CrossMark

Labor supply heterogeneity and demand for child care of mothers with young children

Patricia Apps[1,2] · **Jan Kabátek[2,3,4]** · **Ray Rees[5]** ·
Arthur van Soest[2,4,6]

Abstract This paper presents a structural model of the labor supply and child care choices of partnered mothers with pre-school aged children. The father's time-use decisions are taken as given. The main goal is to analyze the sensitivity of maternal time use to the price of child care, taxes, benefits and child care subsidies. To account for non-convexities in the budget sets, we specify a discrete choice model. We estimate the model on data on couples with young children from the HILDA survey representative of the Australian population, which contains detailed information on time use and bought-in child care. Simulations based on the estimated parameters show that the time decisions of mothers with pre-school children are highly sensitive to changes in wages and the cost of child care. Our results also suggest that lowering effective tax rates faced by partnered mothers as second earners, by switching from family payments that are targeted on joint incomes to payments that are universal and funded by a more

✉ Jan Kabátek
j.kabatek@uvt.nl

[1] Sydney University Law School, Sydney, Australia

[2] IZA, Bonn, Germany

[3] Melbourne Institute of Applied Social and Economic Research, University of Melbourne, Melbourne, Australia

[4] Netspar, Tilburg, The Netherlands

[5] CES, University of Munich, Munich, Germany

[6] Department of Econometrics and OR, Tilburg University, Tilburg, The Netherlands

🕿 Springer

progressive individual-based income tax, would lead to a substantial increase in their labor force participation and hours of work.

Keywords Time use · Income tax · Child care subsidies

JEL classification J22 · J13 · H24

1 Introduction

One of the most striking, and still largely unexplained, facts about female labor supply in the developed countries is its heterogeneity across households, and indeed across countries. In many OECD countries, on average around one third of partnered women work full time in the labor force, one third do various amounts of part-time work, and one third work solely in household production. Very little of the aggregate heterogeneity across all households in any one country is explained by wage rate differences and by the number of children present in the household. Moreover, the correlation between female labor supply and fertility across these countries is strongly positive, even though historically, in any one country, there has been an inverse relationship between them.

Some insight is gained by organizing the data in terms of life cycle phases based on the number and age of children in the household. In the pre-children phase, there is very little difference between male and female labor supply distributions. This changes dramatically when children arrive, and this is when the heterogeneity in female labor supply essentially sets in. Though there is a trend of return to the labor force over subsequent phases of the life cycle as the children reach school age and beyond, the basic pattern of heterogeneity persists. Such findings suggest that for the theoretical and empirical analysis of female labor supply it is fruitful to focus on the life cycle phase in which households have young children.

The dramatic change in female labor supply with the birth of the first child reflects the additional work choice created by that event. At least one parent, typically the mother, can choose between working at home providing her own child care or working in the market and buying in care from formal care providers, such as kindergartens and child care centers, or by engaging other care givers, including relatives and friends. The importance of the availability and cost of child care for the labor supply of mothers with young children has been confirmed by theoretical (Apps and Rees 2009) and empirical studies, including Ribar (1995), Blau (2003), Connelly and Kimmel (2003), Doiron and Kalb (2005), Kalenkoski et al. (2005), Kornstad and Thoresen (2007), Baker et al. (2008), and Blundell and Shephard (2012).

The current paper presents a structural discrete choice model of the time allocation choices of partnered mothers with pre-school aged children. The main advantage of the discrete choice approach is that it can account for the non-convex nature of the household budget sets. Within this model, we analyze the decisions of mothers on hours of market work, time spent on child care and domestic work, and hours of formal child care. The main goal is to assess the sensitivity of choices at the intensive

and extensive margin of female labor supply and to capture underlying substitution patterns between the alternative uses of maternal time.

Similar models are employed by Doiron and Kalb (2005), Kornstad and Thoresen (2007), and Blundell and Shephard (2012). We allow for a more flexible household utility function than previous studies (following Van Soest 1995; Kabátek et al. 2014) and include both formal child care and maternal care in the utility function. Bought-in child care can be incorporated in two ways—either indirectly, subtracting child care costs from disposable household income (Doiron and Kalb 2005; Kornstad and Thoresen 2007), or directly, with the hours of child care taken as an additional argument of the utility function (Ribar 1995; Bernal 2008). We follow the direct approach, implying that formal child care and maternal care can be imperfect substitutes, with their own effects on household utility.

An important aspect of our empirical model is that we incorporate unobserved heterogeneity in the flexible form of latent classes, following Train (2008) and Pacifico (2012). We thus extend the treatment of unobserved heterogeneity beyond the traditional framework of random coefficient models,[1] avoiding restrictive assumptions on the distribution of the population parameters of the utility function, which we show has a pronounced effect on estimated labor supply elasticities.[2]

The model is estimated on data drawn from the Household, Income and Labour Dynamics in Australia (HILDA) survey which provides detailed information on time use and child care use and corresponding prices. Simulations based on the estimated parameters show that the time allocations of partnered mothers with pre-school children are highly sensitive to changes in wages and the cost of child care. A policy simulation also suggests that lowering effective tax rates faced by partnered mothers as second earners, by switching from joint to individual taxation, would lead to a substantial increase in their labor force participation and hours of work.

The paper is organized as follows. In the next section we set out the underlying theoretical model. In Sect. 3 we present the econometric specification that we take to the data. Section 4 discusses our dataset and Sect. 5 presents parameter estimates. Section 6 reports the results of policy simulations. Section 7 concludes.

2 Economic model

We construct a one period model of mother's time use and child care decisions during the preschool phase of the life cycle of a two-parent family. The time-use decisions of the father, taken to be the "primary" earner, are treated as exogenous.

By treating father's choices as exogenous, we depart from the assumption of full Pareto efficiency that underlies for example the collective model. There is, however, a growing literature that seeks to relax this assumption, for example that based on non-cooperative rather than cooperative household equilibria. The assumption of

---

[1] Applications using this approach include Ribar (1995), Doiron and Kalb (2005) and Kabátek et al. (2014).

[2] Several studies of female labor supply allow for more flexible treatment of the unobserved heterogeneity (Bernal 2008; Blau and Hagy 1998 and Tekin 2007), building on seminal works of Heckman and Singer (1984) and Mroz (1999). The latent-class approach can be considered a generalization of these models, allowing for additional flexibility.

exogenous male choices would seem to us to be an acceptable approximation in the light of the results of time-use studies showing relatively little variation in male time choices in the early child rearing years, with the vast majority working full time.

In a one period model, potentially important intertemporal effects, such as the anticipated loss of future human capital and employment possibilities from reducing current labor supply, cannot be incorporated explicitly. A mother may continue to work throughout the preschool phase despite facing a very low net wage or negative net earnings after child care costs, as an investment in her long-term career prospects. We can, however, partially capture these effects in a reduced form sense, through their impact on the marginal utility of market work *vis á vis* leisure or home child care and domestic work. We also take the number of children in the household as exogenous and therefore do not model fertility decisions.

Household $h = 1, 2, \ldots, H$, chooses:

- its consumption of a market good x_{ih}, with $i = 1, 2, \ldots, n$ denoting the individuals within the household;
- the mother's leisure consumption l_{2h};
- consumption of a composite household good, y_h, representing child care and domestic work;
- the mother's time input to the production of the household good, t_{2h}^y;
- purchases of the market child care good m_h^c.

Consumption is a composite market good with price 1, the mother's gross wage rate is w_{2h}, and the price of the market child care good is p_h^c, which varies across households.[3] The father's leisure and time allocation to household production are taken to be exogenous and therefore denoted by $\hat{l}_{1h}$ and $\hat{t}_{1h}^y$. Given the time endowment constraint, his market labor supply, $L_{1h} = \hat{L}_{1h}$, is also exogenous. The sum of the parents' gross incomes from market supply, $\sum_i w_{ih} L_{ih}$, is denoted by $I_h(w_{1h}, w_{2h})$. Their utility functions are $u_{ih}(x_{ih}, y_h, l_{ih}), i = 1, 2$. The remaining utilities $u_{ih}(x_{ih}, y_h), i = 3, \ldots, n$ correspond to children.

The household is assumed to maximize a unitary household welfare function,[4] concave in utilities,

$$W_h = \Psi_h(u_{1h}(.), \ldots, u_{nh}(.); \mathbf{e}_h) \quad h = 1, 2, \ldots, H \tag{1}$$

where $\mathbf{e}_h$ is a vector of exogenously given "environmental" or "distributional" factors which can be interpreted as determining the household's preferences over the utility

---

[3] Every variable or function with subscript h can vary across households. Each of these is therefore in principle a contributor to across-household heterogeneity in choices.

[4] The fact that we restrict ourselves to the unitary framework and do not develop a collective model is essentially due to data limitations that constrain what we can estimate. Furthermore, non-convexity of budget constraints is a central aspect of the motivation of the modeling in this paper, and this non-convexity poses a problem for the possibility of the existence of a decentralized sharing rule, which is a central aspect of the collective approach.

profiles of its members.[5] This function is based upon some household choice process which need not be further specified.[6]

The household's budget constraint can be written as

$$\sum_i x_{ih} + p_h^c m_h^c \leq I_h(w_{1h}, w_{2h}) - T\left(I_h(w_{1h}, w_{2h}), p_h^c m_h^c; n, \ldots\right) \quad h = 1, 2, \ldots, H$$

(2)

where $T(.)$ is a tax-benefit function which may contain as arguments demographic variables as well as gross incomes and expenditure on bought-in child care.[7]

The technology of the household production of y_h is expressed by the function

$$y_h = g_h\left(\hat{t}_{1h}^y, t_{2h}^y\right) \quad h = 1, 2, \ldots, H$$

(3)

and there is a time constraint

$$l_{2h} + t_{2h}^y + L_{2h} = T$$

(4)

where T is a given time endowment. Because we will be adopting a discrete optimization approach, directly comparing values of the household welfare function at all choice opportunities (see Van Soest 1995), we do not need to impose conditions of convexity or even differentiability on the function in (3). Thus the household can be thought of as choosing the variables l_{2h}, t_{2h}^y and m_h^c that determine consumptions, market labor supplies and income *via* the constraints (2)–(4) in such a way as to yield a global maximum of the function $\Psi_h(.)$. We can obtain a reduced form of this function by substituting from (2)–(4) into (1) to obtain a utility function that depends on these three choice variables as well as net household income Y. This then forms the basis for the empirical model specification.

3 Econometric specification

In order to specify a discrete choice model, we restrict the values of the three choice variables, the mother's labor supply, L_{2h}, her time allocated to household production, t_{2h}^y, and the hours of bought-in child care, m_h^c, to take one of five possible values which can be characterized as "low", "low-medium", "medium", "high-medium" and "high" according to their observed distributions.[8] The five values of each variable yields a grid of $5^3 = 125$ possible discrete choice points from which the household can choose its optimal allocation. The only restriction we impose on the household-specific choice set is that we exclude alternatives which would imply bought-in child

---

[5] In principle, the distributional factors could also include the wage rates, but this will not be allowed for in the empirical model.

[6] For a detailed exposition of the economics of this type of household model, see Apps and Rees (2009), Ch 3.

[7] For example there may be tax offsets for expenditure on market child care.

[8] For detailed discussion and applications of the discrete approach adopted here see, for example, Van Soest (1995), Van Soest et al. (2002) and Pacifico (2012).

care costs exceed family income. This restriction applies mainly to households with the lowest disposable incomes and long hours of formal care.[9]

Dropping the household subscript, we specify for the purposes of our model the vector $\mu = [l_2, t_2^y, m^c, Y]$. The leisure variable, l_2, is the residual of the daily time constraint in (4) with $T = 24$. The mother's household production time, t_2^y, is computed as the sum of hours allocated to child care and to other home production activities, many of which may simultaneously include child care.[10] Net household income, Y, is calculated as gross income net of taxes, family tax benefits and expenditure on child care. Gross income is the sum of each partner's earnings and the family's non-labor income. Since household income does not include the implicit value of household production it does not depend on t_2^y. There are therefore 25 possible values of net household income for each household, corresponding to combinations of the five choices of L_2 and the five choices of m^c.

The mother's gross earnings are calculated as the product of her gross wage and hours of market work. Unobserved wages are predicted by a Heckman selection model (Heckman 1979), with the exclusion restrictions being number of children in the household and the sum of husband's income and family non-wage income.[11] Expenditure on child care is calculated as the product of a household-specific child care price and the household's choice of formal child care hours. To account for families who do not use formal child care (and therefore do not report a corresponding price), we follow Connelly (1992) and use a predicted price derived from a Heckman selection model with the exclusion restrictions being number of adults in the household (excluding spouses) and distance from grandparents.[12] Sample selection criteria and regression results for both selection models are presented in "Appendix 1".

3.1 Baseline model without unobserved heterogeneity

We first present the model without unobserved heterogeneity. We take a reduced form of the household welfare function introduced in the previous section, specified as a flexible quadratic function

$$\Psi(\mu) = \mu'\mathbf{A}\mu + \mathbf{b}'\mu \tag{5}$$

---

[9] In other empirical studies, additional household-specific restrictions are often needed to account for infeasibility of certain choices. Kornstad and Thoresen (2007), for example, constrain choice sets of selected households to account for high degree of rationing in Norwegian day care centers.

[10] A limitation of the HILDA time-use data is that only one activity is reported for each episode. In contrast, the Australian Bureau of Statistics (ABS) Time Use Surveys report a second activity when relevant. Child care is almost always a second activity during housework and related activities.

[11] Similar income-based exclusion restrictions are used by, e.g., Blundell et al. (2007b) and Sorensen (1993).

[12] The imputation of child care prices should be approached with caution because the observed prices can reflect variation in quality of the service. The quality of child care can be endogenous to the regressors used in the Heckman selection model and hence can distort reliability of the imputed prices. To address these concerns, we estimate an alternative specification of the model which uses imputed prices of child care for all families in the sample. Relative differences in the predictions made by the original model and the alternative specification can be used to assess whether the endogeneity is likely to play a role here.

where $\mathbf{A}$ is a symmetric 4×4 coefficient matrix, and $\mathbf{b}$ is a 4-component vector. The first three components of $\mathbf{b}$, corresponding to the time-use variables l_2, t_2^y, m^c, are defined as

$$b_j = \sum_{k=1}^{K} \beta_{kj} X_k, \quad j = 1, \ldots, 3 \tag{6}$$

where the X_k denote, respectively, a constant term and variables representing observed household characteristics: wife's age; wife's age squared; number of pre-school age children; number of school-age children; and hours of informal child care provided by relatives, friends or the husband. These represent sources of observed heterogeneity. The elements of the matrix $\mathbf{A}$ as well as the component b_4 are assumed the same for all households.[13]

The household welfare function in reduced form does not explicitly separate the parameters of the household production function, the utility functions of the household members, or the household process which combines the utilities of the members. This should be kept in mind when interpreting the parameters. For example, the partial derivative of $\Psi(.)$ with respect to t_2^y is the marginal change in household welfare with the other components of (μ) (that is, (l_2, m^c) and Y) held constant, that is, when an hour of market work is replaced by an hour of work at home without changing income. This captures the (positive) effect of additional home production as well as the potential (positive or negative) effect of a higher or lower preference for home rather than market work, not accounting for the value of home production or the wage for market work. Differences in b_1 across households may therefore reflect either differences in productivity in household production or differences in preferences, or both. Conceptually, these are of course two quite distinct sources of heterogeneity, but they cannot be separately identified in the available data.

We introduce randomness in the value of the household welfare function at each possible choice point (l_2, t_2^y, m^c, Y) by specifying:

$$\Psi_r = \Psi(.) + \varepsilon_r, \quad r = 1, 2, \ldots, 125 \tag{7}$$

We can rationalize these errors as being errors of optimization or as being due to unobserved alternative specific characteristics that make each alternative more or less attractive than predicted by the systematic part. They can be due to factors that make a specific alternative more (less) attractive because of high (low) productivity or other, possibly preference-related, factors. The ε_r are assumed to be independent of each other and identically distributed and to follow the Type 1 Extreme Value Distribution. This implies that the conditional probability that point r^* is chosen as the optimal point is

$$P\left[\Psi_{r^*} > \Psi_r, \ \forall r \neq r^* \mid \mu, \mathbf{A}, \mathbf{b}\right] = \frac{\exp \Psi(\mu_{r^*}, \mathbf{A}, \mathbf{b})}{\sum_{r=1}^{125} \exp \Psi(\mu_r, \mathbf{A}, \mathbf{b})} \tag{8}$$

Finally, to guarantee that household welfare always increases with household income (an assumption which is needed for economic interpretation of the estimates), we

[13] This helps to reduce the computational complexity of the problem. Given that the utility function is identified up to a monotonic transformation only, it does not seem overly restrictive.

penalize the likelihood when necessary by adding points inside the budget frontier as additional choices that are never chosen by the household.[14]

3.2 Unobserved heterogeneity

It is likely that different households within the selected sample of families with young children have different unobserved attributes, for example in human and physical capital, which may impact on home productivity, measured, for example, by child outcomes. There may also be unobserved variation in the quality of market child care. Unobserved heterogeneity, whether in home productivity, in market child care or in preferences, is captured by the specification of error terms ε_r in the model as interdependent across alternatives. This contrasts with the basic model in which the errors are alternative-specific, which implies independence of irrelevant alternatives.

Several alternative approaches have been developed to allow for unobserved heterogeneity in the context of discrete choice labor supply models. The most prominent one is the parametric random coefficients model (see Van Soest 1995; or Keane and Moffitt 1998). This method has been criticized for the restrictive assumptions imposed on the distribution of stochastic terms (see Burda et al. 2008; Train 2008; Pacifico 2012). The distributions are predominantly assumed to be multivariate normal or lognormal, which implies that the corresponding density of parameter values is unimodal, that is, it has one peak characterizing the most frequent household welfare function. The restrictiveness of the unimodality assumption is well documented in Burda et al. (2008) who show that the standard random coefficients models perform poorly when the distribution of unobserved heterogeneity has multiple modes. This is not well captured by standard models, rendering the resulting preference ordering too uniform. This issue is of particular importance for our analysis, because previous theoretical work (Apps and Rees 2009) suggests that multimodal parameter distributions might well be present in the context of female labor supply.

A small body of literature on female labor supply allows for more flexible treatment of unobserved heterogeneity (Bernal 2008; Blau and Hagy 1998 and Tekin 2007). These studies draw on Heckman and Singer (1984) and Mroz (1999), using a step function to model the unknown distribution of the key random coefficient. This random coefficient therefore follows a discrete distribution and if enough mass points are allowed for, this distribution is very flexible and can approximate any underlying distribution.

The latent-class model can be seen as a tractable generalization of this approach, allowing for flexible discrete distributions of all parameters of the utility function. The underlying assumption is that the population consists of a number of different homogeneous populations (or classes) $K_c, c = 1, \ldots, C$, characterized by utility

---

[14] The penalized choices are identical to the standard choices in all respects other than household income, which is lower for penalized choices. Negative marginal utility would imply that penalized choices in the interior would be favored by the decision makers, but the observed choices are always on the budget frontier. Maximum Likelihood estimates maximize the probability of the observed outcomes and will therefore give parameters that make internal choices unlikely, that is, parameters that imply a positive marginal utility of income.

functions with parameters $\mathbf{A}_c$, $\mathbf{b}_c$ (see Train 2008). The parameterization is class-specific, implying that the probability mass is assigned to the whole set of parameters. This allows individual random coefficients to be correlated, although the correlation structure is not explicitly modeled.

Given the probability $P(h \in K_c)$ that a household $h = 1, \ldots, H$ is in the class K_c, $c = 1, \ldots, C$, and writing the probability that point r^* is chosen by this household as

$$P[\Psi_{r^*} > \Psi_r, \ \forall r \neq r^* \mid \mu, \mathbf{A}_c, \mathbf{b}_c, \mathbf{X}] = \frac{\exp \Psi(\mu_{r^*}, \mathbf{A}_c, \mathbf{b}_c, \mathbf{X})}{\sum_{r=1}^{125} \exp \Psi(\mu_r, \mathbf{A}_c, \mathbf{b}_c, \mathbf{X})} \qquad (9)$$

the unconditional probability that alternative r^* is chosen by household h is

$$\sum_{c=1}^{C} P(h \in K_c) \times P[\Psi_{r^*} > \Psi_r, \ \forall r \neq r^* \mid \mu, \mathbf{A}_c, \mathbf{b}_c, \mathbf{X}], \ c = 1, \ldots, C \qquad (10)$$

Allowing for multiple latent classes makes the model more difficult to estimate, with the traditional maximum likelihood optimization methods often failing to converge. Train (2008) and Pacifico (2012) show that in such cases we can take advantage of the well-known EM algorithm. This estimation procedure is considerably faster and more stable than the traditional methods, which makes it feasible to estimate flexible models even with a large number of latent classes.

4 Data

The HILDA survey provides data on a wide range of variables for a representative sample (17,000 respondents) of the Australian population interviewed annually since the year 2001. Particularly relevant to this study are the detailed data on time use and cost and utilization of formal and informal child care.

Mothers with pre-school aged children represent only a small fraction of each wave of the HILDA survey. To increase sample size, we construct a pooled cross-section using the four consecutive waves of HILDA from 2005 to 2008. From each wave we select partnered mothers with pre-school children. We exclude couples in which a partner is disabled, retired, or a full-time student, the husband is unemployed or the family lives in a multi-family household. We also exclude records with incomplete or implausible survey responses (usually on the relevant time-use variables).

The final sample contains 1465 records. Descriptive statistics for the dependent variables and the socioeconomic and demographic characteristics entering as independent variables in X are reported in Table 1. To enable comparisons by gender, the table also includes descriptive statistics for male wage rates and time use.

On average, parents of pre-school children are in their early thirties, with the father around two years older than the mother. Only 56 % of mothers in the sample are employed and, as we would expect, market hours distributions differ dramatically by gender, as shown in Fig. 1. The result is a gap of over 30 h per week between average female and male labor supplies. The vast majority of men work full-time (more than

Table 1 Summary statistics, sample of couples with preschool children

Variable	Mean	SD	Min.	Max.
Mother's age	32.9	5.6	16	48
Father's age	34.9	6.3	18	58
Marital status (dummy)	0.82	0.4	0	1
Mother's employment status (dummy)	0.55	0.5	0	1
Father's employment status (dummy)	1	0	1	1
Number of children aged 0–4	1.39	0.6	1	4
Number of children aged 5–9	0.47	0.7	0	4
Total number of children	2.01	0.96	1	6
Mother's market work, weekly hours	13.9	15.8	0	80
Father's market work, weekly hours	44.5	11.3	0	128
Mother's household production, weekly hours	71.6	31.2	0	166.8
Father's household production, weekly hours	31.2	17.5	0	120
Mother's leisure, weekly hours	82.3	27.6	0	160
Father's leisure, weekly hours	91.9	18.01	10	150.5
Formal child care, weekly hours	8.3	12.8	0	100
Formal child care price, in AUD, hourly	8.7	3.5	1.583	23.8
Informal child care (other relatives and friends), weekly hours	6.9	13.6	0	120
Mother's annual earnings in AUD	18,514	23,570	0	182,256
Father's annual earnings in AUD	63,373	35,736	5136	357,216
Annual non-labor family income in AUD	6617	31,778	0	683,974
Number of observations	1465			

35 h per week[15]), while women have a distribution of market hours that is relatively uniform apart from a large spike at zero hours. There are 83 mothers who report working more than 18 h a day for seven days a week.[16] In these cases we scale down the reported hours to satisfy a time constraint of 18 h of market work and housework per day while retaining the same relative time allocations as in the original data.[17]

Figure 2 compares hours spent on household production activities by gender. As noted above, household production is defined to include the allocation of time to activ-

---

[15] "Full-time" employment is defined by the ABS as 35 hours or more per week.

[16] The time-use data are collected by questionnaire and reported as weekly time uses. Unlike diary data, questionnaire data are typically subject to larger reporting errors, and as a result the sum of individual time allocations to the various activities sometimes fails to satisfy the time constraint.

[17] As a robustness check, we also estimated the structural model excluding these 85 women. This gave virtually identical estimates of elasticities and policy effects.
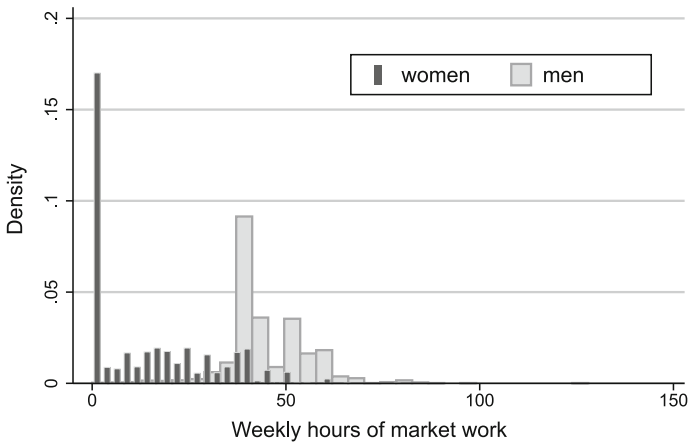
Fig. 1 Distribution of weekly hours of market work in families with preschool children

Fig. 2 Distribution of weekly hours of household production in families with preschool children

ities involving direct interaction with children, such as "playing with your children", and domestic work, much of which may also involve supervision of children aged 0–4. As we would expect, household production hours are higher for females than for males, as shown in Fig. 3, and their leisure hours[18] are more dispersed, with substantially higher frequencies at the lower levels of weekly leisure time. It is clear that for this group of households with young children, the total work burden is on average greater for mothers than for fathers.

---

[18] Leisure is computed as the remainder of the daily time endowment after subtracting market work and household production hours, which may be adjusted to satisfy the total time constraint. The 42-h per week threshold follows from assuming at least 6 h per day for sleep and personal care.
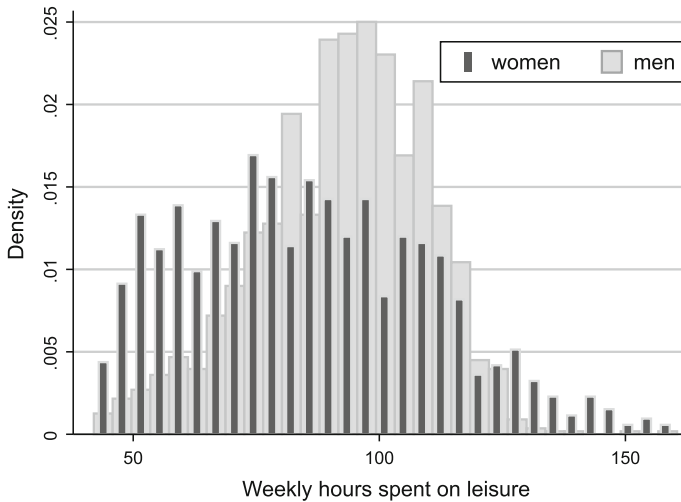
Fig. 3 Distribution of weekly hours spent on leisure in families with preschool children

We differentiate between formal child care provided by recognized institutions, such as kindergartens and care centers, and informal care provided by the father, grandparents or other relatives, and friends, for two reasons. First, formal child care differs from informal child care in that it is recognized as incurring costs by the Australian fiscal authorities, and the family is eligible for reimbursement of a considerable part of these costs. Second, the price data on informal care are rather unreliable. The price of formal child care is reported for all children in registered care. In contrast, informal child care is often provided with no charge, or at a price that implies an unobserved subsidy from the carer. The lack of more detailed information about the costs of informal child care makes any effort to impute corresponding prices infeasible. Therefore, we consider the choice of formal care only, treating informal care as exogenous.[19] Informal care enters the utility function as one of the interaction terms X in (6), measured in hours, without a specified price.

Formal care is used by 43 % of the families, while the use of informal child care is almost universal (only 9 families report that they used no form of informal child care). The distributions of the weekly hours of child care are presented in Fig. 4. The profiles for both types of care are relatively similar, although the formal care distribution does not go far above 60 h per week. This reflects the fact that formal care centers are closed on weekends.

Annual labor incomes are derived from reported weekly gross salaries from all jobs. The annual non-labor income of the couple is computed as the sum of each partner's business income, investment income, private domestic pensions and overseas pensions. Figure 5 presents distributions of male and female labor incomes and household non-labor income. According to these data, around 45 % of mothers have zero labor income,

---

[19] An economic rationale for this would be that informal child care is quantity-rationed and has a lower cost than formal child care, the price of which determines demand for child care at the margin.
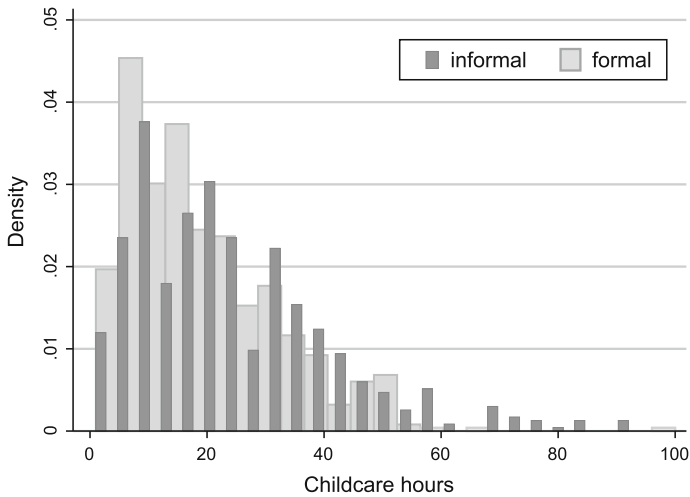
Fig. 4 Distribution of weekly hours of informal and formal child care, families with preschool children using child care

while 54 % of families in the sample have zero non-wage income. The distribution of non-labor income for the subsample of families with non-negative incomes is skewed toward zero. At the same time, several outliers report very large incomes from business and investments.

These income data are used to derive the set of 25 family incomes, net of the taxes and benefits and cost of child care, associated with the discrete time-use choices. All incomes are deflated to 2005, the selected base year, using the Australian consumer price index.

4.1 Income taxes and family benefits

The rate scale of the formal Australian income tax, the Personal Income Tax, is strictly progressive and applies to individual taxable income. However, strict progressivity is lost with the phasing out of an offset, the Low Income Tax Offset, also based on individual taxable income. While tax rates and the offset vary across the four waves of HILDA, the basic structure of the system is essentially the same in each year. For the purpose of illustration, Fig. 6 plots the profiles of marginal and average tax rates with respect to individual taxable income for the 2007–2008 financial year. Details of the rate scale and offset for that year are provided in "Appendix 2".

While the Australia income tax, the Personal Income Tax combined with the Low Income Tax Offset, is based on individual incomes, families are taxed effectively under a system of "quasi-joint" taxation. This is due to the withdrawal of child payments at various thresholds defined on the combined income of partners under a complex family payment system labeled "Family Tax Benefit Part A". The effective marginal tax rate, obtained by adding the withdrawal rate of payments to the marginal income tax rate, varies widely across the distribution of earnings and can be well above the top rate
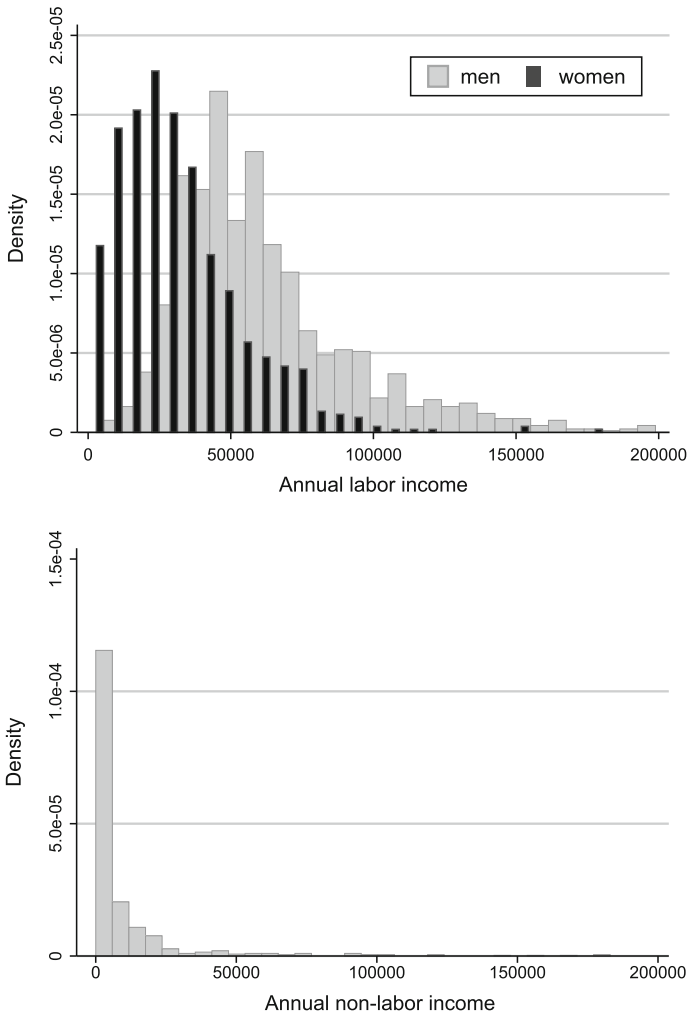
Fig. 5 Annual labor and non-labor gross incomes of families with preschool children, 2008, AUD
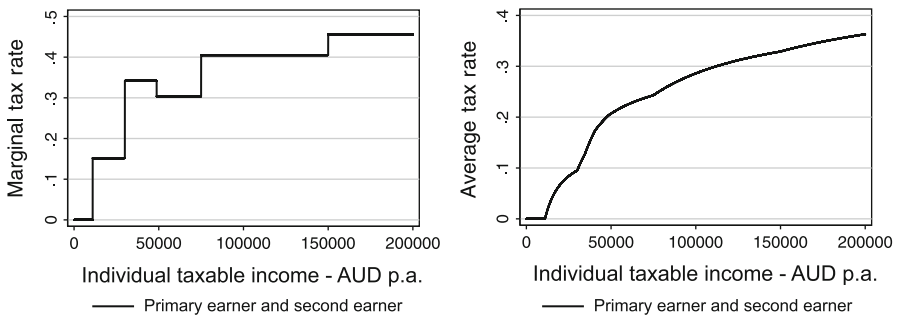

Fig. 6 Marginal and average tax rates of the 2007–2008 individual income tax
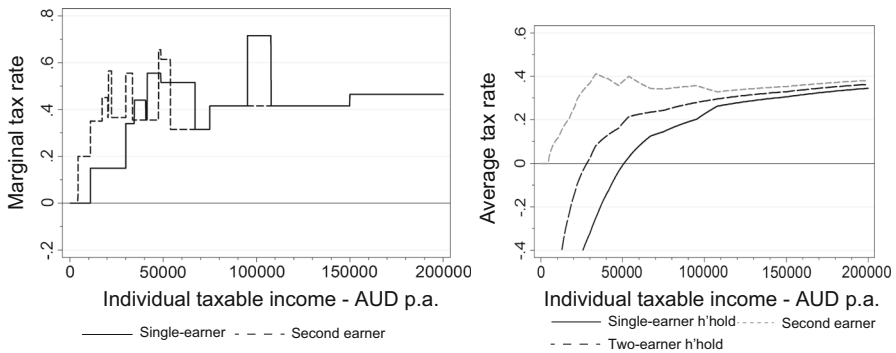
Fig. 7 Marginal and average tax rates of the 2007–2008 family tax system (including family payments)

of the Personal Income Tax scale at relatively low incomes levels. This is illustrated in Fig. 7 for the 2007–2008 financial year, for a family with two children aged under 13, with one under 5 years.[20] The figure plots the profiles of effective marginal and average tax rates with respect to the income of the primary earner for two limiting cases: a single income family and a two income family in which both partners earn the same income. For the latter case, the figure plots the effective marginal and average tax rates applying to the income of the second earner.[21] The higher rates indicate the tax penalty[22] married mothers as second earners can face on entering the workforce under a system of joint or "quasi-joint" taxation.

The pattern of marginal tax rates in Fig. 7 implies a budget set with many non-convexities and would make the traditional approach of finding the optimum in the complete budget set infeasible. This makes the discrete approach, approximating the complicated budget frontier with a small finite set of points, particularly useful.

5 Results

We first report the results for the baseline homogeneous specification presented in Sect. 3.1 and then discuss those for the model with unobserved heterogeneity introduced in Sect. 3.2.

5.1 Baseline model without unobserved heterogeneity

The estimated parameters of the baseline model are reported in Table 2. If the homogeneity assumption were found to be valid, the results would be consistent and more

---

[20] The details of the tax rates, family payments and income thresholds on which the figure is based are set out in "Appendix 2".

[21] For a graphical analysis of cases in which the second earner has lower income than the primary earner, see Apps and Rees (2009), Ch 6.

[22] This "tax penalty" on the second earner's income under joint taxation explains why switching from a joint to a purely individual-based system stimulates the participation of the second earner (see, for example, Steiner and Wrohlich 2004).

Table 2 Regression results for the baseline homogeneous model

Matrix A		Vector b	
Income	.199	Formal care*log(age)	−.613
	(.092)**		(3.63)
Formal care	2.23	Formal care*log(age)^2	.130
	(6.20)		(.529)
Household production	4.71	Formal care*married	−.031
	(3.42)		(.057)
Leisure	7.59	Formal care*no. dependent children	−.127
	(3.52)**		(.037)***
Formal care^2	.128	Formal care*children aged 0–4	.329
	(.014)***		(.056)***
Household production^2	.066	Formal care*children aged 5–9	.093
	(.005)***		(.049)*
Leisure^2	.022	Formal care*informal care	−.009
	(.006)***		(.001)***
Income*formal care	.011	H'hold prod.*log(age)	−4.07
	(.004)**		(2.01)**
Income*h'hold prod.	.019	H'hold prod.*log(age)^2	.607
	(.002)***		(.294)**
Income*leisure	.020	H'hold prod.*married	.026
	(.002)***		(.032)
Formal care*h'hold prod.	−.066	H'hold prod.*no. dependent children	−.112
	(.004)***		(.020)***
Formal care*leisure	−.064	H'hold prod.*children aged 0–4	.336
	(.005)***		(.032)***
H'hold prod.*leisure	−.050	H'hold prod.*children aged 5–9	.148
	(.005)***		(.027)***
		H'hold prod.*informal care	−.008
			(.001)***
		Leisure*log(age)	−4.87
			(2.07)**
		Leisure*log(age)^2	.685
			(.305)**
		Leisure*married	.012
			(.034)
		Leisure*no. dependent children	−.053
			(.023)**
		Leisure*children aged 0–4	.231
			(.034)***
		Leisure*children aged 5–9	.114
			(.031)***

Table 2 continued

Matrix A	Vector b	
	Leisure*informal care	−.009
		(.001)***
n		1465
Log-likelihood		−6076.44

Standard errors in the parentheses, significance levels: * 90, ** 95, *** 99

Table 3 Average marginal utilities of the main regressors and fraction of the population sample with negative marginal utilities, homogeneous model

	Average marginal utility		Negative fraction	
	Full sample	Child care users	Full sample	Child care users
Income	1.06	1.04	0	0
Formal care	−0.69	−0.05	0.83	0.58
H'hold prod.	0.63	0.26	0.17	0.37
Leisure	0.64	0.33	0.24	0.28

efficient than the latent-class model. The coefficients indicate that several of the interaction terms yield intuitively plausible results. An increase in the number of pre-school aged children in the household raises the marginal utility of formal child care and therefore strengthens the demand for it. On the other hand, an increase in the (assumed exogenous) availability of informal child care weakens it. The same is true for the allocation of time to household production.

The estimated marginal utilities of the choice variables, the components of the vector μ, are central to our analysis, but their evaluation is more complex than consideration of the simple regression coefficients in isolation, since the marginal utilities depend upon the entire matrix $\mathbf{A}$ and the vector $\mathbf{b}$. They also vary with the household-specific socio-demographic characteristics, $\mathbf{X}$, and with the values of the choice variables μ themselves. In Table 3 we summarize the distribution of the estimated marginal utilities at the observed choices, presenting first their sample averages and second, the proportion of households that have negative marginal utilities. We do this for the full sample as well as for the subsample of households that actually buy formal child care.

As expected, marginal utilities of income, household production and leisure are on average positive, with only a very small fraction of households having a negative value in each case. On the other hand, around 90 % of households have a negative marginal utility of formal child care. This is of course not a problem for those households that do not use formal child care, but the last column of the table shows that households that do buy formal care have, on average, negative marginal utilities. This implies that this model is not successful in explaining the use of formal child care from economic arguments. For most households, the use of formal child care can only be predicted with the inclusion of error terms ϵ_r, reflecting optimization errors or unobserved factors that make specific choices more or less attractive.

Table 4 Bayesian information criteria for multi-class models

No. of classes	Log-likelihood	BIC
1	−6076.44	12,415.31
2	−4921.44	10,367.73
3	−4676.13	10,139.53
4	−4226.96	9503.61
5	−3963.88	9239.88
6	−3748.64	9071.83
7	−3584.41	9005.79
8	**−3319.71**	**8738.83**
9	−3236.49	8823.87

The bold emphasized that the 8-class model returned lowest BIC (3rd column)

This counter-intuitive result could be due to unobserved heterogeneity. Our sample contains a large proportion (57 %) of households that do not use formal child care. This can be problematic for the homogeneous model if the decision to use formal child is influenced by, for example, unobserved differences in home productivity. The model tries to explain this relation in terms of the variables included in the utility function, assigning strong disutility to formal child care. Since the majority of families do not use formal care, the failure to take account of unobserved heterogeneity forces the common coefficient to be negative. Introducing unobserved heterogeneity may help to solve this problem.

5.2 Latent-class models

A key step in the EM estimation procedure is the initial selection of the number of latent classes. This decision involves a trade-off. On the one hand, the higher the number of heterogeneous groups, the better is the fit of the model because we account for unobserved heterogeneity in a more flexible form. On the other hand, more stratified models are bound to be estimated less precisely because the number of unknown parameters rises proportionally to the number of allowed latent classes. The determination of the optimal number of classes is therefore crucial.

Following Train (2008), we compare the models with varying classification choices on the basis of their Schwarz–Bayesian information criteria (BIC)

$$BIC = -2\log(L) + k\log(n) \tag{11}$$

where L is the likelihood, k is the number of free parameters in the model and n is the number of observations in our sample. The multiple-class models yield the statistics in Table 4. The table shows that the 8-class model attains the lowest BIC, and should therefore be considered as the most reliable specification for further analysis.

In order to examine whether our models actually fit the data, we simulate individual time-use allocations using the estimated models and compare the simulated aggregated distributions to their observed counterpart. Figure 8 presents this comparison both for the baseline model and the model with eight latent classes.
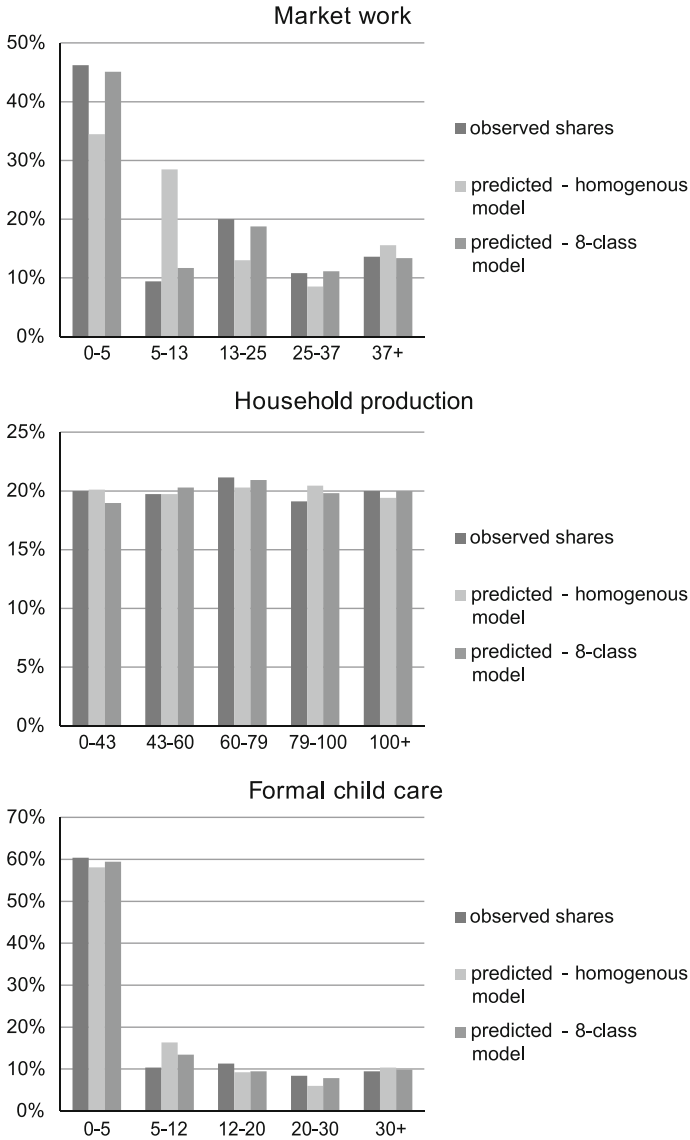
Fig. 8 Distribution of time-use variables into intensity levels, observed and predicted shares

As expected, the 8-class model replicates the empirical distributions very well, attaining almost identical shares of intensity levels among all three time-use choices. The homogeneous model performs much worse and essentially fails to capture the distribution of market work hours. In particular, the model underestimates the proportion of mothers with zero market hours and overestimates the proportion with low hours in part-time work. The distributions of the other two choice variables are replicated well even by the homogeneous model, though the latent-class model still provides more precise approximations.

Table 5 Fraction of the population sample with negative marginal utilities of the main regressors, model with eight latent classes

	Whole sample	Child care users
Income	0	0
Formal care	0.53	0.31
Household production	0.24	0.41
Leisure	0.23	0.38

We do not present the regression coefficients for the 8-class model because the class-level stratification makes their interpretation practically impossible. However, one statistic which can be readily interpreted is the fraction of the sample with negative marginal utilities (see Table 5).

The only result which exhibits a substantial change compared to the baseline specification (see Table 3) is that for formal child care. The proportion of mothers with disutility from additional formal child care drops by 30 % points, to 53 % in total and to 31 % when we restrict the sample to mothers who are using formal child care.[23] This is a considerable improvement over the homogeneous specification, though the 31 % is still substantial. The reason for this is that the majority of all mothers do not use any formal child care. As a consequence, the estimates imply that most mothers do not attach positive utility to formal child care. Due to the important role played by unobservables (error terms and unobserved heterogeneity), the model is not able to perfectly predict who will and who will not attach utility to formal child care.

The relative performance of the models with varying numbers of latent classes is further tested through a series of simulations in the next section. The aim of these simulations is to predict how people respond to selected changes within their economic environment. By predicting (and comparing) the behavioral responses for different model specifications, we can analyze the importance of unobserved heterogeneity and assess the limitations of the homogeneity assumption.

6 Microsimulations

First, to analyze the sensitivity of choices to wages and prices, we simulate a 10 % increase in the wages of all mothers, and a 10 % increase in the prices of formal child care. Second, we carry out a policy simulation in the spirit of Apps and Rees (2009), building on their critique of joint taxation (as discussed in the previous section). We propose an alternative system of taxes and benefits designed to have a less distortionary effect on female labor supply than the actual system and we estimate its impact on the choices of the type of households we consider.

---

[23] The latter is obtained by taking weighted means over all classes, where the weights are the class probabilities given the observed choice.

6.1 Changing wages and child care prices

The impact of wage and price changes is measured in terms of aggregate elasticities. We compute the percentage changes in total hours of market work, total hours of household production, and total hours of formal child care with respect to changes in the wage rates of all mothers, or all child care prices, holding all other variables constant. We present both gross and net elasticities, which correspond to changes being applied before and after subjecting families to all taxes and benefits. Accordingly, changing gross wages will have an impact on applicable tax rates and benefit eligibility criteria. Due to the progressive features of the income tax system, the 10 % increase in gross wages generally leads to lower increase in net wages. Changing net wages has the advantage of circumventing secondary effects caused by changes in the effective tax rates: increasing net wages by 10 % results in 10 % higher disposable incomes from the mother's market work across all households.

The resulting income changes are proportional to the earnings of mothers so that those working earn more while non-participants retain their original disposable incomes. Since the 10 % increase in the wage makes participation more attractive, we can expect an increase both in market hours of employed mothers and in the labor market participation rate. Similarly, an increase in child care prices results in an income reduction that is proportional to the cost of bought-in child care, and we can expect that it leads to a reduction in bought-in hours and in the fraction of the sample using formal child care.

We compute aggregate elasticities as the ratios of percentage changes in the relevant time or care use to the percentage changes in the wages and prices (where the latter are 10 %, by construction). This is done as follows. We first derive the benchmark time-use allocations (using the same wages and prices that are used for estimation) by averaging individual choice probabilities predicted by the model and using these to compute the average hours of each activity.

A similar procedure is applied to calculate the average hours of activities after the wage or price increase. The only difference is that the choice probabilities are derived using adjusted disposable income for each alternative in the choice set. This changes the utility values for some of the choice alternatives but not for others and, as a consequence, changes the probabilities of all choices. Using the new probabilities we recompute average hours. Finally, we compute the percentage deviations in the new averages compared to the benchmark. The elasticities for the homogenous model and our preferred specification are provided in Table 6.[24]

The first panel gives the responses to the increase in all mothers' net wage rates. The first thing to note is the large difference between the homogeneous (one class) model and the model with unobserved heterogeneity, demonstrating the importance of controlling for unobserved heterogeneity. When we allow for unobserved heterogeneity, the predicted responses fall substantially. Their sizes remain relatively stable among models with different numbers of classes.

---

[24] Standard errors on the elasticities were computed through 199 Monte Carlo simulations, recomputing the percentage changes with simulated sets of parameters determining **A** and **b**. These parameters were drawn from the estimated (multivariate normal) distribution of the ML estimates. See Ruud (1991).

No. of classes	One class	Eight classes
Table 6 Elasticities of time-use allocations with respect to changes in wages and child care prices, partnered mothers with pre-school children		
Mothers' net wage increased by 10%		
Formal care hours	1.01	0.42
	(0.051)***	(0.065)***
Market work hours	1.35	0.43
	(0.043)***	(0.086)***
Household production hours	−0.23	−0.08
	(0.012)***	(0.015)***
Mothers' gross wage increased by 10%		
Formal care hours	0.70	0.25
	(0.031)***	(0.069)***
Market work hours	1.02	0.31
	(0.031)***	(0.099)***
Household production hours	−0.17	−0.05
	(0.008)***	(0.012)***
Net child care price increased by 10%		
Formal care hours	−0.51	−0.42
	(0.017)***	(0.081)***
Market work hours	−0.17	−0.08
	(0.01)***	(0.021)***
Household production hours	0.03	0.02
	(0.003)***	(0.008)***
Gross child care price increased by 10%		
Formal care hours	−0.78	−0.71
	(0.027)***	(0.159)***
Market work hours	−0.25	−0.16
	(0.014)***	(0.060)***
Household production hours	0.04	0.03
	(0.005)***	(0.013)***

Standard errors in the parentheses, significance levels: *90, **95, ***99

Standard errors tend to increase as we allow for latent classes. This reflects the fact that this model is more flexible and therefore requires more data for accurate estimation.

Focusing on the latent-class model, the net wage increase leads to time-use shifts that correspond to intuition. A 10 % increase in all net wage rates results in a (significant) 4.3 % rise in average working hours, implying a positive uncompensated own labor supply elasticity of 0.43 for this group of mothers with young children. This is well in line with the large literature on female labor supply. The positive substitution effect (the price of leisure increases) dominates the negative income effect. Moreover, the 10 % wage increase leads to a (significant) 4.2 % increase in hours of formal child care (a "cross" elasticity of 0.42). First, the higher demand on time due to increasing hours of market work leads to substitution of own child care for bought-in child care. Second, higher earnings lead to higher family income, increasing the demand for formal child care if this is a normal good.

The elasticity of time allocated to household production is significantly negative, at -0.08. The negative sign implies that higher wages lead mothers to work less in the household. However, the actual change in home production hours is not large enough to compensate for the increase in market hours, implying that mothers also reduce their leisure in order to do more market work.[25] The second panel reports mothers' gross wage elasticities. The responses induced by 10 % gross wage change are weaker than those reported for the net wage change. This follows from the fact that gross stimuli are still subject to tax, so that the resulting net wage changes will be smaller than 10 %.

Turning to the impact of the rise in child care prices in the third panel, it is not surprising that the highest elasticity is that of formal child care itself. With a 10 % rise in child care prices, the demand for formal child care falls significantly, by 4.2 %. This in turn causes mothers to work less in the market–market hours drop significantly, by 0.8 %, as they have to substitute their own time for bought-in services.[26] Accordingly, the hours of household production increase by 0.2 %, replacing almost all of the forgone time formerly spent on market work.[27] The gross child care price elasticities presented in the fourth panel are higher compared to their net counterparts. This is a direct consequence of the Child Care Rebate (CCR). CCR imposes an upper cap on rebatable child care expenditures (see "Appendix 2"), and this cap is binding for many families in the sample. By increasing the gross child care prices, these families will be facing additional child care expenditures which cannot be rebated. As a result, they will be subject to net child care costs higher than the ones induced by the net child care price increase. For families with expenditures below the upper cap, the effects of gross and net price changes will be equivalent.

6.2 Simulation of a tax and benefit reform

As discussed in Sect. 4, the phasing out of family benefits on household income creates high effective marginal tax rates for many mothers as secondary earners. To investigate the impact of these high rates on their labor supply and participation, and also on the demand for formal child care, we simulate the effects of switching to an individual-based income tax with universal payments. The reform replaces the marginal rate scale depicted in Fig. 7 with one that applies to individual taxable incomes, as illustrated in Fig. 6. To fund the increase in benefit payments, we increase proportionally (by 26.76 %) all marginal tax rates of the system in Fig. 6 to achieve a reform which is ex ante (that is, before behavioral responses) revenue neutral.[28]

---

[25] In absolute terms, the wage increase induces the average mother to spend about 0.55 h per week more on market work, 0.40 h less on household production, and 0.15 h less on leisure.

[26] This elasticity is well in line with that of Gong and Breunig (2011) but smaller than that in Breunig et al. (2012).

[27] On average mothers spend about 0.10 h per week longer on household production and reduce their market work by 0.10 h and bought-in child care by 0.05 h.

[28] It should be noted that these results are specific to the population of partnered mothers with preschool children. The revenue neutrality is also achieved by raising income taxes only for this group.
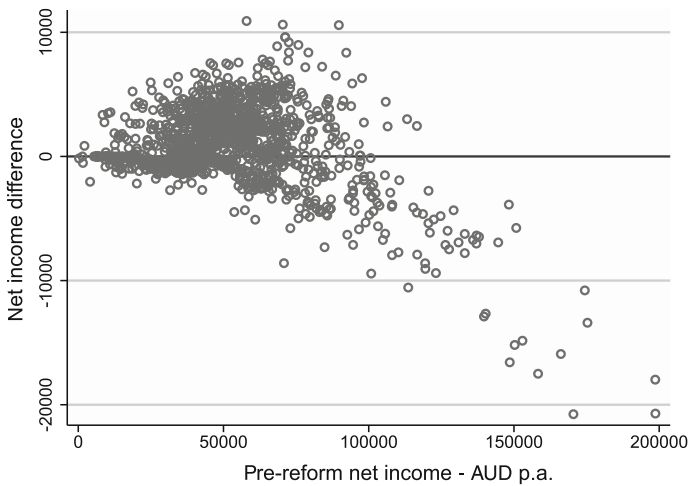
Fig. 9 Post-reform differences in the net tax positions of families, ordered by pre-reform net household incomes

Figure 9 shows graphically the differences in the net tax positions of households resulting from the reform (assuming no behavioral responses). The differentials are ordered by the corresponding pre-reform net household incomes, so that we can see how the shift in the tax burden varies with household income. Since the reform is ex ante revenue neutral, the changes for all families in the sample sum to zero. The figure shows that when benefits are universal families with average joint incomes gain. It is important to keep in mind that this is due to gains for relatively low to average two earner families who previously lost the joint-income-tested benefits. The proportionally higher marginal tax rates shift the tax burden toward the higher income groups, in effect shifting the burden from lower wage two earner families to those with higher wage rates.

Table 7 summarizes the simulated changes in time allocations and hours of formal child care in response to the reform. As in the previous simulations, we observe a large discrepancy between the changes predicted by the homogeneous model and those predicted by the models with more than one latent class, with the results of the latter proving relatively stable across different specifications. Again this implies that accounting for unobserved heterogeneity is important not only to improve the fit of the model but also from a substantive point of view.

We again focus on the outcomes for the 8-class specification. We observe that the reform would lead to a 3.11 % increase in average hours of work (about 0.43 h per week, using the average hours in Table 1), a 1.75 % increase in average hours of formal child care (0.15 h per week), and a 0.63 % decrease in the average hours of home production (about 0.45 h per week). All these effects are statistically significant. On average, the positive effect on market work of not phasing out family benefits is more important than the negative effect due to the increase in the marginal tax rates. Market hours of work therefore increase and, as a consequence, hours of home production fall and demand for formal child care increases.

Table 7 Percentage changes in time allocations after ftb reform, partnered mothers with pre-school children

No. of classes	One class	Eight classes
Formal care hours	3.90 %	1.75 %
	(0.243)***	(0.38)***
Market work hours	6.49 %	3.11 %
	(0.337)***	(0.684)***
Household production hours	−1.11 %	−0.63 %
	(0.091)***	(0.133)***

Standard errors in the parentheses, significance levels: *90, **95, ***99

Table 8 Percentage changes in time allocations after FTB reform by household income quartiles, model with eight classes, partnered mothers with pre-school children

Household income quartile	1st (%)	2nd (%)	3rd (%)	4th (%)
Formal care hours	4.28	4.22	2.71	−1.65
Market work hours	12.43	5.33	2.79	−0.92
Household production hours	−0.96	−0.80	−0.83	0.09

6.2.1 Heterogeneity of the behavioral responses

The behavioral effects induced by the reform appear to be highly heterogeneous across population groups and latent classes. Closer analysis of our results reveals positive effects at the extensive margin of female market labor supply, with the predicted labor market participation rate rising by 4.4 % (to about 58 %). On the other hand, these effects are mitigated by responses at the intensive margin, with some employed mothers choosing to work fewer hours under the reform. Average hours of market work (conditional on being employed) fall by 1.3 %, with individual responses showing considerable variation. In fact, expected hours of market work increase for 69 % of all women in the sample. The 1.3 % decline in the aggregate work at the intensive margin is driven by the response of mothers on higher wages and in full-time employment. For these women, the generic increase in the marginal tax rate needed to finance the reform leads to a negative substitution effect that reduces their market work and increases non-market time uses. This effect dominates the positive income effect stemming from the loss of disposable income. This intuition is supported by the results presented in Table 8 where we divide households into four groups corresponding to quartiles of the observed household income distribution.

Households in the first three income quartiles exhibit responses with signs in line with the aggregate effect, whereas the high-income households in the fourth quartile are predicted to adjust their time-use allocations in the opposite direction. For the latter group, the female labor force participation drops by 0.92 %. This decline is smaller than the one derived for the aggregate intensive margin of female labor supply,[29] since

---

[29] The percentage changes presented in Table 8 use quartile-specific sample means of time-use allocations as a point of reference. If we would use unconditional means, the percentage change for female labor supply in the fourth quartile would be 1.26 %.

high household income does not necessarily imply that a woman in such household would be a high earner.

The behavioral heterogeneity is crucial for successful targeting of policy reforms, as it helps to identify the potential impact on different subsamples of the population. It is also interesting from the perspective of economic modeling, as we can compare the relative performance of homogeneous and latent-class models. In order to do so, we split the sample into two groups according to actual employment status and compute the elasticities separately for the two groups, using both the homogeneous model and the latent-class model with eight classes. Using the homogeneous specification, the effects prove to be almost identical for both groups, as this model captures only a small part of the differences in productivity and preferences between the groups (the "observed heterogeneity" part captured by the covariates in the model). According to the eight class model results, the simulated increase in aggregate working hours is much stronger for non-employed mothers, with the absolute increase in market work hours being 28 % larger. As for the change of formal child care hours, the non-employed mothers exhibit a rather modest increase in absolute terms (70 % lower than employed mothers), but in relative terms their bought-in child care rises more than for employed women (the initial level of formal child care utilization is substantially lower for non-employed mothers).

The failure to capture heterogeneity in responses of the homogeneous model is further illustrated by the fact that this model cannot replicate observed differences in reported time-use allocations between the two groups, overestimating work and formal child care allocations of non-employed mothers and underestimating them for employed mothers. On the other hand, the 8-class model produces almost identical time-use patterns as observed in the data. For these reasons, it is hard to maintain that the homogeneous model would be able to provide reliable predictions of the responses to proposed policy changes.

6.2.2 Net fiscal effect of the reform

We also analyze the net revenue effect of the reform taking account for behavioral changes predicted by our 8-class model. Changes in time allocations can affect government revenue through two distinct channels: by increasing (reducing) their hours of work mothers are also increasing (reducing) income tax revenues, and by buying in longer (shorter) hours of formal child care, child care benefits rise (decline).

The key result in this context is that the government marginally improves its net fiscal position. Income tax revenue from mothers rises by only 0.5 %, which seems low compared to the 3.1 % increase in aggregate working hours. The reason is the heterogeneity in responses discussed above: mothers with higher wages tend to reduce their hours of market work, and the progressive nature of the income tax system makes the fall in tax revenues from this group relatively large, substantially offsetting the additional revenue from low and middle income households. More specifically, mothers who increase their market hours (69 % of the sample) are predicted to pay an average of \$151 more in annual income taxes (a 5.6 % increase), whereas those who reduce their hours reduce their income tax liabilities by a predicted average of \$261 per annum (a 2.3 % reduction). The net result is an aggregate increase in the income

taxes by $25 per household (which translates into the aforementioned 0.5 % revenue gain).

The situation is very similar for the child care benefits, which increase only slightly in aggregate: on average, a household gets an additional $2 (0.1 % of the initial payments), which is small compared with the 1.75 % change of formal child care hours. Analogously to the income tax effects, this outcome reflects the heterogeneity in behavioral responses.[30] Combining the two effects, we estimate that on average households will contribute an additional $23 to government tax revenue, an increase that represents 0.2 % of their original contribution.

6.3 Robustness checks

In order to assess the stability of our results, we run a series of sensitivity checks, altering the econometric specification of our model in the following ways. First, to achieve a more flexible specification, we divide the time-use variables into a finer grid (6^3) of discrete points, allowing a greater degree of choice in household decision making. Second, we experiment with the composition of time-use variables, reducing the mother's household production decision to a single maternal child care choice.[31] A third extension augments the model by a fixed disutility of work which is modeled as an additional term in the utility function (a dummy for positive hours of paid work times an additional parameter to be estimated).[32] In a fourth extension, in order to account for potential dependence of child care prices on the quality of the service, and for misreporting in the individual household accounts, we estimate a model with imputed wages and child care prices for everyone (instead of just for the households where wages or prices are not observed). A fifth alternative model investigates plausibility of the assumption that post-reform informal child care use remains fixed at the pre-reform levels. We do so by estimating a model which does not allow for interactions of choice variables with informal child care. This model does not assume that the provision of informal child care remains unchanged, and the potential effects of substitution between informal and formal child care are contained within the formal child care coefficients included in the utility function. The sixth robustness check generalizes the model using the random opportunity specification of Aaberge et al. (1999). This model allows part-time jobs to differ in their availability from full-time jobs, reflecting the fact that women may be facing more (or less) job opportunities depending on the desired work hours. Unlike Aaberge et al. (1999), we do not allow for availability of

---

[30] In our sample, 65 % of mothers are predicted to increase their hours of formal child care (typically the mothers who increase their hours of market work). Once again, a key role is played by mothers with higher wages who reduce their market work hours and also their utilization of formal care. The relatively large fall in their claimed benefits is sufficient to offset most of the rise in benefit claims by other households.

[31] This allows us to examine direct substitution effects between maternal and formal child care.

[32] There is no clear consensus with respect to which form the working indicator should take on. Blundell et al. (2007a) put the employment dummy into the budget constraint, so that it represents fixed monetary costs of working. Donald and Hamermesh (2009) interact the dummy with time-use variables entering the utility function, referring to the corresponding parameters as shifters of time-use efficiency. We choose to add the employment dummy into the individual utility function in a non-interacted form, which allows us to model fixed disutility from work without substantially increasing the computational burden.

jobs to depend on the corresponding wages, as this adjustment proves cumbersome in our model specification. The final extension evaluates the maintained assumption that women's leisure is consumed privately, yielding the same utility irrespective of the leisure choices of their partners. To check its plausibility, we augment the utility function with an interaction term of male and female leisure time, where male leisure is treated as given. The interaction term allows for a preference for shared spousal leisure time.

Table 9 shows that changes in the econometric specification induce changes in the values of the elasticities, but their relative sizes and signs remain similar to those in the original model. Most of the values remain within the 95 % confidence interval of the corresponding baseline elasticities. This finding is particularly important in the context of child care quality concerns, as it suggests that the differences in service quality reflected by variation in observed prices are unlikely to distort our estimates.

The stability of the elasticities is also interesting in the context of the model containing maternal child care decisions, as it suggests that changes in the hours of home production are proportional, irrespective of the distinction between child care-related and other household activities. Women who engage in the labor market will therefore work less in the household, delegating part of their chores either to the husband or buying in the services from the market.

We also check the validity of standard errors corresponding to the measured elasticities (without changing the specification of the model itself), by calculating standard errors in a robust way, controlling for general heteroskedasticity and household-specific clustering (considering the estimates as pseudo maximum likelihood estimates). In both cases the newly derived standard errors preserve the significance levels attained by the benchmark approach, suggesting that heteroskedasticity or clustering does not distort our results.

7 Conclusions

In this paper we have analyzed the time allocation decisions of partnered mothers with pre-school children, with emphasis on the influence of a non-convex tax and benefit system on labor supply, household production and the use of formal child care. We have focused on incorporating unobserved heterogeneity, originating possibly from differences in productivities and preferences. Our findings show this plays a dominant role in analyzing the mothers' decisions in our data. Our results cast strong doubts on the usefulness of the homogeneous model with no unobserved heterogeneity. The parameters fail to capture the true effects of factors driving household decision making, and hence simulations based on the baseline homogeneous model specification give misleading results.

To control for unobserved heterogeneity, we estimated a series of latent-class models, among which the 8-class model was found to perform best, balancing goodness of fit against parsimony. To assess the responsiveness to changes in the family tax and benefit system and in child care prices, we conducted several simulations based upon our estimated models, increasing net wages of mothers or net child care prices, and

Table 9 Robustness check—elasticities and reform responses derived by alternative model specifications with eight latent classes, partnered mothers with pre-school children

	Original model	Six brackets	Maternal care	Fixed disutility
Mothers' net wage increased by 10%				
Formal care hrs.	0.42 (0.065)***	0.51	0.43	0.42
Market work hrs.	0.43 (0.086)**	0.61	0.57	0.57
H'hold prod. hrs.	−0.08 (0.015)***	−0.12	−0.07	−0.09
	Imputed	No inf. care	Random opportunity	Interacted leisure
Formal care hrs.	0.36	0.35	0.51	0.55
Market work hrs.	0.47	0.51	0.52	0.76
H'hold prod. hrs.	−0.06	−0.10	−0.14	−0.11
	Original model	Six brackets	Maternal care	Fixed disutility
Net child care prices increased by 10%				
Formal care hrs.	−0.42 (0.081)***	−0.45	−0.44	−0.35
Market work hrs.	−0.08 (0.021)***	−0.11	−0.09	−0.11
H'hold prod. hrs.	0.02 (0.008)***	0.02	0.01	0.01
	Imputed	No inf. care	Random opportunity	Interacted leisure
Formal care hrs.	−0.40	−0.42	−0.43	−0.39
Market work hrs.	−0.10	−0.07	−0.09	−0.11
H'hold prod. hrs.	0.01	0.02	0.02	0.02

Table 9 continued

	Original model	Six brackets	Maternal care	Fixed disutility
FTB reform—percentage changes				
Formal care hrs.	1.75% (0.380)***	1.94%	2.29%	2.68%
Market work hrs.	3.11% (0.684)***	3.18%	3.01%	4.26%
H'hold prod. hrs.	−0.61% (0.133)***	−0.77%	−0.37%	−0.96%
	Imputed (%)	No inf. care (%)	Random opportunity (%)	Interacted leisure (%)
Formal care hrs.	1.92	1.14	2.58	2.18
Market work hrs.	3.25	3.45	4.23	3.73
H'hold prod. hrs.	−0.58	−0.83	−0.87	−0.76

Standard errors in the parentheses, significance levels: * 90, ** 95, *** 99

altering the joint-income structure of the existing tax and benefit system in the third reform.

The simulations show that mothers with pre-school children are responsive to changes in wages as well as changes in child care prices. The results suggest that market work and formal child care tend to be complements, and respond significantly to wage and price changes. The results also indicate that significant changes in labor supply and child care demand can remain unidentified when the unobserved heterogeneity is not accounted for, since the homogeneous model leads to significantly distorted female labor supply elasticities.

In the third simulation, we show that the phasing out of family benefits on the basis of joint income increases marginal tax rates on the incomes of mothers as second earners, with a negative impact on their labor supply. The tax system can be made more favorable for mothers with pre-school children by switching to a fully individual based system. In such a reformed setting, mothers are predicted to increase their labor supply and use of formal child care. The net budgetary effect of these behavioral responses is to raise additional tax revenue which could be used to lower tax rates and therefore achieve efficiency gains. The gains from the reforms we have simulated arise from changing the structure of effective marginal tax rates under the Australian "quasi-joint" family tax and benefit system. Running a similar policy simulation on data for countries with full joint taxation may yield considerably stronger behavioral responses.

It should be noted that these results are derived from the responses of partnered mothers with preschool children, and so they are specific for this population. However, in the context of the Australian family tax and benefit system for two-parent families with dependent children, the reform which we pursue could be potentially viewed as the one that needs to be considered for the full-population. The extremely large child payments increase further for older children, extending the problem of high effective marginal tax rates on secondary earners across the full-population of families with dependent children. For that reason, it is plausible that the proposed reform would prove beneficial for much broader set of Australian families than the one analyzed in this paper.

A number of improvements and extensions are of course possible. First, our analysis would benefit from exploiting the panel structure of the HILDA dataset, controlling for time-stable individual effects. Secondly, although we consider the current method of treating unobserved heterogeneity to perform well, it could be worthwhile to assess the stability of our results by using alternative ways of controlling for unobserved heterogeneity, such as the random coefficient mixed logit model, or the approaches utilizing Bayesian nonparametric methods. Thirdly, with sufficient information on disposable income of single mothers, the model could be extended to account also for this demographic. It is likely that we would then observe even stronger dependence of market work on child care, since single mothers are more restricted in their informal child care choices. An interesting extension would be to model fertility as a choice, since the family tax benefits are directly influencing the costs of childbearing. This would be, however, difficult in the current static setup of our model.

Appendix 1: Heckman selection models

In Tables 10 and 11 below we present details of the Heckman selection models used to predict missing wages and child care prices. The models are estimated on a sample of partnered women aged 55 years and under, with those reported as full-time students or disabled excluded.[33] The exclusion restrictions used in the participation equation are non-wage income and number of children.

The child care price equations are estimated on a sample limited to mothers with pre-school children. This subsample is larger than the sample used for estimating the structural time-use model because the time-use data in HILDA was collected for a randomized subsample only. The exclusion restrictions in the child care participation equation are number of adults in the household (excluding the spouses) and residential distance from grandparents (the base group represents families without grandparents).

Appendix 2: Australian family income taxes and child care subsidies

Net household income is calculated as gross income net of tax liabilities and family payments. We compute tax liabilities under the Personal Income Tax (PIT), Low Income Tax Offset (LITO), Medicare Levy (ML),[34] and family payments under Family Tax Benefit Part A (FTB-A) and Family Tax Benefit Part B (FTB-B). The calculation of the net price of formal child care takes account of the two main subsidies for child care, Child Care Benefit (CCB) and the Child Care Rebate (CCR). Each of these components of the overall system is described below. We also report details of the tax-benefit system used to construct Figs. 6 and 7 for a family with two children under 13, with one under 5, in the 2007–2008 financial year.

Personal Income Tax and LITO

The 2007–2008 marginal rate scale of the PIT begins with a zero rated threshold of $6,000, followed by rates of 15, 30 and 40 % up to an income of $150,000, and thereafter a top rate of 45 %. The LITO in the same financial year provided a tax offset of $750, phased out at 4 cents in the dollar on individual incomes above $30,000. The resulting *effective* rate scale was therefore a zero rated threshold of $11,000 and a higher rate of 34 cents in the dollar on incomes from $30,001 to $48,750, as depicted in Fig. 6. Because the

Medicare Levy

A Medicare Levy (ML) applies at a rate of 1.5 % of income, with exemptions defined on family income and varying with the number of children. In 2007–2008 the family income limit for a full reduction for a two-parent family was $29,207, plus $2,682 for each dependent child. The exemption was withdrawn at a rate of 8.5 cents in the

---

[33] Restricting the age range to women aged 48 and under hardly affects the results—the correlation coefficient between the two sets of wage predictions is 0.99.

[34] Despite its title, the ML is a tax on income and is not tied to funding any aspect of the health system.

Table 10 Mothers' wage estimation, Heckman selection model

	Participation equation	Wage equation
Constant	−.027	3.984
	(.307)	(7.380)
Married	−.048	1.057
	(.039)	(.925)
Urbanization index	.100	1.086
	(.028)***	(.730)
Non-English ethnicity	−.287	−4.351
	(.046)***	(1.176)***
Mother's age	.033	.265
	(.015)**	(.366)
Mother's age squared	−.0009	−.006
	(.0002)***	(.005)
Mother's tenure	.032	.404
	(.004)***	(.113)***
Mother's tenure squared	.0005	.0001
	(.0001)***	(.003)
Other household income (log)	−.100	
	(.012)***	
No. of children aged 0–4	−.574	
	(.024)***	
No. of children aged 5–9	−.226	
	(.025)***	
No. of children aged 10–14	−.149	
	(.026)***	
No. of children aged 15–18	.019	
	(.032)	
Inverse Mills ratio		13.121
		(1.994)***
Obs.	9324	9324

The dependent variable is gross hourly wage

Standard errors in the parentheses, significance levels: *90, **95, ***99

Additional controls include regional, educational and yearly dummies

dollar above this limit, with the effect of raising the marginal rate above that limit to 44 cents in the dollar. Thus the ML introduces a further nonconvexity in the *effective* rate scale and also shifts the tax base toward joint income.

FTB-A and FTB-B

FTB-A provides a payments for each dependent child. The size of the payment varies with the age of the child. The "Maximum Rate" of FTB-A in 2007–2008 for a child under 13 years was $4,460.30. This maximum payment was withdrawn at 20 cents

Table 11 Child care price estimation, Heckman selection model

	Participation equation	Price equation
Const.	−2.645	8.090
	(.968)***	(6.187)
Mother's gross hourly wage	.003	−.008
	(.0009)***	(.008)
Other household income (log)	−.010	.049
	(.030)	(.121)
Married	−.050	.409
	(.094)	(.414)
Urbanization index	.096	.354
	(.068)	(.288)
Non-English ethnicity	−.197	−.195
	(.125)	(.586)
Mother's age	.069	.001
	(.055)	(.273)
Mother's age squared	−.0003	−.0002
	(.0008)	(.004)
No. of children aged 0–4	.174	−.276
	(.059)***	(.294)
No. of children aged 5–9	−.084	−.130
	(.051)*	(.229)
No. of children aged 10–14	−.274	−.981
	(.087)***	(.471)**
No. of children aged 15–18	−.277	−1.683
	(.143)*	(.652)***
No. of other adults in the household	.015	
	(.131)	
Distance to grandparents: Same household	−.730	
	(.376)*	
Distance to grandparents: Less than 1 km	−.408	
	(.147)***	
Distance to grandparents: 1–4 kms	−.243	
	(.144)*	
Distance to grandparents: 5–9 kms	−.091	
	(.135)	
Distance to grandparents: 10–19 kms	−.086	
	(.138)	
Distance to grandparents: 20–49 kms	−.297	
	(.153)*	
Distance to grandparents: 50–99 kms	−.025	
	(.210)	

Table 11 continued

	Participation equation	Price equation
Distance to grandparents: 100–499 kms	−.068	
	(.166)	
Distance to grandparents: 500 kms or more	−.030	
	(.154)	
Distance to grandparents: Overseas	−.413	
	(.218)*	
Inverse Mills ratio		.932
		(1.498)
Obs.	1725	1725

The dependent variable is gross hourly price of formal child care

Standard errors in the parentheses, significance levels: *90, **95, ***99

Additional controls include regional, educational and yearly dummies

in the dollar on a family income over $41,318 up to the "Base Rate" of $1,890.70 per annum. The Base Rate was withdrawn at 30 cents in the dollar at a higher family income threshold that depends on the number of dependent children. For a family with two dependent children, the income threshold for the Base Rate was $95,192. Given the size of the payments and withdrawal rates of 20 and 30 cents in the dollar, married mothers can face effective marginal rates on earnings that are well above the top rate of the PIT scale, as illustrated in Fig. 7.

FTB-B provides an annual payment of $3,584.30 in 2007–2008 for a family with a child under 5 years. The payment was withdrawn at a rate of 20 cents in the dollar on a second income above $4,380. It can therefore be classified as a "gender based tax" (see Alesina et al. 2011) with, paradoxically, the higher rate applying to the income of the mother as second earner.

Child Care Benefit and Child Care Rebate

Child Care Benefit depends (among other things) on the ages of children, number of children, type of child care and the hours of child care used. The benefit is phased out with rising family income according to the age of the child and the number of children receiving child care.

The Child Care Rebate reimburses families for their claimed child care expenses. It can cover up to 50 % of the net child care expenses (that is, after subtracting CCB). The CCR rate is not income-tested, but it has an upper cap on the amount of expenses which can be reimbursed. For the year 2008, this cap was $4,354 per year.

References

Aaberge R, Colombino U, Strom S (1999) Labour supply in Italy: an empirical analysis of joint household decisions, with taxes and quantity constraints. J Appl Econ 14:403–422

Alesina A, Ichino A, Karabarbounis L (2011) Gender-based taxation and the division of family chores. Am Econ J Econ Policy 3:1–40

Apps P, Rees R (2009) Public economics and the household. Cambridge University Press, Cambridge

Baker M, Gruber J, Milligan K (2008) Universal child care, maternal labor supply, and family well-being. J Polit Econ 116:709–745

Bernal R (2008) The effect of maternal employment and child care on children's cognitive development. Int Econ Rev 49:1173–1209

Blau DM (2003) Child care subsidy programs, in Means-tested transfer programs in the United States. National Bureau of Economic Research, Inc, NBER Chapters, pp 443–516

Blau DM, Hagy AP (1998) The demand for quality in child care. J Polit Econ 106:104–146

Blundell R, Gosling A, Ichimura H, Meghir C (2007a) Changes in the distribution of male and female wages accounting for employment composition using bounds. Econometrica 75:323–363

Blundell R, MaCurdy T, Meghir C (2007b) Labor supply models: Unobserved heterogeneity, nonparticipation and dynamics. In: Heckman J, Leamer E (eds) Handbook of econometrics, chap. 69, vol. 6. Elsevier, Amsterdam

Blundell R, Shephard A (2012) Employment, hours of work and the optimal taxation of low-income families. Rev Econ Stud 79:481–510

Breunig R, Gong X, King A (2012) Partnered women's labour supply and child-care costs in Australia: measurement error and the child-care price. Econ Rec 88:51–69

Burda M, Harding M, Hausman J (2008) A Bayesian mixed logit & probit model for multinomial choice. J Econom 147:232–246

Connelly R (1992) The effect of child care costs on married women's labor force participation. Rev Econ Stat 74:83–90

Connelly R, Kimmel J (2003) Marital status and full-time/part-time work status in child care choices. Appl Econ 35:761–777

Doiron D, Kalb G (2005) Demands for child care and household labour supply in Australia. Econ Rec 81:215–236

Donald SG, Hamermesh DS (2009) A structural model of the fixed time costs of market work. Econ Lett 104:125–128

Gong X, Breunig R (2011) Estimating net child care price elasticities of partnered women with pre-school children using a discrete structural labour supply-child care model. CEPR Discussion Papers 653, Centre for Economic Policy Research, Research School of Economics, Australian National University

Heckman J (1979) Sample selection bias as a specification error. Econometrica 129:153–161

Heckman J, Singer B (1984) A method for minimizing the impact of distributional assumptions in econometric models for duration data. Econometrica 52:271–320

Kabátek J, van Soest A, Stancanelli E (2014) Income taxation, labour supply and housework: a discrete choice model for French couples. Lab Econ 27:30–43

Kalenkoski CM, Ribar DC, Stratton LS (2005) Parental child care in single-parent, cohabiting, and married-couple families: time-diary evidence from the United Kingdom. Am Econ Rev 95:194–198

Keane M, Moffitt R (1998) A structural model of multiple welfare program participation and labor supply. Int Econ Rev 39:553–589

Kornstad T, Thoresen T (2007) A discrete choice model for labor supply and childcare. J Popul Econ 20:781–803

Mroz TA (1999) Discrete factor approximations in simultaneous equation models: estimating the impact of a dummy endogenous variable on a continuous outcome. J Econom 92:233–274

Pacifico D (2012) On the role of unobserved preference heterogeneity in discrete choice models of labour supply. Empir Econ 45(2):929–963

Ribar DC (1995) A structural model of child care and the labor supply of married women. J Labor Econ 13:558–97

Ruud P (1991) Extensions of estimation methods using the EM algorithm. J Econom 49:305–341

Sorensen E (1993) Continuous female workers: how different are they from other women? East Econ J 19:15–32

Steiner V, Wrohlich K (2004) Household taxation, income splitting and labor supply incentives: a microsimulation study for Germany. Tech. rep

Tekin E (2007) Childcare subsidies, wages, and employment of single mothers. J Hum Resour 42(2):453–487

Train K (2008) EM algorithms for nonparametric estimation of mixing distributions. Journal of Choice Modelling 1:40–69

Van Soest A (1995) Structural models of family labor supply: a discrete choice approach. J Hum Resour 30:63–88

Van Soest A, Das M, Gong X (2002) A structural labour supply model with flexible preferences. J Econom 107:345–374