



CENTRE
FOR
HOUSEHOLD, INCOME, LABOUR AND DEMOGRAPHIC ECONOMICS



Working Papers

On Modeling Household Labor Supply With Taxation

O. Bargain

Child n. 18/2007

e-mail: de-child@unito.it
Web site: <http://www.child-centre.it>

On Modeling Household Labor Supply With Taxation*

Olivier Bargain (UCD, IZA and CHILD)

August 2007

Abstract

Discrete choice models of labor supply easily account for nonlinearity and nonconvexity in budget sets caused by tax-benefit systems. As a result, they have become very popular for *ex ante* evaluations of policy reforms. In this paper, we question whether the degree of flexibility and the implicit household representation in these models are satisfying when confronted to the data. First, we show that attempts to interpret discrete models structurally lead to unnecessary parametric restrictions in most studies. We suggest instead a fully flexible model that retains usual assumptions on economic rationality except regularity conditions on leisure. Indeed, coefficients may account for both tastes and costs of work, possibly making ‘preferences’ appear nonconvex. Second, we show that the static unitary representation, implicit in most tax policy analyses, is rejected against a more general model with price- and income- dependent preferences. The latter can be rationalized in terms of collective or intertemporal models and offers promising perspectives in these directions. Simulations show that the magnitude of predicted labor supply responses to tax-benefit reforms is sensitive to the underlying household representation.

Key Words : multinomial logit, household labor supply, tax reform, unitary model, collective model.

JEL Classification : C25, C52, H31, J22.

***Acknowledgements** : I am indebted for useful comments and advice to François Bourguignon, Martin Browning, Pierre-André Chiappori, Olivier Donni, Paul Devereux, Alan Duncan, François Laisney, Thierry Magnac, Martin Ryan, Arthur van Soest, Alain Trannoy, Frederic Vermeulen and participants of seminars at Tilburg, Royal Holloway, PSE, CAM, IZA, ULB, UCD, Galway, Laval, Sherbrooke and UQAM. All errors or omissions remain mine. Correspondence: O. Bargain, UCD, Newman Building, School of Economics, Dublin 4, Ireland. Email: olivier.bargain@ucd.ie

1 Introduction

The understanding of household labor supply behavior continues to attract considerable research interest, the main motivation for it being the recurring importance placed on responses to tax and benefit reforms (see Blundell and MaCurdy, 2000, for a comprehensive survey). The recent literature relies heavily on discrete choice modeling. This approach requires the explicit parameterization of consumption-leisure preferences and maximization is reduced to choosing among a discrete set of possibilities. This allows dealing easily with complex tax-benefit systems that yield nonlinearities and nonconvexities in agents' budget sets. The simultaneity of the participation decision and the choice of work duration is also handled in a straightforward way. Finally, the joint decision in couples is easily modeled as a natural extension of the single case, as in Aaberge et al. (1995) and van Soest (1995). Under these conditions, the success of the discrete approach among labor supply modelers came as no surprise.¹

Also in contrast with the Hausman approach, discrete-choice models impose in principle little constraint on preferences. However, it is not sure whether constraints on functional forms are totally relaxed in practice. Moreover, the 'standard' household representation most often imposed in current applications – i.e. unitary households taking decision in a static environment – may be rejected when brought to the data. In this paper, we explore these two potential limits. To do so, we estimate the labor supply of French married women and suggest a series of nested models that can be tested one against the others. These models relax step by step the restrictions usually imposed on functional forms and household representation.

The first question we address is whether flexibility is achieved in practice. We survey the recent literature and show that the model usually at use – a structural model with fixed costs of work – imposes unnecessary parametric restrictions. Instead, the utility function may be specified in a very general way, with coefficients that vary with the labor supply alternatives. This unrestricted model achieves the best possible fit while imposing usual requirements on economic rationality. In particular, it relies on a utility-maximizing interpretation, with quasi-concave and increasing utility of income, so that traditional efficiency-equity analysis of tax reforms can be performed. While costs of work are generally identified from preferences only at the price of parametric restriction, they are implicitly incorporated in the coefficients of the flexible model and regularity conditions on leisure need not to hold.

Under this more flexible form, choices still depend on household disposable income achieved at each discrete hour along with household characteristics. This specification thus maintains the 'standard' representation of household decisions; in particular, it still assumes income pooling, a necessary condition of the unitary model. A second generalization then consists in a model that depends on wage rates and disaggregated exogenous incomes, as if preferences were price- and income-dependent. Such a model unambiguously rejects the standard approach, suggesting that the unitary and static representation may not be the most appropriate one to approximate 'true' behaviors. Interestingly, the general model can be rationalized along the lines of intertemporal or bargaining models, providing a basis for further tests

¹The traditional continuous approach presented in Hausman (1981) is usually restricted to the case of piecewise linear and convex budget sets. To account for nonconvexities, as in Hausman and Ruud (1984), labor supply must be specified parametrically together with the corresponding direct utility function, which implies rather restrictive forms for preferences. In addition, MaCurdy et al. (1990) have emphasized that the model requires the global satisfaction of Slutsky conditions by the labor supply function, and hence imposes undesirable *a priori* restrictions on estimated behavior.

in these directions.

Finally, we compare the predictions of the different models to a tax-benefit reform. The magnitude of labor supply responses is sensitive to the model at use. Future *ex post* evaluations may allow discriminating between the different household representations on this basis.

The layout of the paper is as follows. In Section 2, we introduce a series of nested discrete-choice models with different levels of flexibility and briefly discuss their economic interpretation; likelihood ratio tests are conducted in Section 3. The *ex ante* evaluation of a tax-benefit reform is performed in Section 4 using the different models. Section 5 concludes.

2 Discrete-choice Models of Labor Supply with Taxation

2.1 Multinomial Logit

The representation of discrete choices through the multinomial logit specification is the basis of this paper. The choice of working hours is supposed to be made between a finite number of alternatives, corresponding to commonly agreed durations of work, e.g. part-time, full-time and overtime. If household i can choose among J discrete alternatives, the utility it may derive from alternative j ($= 1, \dots, J$) is assumed to be given by:

$$V_{ij} = U(H_j, C_{ij}; Z_i, v_i, \theta) + \epsilon_{ij}. \quad (1)$$

In that expression, U stands for the household utility derived from working H_j hours per week and from the corresponding level of weekly household consumption C_{ij} , conditionally on a vector Z_i of demographic characteristics, unobserved heterogeneity v_i and a vector θ of common preference parameters. Let the first alternative represent the choice of non-participation, i.e. $H_1 = 0$. The deterministic utility is completed by an i.i.d. error term ϵ_{ij} assumed to represent possible observational errors, optimization errors or transitory departures from best choice by agents (see van Soest, 1995). As justified later, we assume in our empirical application that men in couple have fixed labor supply and we focus on the work behavior of married women, for whom we model J alternatives. In principle, it is also possible to model joint decisions of couples by simply extending labor supply alternatives to J combinations of partners' discrete hours.

In the present static framework, consumption coincides with disposable income as given by the following budget constraint:

$$C_{ij} = D(w_i H_j, y_i^m, y_i^K, \zeta_i). \quad (2)$$

The arguments of function D are some socio-demographic characteristics ζ_i of household i as well as the various sources of gross income, namely the female worker's labor income $w_i H_j$, with w_i her wage rate, her husband's labor income y_i^m (treated as exogenous), and the household unearned income y_i^K . The mapping of gross income into disposable income, D , stands for a fairly complex set of tax-benefit rules, typically approximated by microsimulation. Costs of work may also be taken out of total disposable income, as further discussed below.

Under the assumption that error terms ϵ_{ij} follow a I-extreme value (I-EV) distribution, and for a given

type v_i , the (conditional) probability that household i choose alternative k has the following explicit form:

$$P_{ik} = \Pr(V_{ik} \geq V_{ij}, \forall j = 1, \dots, J; v_i) = \frac{\exp U(H_j, C_{ij}; Z_i, v_i, \theta)}{\sum_{j=1}^J \exp U(H_j, C_{ij}; Z_i, v_i, \theta)}. \quad (3)$$

One of the preference parameters θ is assumed to be affected by a normally distributed term v_i , allowing for random taste heterogeneity and unrestricted substitution patterns between alternatives. The unconditional probability of choosing alternative k is then obtained by integrating P_{ik} over all values of v_i .² In practice, this is obtained by averaging the conditional probability over a large number of draws. We thus derive a simulated likelihood function that can be maximized to obtain estimates of parameters θ . Measurement errors due to the discretization can be handled in the way described by MaCurdy et al. (1990).

2.2 Structural and Unconstrained Models with Discrete Choices

The preceding discrete-choice model is fully *structural* in the sense that it completely specifies consumption-leisure preferences for household i and choice j . Thus it shall be referred to as model (S) hereafter. In practice, a certain number of restrictions are imposed implicitly or explicitly on this model. Firstly, the disposable incomes obtained for choices $k \neq j$ do not enter the utility of choosing alternative j . This is a usual restriction made for obvious identification reasons. Secondly, a given functional form is to be chosen for U , in which the set of parameters θ that describe preferences is usually not specific to the quantity of labor that is supplied, that is, $\theta_j = \theta$ for all choices j . This must introduce parameter restrictions across alternatives. Thirdly, well-behaved preferences require the usual properties of monotonicity and quasi-concavity of U with respect to hours of work and consumption. In fact, the discrete-choice approach consists of utility maximization over a finite budget and does not require tangency conditions to hold. They may nonetheless be imposed to comply with economic theory.³ A good reason not to force models to verify regularity conditions on leisure comes from the introduction of other structural components. Precisely, flexibility is often achieved by the addition of fixed costs of work or dummies for part-time options. Those are identified only under parametric assumptions on preferences.⁴ Heim and Meyer (2004) thus point at the difficulty of incorporating a realistic rendering of these costs and argue that they can make ‘preferences’ appear nonconvex.

Acknowledging these limitations, we suggest a model where coefficients vary freely with labor supply alternatives. This ‘unconstrained’ model (U) is written:

$$V_{ij} = U(C_{ij}; Z_i, v_i, \theta_j) + \epsilon_{ij}. \quad (4)$$

²Theoretically, it is possible to vary randomly all the coefficients. This may become enormously complex however, as multiple integrals have to be solved (Train, 2003).

³Often in practice, they are simply checked *a posteriori* to avoid the MaCurdy critique (see MaCurdy et al., 1990).

⁴We could not make this point better than van Soest et al. (2002): “...the intuitive explanation why fixed costs are identified is the lack of observation with a small positive number of working hours. While this argument is valid for a restrictive specification of the utility function that limits the way in which utility can vary with working hours, the argument would no longer hold if the specification of the utility function were fully non-parametric. For such a specification, the utility function itself could pick up the gap in the distribution at few hours, assigning lower utility to such hours values. Thus it seems that the fixed costs are nonparametrically unidentified.”

Naturally, the way utility varies with the level of labor is entirely counted for by alternative-specific coefficients θ_j , i.e. H_j does not need enter the specification any longer. Think, for instance, of a specification where disutility of work takes the linear form αH_j in the structural model. The difference in disutility between alternatives k and l is just a function of the difference in hours, i.e. $\alpha(H_k - H_l)$, so that moving from 20 to 30 hours/week has the same effect as moving from 30 to 40 hours. In model (U), this is replaced by the difference of two parameters, say $\alpha_k - \alpha_l$, so that the (marginal) disutility varies in a flexible way across alternatives.

Model (U) perfectly illustrates the point made above about parametric identification of the different structural components. In (U), choice-specific parameters necessarily capture preferences together with costs of work and several other aspects that determine work duration. In these circumstances, there is no reason to impose regularity conditions on leisure. Nonetheless, the model requires natural restrictions to comply with basic economic rationality. Clearly, it still relies on the fundamental assumption of utility-maximizing behavior. The fact that utility functions increase with income C is also a minimum requirement for meaningful policy analysis using model (U). Finally, quasi-concavity of preferences does not necessarily need to be imposed *a priori*, as discussed previously and in van Soest et al. (2002). In the empirical part, we shall use (U) to test the parameter restrictions across alternatives imposed by the structural model (S).

2.3 A Non-standard Model

In previous models, wages and exogenous incomes influence labor supply only through household disposable income, which supposes income pooling – a necessary condition of the unitary model – and a static framework. In that sense, both (U) and (S) are *standard* models. To go one step further, we suggest a model where the set of explanatory variables has been extended as follows:

$$V_{ij} = U(C_{ij}, w_i, y_i^m, y_i^K; Z_i, v_i, \theta_j) + \epsilon_{ij}. \quad (5)$$

In this general model (G), utility at alternative j now depends on disposable income together with the female wage rate, male earnings and non-labor income, as if, in some sense, preferences were price- and income-dependent. Importantly, a rejection of (U) against (G) would signify that the standard representation is not the best approximation of actual behaviors.

A very appealing structural interpretation of model (G) is the collective model of labor supply (Chiappori, 1988). This approach accounts for several decision-makers in the household and only assumes efficiency of spouses' decisions. Let us consider a collective model with purely private consumption (the reasoning can be extended to account for public goods within the household). We drop subscript i and heterogeneity to simplify notation hereafter. For alternative j , the private consumptions of the wife and the husband, c_j^f and c_j^m , sum up to the total household disposable income C_j . Denote $u^s(H_j, c_j^f, c_j^m)$ the utility function of spouse $s = f, m$ conditionally on the wife working H_j hours a week (the husband's labor supply is exogenously fixed). Under minimal regularity conditions, the collective decision problem can be represented as the maximization of a household welfare index:

$$\begin{aligned} & \text{Max}_{c_j^f, c_j^m} \quad \lambda u^f(H_j, c_j^f, c_j^m) + (1 - \lambda) u^m(H_j, c_j^f, c_j^m) \\ \text{s.t.} \quad & c_j^f + c_j^m = C_j, \end{aligned}$$

where the Pareto weight λ represents the balance of power in the household and depends on wages and exogenous incomes, i.e. $\lambda = \lambda(w, y^m, y^K)$. If we first take j as fixed and maximize with respect to individual consumptions, the solution can be denoted by $c^s = c^s(H_j, C_j, \lambda)$ for $s = f, m$ and substituted in the household welfare index to yield the (reduced-form) welfare:

$$U(C_j, H_j, \lambda)$$

expressed conditionally on j . Then this index is simply maximized with respect to labor supply. When using discrete alternatives, this second step boils down to the discrete labor supply model under study. With the flexible approach previously introduced, the index can be rewritten

$$U(C_j, \lambda(w, y^m, y^K); \theta_j)$$

Clearly, such a collective model is not identified; this is simply a reduced form of the household welfare function.⁵ The important point here is that it resembles model (G) very closely. The only difference is that female wage rate and exogenous incomes enter the household objective function through the same function λ for all choices $j = 1, \dots, J$.⁶

Model (G) can be rationalized in other ways and in particular along the line of an intertemporal unitary model. Indeed, any (static) structural model would be consistent with utility maximization in a life cycle framework with inter-temporally separable preferences if disposable income could be replaced by total expenditures (see Blundell and Walker, 1986). Since households cannot save under the static interpretation, the coefficients on female wage and exogenous incomes in model (G) may capture smoothing or precautionary decisions.

Discriminating amongst the various possible interpretations embedded in the reduced-form utility of (G) is not the purpose of this paper. However, each alternative interpretation (intertemporal, collective, etc.) would imply specific additional restrictions that could be tested. Note finally that previous interpretations assume that tax reforms only change total disposable income but do not affect other features which are not modeled. In particular, we implicitly rule out the effect of tax reforms on wage dynamics over the life cycle or on the (reduced-form) power index λ in a collective model. In this respect, we follow the bulk of the literature.⁷

2.4 Specifications

Before suggesting some specifications for models (S), (U) and (G), we report in Table 1 a non-exhaustive review of recent studies using discrete choice models for policy analysis. It clearly illustrates that the

⁵This is not a problem here since our main objective is to use the model for labor supply analysis, and not to measure welfare analysis at the individual level. Identification of individual preferences would necessarily require additional restrictions.

⁶Preliminary results show that this necessary condition of the collective model is rejected. However, these tests deserve more attention and are the subject of future research.

⁷Outside options or power indices usually depend on gross incomes or gross wages in the literature on bargaining models. Laisney (2002) is among the rare exceptions. In the real world, there are some examples of very specific tax-benefit reforms that imply a net transfer from one spouse to the other, and may therefore affect the balance of power in the household (see for instance Lundberg et al., 1997). The reform simulated in the present paper is unlikely to generate such intrahousehold redistribution.

discrete approach is increasingly used in many OECD countries to study a large variety of tax-benefit policies. Quadratic or translog functional forms are typically used and additional flexibility is often achieved by the introduction of fixed cost of work or part-time dummies – the latter are justified as search cost or rationing of less common working hours.⁸ To our knowledge, the only discrete-choice model that accounts for variable costs is that of Blundell et al. (2000). Variable childcare costs, estimated in a first stage on the sub-sample of households buying childcare, are withdrawn from disposable income using a deterministic relationship between mothers' working hours and these costs.

The only structural model with flexible preferences is that of van Soest et al. (2002).⁹ Their model attains non-parametric flexibility by using higher order polynomial forms of the utility functions than the usual quadratic form. This approach and the model (U) introduced in the present paper are compared in Nyffeler (2005) using estimates on Swiss data. In (U), flexibility is obtained directly, in a way that makes the underlying indifference curves discrete rather than continuous. This is after all consistent with a framework where the budget set itself is discrete.

2.4.1 Structural Model

We then suggest several specifications of the structural model (S), from the popular forms found in the literature to more flexible variants which allow a fair comparison with model (U). First, a frequent specification, as seen in Table 1, consists of a quadratic utility completed with fixed costs of work F_{ij} . A translog specification could alternatively be chosen but log terms would not allow negative net-of-cost income. This first structural model (S1) is written for choice $j = 1, \dots, J$:

$$U_{ij} = \alpha^{cc}(C_{ij} - F_{ij})^2 + \alpha^{hh}(H_j)^2 + \alpha^{ch}(C_{ij} - F_{ij})H_j + \alpha_i^c(C_{ij} - F_{ij}) + \alpha_i^h H_j, \quad (\text{S1})$$

with heterogeneity:

$$\begin{aligned} \alpha_i^c &= \alpha^{c0} + \alpha^{cL} Z_i + v_i \\ \alpha_i^h &= \alpha^{h0} + \alpha^{hL} Z_i, \end{aligned}$$

and vectors $\alpha^{cL} = (\alpha^{c1}, \dots, \alpha^{cL})$, $\alpha^{hL} = (\alpha^{h1}, \dots, \alpha^{hL})$. Preference variation across households is enabled by the introduction of observed heterogeneity in the vector $Z_i = (z_i^1, \dots, z_i^L)'$, which corresponds to $L = 7$ characteristics: to live in the Paris area, the number of children respectively between 0 and 2, 3 and 5, and 6 and 11, total number of children and the parents' age. The coefficient on disposable income, α_i^c , is assumed to be the random one – a natural choice to make models (S) and (U) easily comparable. The random component v_i is modeled as σu_i , with u_i following a standard normal distribution and σ the standard error to be estimated. Costs F_{ij} are to be paid if the wife starts to work. They vary with four household characteristics and are assumed non-stochastic:

$$\begin{aligned} F_{i1} &= 0 \\ F_{ij} &= f^0 + f^1 \text{Paris} + f^2 \text{Child02} + f^3 \text{Child35} + f^4 \text{Child611} \quad \text{if } j > 1. \end{aligned}$$

⁸In the case of the Netherlands, for instance, part-time is typically overpredicted at the expense of non-participation. This is solved by adding dummies for part-time in van Soest (1995) or fixed costs of work in Das and van Soest (2001) and Van Soest et al. (2002). Thereby, the interpretation of structural components appears to be somewhat *ad hoc*.

⁹Blomquist and Newey (2002) estimate a non-parametric labor supply function in the presence of taxation. However, some of the limitations of the Hausman approach equally apply to this generalization.

Table 1: Discrete Choice Models of Labor Supply: Recent Applications

	Country	Data	Selection	Functional form	Structural Flexibility	Unobs. Heterog.	Other features	simulation / focus
			(1)	(2)	(3)	(4)	(5)	
van Soest (1995)	Netherlands	Dutch SEP 1987	C	T	PT	RP		topical tax reforms
Aaberge, Dagsvik & Strøm (1995)	Norway	n.a.	C	#	#		#	progressive tax system
Callan & van Soest (1996)	Ireland	IDS 1987	C	T	FC	RP	rationing	individualized taxation
Euwals & van Soest (1996)	Netherlands	Dutch SEP 1988	S, C	T	FC	RP	rationing	change in income support/assess importance of rationing
Hoynes (1996)	US	SIPP 1984	C	S	FC	MX	stigma	employment effect of the AFDC
Bingley & Walker (1997)	UK	FES 78-92	SM	linear		RP	rationing, stigma	variations in the Family Credit
Keane & Moffitt (1998)	US	SIPP 1984	SM	Q		RP	simult. wage, stigma	actual increase in AFDC
Aaberge, Colombino & Strøm (1999)	Italy	SHIW 1987	C	#	#		#	
Duncan & McCrae (1999)	UK	FRS 1994-96	SM, C	Q	FC	RP	childcare, rationing	from FC to WFTC
Blundell, Duncan, McCrae, Meghir (2000)	UK	FRS 1995-6	S, C	Q	FC	RP	childcare, stigma	from FC to WFTC
Aaberge, Colombino & Strøm (2000)	Italy, Norway, Sweden	n.a.	C	#	#		#	hypothetical flat tax reform
Das and Van Soest (2001)	Netherlands	Dutch SEP 1995	C	T	FC	RP	rationing	proposed tax reforms
van Soest, Das & Gong (2002)	Netherlands	Dutch SEP 1995	C	Q, P	FC	RP	simult. wage	test flexible specifications
Gong & van Soest (2002)	Mexico city	Urban Employment Survey	C	T	FC	RP	simult. wage	estimation and sensitivity analysis
Bonin, Kempe & Schneider (2002)	Germany	GSOEP 2000	C	T	PT			alternative low-wage subsidies
Duncan (2002)	Australia	IDS	SM	Q	FC	RP		announced earned income tax credit
Flood, Hansen & Wahlberg (2003)	Sweden	HIS 1993	C	T		MX	stigma	hypothetical flat tax reform
Kalb & Scutella (2003)	New Zealand	HES 1991-2001	S, C	Q	FC	RP		estimation
Gerfin & Leu (2003)	Switzerland	Swiss IES 1998	S, C	Q		RP		alternative in-work policies
Steiner & Wrohlich (2004)	Germany	GSOEP 2002	C	T	PT			hypothetical income tax reforms and actual mini job reform
Labeaga, Oliver & Spadaro (2005)	Spain	ECHP 1995	S, C	Q	FC			hypothetical flat tax scheme and recent tax reforms
Brewer, Duncan, Shephard, Suarez (2006)	UK	FRS 1995-6 to 2002-3	S, C	Q	FC	RP	childcare, stigma	evaluate the WFTC
Bargain & Orsini (2006)	France, Germany, Finland	HBS 95, GSOEP 98, IDS 98	S, C	Q	FC			hypothetical in-work policies
Orsini (2007)	Belgium	Belgian PSBH 2001	C	Q	PT			recent Belgian reforms

Note: all studies are based on the multinomial logit model except Bingley and Walker (multinomial probit); only static models of hour choice are reported here -- models of mere participation decisions or models mixing labor supply with other types of decisions (fertility, childcare types, etc.) are not reviewed due to lack of space.

(1) **Selection:** Single male and female (S), Single Mothers (SM), Couples (C)

(2) **Functional form:** translog (T), quadratic (Q), higher polynomial form (P), Stone Geary (S)

(4) **Structural flexibility** is achieved using: part-time dummies (PT), fixed cost of work (FC) or, equivalently, fixed revenues of not working.

(5) **Unobserved heterogeneity:** normally distributed component (random parameter logit: RP), mass points à la Heckman-Singer (mixed logit: MX), none (conditional logit).

(6) **Other features:** first-step estimation of childcare expenses in function of female hours (childcare), simultaneous wage estimation (simult. wage), accounting for rationing using double hurdle model or information on desired hours (rationing), modeling of welfare program non-participation due to informational/search costs or stigma (stigma)

: This series of paper departs from the rest of the literature; they suggests the estimation of joint distributions of wage and hours and the sets of hour-wage opportunities vary across individuals.

As discussed above, fixed costs of work are not non-parametrically identified in (S1). Identification relies on the fact that the utility function in (S1) is not fully flexible and on exclusion restrictions regarding household characteristics placed on preferences and fixed costs respectively. Model (S1) thus corresponds to the type of specification most frequently used in the literature (cf. Table 1).

In a variant (S2), we would like to account for variable work costs due to childcare for households with children aged 0-2. For older children, full day childcare provided by public kindergarten (*maternelles*) is basically free, with unlimited supply and accessible geographically to most households. To account for variable care costs for the youngest children, we make the corresponding coefficient f_j^2 vary freely with the labor supply alternative.¹⁰ This conforms to the fact that childcare costs typically increase with the working time of the mothers for children in that age group.

In a last variant (S3), the average cost of work is also made variable across alternatives, i.e. f_j^0 , to account for the fact that other types of costs (e.g., transportation costs) may also depend on the number of hours worked per week. These costs are clearly not identified, even parametrically. In fact, the purpose of this model is rather illustrative: variable costs are a device to achieve more flexibility and make (S3) the most comparable version to model (U).

2.4.2 Unconstrained Model

The unconstrained model (U) is made comparable to (S) by use of the following quadratic form:

$$U_{ij} = a_j C_{ij}^2 + b_{ij} C_{ij} + c_{ij} \quad \text{for } j = 1, \dots, J. \quad (\text{U})$$

Heterogeneity is written as:

$$\begin{aligned} b_{ij} &= b_j^0 + b_j' Z_i + v_i \\ c_{ij} &= c_j^0 + c_j' Z_i + \sum_{l=1}^L \sum_{m=l}^L c_j^{lm} z_i^l z_i^m, \end{aligned} \quad (6)$$

with vectors $b_j' = (b_j^1, \dots, b_j^L)$ and $c_j' = (c_j^1, \dots, c_j^L)$.¹¹ Probability of choice j is written as:

$$\Pr(a_j C_{ij}^2 - a_k C_{ik}^2 + b_{ij} C_{ij} - b_{ik} C_{ik} + c_{ij} - c_{ik} > \epsilon_{ik} - \epsilon_{ij}, \forall k \neq j).$$

Because disposable income C_{ij} differs across alternatives, all coefficients a and b can be identified. The econometric indeterminacy on the last coefficient is removed by setting it to zero for the first alternative ($c_{i1} = 0$). Model (U) nests all types of structural models (S) introduced above. For instance, model (S1) imposes 156 restrictions on model (U), which corresponds to the difference in the numbers of coefficients of the two models. The 4 constraints $a_j = a_1 (= \alpha^{cc})$ for $j = 2, \dots, 5$ are straightforward. It is very simple to derive all the other constraints analytically – a formal proof is available upon request to the author.

¹⁰Performing a first step estimation on households purchasing formal childcare, as in Blundell et al. (2000), did not provide satisfying results.

¹¹Notice that unobserved heterogeneity v_i is placed on the coefficient of consumption, here b_{ij} . As in van Soest (1995), it is not alternative-specific, which makes this part of the specification directly comparable to (S)-models. In other words, additional flexibility in (U) is placed only on the deterministic part of the model.

2.4.3 General Model

Going one step further and keeping with the quadratic specification, a natural form for model (G) is as follows, for $j = 1, \dots, J$:

$$\begin{aligned}
 U_j = & \beta_j^{cc} C_j^2 + \beta_{ij}^c C_j + \beta_j^{ff} (w^f)^2 + \beta_j^{mm} (y^m)^2 \\
 & + \beta_j^{KK} (y^K)^2 + \beta_j^{fm} w^f y^m + \beta_j^{fK} w^f y^K + \beta_j^{mK} y^m y^K \\
 & + \beta_{ij}^f w^f + \beta_{ij}^m y^m + \beta_{ij}^K y^K + \beta_j^{fc} w^f C_{ij} + \beta_j^{mc} y^m C_{ij} + \beta_j^{Kc} y^K C_{ij} + \beta_{ij},
 \end{aligned} \tag{G}$$

with:

$$\begin{aligned}
 \beta_{ij}^c &= \beta_j^{c0} + \beta_j^{c'} Z_i + v_i \\
 \beta_{ij}^s &= \beta_j^{s0} + \beta_j^{s'} Z_i \quad \text{for } s = f, m, K \\
 \beta_{ij} &= \beta_j^0 + \beta_j' Z_i + \sum_{k=1}^L \sum_{l=k}^L \beta_j^{kl} z_i^k z_i^l.
 \end{aligned}$$

It is straightforward to see that this specification nests model (U). In (G) the usual indeterminacy for determinants which are not alternative-specific is removed by setting these coefficients to zero for the first alternative. Note that disposable income is itself a function of female wage rate, male earnings, unearned income and household characteristics. Identification in model (G) then necessarily relies on the strong nonlinearities of the tax-benefit function D in (2).¹²

3 Data, Selection and Discretization

The data used are selected from the French Household Budget Survey 1994-95. We have kept only married or cohabiting couples where adult members are in the age bracket 25 – 64 and where the wife is available for the labor market, i.e. not disabled, retired or a student. Households where the wife is self-employed are also withdrawn since they are subject to different income tax rules from those applying to salary workers and require unavailable additional information. Extreme households are selected out, notably those receiving important levels of non-labor income. To be consistent with a pure supply side approach and because we do not model rationing, we withdraw households with job seekers.

At this stage, 97% of the men in our selection are in work. Indeed, most of men in couple are employed and those out of work are usually involuntary unemployed, and hence not in our selection. Then we withdraw the few households with inactive men and focus purely on female labor supply, i.e. male hours are assumed fixed at observed values. This is a usual choice in the labor supply literature using French data (see Laroque and Salanié, 2002, Donni and Moreau, 2007). The final selection contains 3,397 couples; corresponding descriptive statistics are presented in Table 2.

In France, institutional norms and demand-side constraints imply concentrations around a limited number of working time arrangements. This is illustrated by the distribution of female hours reported in Figure 3. We simply base our discretization on the main concentration points, that is $H_j = 0, 20, 30, 39, 45$

¹²In all models above, non-parametric identification is not obtained for parameters of household/individual characteristics which are present both as taste/cost shifters and as determinants of the tax-benefit rules in function D (i.e. for $Z_i \cap \zeta_i$); again, parametric identification relies on the nonlinearities of the tax-benefit system.

hours a week, with corresponding intervals $[0 - 10[$, $[10 - 25[$, $[25 - 34[$, $[34 - 42[$ and over 42 respectively. For the year under consideration, full-time and part-time are institutionally fixed at 39 and 20 hours in France; three-quarter of a full-time (30 hours) is an option frequently offered by firms, especially to women; the 45 hours option corresponds to overtime. The proportion of households in each bracket is 24%, 12.6%, 10.9%, 44.7% and 7.8% respectively.¹³

Disposable income at each discrete choice of hours is computed using the French tax-benefit microsimulation SYSIFF98. This program allows the simulation of all direct taxes and benefits of instruments (see Bargain and Terraz, 2003). Wages for inactive women are predicted using the traditional Heckman correction. Because the labor supply models are nonlinear, it is necessary to take the wage rate prediction errors explicitly into account for a consistent estimation of the models, by integrating out the disturbance term of the wage equation in the likelihood. Practically, this is done by approximating the integral by a simulated mean for a tractable number of draws (20).¹⁴ Wage prediction for active women shows that the fit is reasonably good, even though the predicted distribution is, as often, more concentrated (results are available upon request). Following the bulk of the literature, we implicitly assume that gross hourly wage rates do not depend on working duration (this assumption is relaxed in Ilmakunnas and Pudney, 1990).

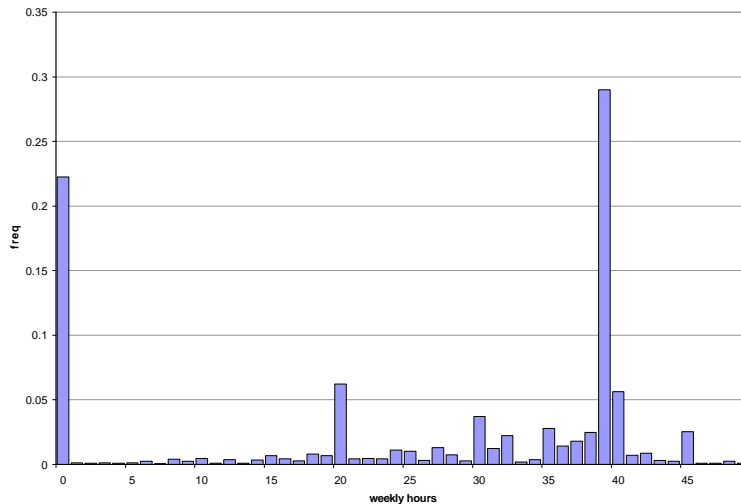
Table 2: Descriptive Statistics for Selected Couples

	Women	Men
Participation	0.77	1
Working time of participants (hours/week)	34.7	41.9
Gross wage rate - participant (euro/hour)	11.7	13.7
Gross wage rate - all potential workers (euro/hour)*	11.0	
Age	38.9	41.1
Primary education	0.31	0.18
Vocational training	0.38	0.46
High school diploma	0.15	0.18
University studies	0.17	0.18
Average number of children		1.43
Presence of child 0-2		0.17
Presence of child 3-5		0.19
Presence of child 6-11		0.33
Number of observations		3,397
Weighted size of the sample		6,369,455
Size in % of total population		0.28

* Those include predicted wages for non-workers.

¹³A large number of alternatives could become intractable, especially for flexible models. Five categories seem reasonable to capture the main peaks in the actual distribution without increasing too much the number of parameters. Nonetheless, models (U) and (G) may be limited by the fact that they require large enough samples and a limited number of alternatives. This point is discussed in the next section.

¹⁴Since the tax-benefit simulations are not performed in an econometric software, it was not possible to estimate the wage equation jointly with the labor supply model, as for instance in Laroque and Salanie (2001).



Distribution of Working Hours for Women with Employed Partners (selection)

4 Empirical Results

4.1 Estimation

Table 3 presents the results of the estimations for the structural models (S). We follow the traditional way to interpret coefficients in terms of consumption-leisure preferences and costs of work. However, as argued in the text, (S1) is identified thanks to functional form assumptions while (S3) is not identified, even parametrically. Interpretations should then be made with much caution. The only significant parameters for disposable income are the constant term and some of the variables related to children. Conversely, many of the estimates for hours are significant. As could have been expected, the marginal utility of work decreases with the presence of children. Women prefer to work significantly more if located in the Paris area and less if in older couples, suggesting a move toward single-earner couples as the household ages or simply a cohort effect.

The coefficient for average cost of work is significant in (S1) and (S2). However, non-identification in (S3) translates in the fact that coefficients on H , H^2 and $C \times H$ are not significant – in contrast to (S1) and (S2) – and neither are alternative-specific coefficients for variable costs. Working in Paris region seems to increase work costs. They decrease with the number of young children in (S1). This counter-intuitive result, also found in van Soest et al. (2002), can probably be attributed to identification problems, since the variable also enters the marginal utilities of consumption and leisure. In model (SC2), however, the variable cost due to the presence of children aged 0-2 is significant and increases as expected with mothers' work duration (excepted for the full-time option). Overall, estimates yield implausible values for the cost of work, equal on average to 41% of the average earnings of working wives. This seemingly deceiving result actually reflects identification issues; coefficients are also likely to capture more than supply-side dimensions and in particular rationing at certain hours. This point is discussed in detail by Bourguignon and Magnac (1990). The non-significance of 'childcare costs' at full-time ($j = 4$) in model (S2) might reflect the overwhelming presence of full-time contracts and the possible lack of alternative

options for women with children. Yet, only 40% of working women work full time (34 to 42 hours) so that the rationing hypothesis cannot be maintained for the whole sample.

Table 3: MNL estimation of the structural models

Variable	model (S1)		model (S2)		model (S3)	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
income ²	-22.640 ***	6.219	-26.376 ***	6.665	-64.852 ***	10.747
female hours ²	-10.684 ***	0.698	-9.831 ***	0.764	1.231	1.021
female hours x income	-14.759 ***	1.656	-14.818 ***	1.548	-4.394	9.425
income	26.451 *	13.517	36.786 ***	13.728	8.286	20.679
x age	0.104	0.653	-0.318	0.672	1.734	1.098
x age square	0.001	0.008	0.006	0.008	-0.017	0.014
x # children	-0.380	0.752	-0.373	0.808	-1.912	1.278
x # children 0-2	-3.449	3.117	-6.661 **	2.639	-3.664	4.068
x # children 3-5	6.047 **	2.698	5.102 **	2.464	5.717 **	2.533
x # children 6-11	2.926	2.328	3.261	2.132	1.644	1.853
x 1(Paris region)	0.447	3.214	-0.275	2.696	-2.952	2.151
female hours	10.562 ***	2.486	10.759 ***	2.280	-2.530	1.738
x age	0.278 **	0.123	0.220 *	0.113	0.212	0.140
x age square	-0.004 ***	0.002	-0.003 **	0.001	-0.004 ***	0.001
x # children	-0.566 ***	0.142	-0.557 ***	0.134	-0.455 ***	0.159
x # children 0-2	-0.407	0.310	-0.852	0.516	0.068	0.531
x # children 3-5	-0.211	0.239	-0.327	0.220	-0.829 *	0.480
x # children 6-11	-0.335 *	0.186	-0.337 **	0.171	-0.624 ***	0.207
x 1(Paris region)	1.790 ***	0.233	1.742 ***	0.207	1.849 ***	0.334
fixed costs/40000	0.190 ***	0.027	0.169 ***	0.024	-	-
x 1(Paris region)	0.029	0.018	0.031 **	0.014	0.027 ***	0.008
x # children 0-2	-0.022 **	0.011	-	-	-	-
x # children 3-5	-0.014	0.010	-0.017 *	0.009	-0.005	0.005
x # children 6-11	0.190 ***	0.027	-0.013	0.009	-0.005	0.004
variable costs/40000						
x # children 0-2 / j=2	-	-	0.034 **	0.016	0.021 **	0.010
x # children 0-2 / j=3	-	-	0.063 ***	0.021	0.015	0.013
x # children 0-2 / j=4	-	-	0.000	0.025	0.017	0.016
x # children 0-2 / j=5	-	-	0.085 ***	0.031	0.041 **	0.019
x j=2	-	-	-	-	0.021	0.032
x j=3	-	-	-	-	0.036	0.048
x j=4	-	-	-	-	0.003	0.064
x j=5	-	-	-	-	0.070	0.071
Nb of observations	3397		3397		3397	

Level of significance: *=10%, **=5%, ***=1%

Flexible specifications (U) and (G) capture broader heterogeneity in preferences but may also cause a larger number of households not to respect regularity conditions. As stated above, positive monotonicity and quasi-concavity in consumption seem natural requirements to perform meaningful policy analysis, albeit unnecessary for coherency of the econometric model. The first condition is respected at more than 99% in the (S)-models, at 96% in (U) but only at 77% in (G). It is written as:

$$2a_j C_{ij} + b_{ij} \geq 0 \text{ in (U)}$$

$$\beta_j^{cc} C_j^2 + \beta_{ij}^c + \beta_j^{fc} w^f + \beta_j^{mc} y^m + \beta_j^{Kc} y^K \geq 0 \text{ in (G)}$$

and is easily imposed as a constraint in the likelihood maximization.¹⁵ Quasi-concavity in C is simply

¹⁵In practice, and for both (U) and (G), Lagrangian multipliers need to depart only very slightly from zero to guarantee C -monotonicity for nearly 100% of the households.

checked *a posteriori* and is verified in all (S)-models, as well as in (U) and (G) once monotonicity is imposed.

Considering the large list of coefficients and the difficulty to interpret the results, regression tables are omitted for (U) and (G) – they are available upon request. Nonetheless, to give a feel of the main aspects captured in the unconstrained model, Table 4 provides the estimates of model (u), i.e., the basic version of (U) where all heterogeneity is withdrawn. With these estimates, monotonicity in consumption is verified up to a ‘point of satiety’ of 7,012 EUR/month for choice $H_1 = 0$ and up to 8,873 EUR/month for choice $H_5 = 45$ hours. Estimates show that the constant term is much larger for the institutional full-time work duration ($H_4 = 39$ hours). Clearly, model (U) picks other dimensions than pure preferences, including availability of certain types of job, in the same way as (S3) does through variable costs.

Table 4: MNL estimation of the unconstrained model without heterogeneity

		model (u)	
Variable		Coef.	Std. Err.
income ²	x j=1	-142.4860	6.1620
	x j=2	-125.7030	6.3610
	x j=3	-142.3500	7.1610
	x j=4	-113.3910	5.9560
	x j=5	-102.4150	6.0470
income	x j=1	75.6200	17.4710
	x j=2	72.4640	15.9410
	x j=3	79.5140	18.4710
	x j=4	67.1170	13.8650
	x j=5	68.7850	12.9910
constant	x j=2	-1.5260	0.2200
	x j=3	-2.6940	0.3470
	x j=4	-0.6510	0.2340
	x j=5	-3.1360	0.2720
Nb of observations		3397	

All parameters are significant at the 1% level. Income is divided by 40,000

4.2 Fit and Tests

We first compare the different models according to their within-sample fit, as summarized in the left part of Table 5. The McFadden pseudo- R^2 , written $1 - \ln L / \ln L_0$, gives the distance between the maximized value of the log-likelihood ($\ln L$) and the log-likelihood when all parameters are set to zero ($\ln L_0$). The measure suggests a ranking of the different models which is line with expectations. In particular, model (S1) and (S2) are outperformed by flexible models. The fit of model (S3) is almost as good as for (U). Yet, a drawback of model (U) is the large number of parameters due mainly to the introduction of household characteristics Z_i in a flexible way. We suggest a variant (U') which penalizes the model by forcing the interaction terms – the double sum in (6) – not to vary with the labor supply alternatives. This restriction can be justified by practical limitations due to curse of dimensionality problems. In this case, the log-likelihood, and consequently the pseudo- R^2 , are very similar in both (S3) and (U). We also

balance fit and parsimony by using the Akaike’s information criterion (AIC), written $2(-\ln L + 2K)/N$ with K the number of model parameters and N the number of observations. In this case, structural models (S1) and (S2) are still dominated by other specifications. While the general model (G) beats all the other models on the basis of its likelihood, it does not perform better than (S3) when penalizing for the number of parameters in AIC values.

Table 5 also reports a series of likelihood ratio (LR) tests. The first set of results concerns the tests on functional forms. The structural model most frequently used in the literature, (S1), is rejected at the 1% level against the unconstrained specification (U). The extended version (S2), with flexibility introduced through the coefficient related to the presence of young children, is also clearly rejected. As expected from the previous measures of fit, the third version (S3) gains enough flexibility and is not rejected against (U). Conclusions are straightforward. First, discrete models of labor supply currently at use impose *unnecessary restrictions on the form of household preferences*. Second, flexibility can be attained by introducing ‘variable work costs’ (or any sort of choice-specific dummies for all alternatives) or by opting for the fully flexible specification (U). These models are parametrically unidentified but fit the data much better and maintain sufficient assumptions on agents’ rationality to be used for meaningful tax reform analysis.

The second set of tests reports a clear *rejection of the standard approach*, i.e. a rejection at the 1% level of (U) versus (G). Thus the ‘true’ model underlying observed behaviors may be much more complex than the unitary and static approach most often assumed for policy analysis. Several studies have previously rejected some necessary conditions of the unitary model (income pooling and Slutsky conditions). For instance, Fortin and Lacroix (1997) reject the pooling of non-labor income in the determination of labor supply choices. The present test is different for at least two reasons. This is a rejection of the standard approach, not of the unitary model alone, since model (G) possibly captures intertemporal aspects. Moreover, and most importantly, the test does not rely on a restrictive parameterization of preferences.

At this stage, the performance of each model has been assessed on the sample used to estimate the models. Yet, richer specifications bear the risk to capture idiosyncrasies of the data at use. To check for overfitting, we then use a standard validation method that consists in estimating each model on a random 60% of the sample (the ‘training’ sample) and check model performance on the 40% holdout sample (the ‘evaluation’ sample). For all models, pseudo- R^2 computed on the evaluation sample are in the order of those previous reported. Importantly, structural models (S1) and (S2) are still rejected against (S3) and (U). We have repeated the exercise for smaller evaluation samples. In this case, sample size may be too small to assume a chi-squared distribution of the the likelihood ratio statistic. Comparisons of pseudo- R^2 nonetheless confirm that richer models outperform (S1) and (S2). Precisely, pseudo- R^2 are around twice as large in (U) and (S3), roughly the same order as in the within-sample validation. These results seem to indicate that better performances are not driven by sampling-error overfitting but, instead, that richer specifications (U) and (S3) better capture the complexity of labor supply behavior.¹⁶

¹⁶We detect however some overfitting in model (U) compared to model (S3): pseudo- R^2 evaluated on the holdout sample are 15% and 19% respectively. While using the penalized version (U’) suggested above, the two models become very similar and have the same pseudo- R^2 , around 19%.

In what follows, we illustrate the potential differences in model predictions when evaluating labor supply elasticities and the effects of a tax-benefit reform. For this purpose, we focus on the structural model (S2), which captures interesting features linked to variations in the cost of care for young children, and on the flexible models (U) and (G). Unsurprisingly, model (S3) yields very similar policy conclusions to the penalized version of (U).

Table 5: Log-likelihoods and LR tests

Model	# coefficients	Fit			LR tests		
		Log-likelihood	Pseudo-R2	AIC	nested by model:	LR	chi2 (1%)
S1	25	-4964	9%	2.94	U	1336	201
S2	28	-4876	11%	2.89	U	1160	197
S3	31	-4353	20%	2.58	U	114	193
U	181	-4296	21%	2.64	G	444	177
U'	100	-4345	20%	2.62	G	542	269
G	316	-4074	25%	2.58	-	-	-

Note: For U and G, the number of coefficient is reduced by 5 (i.e. the number of constraints imposed to guarantee C-monotonicity). Model U' is a variant of U which is penalized on the way taste shifters enter the model. 'LR' is the likelihood ratio statistic and 'chi2(1%)' gives the chi-squared value for the LR test at the 1% significance level.

5 Simulations

We first suggest a comparison of the average wage-elasticities obtained with the different models. A closed-form expression of elasticities is not available in our setting but wage-elasticities can be computed numerically by evaluating the change in female labor supply subsequent to a uniform 1% increase in the wage rate of all women in the sample. Because of the progressivity of the tax-benefit system, this approach has the drawback that net wages will change differently for women facing different effective marginal tax rates (and by slightly less than 1%). However, this definition is in line with the fact that the actual tax-benefit system is the benchmark for policy analysis.

We use a calibration method to simulate transition matrices.¹⁷ An alternative way consists simply in averaging predicted frequencies over all observations at each hour option. We find that the magnitudes of labor supply responses are very similar with both methods. Confidence intervals for each transition cell are simulated by drawing 500 times from the estimated asymptotic distribution of the parameter estimates and by applying the calibration method for each of those parameter draws.

The left part of Table 6 reports elasticities of working hours (intensive margin) and participation (extensive margin) respectively. Elasticities stand in a [.17, .45] range over the confidence intervals of the

¹⁷First, for the pre-reform situation (baseline) to display actual choices, we repetitively draw series of pseudo-residuals $\hat{\epsilon}_{ij}$ ($j = 1, \dots, J$) from a EV-I distribution together with unobserved heterogeneity. For each household, we keep a series that indeed leads to a perfect match between observed and predicted hours when using estimates of the deterministic model. Post-reform optimal choices are then computed using the new disposable income values at each alternative and retaining draws from the previous calibration step. The procedure is repeated 500 times to obtain transition frequencies for each household. Transition tables result from averaging over the whole population. As usually done in simulation studies, we assume that the policy change does not affect the random terms.

different models. These modest values are in line with recent findings for France (cf. Choné et al., 2003, Donni and Moreau, 2007) and for several other countries (cf. Blundell and MaCurdy, 2000). Compared to (S2), labor supply responsiveness is significantly larger with the flexible form (U) and larger still with the general model (G). In the latter, an increase in female wage could capture a higher bargaining position in a collective model, while the intertemporal interpretation rather suggests the positive effect of a higher wage profile on labor supply.

Table 6: Wage Elasticities and Policy Effects

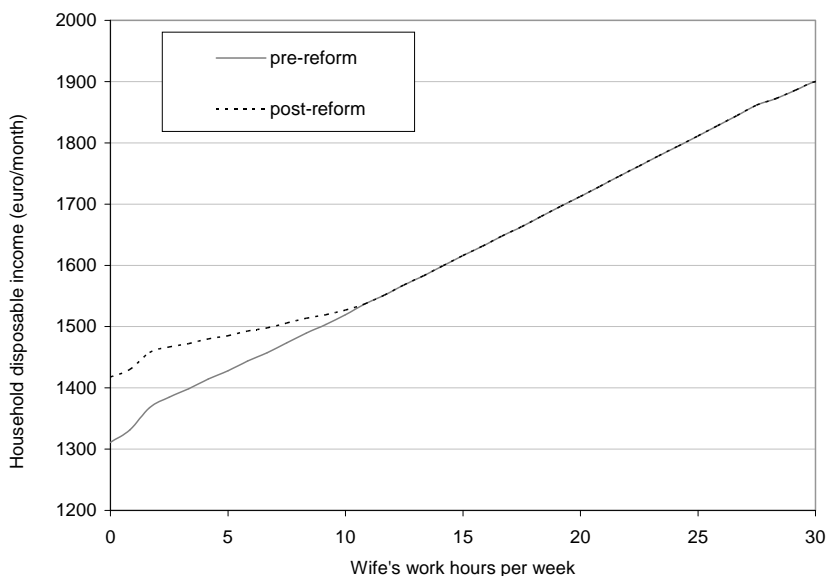
Model	1% increase in own wage		In-work benefit reform	
	Change in participation rate	Change in average work hour	Change in participation rate	Change in average work hour
	(in %-points)	(in %)	(in %-points)	(in %)
S2	0.14 [0.11; 0.17]	0.20 [0.17; 0.25]	-0.46 [-0.54; -0.37]	-0.62 [-0.74; -0.50]
U	0.20 [0.17; 0.24]	0.31 [0.27; 0.37]	-0.79 [-0.91; -0.65]	-1.08 [-1.24; -0.88]
G	0.30 [0.26; 0.34]	0.37 [0.32; 0.45]	-2.23 [-2.79; -2.06]	-3.08 [-3.75; -2.69]

Note: labor supply responses are computed using averaged simulated transitions; figures in brackets give bootstrapped confidence intervals.

The reform we examine next is a scenario of in-work benefit that has been discussed lately in France, namely the Earned Income Supplement (EIS) or *Allocation Compensatrice de Revenu*. It is similar to the US earned income tax credit or the British working family tax credit (WFTC). It has been advocated in France to offset the disincentive effects of social assistance (SA) by complementing earnings of low-wage households (see Godino et al., 1999). As a result, the EIS cumulated to SA corresponds to a transfer (350 euro/month for a single person plus increments to account for family size) reduced by 42% of household income, instead of the current 100% taper rate in the SA scheme. There is no condition on minimum work duration. Since this exercise merely aims to compare labor supply predictions across models, we do not attempt to reach revenue neutrality when simulating the reform.

This type of policy measure strongly encourages single individuals and lone parent to take up a job. However, it is conditional on household income and has therefore a well-known disincentive effect on secondary earners in couples (cf. Blundell et al., 2000, in the case of the WFTC). In effect, means-testing on joint incomes makes that the transfer is phased out as the second-earner increases her labor supply. This is illustrated by the budget constraint of an hypothetical household in Figure 1. Using a structural model similar to (S2), and assuming a 100% take-up, Blundell et al. (2000) find that around 28,000 women with an employed partner would stop working after the replacement of the old family tax credit by the WFTC.

Responses to the introduction of the EIS in France are summarized in the right part of Table 6. Most of the effects are negative – hardly any women are encouraged to take up a job or to increase their labor supply. The proportions of women who would potentially leave the labor market are 0.46% using (S2) and 0.79% with (U), which corresponds to 29,000 and 50,000 women respectively. While confidence intervals confirm that there is a significant difference in the predictions of restricted and flexible models,



Note: the budget curve is illustrated for a representative couple with two children. The husband is assumed to work 39 hours/week paid 1.33 times the minimum hourly wage and the wife is paid at 1.2 times the minimum hourly wage.

Figure 1: Effect of the Reform on a Hypothetical Budget Constraint

both results conclude that the disincentive effect of the reform is only moderately sized. Behavioral effects are significantly larger with the general model (G) as 2.23% of the women in the sample would be discouraged to work, i.e. around 142,000. Thus predictions turn out to be sensitive to the underlying household representation. In particular, the fact that gross wages come into play directly in model (G) – and not only through disposable income as in standard models – affects the sensitivity of the model to shocks upon the budget constraint, especially those due to policy reforms. Yet, the reform considered here is obviously just one example and it affects primarily the budget constraint of low-wage households. It would be interesting to confirm the present results for several reforms targeting different income or demographic groups in the population.

Conclusion

This paper first questions whether the possibility to use very general functional forms in discrete labor supply models is exploited in practice. It seems that structural models currently used for policy evaluation impose unnecessary constraints on leisure-consumption preferences. In these models, the identification of preferences from costs of work rests on weak ground. Acknowledging these limitations, we suggest a model where utility associated with the various hour choices depends on disposable income in a way that is totally independent across alternatives. Coefficients implicitly account for preferences together with costs of work and other structural aspects which influence hour choices. A structural model with equal flexibility would require enough variability across alternatives, for instance through the introduction of variables costs of work. In both cases, parametric identification is no longer possible but these models attain a significantly

better fit to the data. At the same time, they maintain a strict utility-maximizing interpretation and usual regularity conditions on consumption, which make them suitable for meaningful policy analysis. Going one step further, we introduce a model with wage- and income- dependent preferences. This model clearly departs from the usual standard representation used for policy analysis based on cross-sectional data, that is, the unitary and static assumption. In effect, this model can well nest specifications of intertemporal or collective models. Nested standard models are strongly rejected against this general setting.

Two paths could be followed in future research. First, it seems important to compare the predictions of the different models used for *ex ante* evaluation of reforms to actual changes following the reforms. Since the magnitude of labor supply responses is sensitive to the structural model at use, *ex post* evaluations of actual reforms may help to discriminate between the various models compared in this paper. Second, additional restrictions could be put on the reduced-form utility of the general model in order to construct and test alternative household representations, like the collective model of labor supply. In particular, this could help to understand what information is carried by (non-pooled) income sources. Additional variables may also be used for identification, as in Donni and Moreau (2007).

Two final aspects are worth mentioning. Attempts should be made to rule out demand-side aspects. A straightforward solution consists in using data on desired rather than observed hours, as done for instance in Ilmakunnas and Pudney (1990); it is nonetheless difficult to make sure that individuals' answers to the preferred-hours question only reflect preferences and are not themselves affected by labor market constraints. Other ways to account for labor market inflexibilities with regard to the available options for hours of work are suggested by Duncan and Harris (2002) and Aaberge et al. (1995, 1999, 2000). Also, the questions asked in this paper concerning the flexibility and identification of the deterministic part of structural models may also apply to the stochastic components. Indeed, distributional assumptions (e.g. extreme value) are potentially significant and random coefficients, introduced to render unobserved heterogeneity, may capture other dimensions (e.g. measurement errors due to the model discretization).

References

- [1] Aaberge, R., J.K. Dagsvik and S. Strom (1995): "Labor Supply Responses and Welfare Effects of Tax Reforms", *Scandinavian Journal of Economics*, 97(4), 635-59.
- [2] Aaberge, R., U. Colombino and S. Strom (1999): "Labour Supply in Italy: An Empirical Analysis of Joint Household Decisions, with Taxes and Quantity Constraints", *Journal of Applied Econometrics*, 14, 403-422.
- [3] Bargain, O. and K. Orsini (2006): "In-work policies in Europe: killing two birds with one stone", *Labour Economics*, forthcoming.
- [4] Bargain, O. and I. Terraz (2003): "Evaluation et mise en perspective des effets incitatifs et redistributifs de la Prime pour l'Emploi.", *Economie et Prévision*, 160/161, 121-148.
- [5] Bingley, P. and I. Walker (1997): "The Labour Supply, Unemployment and Participation of Lone Mothers in In-Work Transfer Programmes", *Economic Journal*, 107, 1375-1390

- [6] Blomquist, S. and W. Newey (2002): “Nonparametric estimation with nonlinear budget sets”, *Econometrica*, 70, 6, 2455-2480.
- [7] Blundell, R.W. and T. MaCurdy (2000): “Labor Supply: A Review of Alternative Approaches”, in Ashenfelter and Card (eds), *Handbook of Labor Economics*, Elsevier North-Holland.
- [8] Blundell, R.W., A. Duncan, J. McCrae and C. Meghir (2000): “The Labour Market Impact of the Working Families’ Tax Credit”, *Fiscal Studies*, 21, 1, 75-103.
- [9] Bonin, H., W. Kempe and H. Schneider (2002): “Household Labor Supply Effects of Low-Wage Subsidies in Germany”, *Journal of Applied Social Science Studies*.
- [10] Bourguignon, F. and T. Magnac (1990): “Labour supply and taxation in France”, *Journal of Human Resources*, 25, 358-389.
- [11] Brewer, M., A. Duncan, A. Shephard, M.J. Suarez (2006): “Did the Working Families’ Tax Credit work? The impact of in-work support on labour supply in Great Britain”, *Labour Economics*, forthcoming.
- [12] Callan, T and A. van Soest (1996): “Family Labour Supply and Taxes in Ireland”, ESRI working paper.
- [13] Chiappori, P.-A. (1988): “Rational Household Labor Supply ”, *Econometrica*, 56, 63-89.
- [14] Choné, P., D. Le Blanc and I. Robert-Bobée (2003): “Female labor supply and child care in France”, *CREST working paper*.
- [15] Dagsvik, J.K. and A. Karlstrom (2003): “Compensating variations and Hicksian choice probabilities in random utility models that are nonlinear in income”, working paper Royal Institute of Technology, Stockholm, Sweden.
- [16] Das, M. and A. van Soest (2001): “Family labor supply and proposed tax reforms in the Netherlands”, *De Economist*, 149(2), 191-218.
- [17] Donni, O. and N. Moreau (2007): “Collective Labor Supply: A Single-Equation Model and Some Evidence from French Data.”, *Journal of Human Resource*, 42(1), 214-246.
- [18] Duncan, A. and M. Weeks (1998): “Simulating transitions using discrete choice models”, *Papers and Proceedings of the American Economic Statistical Association*, 106, 151-156.
- [19] Duncan, A. and J. MacCrae (1999): “Household Labour Supply, Childcare Cost and In-Work Benefits: Modeling the Impact of the Working Families Tax Credit in the UK”, University of York.
- [20] Duncan, A. and M. Harris (2002): "Intransigencies in the Labour Supply Choice", Melbourne Institute working paper.
- [21] Duncan, A. (2002): “Work incentives and labor supply of lone parents: Employment credits in Australia”, Melbourne Institute working paper.

- [22] Euwals, R. and A. van Soest (1996): “Desired and actual labor supply of unmarried men and women in the Netherlands”, *Labour Economics*, 6, 95-116
- [23] Flood, L., J. Hansen & R. Wahlberg (2003): “Household Labour Supply and Welfare Participation in Sweden”, IZA discussion paper 769.
- [24] Fortin, B. and G. Lacroix (1997): “A Test of the Neo-Classical and Collective Models of Household labor Supply”, *Economic Journal*, 107, 933-955.
- [25] Gerfin, M. and R. Leu (2003): “The impact of in-work benefits on poverty and household labor supply: a simulation study for Switzerland”, IZA discussion paper 762.
- [26] Godino R., R. Castel, M. Jalmain and T. Picketty (1999): “Pour la création d’une allocation compensatrice de revenu”, *Note de la Fondation Saint Simon*, 104, 7-20.
- [27] Gong, X. and A. van Soest (2002): “Family Structure and female Labor Supply in Mexico City”, *Journal of Human Resources*, 37(1), 163-191
- [28] Hausman, J.A. (1981): “Labor Supply”, in *How Taxes Affect Economic Behavior*, ed.H. Aaron and J.Pechman, The Brookings Institution, Washington D.C., 27-72.
- [29] Hausman, J.A. and P. Ruud (1984): “Family labor supply with taxes”, *American Economic Review*, 74, 242-248.
- [30] Heim, B.T. and B.D. Meyer (2004): “Work Costs and Nonconvex Preferences in the Estimation of Labor Supply Models”, *Journal of Public Economics*, 88, 2323-2338.
- [31] Hoynes, H. (1996): “Welfare Transfers in Two-Parent Families: Labor Supply and Welfare Participation Under AFDC-UP”, *Econometrica*, 64, 295-332.
- [32] Ilmakunnas, S. and S. Pudney (1990): “A model of female labor supply in the presence of hours restrictions”, *Journal of Public Economics*, 41, 183-210.
- [33] Kalb, G. and R. Scutella (2003): “Wage and Employment Rates in New Zealand from 1991 to 2001”, Treasury Working Paper Series 03/13
- [34] Keane, M. and R. Moffit (1998): “A Structural Model of Multiple Welfare Program Participation and Labor Supply”, *International Economic Review*, 39, 553-589.
- [35] Laroque, G. and B. Salanié (2002): “Labor Market Institutions and Employment in France”, *Journal of Applied Econometrics*, 17(1), 25-48.
- [36] Labeaga, J.M., X. Oliver and A. Spadaro (2005): “Discrete choice models of labour supply, behavioural microsimulation and the Spanish tax reform”, PSE Working Papers 2005-13.
- [37] Laisney, F. (2002, ed.): “Welfare Analysis of Fiscal and Social Security Reforms in Europe: Does the Representation of the Family Decision Process Matter?” *Final Report on EU-project VS/2000/0778*, Mannheim, ZEW.

- [38] Lundberg, S., R. Pollak and T. Wales (1997): "Do husbands and wives pool their resources? Evidence from the UK child benefit", *Journal of Human Resources*, 32, 463-480.
- [39] MaCurdy, T., D. Green and H. Paarsch (1990): "Assessing Empirical Approaches for Analysing Taxes and Labor Supply", *Journal of Human Resources*, 25, 415-490.
- [40] Nyffeler, R. (2005): "Different Modeling Strategies for Discrete Choice Models of Female Labour Supply: Estimates for Switzerland", discussion paper, University of Bern.
- [41] Orsini, K. (2007): "Tax benefit reforms and the labor market: Belgium in international perspective", Leuven University, working paper.
- [42] Steiner, V. and K. Wrohlich (2004): "Household Taxation, Income Splitting and Labor Supply Incentives - A Microsimulation Study for Germany", *CESifo Economic Studies*, 50, 3, 541-568.
- [43] Train, K. (2003). *Discrete Choice Models using Simulation*. Cambridge, UK: Cambridge University Press.
- [44] van Soest, A. (1995): "Structural Models of Family Labor Supply: a Discrete Choice Approach", *Journal of human Resources*, 30, 63-88.
- [45] van Soest, A., M. Das and X. Gong (2002): "A structural labor supply model with non-parametric preferences", *Journal of Econometrics*, 107 (1-2), 345-374.