

IZA DP No. 1455

On Modeling Household Labor Supply with Taxation

Olivier Bargain

January 2005

On Modeling Household Labor Supply with Taxation

Olivier Bargain

*IZA Bonn
and DELTA, Paris*

Discussion Paper No. 1455
January 2005

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
Email: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

On Modeling Household Labor Supply with Taxation*

Discrete-choice models provide a simple way of representing utility-maximizing labor supply decisions in the presence of highly nonlinear and possibly non-convex budget constraints. Thus, it is not surprising that they are so extensively used for ex-ante evaluation of tax-benefit reforms. The question asked in this paper is whether it is possible and desirable to get still more flexibility by relaxing some of the usual constraints imposed on household preferences and rationality. We first suggest a model which attains flexibility by making parameters vary freely across hours choices. By embedding the traditional structural approach in this specification, it is shown that the restrictions on underlying well-behaved leisure-consumption preferences are rejected. More fundamentally still, the standard approach, i.e., the assumption of unitary households optimizing statically, is strongly rejected when tested against a general model with price- and income-dependent preferences. In a static environment, the result boils down to a rejection of the unitary model. Interestingly, restrictions from both structural and standard models also imply important discrepancies in estimated elasticities and simulated predictions of responses to a tax reform. In particular, large differences appear between standard models and the general model which possibly encompasses several interpretations including dynamic aspects and intrahousehold negotiation. These findings illustrate the difficulty to conduct policy analysis in a way which reconciles the best explanatory power and a framework consistent with economic theory. The general model we suggest may provide future research with an interesting setting to test some of the dimensions of household behavior.

JEL Classification: C25, C52, H31, J22

Keywords: multinomial logit, household labor supply, taxation, microsimulation, unitary model, collective model

Olivier Bargain
IZA Bonn
P.O. Box 7240
53072 Bonn
Germany
Email: bargain@iza.org

* I am grateful above all to François Bourguignon for initiating this project and to François Laisney for an in-depth review of the paper. I would also like to thank Dan Anderberg, Denis Bolduc, Holger Bonin, Martin Browning, Pierre-André Chiappori, Olivier Donni, Alan Duncan, Bernard Fortin, Mario Fortin, Thierry Magnac, Nicolas Moreau, Thomas Picketty, Alain Trannoy and Jonathan Wadsworth for useful comments and advice. This paper has been presented at RHUL (London), DELTA (Paris), CAM (Copenhaguen), IZA (Bonn), ULB (Brussels), Laval, Sherbrooke and UQAM (Canada). All errors or omissions remain mine.

1 Introduction

The understanding of labor supply behavior continues to attract considerable research interest, the main motivation for it being the recurring importance placed on labor supply responses to tax and benefit reforms.¹ The recent literature on tax policy analysis relies heavily on discrete choice modeling that permits escaping the computational and analytical difficulties linked to utility maximization in a continuous setting. Discrete models still rely on an explicit parameterization of consumption-leisure preferences. Yet, reducing maximization to choosing the optimal alternative among a discrete set of possibilities considerably simplifies the problem and allows to rather easily account for the nonconvexities implied by actual tax-benefit systems or by fixed costs of work.² In addition, it simultaneously explains participation decision and choice in work hours and enables joint labor supply decisions of spouses to be dealt with in a straightforward way. Under these conditions, the success of this discrete approach among labor supply modelers came as no surprise.³

However, we argue in this paper that the discrete choice framework and its direct utility maximizing interpretation allow a very general representation of household preferences which is rarely exploited in current applications. Indeed, such a framework makes the explicit reference to ‘well-behaved’ preferences or the use of restrictive functional forms unnecessary. Recall that a fundamental critique of Hausman’s continuous approach was precisely the imposition of behavioral restrictions – the model required the global satisfaction of Slutsky conditions by the labor supply function (see MaCurdy et al., 1990, and MaCurdy, 1992). More fundamentally still, most policy analyses rely on the assumption of unitary households optimizing statically – what we shall refer to as the *standard* approach – whose validity can be seriously questioned.

In the present paper, we suggest a series of nested discrete-choice models which relax step by step the restrictions usually imposed on household preferences and rationality. All these models maintain a strict utility maximizing interpretation of observed and simulated choices so that positive and normative analyses can be performed in the usual way. The first suggestion consists precisely of broadening considerably the functional form of the structural model, without requiring that the implied utility function satisfies specific properties. This is achieved simply by allowing the utility associated with the various labor supply choices to depend on disposable income in a way that is totally independent across alternatives. The structural model thus appears as a particular case of this ‘unconstrained’ model and the corresponding set of constraints is easily identified. Using the quadratic specification from Blundell et al. (2000), we find that these restrictions are clearly rejected by French data.

The only assumption in the unconstrained model is that the utility of a particular labor supply alternative depends solely on the disposable income that it generates, along with household characteristics. In this discrete framework, wage rates and non-labor income influence labor choices only through disposable income. This standard approach can be seen as some kind of minimal form of the ‘unitary’ hypothesis about household behavior in a static environment. Then, the second generalization considered in this paper consists of allowing the utility of each alternative to depend on disposable income as well as on

¹Blundell and MaCurdy (2000) present a detailed state of the art.

²The traditional continuous approach presented in Hausman (1981) is usually restricted to the case of piecewise linear and convex budget sets. This limitation applies equally to recent generalizations of the technique to non-parametric estimations (Blomquist and Newey, 2002). To account for nonconvexities, as in Hausman (1985) and Hausman and Ruud (1984), labor supply must be specified parametrically together with the corresponding direct utility function, which implies rather restrictive forms for preferences. See the discussion in Van Soest and Das (2000).

³See Van Soest (1995), Hoynes (1996), Bingley and Walker (1997), Keane and Moffitt (1998) or Blundell et al. (2000) among others.

wage rates and non-labor income, as if, in some sense, preferences were price- and income-dependent. This ‘general’ model may actually be rationalized in various ways, including dynamic, labor demand-side or intra-household bargaining aspects. Again, it turns out that the standard models considered before are unambiguously rejected by this general specification.

Finally, the three models (structural, unconstrained and general) are used to simulate the potential response to a tax-benefit reform, namely the introduction of an in-work benefit aimed at ‘making work pay’ in France. Interestingly enough, serious discrepancies are found in the predicted effects of the reform. This finding throws doubts upon the accuracy of policy evaluations based on traditional structural and standard models and encourages the study of more general models incorporating both dynamic and bargaining dimensions.

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. In Section 3, a series of likelihood ratio tests discriminates among these models. Section 4 outlines the structure of the reform and analyzes the potential labor supply responses and welfare changes. Likely discrepancies coming from the use of structural and standard models are assessed. Section 5 concludes.

2 Three Discrete-choice Models of Labor Supply with Taxation

In this section, we show that the direct utility maximizing interpretation of multinomial discrete-choice models allows a very general representation of household preferences, which leads to question the generality of the standard structural models usually used for tax policy analysis.

2.1 Multinomial Logit with Unobserved Heterogeneity

The representation of discrete choices through a random utility model is the basis of the present paper. If household i is offered to choose one among J alternatives, it is assumed that the utility it may derive from alternative j ($= 1, \dots, J$) is given by:

$$V_{ij} = U(X_{ij}; Z_i, \theta_j) + \epsilon_{ij}. \quad (1)$$

In that expression, U stands for the expected utility of alternative j and is conditional on a vector θ_j of preference parameters. It depends on a vector Z_i of socio-demographic characteristics, accounting for observed heterogeneity across households, and a vector X_{ij} of variables which are specific to alternative j and possibly to household i . The actual utility derived from alternative j for household i , V_{ij} , also includes an error term ϵ_{ij} which is assumed to be identically and independently distributed across alternatives and agents. This residual cannot be interpreted as reflecting random preferences due to unobserved family characteristics. Otherwise, error terms would be correlated across alternatives (see Van Soest, 1995, and Ben-Akiva and Lerman, 1985). It is better to think of this term as describing observational errors, or possibly optimization errors (as in Hausman, 1985b) or transitory departures from best choice by agents. Under these conditions, the set of parameters $(\theta_j, j = 1, \dots, J)$ represents the common household preferences across alternatives.

Throughout the paper, we focus on the conditional/multinomial logit specification (MNL hereafter) for reasons of tractability. In this model, error terms ϵ_{ij} are supposed to follow a I-extreme value distribution. Under this assumption, McFadden (1973) has proved that the probability that alternative k is chosen by

household i is given by:

$$P_{ik} = \Pr(V_{ik} \geq V_{ij}, \forall j = 1, \dots, J) = \frac{\exp U(X_{ik}; Z_i, \theta_k)}{\sum_{j=1}^J \exp U(X_{ij}; Z_i, \theta_j)}.$$

The likelihood for a sample of observed choices can be derived from that expression as a function of the set of parameters $(\theta_j, j = 1, \dots, J)$. Estimates of these parameters may be obtained by maximum likelihood techniques.

The preceding framework can be applied to describe the choice of working hours of an individual or a couple. This problem is discretized in the sense that the choice of hours is supposed to be made between few alternatives. The idea is simply that there generally are commonly agreed durations of work in the labor market – full-time, 3/4 of full-time, half-time, etc. – so that employees have indeed a limited finite set of options, including the possibility not to work at all. This is relatively realistic and particularly appropriate in the case of France where social and institutional norms as well as demand-side rigidities are strong and imply concentrations around a limited number of hours choices. The next section shall provide a more in-depth discussion on the discretization we use.

Thus, the set of alternatives $(j = 1, \dots, J)$ now corresponds to J work durations or to J combinations of spouses' labor supplies in the case of joint decisions by a couple. As often done in the literature and for empirical evidence discussed in the Appendices, we assume here that husbands' working hours are fixed at observed levels and focus on the labor supply of women whose husbands work.⁴ In the following, we denote H_j the number of hours worked by the wife for the discrete alternative j and assume that the first alternative corresponds to non-participation ($H_1 = 0$).

The preceding model exhibits the property of Independence of Irrelevant Alternatives (IIA). This unpleasant property may be avoided by introducing some unobserved heterogeneity in preferences. In practice, some of the parameters are assumed to be randomly distributed to account for unobserved heterogeneity (see Van Soest, 1995). This technique is extensively described in McFadden and Train (2000) and often referred to as 'mixed logit' or 'random parameter logit'. Another simple way of introducing heterogeneity is suggested by Heckman and Singer (1984) and applied to labor supply estimation by Hoynes (1996). Instead of having common preferences represented by a unique set of parameters $(\theta_j, j = 1, \dots, J)$, each household i is assumed to have preferences drawn within a discrete set $(\theta_j^r, j = 1, \dots, J, r = 1, \dots, R)$. Each regime of preferences $r = 1, \dots, R$ is present with a probability π_r and both the sets of parameters and their associated probabilities are estimated in the model. Without any information on the way individual preferences are drawn, the probability that household i chooses alternative k is now given by:

$$P_{ik} = \sum_{r=1}^R \pi_r \frac{\exp U(X_{ik}; Z_i, \theta_k^r)}{\sum_{j=1}^J \exp U(X_{ij}; Z_i, \theta_j^r)}.$$

It can be checked from that expression that the IIA property is not satisfied by the model. Of course, it is not necessary to differentiate all the parameters θ_j^r across sets of preference regimes r . Practically, differentiating the parameters associated with some specific characteristics of the household or the alternative is sufficient. Notice that the non-parametric approach of Heckman and Singer avoids making extra assumptions on the distribution of heterogenous preferences as is done in mixed logits.

2.2 'Structural' vs. 'Unconstrained' Discrete-choice Models of Labor Supply

Usual discrete-choice models are fully structural in the sense that they completely identify consumption-leisure preferences. This way, and with the definitions above, the typical structural model of labor supply

⁴See Killingsworth and Heckman (1986) specifically on female labor supply.

is given by:

$$\begin{aligned} U(X_{ij}; Z_i, \theta_j) &= W(H_j, C_{ij}; Z_i, \theta) \\ \text{with } C_{ij} &= D(w_i H_j, y_i^m, y_i^K, Z_i). \end{aligned} \tag{SC}$$

This model is referred to as (SC) hereafter, which stands for ‘Structural model based on Consumption’. In this setting, $W(H_j, C_{ij}; Z_i, \theta)$ is a conventional utility function representing work (H_j) vs. consumption (C_{ij}) preferences for household i and choice j , conditionally on a vector Z_i of demographic characteristics.

In the present static framework, consumption is given by disposable income, which may be expressed itself as a function D , the arguments of which are again some of the socio-demographic characteristics of household i as well as the various sources of gross income. These are the wife’s labor income $w_i H_j$, with w_i her wage rate, the husband’s labor income y_i^m and the non-earned income of the household, y_i^K . Precisely, D represents the way in which the tax-benefit system transforms gross income into disposable income. This function actually stands for a fairly complex set of rules for the computation of taxes and benefits. Costs of work, including fixed costs or variable costs (e.g. childcare costs), may be taken out of total disposable income.

Our main points can be made at this stage. Firstly, a given functional form is to be chosen for W , in which the set of parameters that describe preferences, θ , is not choice-specific. In view of the general specification (1) of the MNL model with θ_j preference terms, this introduces serious parameter restrictions across alternatives. Most empirical analyses are based on quadratic functions of consumption and leisure (or quadratic translog), as in Van Soest (1995), Van Soest and Das (2000) or Blundell et al. (2000). It is worth testing how restrictive these representations of preferences can be.

Secondly, in order to comply with the usual properties required for well-behaved preferences, some constraints are usually added to the preceding framework, namely monotonicity and quasi-concavity of W with respect to hours of work and consumption. Recall, however, that the fundamental critique of MaCurdy concerning the continuous approach was precisely the imposition of behavioral restrictions, namely the global satisfaction of Slutsky conditions by the labor supply function.⁵ The discrete-choice approach consists of utility maximization over a finite budget and does not require tangency conditions, so that regularity conditions could be relaxed.

Overall, the structural model (SC) seems to impose severe constraints on the parameters that describe how the utility of an alternative depends on its attributes and socio-demographic characteristics. It is fair to say, however, that in practice in related studies, quasi-concavity of the utility function is most often relaxed and simply checked *a posteriori*, thus avoiding the MaCurdy critique that elasticities are largely determined *a priori*. In these conditions, our demonstration will mainly focus on the first argument, i.e. on the restrictiveness of functional forms at use.

For this purpose, we introduce a specification which appears more in line with the general formulation (1) of the MNL applied to discrete choice of working hours, that is:

$$\begin{aligned} U(X_{ij}; Z_i, \theta_j) &= W(C_{ij}; Z_i, \theta_j) \\ \text{with } C_{ij} &= D(w_i H_j, y_i^m, y_i^K, Z_i). \end{aligned} \tag{UC}$$

⁵Heim and Meyer (2002) have recently reinterpreted the MaCurdy critique by saying that this restriction amounts to constraining the estimated parameters to be consistent with the maximization of globally convex preferences. While optimizing behavior is an absolute necessity, nothing precludes from relaxing the convexity assumption. Heim and Meyer (2004) actually justify non-convexities. An obvious reason why preferences may appear to be non-convex is that work costs are not explicitly incorporated in the models. However, Heim and Meyer also emphasize that true preferences might themselves be non-convex.

This model, denoted (UC) for ‘Unconstrained model based on Consumption’, echoes to the criticism voiced above. Firstly, preference parameters are alternative-specific so that the utility can depend on consumption in a fully flexible way across hours of work. Despite its generality, the model relies on a utility-maximizing interpretation and therefore, allows for standard welfare analyses of tax-benefit reforms. Secondly, regularity conditions on leisure can be relaxed. Positive monotonicity in consumption seems to be the only minimum requirements to allow economic interpretation and meaningful policy simulations, as discussed in the next Section.

An alternative way to gain flexibility is to choose a very flexible form for the utility function in (SC). To our knowledge, this is done only in Van Soest et al. (2002), where the authors attain non-parametric flexibility by approximating the direct utility function with series expansions up to the fifth order. In our setting, flexibility is obtained directly, in a way which makes the underlying indifference curves discrete rather than continuous, which is after all consistent with a framework where the budget set itself is discrete.

2.3 ‘Standard’ vs. ‘Non-standard’ Models of Labor Supply

Along with household characteristics, total disposable income is the sole argument of the objective function in (UC). This limited structure still supposes pooling and a static framework. In other words, the fact that wage rates and non-labor income influence labor supply only through disposable income implies that the model (UC), such as model (SC), is of the ‘common preferences’ or ‘unitary’ type and completely ignores dynamic and demand-side aspects. Both (UC) and (SC) are referred to as *standard* models below.

A non-standard model, noted (GC), may be written as follows:

$$\begin{aligned} U(X_{ij}; Z_i, \theta_j) &= W(C_{ij}, w_i, y_i^m, y_i^K; Z_i, \xi_j) \\ \text{with } C_{ij} &= D(w_i H_j, y_i^m, y_i^K, Z_i), \end{aligned} \tag{GC}$$

where expected utility is conditional on a vector ξ_j of preference parameters. The set of explanatory variables has been extended so that alternative j now depends on disposable income as well as on the female wage rate, male earnings and non-labor income, as if, in some sense, preferences were price- and income-dependent. Econometric identification of this model and of the previous ones is discussed in the next sub-section; we primarily discuss the economic interpretation of (GC).

The generalization in model (GC), namely the fact that preferences are wage- and income-dependent, can be rationalized in several ways. Firstly, this can be related to the collective approach of Chiappori (1988, 1992). Collective models account for several decision-makers with specific preferences entering a (non-specified) cooperative negotiation. Assuming only efficiency, this setting encompasses all models of intrahousehold cooperative bargaining and, naturally, the unitary model. Let us drop the subscript i and the vector Z_i to simplify notation and denote c^f and c^m the private consumption expenditures of the wife and the husband respectively. Denote $u^s(H_j, c^f, c^m)$ the utility functions of spouse $s = f, m$ conditionally on the wife working H_j hours a week (we have supposed that the husband work full time exogenously). Under some minimal conditions (see Chiappori, 1992), the collective household decision problem corresponds to the maximization of a household ‘welfare’ index:

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

where the Pareto weight λ can be seen as the ‘power index’ of the wife in the couple. In a collective setting, this weight specifically depends on wage rates and unearned income, together with possible

distribution factors which may influence negotiation between spouses. As before, C_j is the disposable income associated with market income for choice H_j . This problem may be solved in two steps. Take j as fixed and maximize with respect to individual consumptions. The solution can be denoted by $c^f = c^f(H_j, C_j, \lambda)$ and $c^m = c^m(H_j, C_j, \lambda)$ and substituted in the household ‘welfare’ index to yield:

$$\begin{aligned} U_j &= \lambda u^f(H_j, c^f(H_j, C_j, \lambda), c^m(H_j, C_j, \lambda)) + (1 - \lambda) u^m(H_j, c^f(H_j, C_j, \lambda), c^m(H_j, C_j, \lambda)) \\ &= W(C_j, H_j^f, \lambda) \end{aligned}$$

Then, this index, expressed conditionally on j , can be seen as the objective function in a MNL, insofar as all the determinants can be identified. If we adopt the flexible approach introduced in this paper and recall that Pareto weights depend on wage rates and unearned income, the close form of this index can be rewritten directly as:

$$U_j = W_j(C_j, \lambda(w, y^m, y^K)). \quad (2)$$

Interestingly, this index boils down to the expression (GC), with possibly additional restrictions on the way wage rates and incomes enter function W . Whether these restrictions allow a test of the collective approach is an open question left for future research.

Secondly, model (GC) can also be made consistent with utility maximization in a life cycle framework with intertemporal separability. Let us suggest a basic illustration of this point, using excessively simplified notations in a continuous set-up. Ignore interest rates and initial wealth. Denote S_t the net dissavings decision at period $t = 1, \dots, T$, $D_t(w_t H_t, y_t)$ the disposable income obtained for female working hour H_t and male earnings y_t , and W_t the corresponding level of utility. We can resort to the two-stage budgeting formulation. A first stage determines the allocation of wealth across periods, while the second stage corresponds to the resolution of the within period programs. The solution to this problem can be found by reversing the two stages. First, the household maximizes its utility $W_t(H_t, C_t)$ at each period t , subject to the budget constraint $C_t = D_t(w_t H_t, y_t) + S_t$, given a certain level of dissavings S_t . This yields an indirect utility function $V_t(S_t, w_t, y_t, D_t)$ for each period. Then, it is possible to insert V_t 's into the intertemporally separable utility function which is maximized to choose the optimal S_t 's. As a result, dissavings decisions can be expressed as follows:

$$S_t = S_t(w_1, \dots, w_T, y_1, \dots, y_T, D_1, \dots, D_T) \quad \text{for } t = 1, \dots, T,$$

i.e., in function of wage rates, incomes and tax-benefit systems at all periods. This way, the problem at current period $t = 1$ can be rewritten as the maximization of the welfare function:

$$W_1(H_1, D_1(w_1 H_1, y_1) + S_1(w_1, \dots, w_T, y_1, \dots, y_T, D_1, \dots, D_T)).$$

or expressed in terms of our discrete approach as:

$$W_{1j}(D_1(w_1 H_{j1} + y_1) + S_1(w_1, \dots, w_T, y_1, \dots, y_T, D_1, \dots, D_T)).$$

Future wage rates are likely to depend on the current wage w_1 so that W_{1j} is a function of w_1 not only through the current level of disposable income D_1 but also separately and in a complex way through S_1 . The same may apply for male income y_1 . In these conditions, the model above is close to (GC), further imposing additive separability and including future tax-benefit systems via functions D_t . The life-cycle wage profiles are naturally not modelled in (GC) but the flexible specification suggested below, varying

with some individual characteristics, can serve as an approximation. We shall not push the comparison any further.⁶

Thirdly, the presence of wage rates in the objective function can also reflect the way workers are constrained by demand side rigidities. In France, it is a well-known fact that the least productive women are often rationed to part-time jobs (see Laroque and Salanié, 2002). It may also be the case, for instance, that career-oriented executives may have to work extra-hours beyond the institutional 39 hours per week. In (GC), institutional or demand-side rigidities differing in skill level can then be captured somehow by the introduction of the wage rate in the model. In that case, the welfare function is not interpretable as such and can only be seen as a probability to choose a particular working hour, conditionally on household characteristics.⁷

The list may not be exhaustive and we restrain from discriminating amongst the various possible interpretations embedded in model (GC). We do so for two reasons. Firstly, this would be a giant task since each of these interpretations deals with a whole specific branch of the literature. Secondly, model (GC) might well be used to set up theoretical models and to derive related testable necessary conditions from it.⁸ Yet, testing these conditions on observed behaviors would require restricting the model on other accounts.⁹ Instead, we intentionally keep model (GC) as general as possible, without restricting to a particular interpretation, in order to examine the implications – on explanatory power and policy analysis – of a model which potentially captures various dimensions simultaneously. The model is not structural but guarantees the basic requirement for economic rationality, i.e. it is grounded on utility-maximizing behaviors under budget constraint.

2.4 Econometric Specifications

2.4.1 Structural Model

We specify the constrained structural model (SC) in the spirit of Blundell et al. (2000). The functional form is quadratic and the model is refined by the introduction of costs of work F_{ij} , taken out of disposable income, and some coefficients depend linearly on socio-demographic characteristics Z_i .¹⁰ For choice $j = 1, \dots, J$, the structural model is then written:

$$U_{ij} = \alpha_{cc}(C_{ij} - F_{ij})^2 + \alpha_{hh}(H_j)^2 + \alpha_{ch}(C_{ij} - F_{ij})H_j + \alpha_{cir}(C_{ij} - F_{ij}) + \alpha_{hi}H_j, \quad (\text{SC}')$$

with heterogeneity:

$$\begin{aligned} \alpha_{cir} &= \alpha_{cr}^0 + \alpha'_c Z_i \\ \alpha_{hi} &= \alpha_h^0 + \alpha'_h Z_i, \end{aligned}$$

⁶In fact, the model could truly be redesigned and estimated as a dynamic model if we used data on expenditures in order to identify consumption and savings (see Blundell and Walker, 1986).

⁷If this effect matters, then using model (GC) for welfare analysis becomes highly suspect. This is true to a lesser extent for the other models presented in this study since the wage rate was forced to influence labor choices only through disposable income.

⁸For instance, in the collective interpretation (2), compared to the general expression of (GC), some restrictions necessarily stem from the fact that wage rate and incomes enter the W_j function through the same λ function for all choices $j = 1, \dots, J$.

⁹Testing the collective framework, for instance, would require to assume a static and purely supply-side setting. This assumption is systematically made in the literature on collective model of labor supply, as in Chiappori (1998, 1992) and Chiappori et al. (2002). In the same way, note that a rejection of (UC) vs (GC) would be a rejection of the ‘standard’ approach, but to interpret it as a rejection of the unitary approach would require additionally to assume a static and supply-side environment.

¹⁰Notice that log terms are not allowed here as the introduction of costs of work may lead to negative net-of-cost income.

and vectors $\alpha'_c = (\alpha_c^1, \dots, \alpha_c^L)$, $\alpha'_h = (\alpha_h^1, \dots, \alpha_h^L)$. 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. In addition, we allow for unobserved heterogeneity. Each preference regime r corresponds to a mass point α_{cr}^0 placed on the coefficient of disposable income.¹¹

Costs F_{ij} are to be paid if the wife starts to work. They vary with four household characteristics and are assumed not stochastic as in Blundell et al. (2000). In alternative j , we then have:

$$\begin{aligned} F_{i1} &= 0 \\ F_{ij} &= f^0 + f^1 Paris + f_j^2 Child02 + f^3 Child35 + f^4 Child611. \end{aligned} \quad (3)$$

The term related to the number of small children is allowed to vary freely with the alternative. This way, we account not only for the fixed cost of work but also for the variability introduced by Blundell et al. (2000) through childcare costs.¹²

We suggest two variants of this model. We first introduce model (SY), similar to (SC) but where disposable income C_{ij} is replaced by gross income $Y_{ij} = w_i H_j + y_i^m + y_i^K$. To investigate how recent studies render structural models more flexible, we suggest a variant (SC2) for which the basic cost of work is made variable across alternatives, hence for $j \neq 1$:

$$F_{ij} = f_j^0 + f^1 Paris + f_j^2 Child02 + f^3 Child35 + f^4 Child611. \quad (4)$$

The implications of choosing model (SC2) rather than (SC) are discussed in the next Section.

2.4.2 Unconstrained Model

The flexible/unconstrained model (UC) is naturally made comparable to (SC) by use of the following quadratic form:

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

which actually nests model (SC). This is shown in the Appendices where we provide an identification of the constraints imposed by model (SC) on (UC). Heterogeneity is written as:

$$\begin{aligned} b_{ijr} &= b_{jr}^0 + b'_j Z_i \\ c_{ij} &= c_j^0 + c'_j Z_i + \sum_{l=1}^L \sum_{m=l}^L c_j^{lm} z_i^l z_i^m, \end{aligned}$$

¹¹This choice is made by convenience, keeping in mind that models (SC) and (UC) must be rendered comparable and nested while only terms in C remains in (UC).

¹²Blundell et al. (2000) assume a deterministic relationship between hours of childcare and hours of work, estimated on the sub-sample of households buying childcare. This aspect is crucial in the UK considering the high costs of market childcare and the fact that Blundell et al. analyze the childcare credit component of the WFTC. It is less the case in France and data on childcare costs is of poor quality anyway. We have attempted to proceed as in Blundell et al. but the estimation was not robust enough. The important point (with respect to the present study) is that childcare use is allowed to vary with work hours and the demographic composition of the household. Notice that in the present setting, the variable cost due to childcare varies only with the number of children between 0 and 2 in order to conform to the fact that public kindergartens are forced to accept all children above 2.

with vectors $b'_j = (b_j^1, \dots, b_j^L)$ and $c'_j = (c_j^1, \dots, c_j^L)$.¹³ For each preference regime r , unobserved heterogeneity corresponds to a set of mass points b_{jr}^0 ($j = 1, \dots, J$). Probability of choice j is written as:

$$\Pr(a_j C_{ij}^2 - a_k C_{ik}^2 + b_{ij} C_{ij} - b_{ik} C_{ik} + c_{ij} - c_{ik} > 0, \forall k \neq j).$$

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 this coefficient to zero for the first alternative ($c_{i1} = 0$). We also introduce (UY), a model similar to (UC) where disposable income is replaced by gross income.

Also note that in models (SC), (SC2) and (UC), non-parametric identification of parameters for Z_i variables is not obtained since most of these variables are both taste shifters and determinants of the tax-benefit rules in function D . This is a usual feature of static structural models with taxation. As ever so often, parametric identification relies on the strong nonlinearities coming from the tax-benefit system.

2.4.3 General Model

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

$$\begin{aligned} U_j = & \alpha_j^{cc} C_j^2 + \alpha_{ijr}^c C_j + \alpha_j^{ff} (w^f)^2 + \alpha_j^{mm} (y^m)^2 & \text{(GC')} \\ & + \alpha_j^{KK} (y^K)^2 + \alpha_j^{fm} w^f y^m + \alpha_j^{fK} w^f y^K + \alpha_j^{mK} y^m y^K \\ & + \alpha_{ij}^f w^f + \alpha_{ij}^m y^m + \alpha_{ij}^K y^K + \alpha_j^{fc} w^f C_{ij} + \alpha_j^{mc} y^m C_{ij} + \alpha_j^{Kc} y^K C_{ij} + c_{ij}, \end{aligned}$$

with:

$$\begin{aligned} \alpha_{ijr}^c &= \alpha_{jr}^c + \alpha_j^{c'} Z_i \\ \alpha_{ij}^h &= \alpha_j^{0h} + \alpha_j^{h'} Z_i \quad \text{for } h = f, m, K \\ c_{ij} &= c_j^0 + c_j' Z_i + \sum_{k=1}^L \sum_{l=k}^L c_j^{kl} z_i^k z_i^l. \end{aligned}$$

It can be seen at first glance that this specification nests model (UC). In (GC) the usual indeterminacy for determinants which are not alternative-specific is removed by setting these coefficients to zero for the first alternative. A model (G) is also introduced, specified as (GC) with gross income instead of disposable income. Since total gross income is colinear to the three sources of income w_i , y_i^m and y_i^K , it must be taken out, hence the terminology (G) rather than (GY).

Here too, non-parametric identification is not obtained. Parametric identification relies on the strong nonlinearities due to the tax-benefit system in function D . This point requires additional explanation, in particular with regard to the identification of the parameters for the wage rate w_i . Let us locally linearize the budget constraint of household i so that it can be written:

$$C_{ij} = (1 - t_i(w_i H_j)) w_i H_j + \tilde{y}_i,$$

with $t_i(w_i H_j)$ the effective marginal tax rate on female earnings and \tilde{y}_i the implicit residual income (the intercept between the vertical axis and the tangent to the budget curve at the point under consideration). In (GC), then, the hourly wage rate appears twice, once as a gross wage w_i and once as an implicit net wage $(1 - t_i(w_i H_j)) w_i$. For a given work duration, one could argue that the same implicit taxation applies

¹³In the double sum, the 4 coefficients corresponding to the square of the dummy for the Paris region are set to zero.

to all females with the same gross wage rate. This is not the case, however, since for a given level $w_i H_j$, the effective marginal tax rate t_i depends in a complex way on household characteristics, and in particular on household composition and marital status.¹⁴ Moreover, the location on the budget curve – hence the slope faced by the household – depends on the level of capital income and husband’s earnings.¹⁵

3 Empirical Results

3.1 Estimations and Fit of the Different Models

The data at use, the selection process and the procedure of discretization are described in the Appendices. Using the specifications presented in the previous section, models (SC), (UC) and (GC) as well as the variants are estimated by maximum likelihood. A classic explanation of the estimation results for the standard models is given in the Appendices.

A usual approach to measure goodness of fit in a multinomial setting is to compare, for all discrete choices, the observed frequency with its average estimated value over all households. Table 1 shows that the probabilities predicted by all models but (SC) represent correctly the proportions of the sample. The figure also displays generalized R2 for each choice, i.e., the percentage of observed variance explained by the model. Results seem reasonable for (GC) but standard models perform very poorly in predicting part-time jobs.¹⁶ Since part-time activity is mainly imposed in France, this result conveys the idea that the general specification of model (GC), notably the presence of the female wage rate, captures some demand-side aspects.¹⁷

Another indicator of fit presented in Table 1 is the pseudo-R2 or Likelihood Ratio Index of McFadden (1973). This is simply a measure in $[0, 1]$ of the distance between the maximized value of the log-likelihood and the log-likelihood when all parameters are set to zero. This is a convenient indicator as it summarizes the fit in one figure, which allows a ranking of the different models. This is a different way to read the information about the Likelihood Ratios used in the tests which follow. It comes as no surprise that the general model provides a better fit than standard models while the standard models with additional flexibility (UC and SC2) clearly dominate model (SC).

3.2 Testing the Constraints on Preferences

It is important to note that the first series of tests presented in 3.3 and 3.4 ignores the unobserved heterogeneity and the imposition of C-monotonicity and quasiconcavity on (UC) and (GC). Estimation results enable us to compute a series of likelihood ratio statistics in order to test the constraints on preferences, i.e., structural vs unconstrained models. Table 2 displays the log-likelihood values for the models (SC),

¹⁴Cohabiting couples are taxed independently while married ones are jointly taxed. The number and age of children condition a whole set of family benefits, including household benefits, universal child benefit and several means-tested child benefits. Many other characteristics come into play, e.g. the place of residence and housing characteristics for housing benefits. Also note that these tax and benefit rules do not only imply nonlinearities of the budget set but possible nonconvexities and discontinuities.

¹⁵In complement, note that the implicit wage varies with the alternative while the gross wage is some sort of fixed effect across choices.

¹⁶The breakdown of predicted frequencies per hours category, available upon request, confirm these findings.

¹⁷As pointed out by Bourguignon and Magnac (1990) about the French labor market, serious attention should indeed be paid to possible constraints coming from the labor demand-side. See the discussion in the Appendices.

Table 1: Fit of the Models

choice	observed frequencies	SC		SC2		UC		GC	
		average predicted frequencies	generalized R2						
0	0,240	0,245	0,107	0,245	0,116	0,246	0,162	0,246	0,211
20	0,126	0,096	0,013	0,126	0,015	0,125	0,027	0,125	0,074
30	0,109	0,224	0,034	0,109	0,006	0,109	0,025	0,109	0,074
39	0,447	0,284	0,049	0,440	0,076	0,444	0,105	0,443	0,197
45	0,078	0,151	0,095	0,079	0,042	0,077	0,051	0,077	0,080
pseudo-R2		10,1%		18,4%		21,17%		25,4%	

(SC2) and (UC). It also provides the LR statistics and the critical values at the 1% significance level for the corresponding number of constraints.

In addition to the previous tests, we suggest a test comparable to (SC) vs. (UC) where disposable income C is replaced by gross income Y in both specifications, that is, a test of (SY) vs. (UY). The rationale for this test goes as follows. In model (UC), the structure is reduced to the choice of disposable income to explain labor supply decisions, which implies that households fully account for the impact of their decisions on the taxes they (potentially) pay and the transfers they (potentially) receive. This alone is already a strong assumption on agents' rationality, and the test (SY) vs. (UY) relaxes this assumption to some extent. To some extent only since in this variant, taxation is not introduced explicitly but estimated coefficients might capture some of the impact of taxes and benefits on labor supply decisions.

In all cases the structural models are rejected at the 1% level, as proven by the three last lines of Table 2. In the third line, for instance, the LR statistic (1207) is well over the critical value (203) so that model (SC) is rejected against (UC). It turns out that the structural constraints usually imposed on preferences are far too restrictive. This finding holds true even when the flexibility of the structural model increases by addition of the alternative-specific dummies f_j^0 , as proven by the rejection of the test (SC2) vs. (UC) in the fourth line.

Note that regularity constraints in hours are *not* imposed on (SC) and (SC2). This way, the tests above focus specifically on the restrictions imposed by the functional form and by C -monotonicity and quasiconcavity (which are not imposed on UC but verified for almost all households in SC and SC2). Should the monotonicity and quasi-concavity in hours be added, the tests would be rejected *a fortiori*.

Table 2: Structural Models vs Non-constrained Model

	model	max log L	nb coeff*	vs model	df	LR	chi2 (1%)
Flexible models	UC	-4306.1	185	-	-	-	-
	UY	-4333.8	185	-	-	-	-
Structural models	SC	-4909.7	27	UC	158	1207	203
	SC2	-4458.0	30	UC	155	304	198
	SY	-4952.5	27	UY	158	1237	203

* 4 coefficients for the square of the dummy 'Paris region' have been taken out from UC

Note: "max log-L" gives the value of the maximized log likelihood for each model; "nb coeff" gives the number of coefficients of the model; "d.f." is the number of degrees of freedom of the test (i.e., the number of constraints imposed by each model on the model of reference, equals to the difference in the number of coefficients); "LR" is the likelihood ratio statistic; "chi2(1%)" gives the chi-squared value for the LR test at the 1% significance level.

3.3 Testing the Standard Approach

We now test the standard approach, that is (UC) vs. (GC). A total of 135 constraints are imposed by the former model on the latter. Table 3 displays the log-likelihood values for both models and provides the LR statistics as well as the critical value for 135 degrees of freedom (third line). We also present a similar test where disposable income C is replaced by gross income Y in the specifications, i.e., a test of (UY) vs. (G). In both cases, the standard approach is clearly rejected at the 1% level.

Following the discussion above on the nature of model (GC), it appears that this finding can be related to the recent literature dealing with tests of family models. In this literature, tests are most often performed in a static setting, assuming partial equilibrium. Under these assumptions, the exercise presented above boils down to testing the unitary model versus a model with price-dependent preferences, that is, a model which does not verify Slutsky conditions nor pooling (Pollak, 1977). Consequently, we provide here a strong rejection of the unitary approach. Contrary to previous related studies, this test seems all the more robust as it relies on a flexible representation of preferences.¹⁸

Table 3: Test of the Standard Approach

	model	max log L	nb coeff*	vs model	df	LR	chi2 (1%)
General models	<i>GC</i>	-4074.1	320	-	-	-	-
	<i>G</i>	-4134.3	260	-	-	-	-
Tests of the standard approach	<i>UC</i>	-4306.1	185	<i>GC</i>	135	464	176
	<i>UY</i>	-4333.8	185	<i>G</i>	75	399	106

* 4 coefficients for the square of the dummy 'Paris region' have been taken out from the models

3.4 Regularity Conditions, Unobserved Heterogeneity and Robustness of the Tests

Flexible specifications of (UC) and (GC) allow to capture broader heterogeneity in preferences but may also cause a larger number of households not to respect regularity conditions. Indeed, positive monotonicity in consumption is not met for 3.8% (resp. almost 23%) of the households when using model (UC) (resp. model (GC)), while it is respected in almost all households when the structural models are used. Albeit unnecessary for coherency of the econometric model, this regularity condition, along with quasi-concavity in C , seems a natural requirement to perform meaningful policy analysis. Positive monotonicity is written:

$$\begin{aligned}
 2a_j C_{ij} + b_{ijr} &\geq 0 \quad \text{for (UC)} \\
 \alpha_j^{cc} C_j^2 + \alpha_{ijr}^c + \alpha_j^{fc} w^f + \alpha_j^{mc} y^m + \alpha_j^{Kc} y^K &\geq 0 \quad \text{for (GC)}.
 \end{aligned}$$

¹⁸Usually, in related studies, the unitary approach is rejected on the basis of the rejection of income pooling or Slutsky conditions (symmetry and positive semidefiniteness of the Slutsky matrix in the case of labor supplies). In particular, according to the pooling assumption, who controls the resources in the family does not affect household decisions. Most studies, surveyed in Lundberg and Pollak (1996), check if, for a given total family income, spouses' relative earnings affect expenditure patterns. The validity of most tests can be questioned insofar as these studies assume the exogeneity of labor income whereas labor decisions are actually determined simultaneously with the consumption decisions observed for the tests (cf. Browning and Meghir, 1991). Fewer tests are available for labor supply. To my knowledge, only Schultz (1990) and Fortin and Lacroix (1997) suggest a test of the pooling of non-labor income on labor supply choices. In the latter study, a parametric model nests both the unitary and the collective settings which can be tested as special cases. Yet, the tests rely on a parameterization which seems to restrict preferences considerably.

Consequently, we repeat previous tests while imposing these conditions. Practically, this is done by adding them as constraints in the likelihood maximization. Quasi-concavity in C is checked *a posteriori* and turns out to always be respected once monotonicity is imposed. We also proceed with the tests once more, adding unobserved heterogeneity in all models in the way described in the previous section.

These two transformations of the original models and their impact on the tests are presented in Tables 5 and 6, to be compared to previous results gathered in Table 4. The main conclusion is that these transformations do not change the ranking of the models in terms of likelihood performances.

Monotonicity in consumption is already respected in models (SC) and (SC2) so that likelihoods are left unchanged from Table 4 to Table 5. On the other hand, the coefficients of models (UC) and (GC) are now forced to verify C -monotonicity and it follows that the likelihoods of these models naturally decrease, as shown in Table 5. The tests now purely concern the restrictions coming from the functional form of preferences. It can be seen that the ranking across the models is preserved and the rejection of (SC) vs. (UC) and of (UC) vs. (GC) still hold.

The incorporation of unobserved heterogeneity significantly improves the likelihoods of all models – except (GC) – in a rank-preserving way, as shown in Table 6.¹⁹ Constraints on preferences are rejected once again. Despite the fact that the likelihoods of (UC) and (GC) come closer, the test of the standard approach is still rejected.

Table 4: Summary of the LR Tests using Multinomial Logit Models

<i>model</i>	max log L	nb coeff*	<i>vs model</i>	df	LR	chi2 (1%)
SC	-4909,7	27	UC	158	1207	203
SC2	-4458,0	30	UC	155	304	198
UC	-4306,1	185	GC	135	464	176
GC	-4074,1	320	-	-	-	-

* 4 coefficients for the square of the dummy 'Paris region' have been taken out from UC and GC

Table 5: LR Tests using MNL with Monotonicity Imposed on Consumption

<i>model</i>	max log L	nb coeff*	<i>vs model</i>	df	LR	chi2 (1%)
SC	-4909,7	27	UC	153	1068	196
SC2	-4458,0	30	UC	150	164	192
UC	-4375,8	180	GC	135	477	176
GC	-4137,2	315	-	-	-	-

* 4 coefficients for the square of the dummy 'Paris region' have been taken out from UC and GC

Table 6: LR Tests using MNL with Discrete Unobserved Heterogeneity

<i>model</i>	max log L	nb coeff*	<i>vs model</i>	df	LR	chi2 (1%)
SC	-4671,5	29	UC	162	956	207
SC2	-4312,1	32	UC	159	238	203
UC	-4193,2	191	GC	135	238	176
GC	-4074,0	326	-	-	-	-

* 4 coefficients for the square of the dummy 'Paris region' have been taken out from UC and GC

¹⁹When estimating the models with unobserved heterogeneity, two mass points have appeared sufficient ($R = 2$) given the heavy dominance of one of the regimes.

It might seem that model (UC) gains flexibility compared to (SC) and (SC2) mainly owing to the fact that household characteristics Z_i are introduced in a flexible way. To investigate this point, it is desirable to do a last series of tests, once all heterogeneity is withdrawn. Denote (sc), (sc2) and (uc) the variants of previous models in that case, for $j = 1, \dots, J$:

$$U_{ij} = \alpha_{cc}(C_{ij} - F)^2 + \alpha_{hh}(H_j)^2 + \alpha_{ch}(C_{ij} - F)H_j + \alpha_c(C_{ij} - F) + \alpha_h H_j, \quad (\text{sc})$$

$$U_{ij} = \alpha_{cc}(C_{ij} - F_j)^2 + \alpha_{hh}(H_j)^2 + \alpha_{ch}(C_{ij} - F_j)H_j + \alpha_c(C_{ij} - F_j) + \alpha_h H_j, \quad (\text{sc2})$$

$$U_{ij} = a_j C_{ij}^2 + b_j C_{ij} + c_j. \quad (\text{uc})$$

It is very easy to show that for $J = 5$ alternatives, (uc) still nests (sc2) which itself nests (sc).

Compared to model (UC), the likelihood of model (uc) decreases by 297 (−6.9%). Compared to (SC2), the likelihood of (sc2) decreases only by 170 (−3.8%). This comparison reveals that the role of the taste shifters is indeed not innocuous and that the specifications previously retained may have favored (UC) compared to (SC2). On the other hand, the number of degrees of freedom is reduced considerably.²⁰ We conduct the tests as before and results are reported in Table 7. Even in this case the rejection of (sc2) against (uc) occurs, while the rejection of (sc) is extremely strong again.

Table 7: LR Tests with Basic MNL without Heterogeneity

model	max log L	nb coeff*	vs model	df	LR	chi2 (1%)
sc	-5162.8	6	uc	8	1119	20
sc2	-4628.3	9	uc	5	50	15
uc	-4603.3	14	-	-	-	-

Finally, note that the flexibility of (UC) logically increases with the number of alternatives. In the simple models above, it can be easily shown that (uc) nests structural models if the number of alternatives is large enough.²¹ With finite samples, limitations to this approach are due to curse of dimensionality and a large number of alternatives could become untractable. In practice, however, the discretization relies on a small number of alternatives driven by demand-side and institutional rigidities, as discussed in the Appendices.

3.5 Additional Flexibility in Structural Models

It is worth noticing that model (SC2) cannot be rejected at the 1% significance level against model (UC) when regularity constraints on C are imposed in the latter (Table 5). In effect, in this case, neither model imposes regularity conditions on hours, both models verify C -monotonicity and quasiconcavity and both models present rather flexible functional forms. We find then that the structural model can be rejected only for a significance level of 20% at best.

This result leads to a necessary discussion about the modeling strategy in discrete-choice models of labor supply. Often, state-specific dummies are introduced in recent studies to increase flexibility of structural models for a better fit to the data. For instance in the case of the Netherlands, Van Soest (1995)

²⁰There are 14, 9 and 6 coefficients in models (uc), (sc2) and (sc) respectively, compared to 185, 30 and 27 in (UC), (SC2) and (SC). Then, model (uc) has respectively 5 and 8 more coefficients than structural models (sc2) and (sc), while (UC) had 155 and 158 more coefficients than (SC2) and (SC) respectively

²¹Model (uc) nests (sc) if $J \geq 3$ and nests (sc2) if $J \geq 4$.

notices that part-time is overpredicted by his model at the expense of non-participation. This is solved by adding dummies for part-time, which, in a pure supply-side fashion, are interpreted as additional search costs of jobs explained by the relative lack of part-time jobs in the Netherlands. Instead, Van Soest and Das (2000) or Van Soest et al. (2002) prefer to incorporate fixed costs of work to solve the understatement in the number of non-workers.

Thus, additional variability seems to be introduced in a fairly *ad hoc* way in recent studies.²² Note, however, that a model ignoring these structural additions and relying solely on the traditional consumption-leisure specification would be severely restrictive, as proven by the (SC) vs (UC) test. Also note that once a flexible specification of preferences is introduced, as in Van Soest et al. (2002), the utility function itself could pick up the gap in the distribution at certain hours, by assigning lower utility to such hours values. Additional structure, such as fixed costs, is non-parametrically unidentified in that case.

These points are crucial and cast serious doubt on the interpretability of a structural model. Model (UC) implicitly acknowledges these limitations and remains agnostic on the different dimensions that can be brought up to explain concentrations at certain hours, while introducing the appropriate flexibility to the discrete choice approach.²³

Note that even if structural interpretations disappear in (UC), the model maintains a strict utility-maximizing representation of household decision-making so that welfare analysis can still be performed. A natural alternative to this model consists of using flexible polynomial approximation of the utility function. In Van Soest et al. (2002), the model with a second order series expansion is statically rejected against higher order models. Yet, for a finite sample size, the order of the polynomial that can be used is limited as asymptotics requires that the order tends to infinity much slower than the number of observation. In practice, however, estimates of labor supply elasticities and effects of a tax reform suggest that the results do not change much once the order of the series expansion is extended beyond two.²⁴ As we shall see, the same is not true here.

4 Simulations

In this section, we assess the discrepancies of prediction between models when gauging elasticities and the effects of a tax reform on labor supply and welfare. In the Appendices, we describe the strategy retained to compute transition frequencies after a shock on the budget constraint (an increase in wage rates to compute elasticities or a tax reform). In the sequel, we ensure that all models respect C -monotonicity and quasiconcavity.

As a preliminary remark, notice that compared to structural models, (UC) and (GC) contain many more variables, a certain number of which will be imprecisely estimated. This should make these two models more accurate (i.e. fit the data better) but less precise in their prediction.²⁵ This is not a matter

²²In the present setting, consider the constant term f_j^0 in the variable cost of work (4) of model (SC2). This term is interpreted as an average cost of work for choice j , although this state-specific coefficient could well represent disutility from job search, distaste for work, demand-side constraints etc., specific to the j th alternative.

²³The interpretability is all the more difficult as demand-side may contaminate the estimates. Again, in the French case, there are strong suspicions that our labor supply models are affected by strong demand-side constraints in addition to fixed costs or other supply-side aspects. See the discussion in the Appendices.

²⁴Again, this is likely due to the fact that all models incorporate fixed costs of work, so that less flexibility is required of the specification of preferences.

²⁵The trade-off between a good fit and reasonably precise predictions requires to select only significant covariates, which could be done easily if one actually wished to make use of these models for policy evaluation. This selection was not made

of concern since our primary objective is to assess the discrepancies between models, which can be done by checking if confidence intervals for predicted elasticities or responses to a reform overlap significantly or not. These confidence intervals are obtained for each model by bootstrapping of the parameters, using their estimated asymptotic distributions, as described in the Appendices.

As we shall see below, it turns out that (UC) is fairly precise, at least not less than (SC). The general model (GC) is slightly less precise, which seems natural, considering the argument above and the large number of covariates (320). Surprisingly, model (SC2) is as imprecise as (GC). The lack of precision is likely to result from the introduction of state-specific dummies in addition to usual covariates related to female hours.²⁶

4.1 Elasticities

We first suggest a comparison of the average wage-elasticities obtained with the various models. A closed-form expression of elasticities is obviously not available in our setting but wage-elasticities can be computed numerically by evaluating the change in average work duration and in participation rate subsequent to a uniform exogenous increase in the wage rate of all women in the sample (respectively +1 and +10%).²⁷

Notice that with the limited number of choices in the discretized set, the computed wage-elasticities on work hours may account only partly for marginal changes in hours. However, the simulated change in average work duration fully captures changes in participation, which is what fundamentally matters, as argued by Heckman (1993). In addition, empirical evidence presented in the Appendices show that in France workers cannot freely choose their working hours in a continuum while the alternatives selected in the discretization tend to reproduce the set of options actually available.

Tables 8 and 9 present the predicted elasticities of working hours and participation respectively. In each case we compute elasticities following the different approaches described in the Appendices, namely the ‘calibration method’, which aggregate individually simulated transitions, and the ‘aggregate probability method’ which simply averages predicted frequencies. Both yield similar results. Overall, it appears that the magnitude of potential responses is larger with the unconstrained model and larger still with the general model. This conclusion is robust to the sensitivity analysis, as can be seen by comparing confidence intervals. Note, however, that the structural model with additional flexibility (SC2) gives results in between those of (SC) and (UC) for participation, while elasticities of hours are not significantly different from those of (UC).

The elasticities stand in a narrow [.2, .4] range across models, which is in line with the recent labor supply literature (see Blundell and MaCurdy, 2000) and recent findings in France. Choné et al. (2003) find an average wage-elasticity of participation (resp. hours) of around 0.3 (resp. 0.26) when the censorship of the minimum wage is not accounted for. This recent study focuses on couples with children only and we might expect even lower levels of elasticity when considering all women in couples. These small values convey the idea that there may not be as much scope for large incentive effects from reforms as thought

here as we wanted the models to be nested by natural extension of the initial quadratic specification.

²⁶In the estimation results presented in the Appendices, notice that two of the three coefficients relative to hours (hours squared and hours \times consumption) are not significantly different from zero in (SC2), compared to (SC), while the state-specific dummies are not significant either.

²⁷To use models for policy analysis, economists are interested in aggregate elasticities. In many studies, yet, elasticities are computed for the average representative household. As pointed out by Van Soest and Das (2000), this is not very informative - in a highly nonlinear model like ours - on the consequence of wage changes in a heterogeneous population.

in older related studies on the participation of women in France.²⁸

Table 8: Change in Average Work Hours

<i>model</i>	elasticities computed using averaged simulated transitions		elasticities computed using average predicted frequencies	
	wage + 1%	wage + 10%	wage + 1%	wage + 10%
SC	0.20%	2.02%	0.23%	2.26%
SC2	[0.17; 0.25]	[1.7; 2.4]	0.29%	2.76%
	[0.20; 0.35]	[2.1; 3.4]		
UC	0.31%	3.04%	0.33%	3.13%
	[0.27; 0.37]	[2.7; 3.4]		
GC	0.37%	3.50%	0.40%	3.58%
	[0.32; 0.45]	[3.1; 4.0]		

Figures in brackets give a bootstrapped 90% confidence interval of the elasticity

Table 9: Change in Participation Rates

<i>model</i>	elasticities computed using averaged simulated transitions		elasticities computed using average predicted frequencies	
	wage + 1%	wage + 10%	wage + 1%	wage + 10%
SC	0.19%	1.81%	0.21%	2.06%
SC2	[0.14; 0.23]	[1.5; 2.1]	0.24%	2.36%
	[0.17; 0.27]	[1.7; 2.6]		
UC	0.26%	2.52%	0.28%	2.65%
	[0.22; 0.31]	[2.2; 2.8]		
GC	0.39%	3.67%	0.41%	3.74%
	[0.34; 0.45]	[3.2; 4.1]		

Figures in brackets give a bootstrapped 90% confidence interval of the elasticity

4.2 Labor Supply Response to a Tax-benefit Reform

4.2.1 Brief Description of the Reform

The reform we examine here is a scenario of in-work benefit that has been discussed lately in France. This Earned Income Supplement (*Allocation Compensatrice de Revenu*) has been suggested by Godino et al. (1999) and meticulously analyzed by Bargain and Terraz (2003). It is close to the spirit of the British *Working Family Tax Credit* (WFTC) as both schemes are means-tested on joint income in the household. Yet, there is no condition on minimum work duration as there is for the WFTC. The phase-out rate of 42% is computed so that the scheme exhausts when the worker reaches a level of earnings equivalent to a full-time job at minimum wage rate. In fact, the Earned Income Supplement (EIS) has been advocated to offset the disincentive effects of the minimum income (MI). It could well be thought of as an extension

²⁸For married women, Bourguignon and Magnac (1990) use Hausman's technique and obtain a wage-elasticity of 1 which appears implausibly high, as noted by the authors who find their results very sensitive to several aspects of the specification. When fixed costs are added to the model, the wage-elasticity captures only the variations in hours at the mean point and not the changes in participation for the whole sample. It then becomes extremely small (0.05), suggesting that large previous elasticities were driven mainly by changes in participation.

of the MI so that when cumulated, both benefits amount to a minimum income scheme with a 42% taper on earnings, instead of the 100% taper in the current MI system. Denote B the maximum amount of MI and Y total household earnings after social contributions. Ignoring other incomes, the formula of the EIS cumulated to the MI is thus simply:

$$EIS + MI = B - 42\%Y.$$

The maximum amount B is close to 350 euros per month for a single person and increases with the presence of a partner and with the number of dependent children.²⁹

Evaluating the financial gains to work and their change after reform, Bargain and Terraz (2003) highlight the fact that this type of in-work benefit is likely to create disincentive effects on the participation of second earners in couples, essentially the wives. Indeed, while the EIS generously tops up the income of the first earner, the means-tested condition at the family level reduces the amount of transfer as the second-earner starts to work. Therefore, the net gain for her to work shrinks with the introduction of the reform. This is a common feature to in-work benefits which are conditioned on family income, as stressed by Eissa and Hoynes (2004) in the case of the US *Earned Income Tax Credit* or Blundell et al. (2000) in the case of the WFTC.

4.2.2 Labor supply responses

Table 10 presents the predictions from model (SC) regarding the EIS simulation on French data. Each cell indicates the transition frequency and the corresponding number of households. Behavioral responses are moderately sized and can be compared to those of Blundell et al. (2000, Table 8) for the WFTC reform. For the UK these authors conclude that around 28,000 women with an employed partner would stop working (a net effect of $-20,000$). It turns out that there are strong similarities in the sign of the effects, due to resemblances between the reforms as highlighted above. The orders of magnitude are also very close, even though the comparison cannot be straightforward since WFTC is not a new reform in the UK but rather a generous extension of the previous *Family Credit*. The important point for us is that both evaluations rely on very close strategies, using structural models of the (SC) type in both cases and assuming a 100% take-up.

Table 10: Responses to the EIS

pre-reform	post-reform			marginal freq.
	out of work	part time	full time	
out of work	1 529 943 (24,0%)	0	0	24,02%
part time	9 040 (0,1%)	1 489 056 (23,4%)	0	23,52%
full time	20 217 (0,3%)	1 602 (0,0%)	3 319 597 (52,1%)	52,46%
marginal freq.	24,48%	23,40%	52,12%	6 369 455

We now proceed with the comparison of the main models introduced in the paper. No positive effect on the labor supply of married women with employed partners could be found. Consequently, we have extracted from the transitions matrices only the predictions relative to the disincentive effects of the

²⁹By simplicity for the present exercise, we can assume that the reform is made revenue-neutral by increasing indirect taxation or by reallocations in the governmental budget; either way implies no additional effect on labor supply behaviors.

reform. Average statistics as well as 90% confidence intervals are presented in Table 11. It appears that the net proportion of women who would potentially leave the labor market is approximately 0.46%, 0.52% and 0.79% according to models (SC), (SC2) and (UC) respectively. As expected from the predicted elasticities, the response is larger when using the unconstrained model. Notice that the predictions of (UC) and (SC2) are significantly different, which was not the case for hours elasticities.³⁰ When it comes to tax-benefit reforms, then, predictions may be sensitive to the degree of flexibility introduced in the model.

Table 11: Discrepancies in Labor Supply Responses to the Reform

Model	Simulated responses to EIS (%)				
	part-time to non-work	full-time to non-work	full-time to part-time	net effect on employment	number of hh
SC	0,14%	0,32%	0,03%	0,46%	29 257
	[0.11;0.17]	[0.26;0.37]	[0.02;0.03]	[0.37;0.54]	[23 567; 34 395]
SC2	0,16%	0,36%	0,03%	0,52%	33 115
	[0.10;0.22]	[0.22;0.47]	[0.02;0.04]	[0.32;0.69]	[20 382; 43 949]
UC	0,24%	0,55%	0,04%	0,79%	50 326
	[0.19;0.29]	[0.46;0.62]	[0.04;0.05]	[0.65;0.91]	[41 401; 57 962]
GC	0,58%	1,65%	0,08%	2,23%	141 930
	[0.72;0.86]	[1.34;1.93]	[0.06;0.10]	[2.06;2.79]	[131 211; 177 708]

Figures in brackets give a bootstrapped 90% confidence interval of the transition frequency, computed by 100 draws of the parameters in their estimated distribution.

When using the general model (GC), behavioral responses to the reform are much larger in magnitude. It is shown that 2.23% of women in the sample would be discouraged to work, i.e., around 142, 000 women compared to 50, 000 women according to the prediction of (UC) and 29, 000 women according to (SC). The comparison of the confidence intervals for (SC), (UC) and (GC) shows that models are sufficiently precise to confirm the average order of magnitude of these discrepancies.

Larger responses in model (GC) probably result from the fact that preferences are wage- and income-dependent. In effect, standard models impose that wage rates come into play only through disposable income. In other words, gross wage and implicit net wage are forced to have the same role. In model (GC), the gross wage recovers an independent role and so does taxation. Sensitivities of labor supply to tax-benefit incentives on the one hand and to wage rates on the other then increase simultaneously.

The large discrepancies across models appeal for further comments. On the one hand, it might be argued that policy recommendations require a model which provides the best fit to data while still grounded on a utility-maximizing rationality. This is actually what model (GC) provides us with, and it can be used in the usual way to analyze labor supply responses or welfare changes (with some reserve on the latter, however, since previous interpretations of (GC) included demand-side dimensions). On the other hand, (GC) can be seen only as a statistical model which loses touch with economic theory. This general specification thus illustrates the difficulties to reconcile the best possible explanatory power and a structural model in line with theory.

³⁰The predictions of (SC) and (SC2) are not significantly different, due to the fact that the latter model is less precisely estimated.

The ideal model would naturally be estimated on panel data, with information on individual preferences in an intra-household bargaining environment, etc. to reconcile all the aspects mentioned before (dynamic aspects, collective rationality and so on) and test each of them as auxiliary features. We are unfortunately often limited to *ex-ante* evaluations based on cross-sectional data with limited information on household characteristics. In these circumstances, model (GC) has clearly shown that the unitary and static approximation is rejected against a general specification which is likely to better capture multiple explanatory dimensions. Most importantly, it has been shown that this approximation also considerably changes the magnitude of the predictions when evaluating tax reforms.

4.2.3 Welfare Analysis

To complete the policy evaluation, we finally assess the impact of the reform on household welfare, using a money metric measure of utility. Notice that there are only gainers with the reform under consideration. Note C_{ij} (resp. C_{ij}^p) the disposable income for choice j before reform (resp. post-reform). Equivalent variation EV_i for household i is defined as the value of a tax-free transfer that would make the household as well-off as with the tax reform. In the case of (UC), chosen for illustration, this is simply the solution of the following equation:

$$\text{Max}_j [W(C_{ij}^p; Z_i, \theta_j) + \epsilon_{ij}] = \text{Max}_k [W(C_{ik} + EV_i; Z_i, \theta_k) + \epsilon_{ik}]. \quad (5)$$

Note that the optimal post-reform choice j is not necessarily the same as the optimal choice k in the case of the EV transfer.

It follows that EV is a random variable which depends on the distribution of the random terms, on initial and final incomes and on all household characteristics. We do not intend to express explicitly the distribution of EV or, alternatively, of compensating variations (as done by Dagsvik and Karlstrom, 2003). This is all the more difficult as the stochastic part of the utility is a mixture of the type I-extreme value distribution and the unobserved heterogeneity whose distribution is unspecified. Instead, we simply compute the utility after reform – l.h.s. of equation (5) – and search numerically for the value of EV which satisfies this equation. For the stochastic part, we proceed in the same way as for the simulation of transition matrices (see Appendices). More precisely, we generate series of pseudo-residuals ϵ_{ij} (for all choices $j = 1, \dots, J$) which lead to a match between observed and optimal labor supply decisions. As usually done in simulation studies, we assume that the policy change does not affect the random terms so that these terms can be used on both sides of equation (5). Final values for EV are the mean values over a large number of draws.

The empirical distribution of the welfare gain EV in the population of movers (expressed in euros/month) is summarized in Table 12; 90% confidence intervals are also presented. Some discrepancies appear between the predictions of the three main models. The average gain is between 79 and 95 euros/month depending on the model used. Contrary to labor supply responses, these discrepancies do not seem very important, which is actually not surprising considering the fact that movers have fairly close characteristics. As argued before, however, welfare analysis is subject to caution as parameters may have been contaminated by demand-side aspects, especially in the case of (GC). This is a common feature to all studies relying on observed hours.

Table 12: Distribution of Welfare Gain in the Population of Movers (Money Metric Utility, in Eur/month)

<i>model</i>	5%	50%	95%	mean
<i>SC</i>	11.8 [6.6; 19.8]	68.0 [39.6; 92.5]	230.0 [145.2; 303.6]	85.8 [57.7; 113.0]
<i>UC</i>	13.4 [6.6; 26.4]	76.2 [52.8; 99.0]	247.3 [191.4; 323.4]	95.0 [77.2; 118.8]
<i>GC</i>	8.84 [6.6; 13.2]	63.7 [52.8; 75.9]	200.8 [165.0; 237.6]	79.0 [68.9; 89.5]

Figures in brackets give a bootstrapped 90% confidence interval of the utility gain

Conclusion

This paper presents several important findings for the modeling of household labor supply with taxation. The main result suggests that labor supply modelers should take advantage of the very general representation allowed by discrete-choice models, mostly in order to relax some of the usual constraints imposed on household preferences and rationality.

Firstly, the structural model traditionally used for policy evaluation imposes functional constraints on well-behaved leisure-consumption preferences. We suggest a specification which is fully flexible regarding the way utility depends on working hours, while maintaining a strict utility-maximizing interpretation. This unconstrained model enables us to identify the constraints imposed by usual structural models and to test them. They happen to be very clearly rejected by the French data.

Secondly, structural models often gain flexibility by increasing variability across alternatives through state-specific dummies. This additional structure relies on *ad hoc* interpretation, be it costs of work, job search disutility, distaste for work etc., whereas many aspects may well be captured simultaneously by the corresponding parameters.

Thirdly, 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 dynamic or collective models, or demand-side aspects even. Nested standard models are strongly rejected against this general setting. Under the assumption of partial equilibrium in a static environment, this result boils down to a rejection of the unitary model, all the more robust as the models do not rely on restrictive functional forms as in previous papers.

All three models – structural model, unconstrained model and general model with price- and income-dependent preferences – yield significantly different responses to a tax reform. This suggests that both constraints on functional forms and approximations of household rationality change the nature of the predictions when using traditional models for tax policy evaluation.

This raises interesting concerns about the way we model labor supply decisions in order to conduct policy analysis. One might reasonably think that robust recommendations to policy makers should rely on a model which explains actual behavior best, while sticking to a utility-maximizing interpretation. This is actually the case of the general model introduced in the paper. On the other hand, one might argue that this model is purely statistical and lacks economic foundation. The ideal model would offer a comprehensive representation of household behavior, made consistent within a neo-classical framework, from which testable restrictions could be derived. However, at this stage and with the data available, it seems difficult to test multiple dimensions in a unifying framework. For instance, the literature on

collective models has mainly consisted of testing the restrictions resulting from the efficiency hypothesis (i.e. the collective rationality) and the tests have been conducted in a purely supply-side and static framework (see Chiappori et al., 2002).

Finally, it is still an open question whether the general model we have presented may help to discriminate between the various dimensions explaining labor supply decisions. Future research should first attempt to rule out the demand-side interpretation. Basic solutions for this purpose consist of using data on desired rather than observed hours, or simply to use data from a country whose labor market is less constrained than France's. Then, our intuition is that the general model may fruitfully serve to test restrictions from a collective model. This is left to future research.

References

- [1] Bargain, O. and I. Terraz (2003): "Evaluation et mise en perspective des effets incitatifs et redistributifs de la Prime pour l'Emploi.", forthcoming in *Economie et Prévision*.
- [2] Blomquist, S. and W. Newey (2002): "Nonparametric estimation with nonlinear budget sets", *Econometrica*, 70, 6, 2455-2480.
- [3] 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.
- [4] 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.
- [5] Bonin, H. and H. Schneider (2004): "Analytical prediction of transitions probabilities in the conditional logit model", *IZA Discussion Paper*, 1015.
- [6] Bourguignon, F. and T. Magnac (1990): "Labour supply and taxation in France", *Journal of Human Resources*, 25, 358-389.
- [7] Browning, M. and C. Meghir (1991): "The Effects of Male and Female Labor Supply on Commodity Demands", *Econometrica*, 59, 4, 925-951.
- [8] Chiappori, P.-A. (1988): "Rational Household Labor Supply", *Econometrica*, 56, 63-89.
- [9] Chiappori, P.-A. (1992): "Collective Labor Supply and Welfare", *Journal of Political Economy*, 100, 437-467.
- [10] Chiappori, P.-A., B. Fortin and G. Lacroix (2002): "Marriage Market, Divorce Legislation and Household Labor Supply", 110, *Journal of Political Economy*, 37-72.
- [11] Choné, P., D. Le Blanc and I. Robert-Bobée (2003): "Female labor supply and child care in France", *CREST working paper*.
- [12] 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.
- [13] Dickens, W. and S. Lundberg (1993): "Hours restrictions and labor supply", *International Economic Review*, 34, 169-192.

- [14] Duncan, A. and M. Weeks (1998): “Simulating transitions using discrete choice models”, *Papers and Proceedings of the American Economic Statistical Association*, 106, 151-156.
- [15] Eissa, N. and H. Hoynes (2004): “Taxes and the Labor Market Participation of Married Couples: The Earned Income Tax Credit”, *Journal of Public Economics*, 88, 9-10, 1931-1958
- [16] Fortin, B. and G. Lacroix (1997): “A Test of the Neo-Classical and Collective Models of Household labor Supply”, *Economic Journal*, 107, 933-955.
- [17] 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.
- [18] 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.
- [19] Hausman J.A. (1985): “The Econometrics of Nonlinear Budget Sets”, *Econometrica*, 53, 1255-82.
- [20] Hausman, J.A. and P. Ruud (1994): “Family labor supply with taxes”, *American Economic Review*, 74, 242-248.
- [21] Heckman, J. (1979): “Sample selection bias as a specification error”, *Econometrica*, 47, 151-161.
- [22] Heckman, J. (1993): “What has been learned about labor supply in the past twenty years?”, *American Economic Review*, 83, 2, 116-121.
- [23] Heckman, J. and B. Singer (1984): “A method for minimizing the impact of distributional assumptions in econometric models of duration data”, *Econometrica*, 52, 271-320.
- [24] Heim, B.T. and B.D. Meyer (2002): “Structural Labor Supply Models when Budget Constraints are Nonlinear”, Northwestern University Working Paper.
- [25] 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.
- [26] Hoynes, H. (1996): “Welfare Transfers in Two-Parent Families: Labor Supply and Welfare Participation Under AFDC-UP”, *Econometrica*, 64, 295-332.
- [27] 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.
- [28] Killingsworth, M. et J. Heckman (1986): “Female labor supply: a survey”, in O. Ashenfelter and R. Layard (eds), *Handbook of Labor Economics*, vol.1, Amsterdam, North-Holland.
- [29] Keane, M. and R. Moffit (1998): “A Structural Model of Multiple Welfare Program Participation and Labor Supply”, *International Economic Review*, 39, 553-589.
- [30] Laroque, G. and B. Salanié (2002): “Labor Market Institutions and Employment in France”, *Journal of Applied Econometrics*, 17(1), 25-48.
- [31] Lundberg S.J. and R.A. Pollak (1996): “ Bargaining and Distribution in Marriage”, *Journal of Economic Perspectives*, 10, 139-158.

- [32] MaCurdy, T. (1992): “Work Disincentives Effects of Taxes: A Reexamination of some evidence”, *American Economic Review, Papers and Proceedings*, 82, 243-249.
- [33] MaCurdy, T., D. Green and H. Paarsch (1990): “Assessing Empirical Approaches for Analysing Taxes and Labor Supply”, *Journal of Human Resources*, 25, 415-490.
- [34] McFadden, D. (1973): “Conditional logit analysis of qualitative choice behavior”, in P. Zarembka, ed., *Frontiers in Econometrics*, Academic Press, NY.
- [35] McFadden, D. and K. Train (2000): “Mixed MNL Models for Discrete Response”, *Journal of Applied Econometrics*, 15, 447-470.
- [36] Moffitt, R. (1984): “Estimation of a joint wage-hours labor supply model”, *Journal of Labor Economics*, 2, 550-556.
- [37] Pollak, R. (1977): “Price-dependent preferences”, *American Economic Review*, 67, 2, 64-75.
- [38] Tummers, M. and I. Woittiez (1991): “A simultaneous wage and labor supply model with hours restriction”, *Journal of Human Resources*, 26, 393-423.
- [39] Van Soest, A. (1995): “Structural Models of Family Labor Supply: a Discrete Choice Approach”, *Journal of human Resources*, 30, 63-88.
- [40] Van Soest, A. and M. Das (2000): “Family labor supply and proposed tax reforms in the Netherlands”, *De Economist*, forthcoming.
- [41] Van Soest, A., I. Woittiez and A. Kapteyn (1990): “Labor supply, income taxes and hours restrictions in the Netherlands”, *Journal of Human Resources*, 25, 517-558.
- [42] 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.

Appendices

4.3 Identification of the Structural Constraints

Using quadratic specifications chosen in the text, we can easily identify the constraints imposed by the structural models (SC) or (SC2) on the unconstrained model (UC). For simplicity, this shall be done when ignoring unobserved heterogeneity, that is, $\alpha_{cr}^0 = \alpha_c^0$ in the structural models and $b_{jr}^0 = b_j^0$ in (UC). The number of constraints imposed by a nested model on the nesting model is the difference in the number of coefficients between the two models, that is 158 for (SC) vs (UC) and 155 for (SC2) vs (UC).³¹ For the purpose of clarity, we assume that all the coefficients of the cost of work are variable and denote (SC3) this extended model. This is slightly more flexible than (SC2) but was not retained for the empirical applications as the additional variables were not significant. Model (SC3) contains 39 variables whereas model (UC) contains 189 variables, hence 150 constraints imposed by the former on the latter.

³¹For the tests, the coefficients of the 4 variables corresponding to the square of the dummy ‘Paris region’ are set to zero in (UC), hence 185 parameters instead of 189.

The structural model can be rewritten for $j = 1, \dots, 5$:

$$U_{ij} = \alpha_{cc}C_{ij}^2 + (\alpha_{ci} - 2\alpha_{cc}F_{ij} + \alpha_{ch}H_j)C_{ij} \\ + (\alpha_{hh}(H_j)^2 + \alpha_{hi}H_j - \alpha_{ch}F_{ij}H_j + \alpha_{cc}F_{ij}^2 - \alpha_{ci}F_{ij})$$

with

$$\alpha_{ci} = \alpha_c^0 + \alpha'_c Z_i \\ \alpha_{hi} = \alpha_h^0 + \alpha'_h Z_i \\ F_{ij} = f_j^0 + \sum_{k=1}^7 f_j^k z_i^k \quad \text{with } f_j^k = 0 \text{ for } k = 5, 6, 7,$$

to be compared to the unconstrained model:

$$U_{ij} = a_j C_{ij}^2 + b_{ij} C_{ij} + c_{ij} \quad \text{for } j = 1, \dots, 5$$

with

$$b_{ij} = b_j^0 + b'_j Z_i \\ c_{ij} = c_j^0 + c'_j Z_i + \sum_{l=1}^7 \sum_{m=l}^7 c_j^{lm} z_i^l z_i^m,$$

so that identification requires:

$$\begin{aligned} \text{condition I} & : a_j = \alpha_{CC} \quad \text{for } j = 1, \dots, 5 \\ \text{condition II} & : b_{ij} = \alpha_{ci} - 2\alpha_{cc}F_{ij} + \alpha_{ch}H_j \quad \text{for } j = 1, \dots, 5 \\ \text{condition III} & : c_{ij} = \alpha_{hh}(H_j)^2 + \alpha_{hi}H_j - \alpha_{ch}F_{ij}H_j + \alpha_{cc}F_{ij}^2 - \alpha_{ci}F_{ij} \quad \text{for } j = 2, \dots, 5. \end{aligned}$$

Condition I directly leads to 4 restrictions on the coefficients of the unconstrained model:

$$a_j = a_1 (= \alpha_{cc}) \quad \text{for } j = 2, \dots, 5. \quad (6)$$

Condition II gives:

$$\begin{aligned} b_1^0 + b'_1 Z_i & = \alpha_c^0 + \alpha'_c Z_i \\ b_j^0 + b'_j Z_i & = \alpha_c^0 + \alpha'_c Z_i - 2\alpha_{cc}(f_j^0 + f'_j Z_i) + \alpha_{ch}H_j \\ & = (\alpha_c^0 - 2\alpha_{cc}f_j^0 + \alpha_{ch}H_j) + (\alpha'_c - 2\alpha_{cc}f'_j)Z_i \quad \text{for } j = 2, \dots, 5 \end{aligned}$$

so that we obtain the following equations:

$$\begin{aligned} b_1^0 & = \alpha_c^0 \\ b_1^k & = \alpha_c^k \quad \text{for } k = 1, \dots, 4 \\ b_j^0 & = \alpha_c^0 - 2\alpha_{cc}f_j^0 + \alpha_{ch}H_j \quad \text{for } j = 2, \dots, 5 \\ b_j^k & = \alpha_c^k - 2\alpha_{cc}f_j^k \quad \text{for } j = 2, \dots, 5 \text{ and } k = 1, \dots, 4 \end{aligned}$$

and the 12 restrictions:

$$b_j^k = b_1^k (= \alpha_c^k) \quad \text{for } j = 2, \dots, 5 \text{ and } k = 5, 6, 7. \quad (7)$$

From condition III, we can write for $j = 2, \dots, 5$:

$$\begin{aligned}
c_j^0 &= \alpha_{hh}(H_j)^2 + \alpha_h^0 H_j - \alpha_{ch} f_j^0 H_j + \alpha_{cc}(f_j^0)^2 - \alpha_c^0 f_j^0 \\
c_j^k &= \alpha_h^k H_j - \alpha_{ch} f_j^k H_j + 2\alpha_{cc} f_j^0 f_j^k - \alpha_c^k f_j^0 - \alpha_c^0 f_j^k \quad \text{for } k = 1, \dots, 4 \\
c_j^k &= \alpha_h^k H_j - \alpha_c^k f_j^0 \quad \text{for } k = 5, 6, 7 \\
c_j^{lm} &= 2\alpha_{cc} f_j^l f_j^m - \alpha_c^l f_j^m - \alpha_c^m f_j^l \quad \text{with } l = 1, \dots, 3 \text{ and } m = l + 1, \dots, 4 \\
c_j^{lm} &= -\alpha_c^m f_j^l \quad \text{with } l = 1, \dots, 4 \text{ and } m = 5, 6, 7 \\
c_j^{kk} &= (\alpha_{cc} f_j^k - \alpha_c^k) f_j^k \quad \text{for } k = 1, \dots, 4,
\end{aligned}$$

and the 24 following restrictions:

$$c_j^{lm} = 0 \quad \text{for } j = 2, \dots, 5 \text{ with } l = 5, 6 \text{ and } m = l + 1, \dots, 7 \quad (8)$$

$$c_j^{kk} = 0 \quad \text{for } j = 2, \dots, 5 \text{ and } k = 5, 6, 7. \quad (9)$$

We now pool the remaining equations, which give after simplification and for $j = 2, \dots, 5$:

$$\begin{aligned}
b_j^0 &= b_1^0 - 2a_1 f_j^0 + \alpha_{ch} H_j \\
c_j^0 &= \alpha_{hh}(H_j)^2 + \alpha_h^0 H_j^f + f_j^0(-\alpha_{ch} H_j + a_1 f_j^0 - b_1^0) \\
c_j^k &= \alpha_h^k H_j + f_j^k(-\alpha_{ch} H_j^f + 2a_1 f_j^0 - b_1^0) - b_1^k f_j^0 \quad \text{for } k = 1, \dots, 4 \\
c_j^k &= \alpha_h^k H_j - b_1^k f_j^0 \quad \text{for } k = 5, 6, 7 \\
c_j^{lm} &= 2a_1 f_j^l f_j^m - b_1^l f_j^m - b_1^m f_j^l \quad \text{with } l = 1, \dots, 3 \text{ and } m = l + 1, \dots, 4 \\
c_j^{lm} &= -b_1^m f_j^l \quad \text{with } l = 1, \dots, 4 \text{ and } m = 5, 6, 7 \\
c_j^{kk} &= (a_1 f_j^k - b_1^k) f_j^k \quad \text{for } k = 1, \dots, 4 \\
f_j^k &= \frac{b_j^k - b_1^k}{-2a_1} \quad \text{for } k = 1, \dots, 4.
\end{aligned}$$

The four last equations of this system give for $j = 2, \dots, 5$:

$$c_j^{lm} = \frac{b_j^m b_j^l - b_1^l b_1^m}{2a_1} \quad \text{with } l = 1, \dots, 3 \text{ and } m = l + 1, \dots, 4 \quad (10)$$

$$c_j^{lm} = b_1^m \frac{b_j^l - b_1^l}{2a_1} \quad \text{with } l = 1, \dots, 4 \text{ and } m = 5, 6, 7 \quad (11)$$

$$c_j^{kk} = \frac{(b_j^k)^2 - (b_1^k)^2}{4a_1} \quad \text{for } k = 1, \dots, 4, \quad (12)$$

that is, 88 new restrictions. The first equation of the system gives an expression of f_j^0 :

$$f_j^0 = \frac{b_1^0 - b_j^0 + \alpha_{ch} H_j}{-2a_1} \quad \text{for } j = 2, \dots, 5.$$

The first and the third equations of the system give for $j = 2, \dots, 5$:

$$c_j^k = \alpha_h^k H_j + \frac{b_j^k - b_1^k}{2a_1} b_j^0 - b_1^k f_j^0 \quad \text{for } k = 1, \dots, 4,$$

where f_j^0 can be replaced by its expression so that:

$$\frac{1}{H_j} [c_j^k - b_j^0 \frac{b_j^k - b_1^k}{2a_1} - b_1^k \frac{b_1^0 - b_j^0}{2a_1}] = \alpha_h^k + b_1^k \frac{\alpha_{ch}}{2a_1} \quad \text{for } j = 2, \dots, 5 \text{ and } k = 1, \dots, 4,$$

hence the 12 restrictions:

$$\begin{aligned} \frac{1}{H_j} [c_j^k - b_j^0 \frac{b_j^k - b_1^k}{2a_1} - b_1^k \frac{b_1^0 - b_j^0}{2a_1}] &= \frac{1}{H_2} [c_2^k - b_2^0 \frac{b_2^k - b_1^k}{2a_1} - b_1^k \frac{b_1^0 - b_2^0}{2a_1}] \\ &(\quad = \alpha_h^k + b_1^k \frac{\alpha_{ch}}{2a_1}) \quad \text{for } j = 3, 4, 5 \text{ and } k = 1, \dots, 4. \end{aligned} \quad (13)$$

Finally, we can substitute f_j^0 in the fourth equation of the system to obtain:

$$\frac{1}{H_j} [c_j^k - \frac{b_1^0 - b_j^0}{2a_1}] = \alpha_h^k + b_1^k \frac{\alpha_{ch}}{2a_1} \quad \text{for } j = 2, \dots, 5 \text{ and } k = 5, 6, 7,$$

hence the 9 following restrictions:

$$\begin{aligned} \frac{1}{H_j} [c_j^k - \frac{b_1^0 - b_j^0}{2a_1}] &= \frac{1}{H_2} [c_2^k - \frac{b_1^0 - b_j^0}{2a_1}] \\ &(\quad = \alpha_h^k + b_1^k \frac{\alpha_{ch}}{2a_1}) \quad \text{for } j = 3, 4, 5 \text{ and } k = 5, 6, 7, \end{aligned} \quad (14)$$

which amounts to a total of 150 constraints. ■

4.4 Data, Sample Selection and Discretization

The data used are selected from the French Household Budget Survey 1994 and monetary variables have been grossed up to 1998, our year of reference, assuming the demography constant. No structural change has occurred in the tax-benefit system between 1994 and 1998 so that there is no inconsistency between the simulated system of 1998 and observed behaviors (see Bargain and Terraz, 2003).

For the selection, we keep 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. For this purpose, households where the wife is disabled, retired or a student are excluded. So are households where she is self-employed or civil servant. Indeed, self-employed and farmers are subject to income tax rules which are substantially different from the ones applied to earnings and which require additional information not available here. Moreover, the labor supply behavior of independent workers or civil servants, whose job is guaranteed for life, may be rather different from salary workers and would require a different modeling strategy altogether.

Extreme households are selected out, notably those receiving important levels of non-labor income and those with more than 3 children or where children have substantial earnings. Households with more than two decision-makers (other adults than the basic couple) are also discarded. At this stage, our selection contains 3,548 couples and the corresponding descriptive statistics are presented in Table 13.

The number of inactive husbands is too small (3.8%) to estimate a joint labor supply model. The attempt to do so did not provide satisfying results, even when reducing the male choice to a simple participation decision (husbands are indeed very concentrated on full-time jobs around 39 hours). Therefore, we have chosen to exclude couples where husbands do not work and to assume male hours fixed at the observed values in order to focus purely on female labor supply. This way, 3,397 households remain. The discretization is based on the observed peaks of the distribution in Figure 1, that is, $H_j = 0, 20, 30, 39, 45$ hours a week. Corresponding intervals are $[0 - 10[$, $[10 - 25[$, $[25 - 34[$, $[34 - 42[$ and over 42 respectively. The proportion of households in each bracket is 24, 12.6, 10.9, 44.7 and 7.8% respectively.

Notice that the pattern of working hours is very rigid in France, as seen for female workers in Figure 1. Accumulation of observations at certain hours are driven by demand-side and institutional constraints

and a discrete approach seems to suit the fact well that workers are actually constrained to choose among a limited number of available options. Full-time and part-time are institutionally fixed at 39 and 20 hours respectively. Three-quarter of full-time (30 hours) is an option frequently offered by firms, especially to women. Overtime – captured by the 45 hour option – is allowed in some companies or due to a secondary activity. Self-employed may have more freedom to choose in a continuous range of hours but they are not included in the present study.

Other types of labor market constraints are not addressed in the present work. Firstly, workers might – to a certain extent – choose their wage rate together with the number of working hours. We follow the bulk of the literature and assume that gross hourly wage rates do not depend on working duration.³² Secondly, demand-side is not modeled here, even though it may imply strong rationing on the choice of hours. One way to deal with this issue is to use information on desired working hours, in addition to observed hours, in order to disentangle supply and demand sides.³³ Even when available, 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. In fact, it seems that rationing at full time (39 hours) is particularly strong, as discussed below. It is worth mentioning that other limitations to our approach are common to most exercises of this type.

Table 13: Descriptive Statistics for Selected Couples

	Women	Men
Participation	77,2	96,2
Working time (hours/week) / participants	34,7	41,9
Working time (hours/week) / all	26,8	36,7
Gross wage rate (euros/hour) / participants	11,7	13,7
Gross wage rate (euros/hour) / all*	11,0	12,9
Age	38,9	41,1
Primary education	30,69%	17,94%
Vocational training	37,85%	45,97%
High school diploma	14,75%	17,90%
University studies	16,71%	18,19%
Average number of children	1,427	
Presence of child 0-2	16,63%	
Presence of child 3-5	18,74%	
Presence of child 6-11	32,67%	
Nb of selected households	3 548	
Corresponding population	6 658 124	
% of total population	28,9%	

* these include predicted wages for non-workers

4.5 Budget constraint and wage prediction

Disposable income at each discrete choice of hours is computed by the French tax-benefit microsimulation SYSIFF98, symbolized by function D above. This program allows the simulation of all direct taxes and benefits of instruments. It is described in Bargain and Terraz (2003).

³²The assumption is relaxed only in Moffitt (1984), Tummers and Woittiez (1991) and Ilmakunnas and Pudney (1990).

³³This piece of data is unfortunately not provided by the French Household Budget Survey. Information on desired hours is used in Ilmakunnas and Pudney (1990) and Van Soest et al. (1990), among others. In traditional labor supply models estimated on observed hours, estimates of preference parameters are certainly contaminated by demand side aspects in a way that makes welfare analysis problematic.

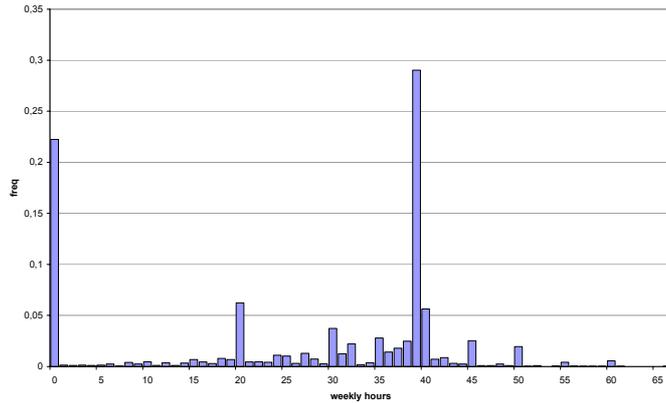


Figure 1: Distribution of Working Hours for Women with Employed Partners

Wage rates are not provided directly and must be computed as weekly gross income divided by the number of working hours. As in Van Soest (1995), wages for inactive women are predicted using the traditional technique (Heckman, 1979).³⁴ A better way would be to follow Van Soest et al. (2002) or Laroque et Salanié (2002) and to estimate the wage equation and the labor supply model jointly. In effect, the participation probit aimed at correcting the selection bias is a (linearized reduced form) approximation to the selection mechanism implied by the labor supply model. In particular, it does not account for the nonconvexities of the tax-benefit system arising from means-tested minimum income. However, proceeding in two steps does not seem an excessive impediment here, since all husbands in our selected data work so that households are unlikely to be eligible for social assistance.

To judge the quality of the estimation, Figure 2 compares the distributions of predicted and observed wages for part-time and full-time female workers respectively. The fit is more satisfying in the second case. Notice that a non negligible share of part-time workers (around 9.5%) displays a wage below the official hourly minimum wage (6 euros in 1998). Overall, the wage equations seem to provide a reasonable representation of the wage distribution, even though the predicted distribution is more concentrated on the mode. The predicted mean (resp. median) for active women is around 10 EUR/hour (resp. 9.2) while it is around 8 EUR/hour (resp. 6.8) among inactive women.

4.6 Estimation Results

Table 14 presents the results of the estimations for the structural models (SC) and (SC2) without unobserved heterogeneity. We follow the traditional way to interpret coefficients in terms of consumption-leisure preferences and costs of work. However, as argued in Section 3.5, this interpretation should be taken very cautiously as the variables in use are likely to capture other aspects which determine the observed work durations.

The only significant taste parameters for consumption (disposable income) are the constant term, female age (in SC2 only) and the variables related to children between 3 and 11. The marginal utility of

³⁴Because 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, for instance 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. However, for a tractable number of draws (20), this correction did not significantly change our results.

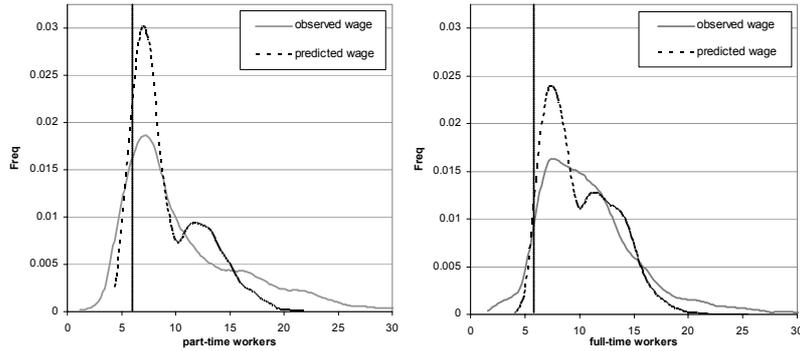


Figure 2: Predicted vs observed wages for female workers (EUR/h)

consumption increases with the number of children in that age bracket. Conversely, almost all 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. Costs of work increase for people in the Paris region. Rather surprisingly, they decrease with the number of children in nursery (3-5) or primary (6-11) school (not significantly in SC2).

Overall, estimates yield implausible values for the cost of work. In model (SC), the constant term is equal to 41% of the average earnings of working wives. This seemingly deceiving result is in fact an illustration of what has been said before about the caution required when interpreting the coefficients. It could well be the case that the implausible cost of work reflects other dimensions, including demand-side aspects. Bourguignon and Magnac (1990) find similar results and argue that a sizeable proportion of working wives are constrained in their choice. In effect, most available jobs are full-time positions (39 hours per week) so that rationing is equivalent to the presence of implicit costs/disutilities which make work sub-optimal for choices other than 39 hours. This intuition is confirmed by the finding that in model (SC2), the ‘cost of work’ is actually minimal for the 39-hour option.³⁵ At the same time, this term increases with the number of working hours (except for 39 hours), which conforms to an interpretation in terms of variable costs.³⁶ Finally in (SC), the variable cost related to the presence of small children rises with the number of working hours of the mother – except for 39 hours for the same reasons as above – which could reflect the increasing need to purchase childcare on the market.

Considering the large lists of coefficients and the difficulty to interpret the results, regression tables are omitted for (UC) and (GC).³⁷ However, to give a feel of the main aspects captured in the unconstrained model, Table 15 provides the estimates of model (uc), i.e., the basic version of (UC) where all heterogeneity is withdrawn.³⁸ 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 ($H_4 = 39$

³⁵It is worth noticing, however, that only 40% of working women in our sample work 39 hours a week. Therefore, the rationing hypothesis cannot be maintained for the whole sample.

³⁶Note, however, that the coefficients are not significant.

³⁷They are available upon request.

³⁸The coefficients of the square terms are larger in absolute value simply because income is smaller than 1 (we consider weekly income divided by 40,000).

Table 14: MNL estimation of the structural models

Variable	SC		SC2	
	Coef.	Std. Err.	Coef.	Std. Err.
income ²	-29.239 ***	7.006	-82.234 ***	11.663
female hours ²	-9.797 ***	.759	.797	.974
female hours x income	-14.819 ***	1.531	-8.080	10.445
income	24.426 ***	5.170	37.996 ***	7.071
x woman age	.231	.142	.486 **	.237
x man age	.059	.133	-.051	.220
x # children	-.755	.738	-1.315	1.290
x # children 0-2	-4.840	3.095	-4.999	4.248
x # children 3-5	6.583 ***	2.730	6.174 **	2.851
x # children 6-11	4.217 *	2.278	1.928	2.067
x 1(Paris region)	.284	2.890	-2.577	2.419
female hours	16.411 ***	1.103	5.351 **	2.545
x woman age	-.026	.0213	-.057 *	.032
x man age	-.040 **	.020	-.047 ***	.013
x # children	-.422 ***	.118	-.326 ***	.105
x # children 0-2	-1.003 *	.535	-.213	.511
x # children 3-5	-.490 **	.221	-.842 *	.444
x # children 6-11	-.388 **	.170	-.636 ***	.192
x 1(Paris region)	1.759 ***	.205	1.736 ***	.280
fixed costs/40000	.162 ***	.022	-	-
x 1(Paris region)	.026 **	.013	.022 ***	.006
x # children 3-5	-.018 **	.009	-.003	.004
x # children 6-11	-.014 *	.008	-.004	.003
variable costs/40000				
x # children 0-2 / j=2	.025 *	.016	.020 ***	.008
x # children 0-2 / j=3	.052 ***	.020	.0170 *	.010
x # children 0-2 / j=4	-.004	.023	.020	.013
x # children 0-2 / j=5	.071 ***	.029	.040 ***	.016
x j=2	-	-	.033	.027
x j=3	-	-	.053	.041
x j=4	-	-	.029	.054
x j=5	-	-	.091	.061
Log-Likelihood	-4909.7		-4458.0	
Nb of observations	3397		3397	

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

hours). This is illustrated by the indifference curves drawn in the (C, H) plane in Figure 3 (note that the specification of the model is discrete and so are the indifference curves). Clearly, the model does not only pick preferences but also a certain number of dimensions, including availability of certain types of job (39-hour jobs), in a similar way as the variable cost of work in (SC2).

Table 15: MNL estimation of the basic unconstrained model

		uc	
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
Log-Likelihood		-4603.3	
Nb of observations		3397	

All parameters are significant at the 1% level

4.7 Simulation Strategy and Sensitivity Analysis

The same procedure is used for each model. Notice first that it seems natural that the baseline (the pre-reform situation) displays actual choices of hours. To generate such a plausible baseline, we then draw some series of pseudo-residuals $\hat{\epsilon}_{ij}$ ($j = 1, \dots, J$) from a type I-extreme value distribution for the random part of the utility at each choice of hours. For each household, series are drawn repetitively together with the preference regime (unobserved heterogeneity) until a perfect match between observed and predicted hours is obtained. Post-reform optimal choices are predicted by the estimated deterministic model plus the retained pseudo-residuals $\hat{\epsilon}_{ij}$ derived from the calibration step. The procedure is repeated 200 times to obtain transition frequencies for each household. Transition tables result from averaging over the whole population.

This ‘calibration method’, suggested also in Duncan and Weeks (1998), can be compared to the more common ‘aggregate probability approach’. With the latter, predicted frequencies of each choice are summed up over the entire sample in order to compute the aggregate probability for this choice, before and after the simulated shock (reform, wage rise, etc.). However, this method ignores the probabilistic nature of the state transitions at the individual level. Bonin and Schneider (2004) ingeniously suggest a way to construct analytical predictions of transitions probabilities. Their approach enables us to avoid the time-consuming process of resampling pseudo-residuals in the ‘calibration method’ suggested here. Yet, as it relies on the specific distribution of the error terms (I-extreme value), this analytical approach

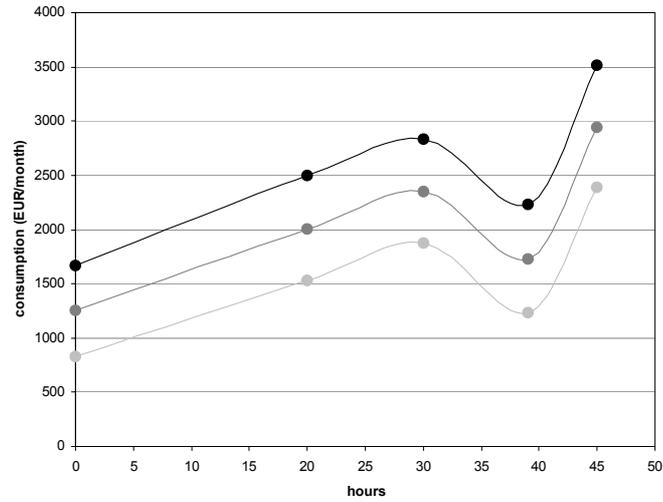


Figure 3: Indifference curves with model (uc)

is limited to the MNL case and cannot be directly applied when unobserved preference heterogeneity is introduced in the model.

As the nonlinearity of the models makes sensitivity analysis fairly complex, we proceed numerically. Confidence intervals (at the 90% level) for each transition cell are simulated by drawing 200 times from the estimated asymptotic (multivariate normal) distribution of the parameter estimates, and for each of those 200 parameter draws, by applying the method described above to build transition matrices.