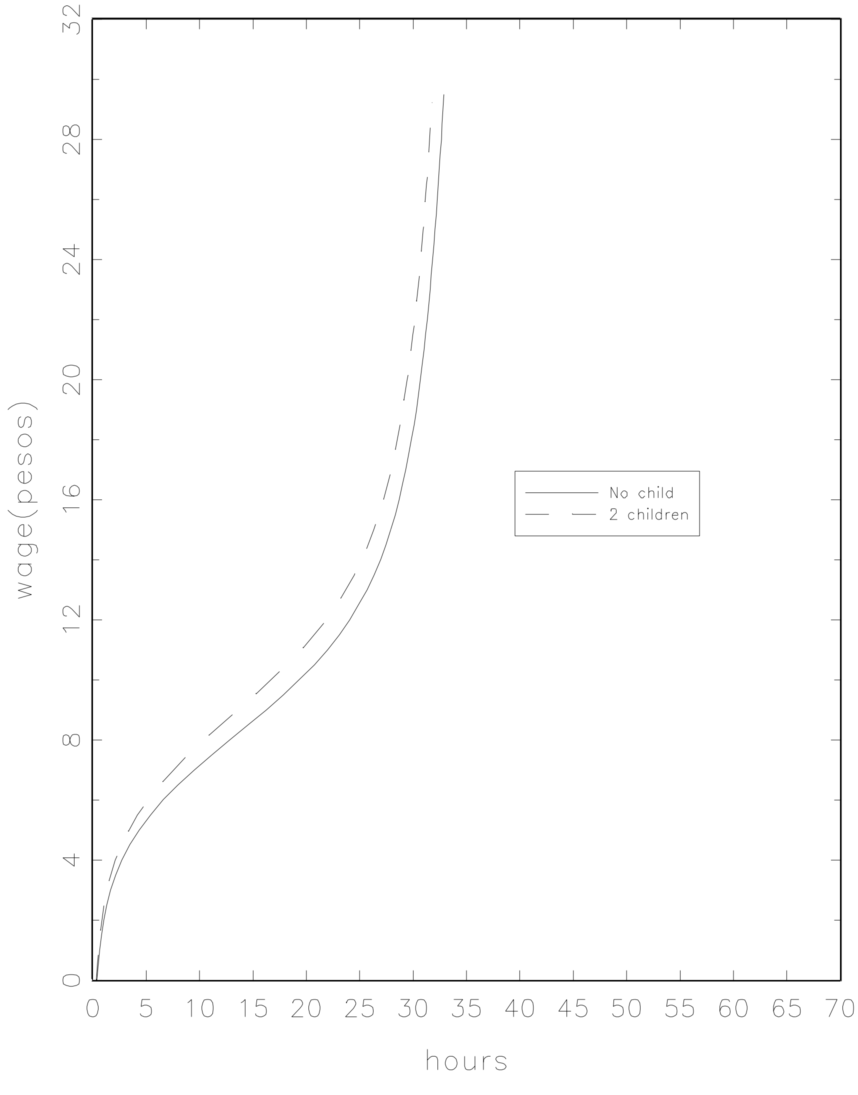

Figure 7. Simulated Labour Supply Curve: Family with one child, without other female, and with three or five adults.

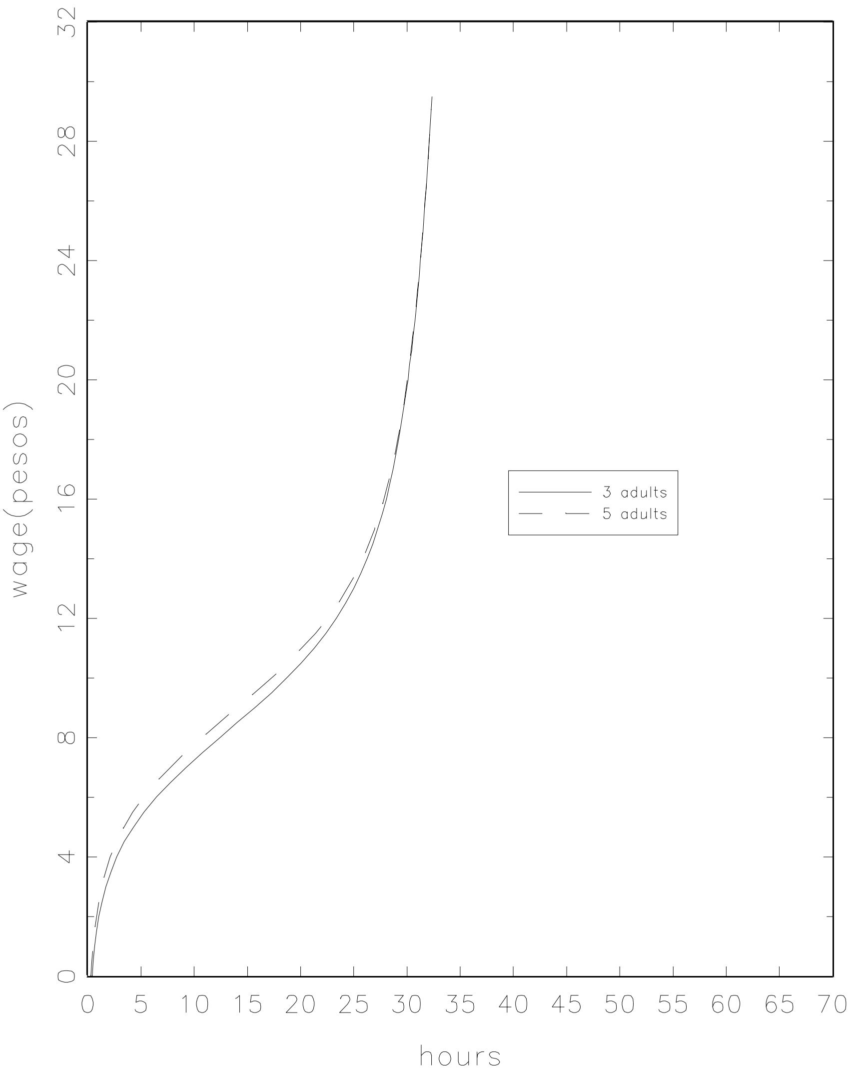

Figure 8. Indifference curves for several type of families

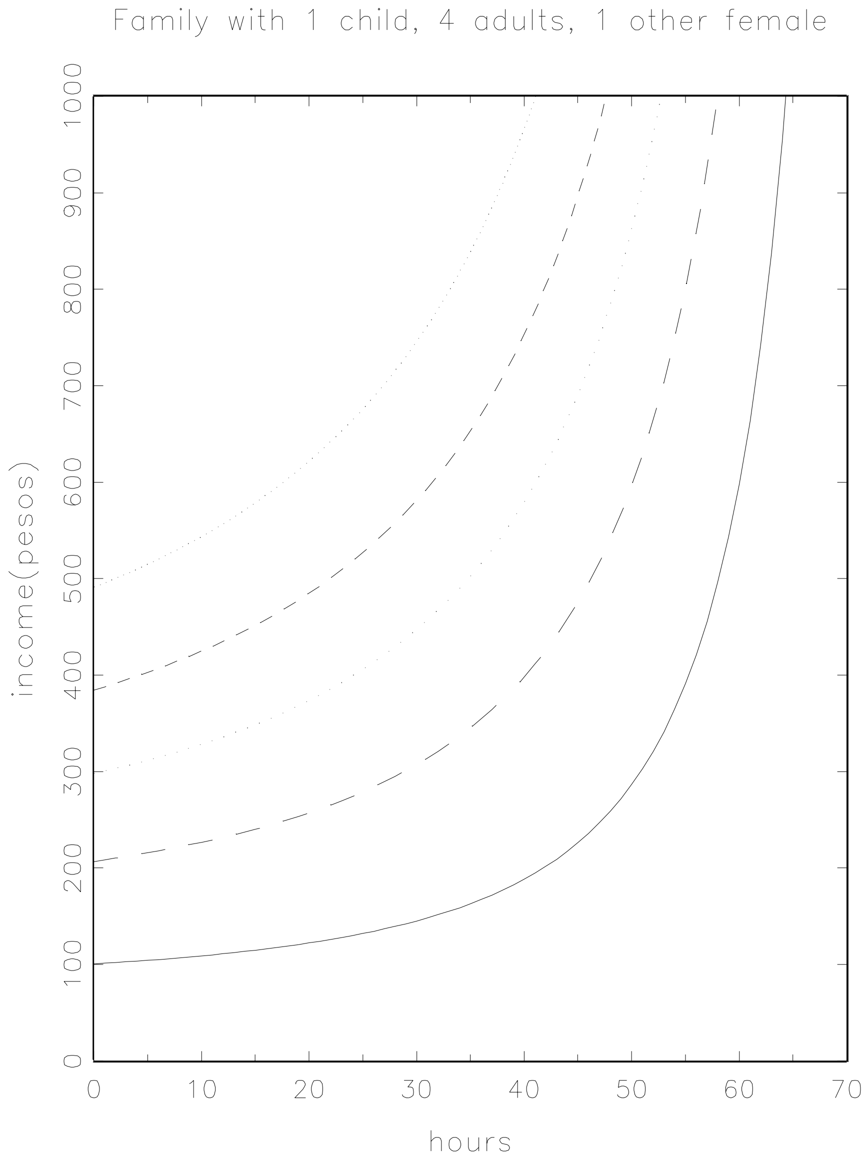

Figure A1. Relationship between after-tax earnings and before-tax earnings (Straight line is the fitted one.)