Labour supply and taxation with restricted choices[☆]

Magali Beffy [a], Richard Blundell [b,c,*], Antoine Bozio [c,d], Guy Laroque [b,c,e], Maxime Tô [b,c]

[a] *Centre de recherche en économie et statistique (CREST), Paris, France*
[b] *University College London (UCL), United Kingdom*
[c] *IFS, 7 Ridgmount Street, WC1E 7AE, London, United Kingdom*
[d] *Paris School of Economics (PSE), France*
[e] *Sciences-Po, France*

ARTICLE INFO

ABSTRACT

A model of labour supply and taxation is developed in which observed hours reflect both the distribution of preferences and restrictions on choices. Hours restrictions are placed in a constrained rational choice setting in which the set of alternative hours on offer is restricted. Choices are made on a random subset of possible hours. We focus on the case where the choice set contains at most two offers. We show that when the choice set distribution is known, preferences can be identified. Conversely we show that, where preferences are known, the choice set distribution can be fully recovered. Conditions for identification of both preferences and the distribution of choice sets are also developed. We illustrate this approach in a labour supply setting with nonlinear budget constraints. Heterogeneity in the budget constraint reveals properties of the choice set. This framework is used to study the labour supply behaviour of a large sample of working age mothers in the UK, accounting for nonlinearities in the tax and welfare benefit system, fixed costs of work and restrictions on hours choices.

1. Introduction

Observed hours of work among working age adults display considerable variation, both over time and across countries, as well as in cross-section, especially among women with children (see, for example, Blundell et al., 2011). One of the key questions in the literature has been to ask under what conditions these variations can be used to identify preferences for labour supply, and what role tax and benefit policies play in framing these variations, see Blundell and MaCurdy (1999), for example.

It has been long recognized that the canonical model of labour supply with unrestricted choices of hours of work is ill-suited to distinguish individual preferences from external constraints. For instance, there is no role for restrictions on hours choices stemming from the demand side of the market. It has also been noted that the peaks in hours distributions at

---

part-time and full-time are hard to fully reconcile with unrestricted hours choice models. This prompted a large literature aiming to introduce hours restrictions in the modelling of labour supply (Aaberge et al., 1999, 2009; Altonji and Paxson, 1992; Bloemen, 2000, 2008; Blundell et al., 2008; Dickens and Lundberg, 1993; van Soest et al., 1990; Ham and Reilly, 2002; Dagsvik and Strom, 2006). Most of this literature rested on identifying hours constraints using either observed individual characteristics (e.g., number and age of children) or stated desired hours of work.

In this paper we exploit the information provided by the nonlinear tax and benefit system. Restricted choices, together with the resulting budget constraints, can give rise to dominated ranges of hours that would never be rationally chosen if a full range of other hours choices had been available. For example, we may observe people working full-time even though they would receive no smaller income if they were to work less, an observation inconsistent with the static neoclassical model of unrestricted hours choices (Saez, 2010; Chetty et al., 2011; Kline and Tartari, 2016).

Observed hours may, however, be consistent with optimal choice given a restricted choice set. As empirical economists we typically do not know the set of alternatives available to individuals. This is similar to the idea of a 'consideration set' in the modern literature on bounded rationality (e.g., Kfir and Spiegler, 2011). In that literature consumers make rational choices from a choice set that is limited by a combination of their own perception of the options and the strategy of firms. Our interpretation, in the labour supply context, is one where rational choices are made from a set of job packages restricted by the hours offered by employers.

We develop and estimate a structural model of labour supply that embeds restrictions on the set of available hours. We place these hours restrictions in a constrained rational choice setting in which the set of alternative choices on offer is restricted. The framework is general and concerns the case where the econometrician does not directly observe the choice set from which the individual has chosen. We suppose that agents make their choices on a random subset of all possible hours. We analyse how this modified model works, and in particular the sets of assumptions under which it still allows to identify the parameters of the underlying structural model. Our approach is akin to the one developed in Chetty (2012) in trying to account for constraints on labour supply choices. Our framework differs, however, in the sense that we are explicit about the sources of the constraints we model, i.e., limited hours offers. We first consider the case where the econometrician knows the probability distribution of offered choices. In the more complete model we generalize this to make the distribution of offers unknown but restrict it to be of finite support or a function of a finite set of unknown parameters.

Our model is placed in a life-cycle setting in which hours of work, employment and savings decisions are made subject to a nonlinear tax and benefit system and fixed costs of work. We draw on the extensive existing literature on labour supply models with nonlinear budget sets (Hausman, 1985; Heckman, 1974a,b), with fixed costs of work (Heckman, 1974a,b, 1979; Cogan, 1981), and with intertemporal choices (Heckman and MaCurdy, 1980). We further develop these models to the case in which individuals face constraints on hours choices.

Here we focus attention on developing a two-offer model in which each individual is assumed to face two independent hours offers — the one at which they are observed to work, if they are working positive hours, and one they turned down. The 'alternative' offer could include the observed hours point in which case the individual would be completely constrained and able to make no other hours choices. We assume the option of not working is always available. As the number of offers increases the specification approaches the standard labour supply model at which observed choices coincide with the fully optimal choice over all hours options.

The policy environment we consider is the labour supply behaviour of women in the UK. We model their decisions in the face of non-linear budget constraints generated through the working of the tax, tax-credit and welfare system. We study the period 1997–2002 when there were a number of key changes to the budget constraint through reforms to the tax-credit and welfare system (see Adam et al., 2010). We use data from the UK Family Expenditure Survey (FES) over this period. The FES is a detailed household survey that records hours worked, earnings and consumer expenditure. For every family in the data we have an accurate tax and benefit model (IFS-Taxben) that simulates the complete budget constraint incorporating all aspects of the tax, tax-credit and welfare systems, given the spouse choices and family composition. We use the consumption measure in the FES to ensure the hours of work model we develop is consistent with a life-cycle model (see Blundell and Walker, 1986).

We present a number of substantive empirical results: First, we provide direct evidence of hours restrictions by recording women working at hours of work that would be strictly dominated by other choices were a full range of hours choices to be available. Second, we estimate a parametric specification of the two-offer model and show that women that appear to be subject to the choice constraints belong to significantly poorer households than average and work shorter hours. Simulating the economy with or without the hours restrictions, we find a lower level of employment in the two-offer model than in the long-run when restrictions on hours are lifted. Together with the estimated preferences for hours and employment we argue that the framework provides a compelling empirical framework for understanding observed hours and employment.

The remainder of the paper is as follows. Section 2 lays out the intertemporal labour supply model with nonlinear budget constraints and fixed costs. We then consider the interpretation of rejections of the standard model and develop a model of labour supply in which individuals face a two-offer distribution over possible hours choices. In Section 3 we show that when the offer distribution is known, preferences can be identified in the standard multinomial choice and random utility models. We are able to show that, where preferences are known, the offer distribution can be fully recovered. We also develop conditions for identification of both the parameters of preferences and of the offer distribution. In section 4 we develop the sample likelihood for the two-offer model and use this model to study the labour supply choices of a large sample of women in the UK, accounting for nonlinear budget constraints and fixed costs of work. In Section 5 we present the estimates of the model and their implications for labour supply behaviour. Section 6 concludes.

2. A model of hours, employment and consumption

We begin by laying out a standard labour supply model with unrestricted hours choice (Section 2.1), then we show the conditions for rejection of this model (Section 2.2) before presenting our restricted choice model (Section 2.3).

2.1. Optimization at the intensive and extensive margins with a non-linear budget set

First, we develop a standard labour supply model in which there are non-linear taxes and fixed costs of work. Decisions are made in an intertemporal setting at the extensive and intensive margins.

At date t, the typical individual, the woman in a household, chooses the household consumption c_t and her own labour supply h_t, maximizing

$$E_t \int_t^T u_t(c_\tau, h_\tau) d\tau$$

subject to an intertemporal budget constraint

$$\int_t^T \exp[-r(\tau - t)]\{c_\tau - R(w_\tau, h_\tau) + b_\tau \mathbf{1}_{h_\tau > 0}\} d\tau \leq S_t,$$

among a set of possible hours choices $\mathcal{H}$. In the standard *unrestricted* case, $\mathcal{H}$ is the positive line. We shall study *restricted* choices, where $\mathcal{H}$ is made of a small number of random points. Here u_t is the unitary instantaneous utility index of the household, a concave twice differentiable function of the vector (c, h) of household consumption and own hours of work. It is increasing in consumption, decreasing in hours. The consumption good is the numeraire.

The function $R(w, h)$ denotes the income after taxes and benefits for someone who works h hours at wage w. It depends on the composition of the household and on the other actions of its members, which are considered fixed in this paper. The extensive margin comes from the fixed costs of being employed, i.e., having a positive h, costs b units of consumption. Accumulated savings at date t are equal to S_t. We denote by λ_t the Lagrange multiplier associated with the budget constraint at date t.

Current consumption maximizes $u_t(c_t, h_t) - \lambda_t c_t$, and therefore at an interior point satisfies the first-order condition

$$\frac{\partial u}{\partial c}(c_t, h_t) = \lambda_t.$$

Also, if the individual works, the optimal hours maximize $u_t(c_t, h_t) + \lambda_t R(w_t, h_t)$ over the set of possible hours $\mathcal{H}$. Let (c^e, h^e), be the optimal choice when working, c^o the household consumption when out of the labour market. The individual will be observed out of the labour market whenever the (revealed preference) inequality

$$u_t(c^e, h^e) - \lambda_t[c^e - R(w_t, h^e) + b_t] < u_t(c^o, 0) - \lambda_t[c^o - R(w_t, 0)]$$

is satisfied.

In the unrestricted choice framework, the choice of hours and employment is made subject to fixed costs of work and nonlinear taxes with *all* hours alternatives available. But observed hours and employment may not be consistent with this choice model.

2.2. Rejections of the unrestricted choice model

We make two standard assumptions: first, we assume that individuals have some disutility of working, i.e., their utility function is decreasing in hours. Second, we assume that, all other things equal, they always prefer more money rather than less. Under these two basic assumptions, we show that there are circumstances where some hours–wage packages are strictly dominated, that is they should not be chosen.

Indeed let h be a length of work such that there exists h' smaller than h, i.e., h' is less tiring than h, but yielding a larger income, i.e., $R(w, h) \leq R(w, h')$. Then h should never be chosen by the individual.

An extension of this argument allows to construct a range of hours that should never been chosen by a rational consumer. Define $S(w, h)$ the maximal income one can get by working at most h hours:

$$S(w, h) = \sup_{x \leq h} R(w, x).$$

The set of hours H^{ir} that the agent under study should not take if she were optimizing over all hours is

$$H^{ir} = \{\text{union of segments } (h_0, h_1) | S(w, h) \text{ is constant on} (h_0, h_1)\}.$$

This dominated set is expected to depend on the household, the number and age of children, eligibility to child benefits. If the after tax income function $R(w, h)$ is increasing in hours, the set H^{ir} is empty, and rationality or optimization does not restrict the choice of hours. If R and/or S has a flat segment, this segment belongs to H^{ir}: only its lower extremity can be chosen by a rational agent.

The previous computation describes a set of hours choices that will never be made by an individual whose utility weakly increases in leisure. It only depends on the shape of the tax function *R*. By construction, this set of 'irrational' choices would be larger if we made stronger assumptions on the structure of the economy than the monotony and separability of the utility index.

One of the contributions of the paper is to specify and estimate a fully parameterized restricted choice model. We use these parameter values to compute the likelihood of the observations under the counterfactual hypothesis of unrestricted choice. All the observations with zero likelihood are rejections of the unrestricted choice model. They include nonparametric rejections where h^e belongs to H^{ir}.

Fig. 1 provides some examples of non-linear budget constraints found in the data. In each case (2.a, 2.b and 2.c), we decompose after-tax income into two categories of income (earnings and other income) and four types of benefit (income support, family credit, rent rebate and local tax rebate).[1] The main non-convexity in the budget set comes from Income Support, that consists in a lump sum transfer to poor households. The level of the transfer depends on the household composition, and on the household total income. This is responsible for the large flat parts in the budget constraints. One of the conditions for women to get income support is to work less than 16 h. Above 16 h, they are no longer entitled to income support, but may be entitled to family credit. However, a women is not always entitled to family credit as illustrated in Fig. 1(b), and family credit may compensate the loss in income support (Fig. 1(a)) or not (1(c)).

The three examples also illustrate the nonparametric rejection of the model. The vertical dotted (red) lines correspond to the actual labour supply decisions of women, whereas the horizontal blue lines correspond to the set of irrational, strictly dominated, choices previously defined. In the first case (Fig. 1(a)), the decision appears to be close to a point that does not belong to the set of irrational choices, but the difference between the actual choice and the closest point that does not belong to that set is much larger in 1(b) and 1(c).

2.3. A model with restrictions on the choice set

To introduce our extension of the standard model, we consider choices over discrete hours. In this framework the typical worker is characterized by a set of parameters β, observed exogenous characteristics Z and unobserved characteristics ε. She chooses *h* that maximizes

$$V(h, Z, \beta, \varepsilon). \tag{1}$$

The possible choices *h* belong to a finite set $\mathcal{H}$ made of *I* elements $\{h_1, \ldots, h_I\}$. Given a subset of possible choices *H* in $\mathcal{H}$, for each β and *Z*, any distribution of ε yields a probability distribution on *H*. We shall denote the probability of choosing h_i in *H* as $p_i(H, Z, \beta)$.

We assume that given *V*, the observation of the family of probabilities $p_i(H, Z, \beta)$ identifies the parameter β, when *Z* varies in the population, and the union of the family of (non singleton) choice sets *H* for which the probabilities are observed covers the whole of $\mathcal{H}$.

The standard choice model has *H* equal to $\mathcal{H}$. For our application this model is not appropriate: because of underlying non-convexities in the budget constraint, for some h_j alternative we have

$$p_j(\mathcal{H}, Z, \beta) = 0,$$

for all (Z, β), while the data contains some observations of h_j.

To tackle this issue, we suppose that individuals do not make their choices over the whole set $\mathcal{H}$, but on a random subset of it. We analyse how this modified model works, and in particular the assumptions under which it still allows identifying the parameters β of the underlying structural model.

2.3.1. The two-offer model

Suppose individual *i* faces a distribution of offers, the probability of being offered h_i being equal to $g_i, g_i > 0, \sum_{i=1}^{I} g_i = 1$. Below we will allow the distribution of offers g_i, for any individual *i*, to depend on observable covariates X_i. First consider the case where individuals draw independently two offers from *g* and choose the one that yields the highest utility.[2] The distribution of the observed choices $\ell_{2i}(Z, \beta)$ (the first index '2' serves to mark that there are two offers) then takes the form

$$\ell_{2i}(Z, \beta) = g_i^2 + 2g_i \sum_{j \neq i} g_j p_i(\{i, j\}, Z, \beta), \tag{2}$$

---

[1] Other income corresponds to any resources that are not related to labour market earnings or benefits. In particular, it includes husband's income, if there is any. Net earnings correspond to labour market earnings of the woman at the observed hourly wage. Thus it increases with the number of hours worked.

[2] Note that as we do not observe past choices we cannot distinguish between an offer that allows the individual to retain their previous hours work rather than choose among completely new offers. In principle we can therefore allow individuals to be offered, and to choose to keep, their existing hours worked. Our assumptions though will imply that the distribution of offers is independent of the past hours worked. This will be an important extension for future work.
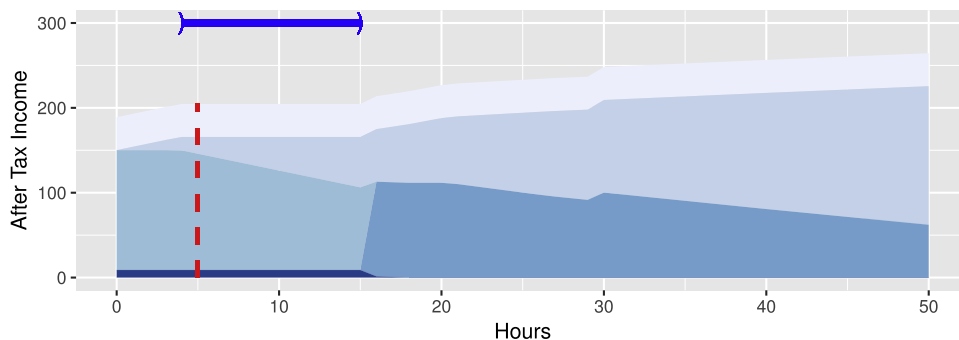
(a) Lone Mother in 1997 (1)
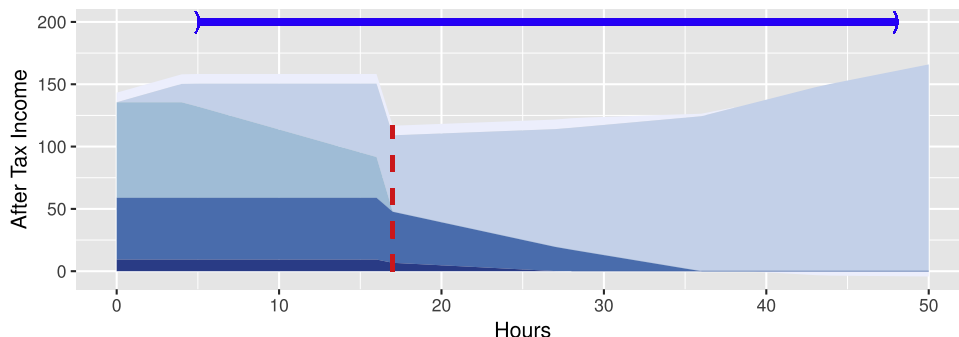
(b) Lone Mother in 1997 (2)
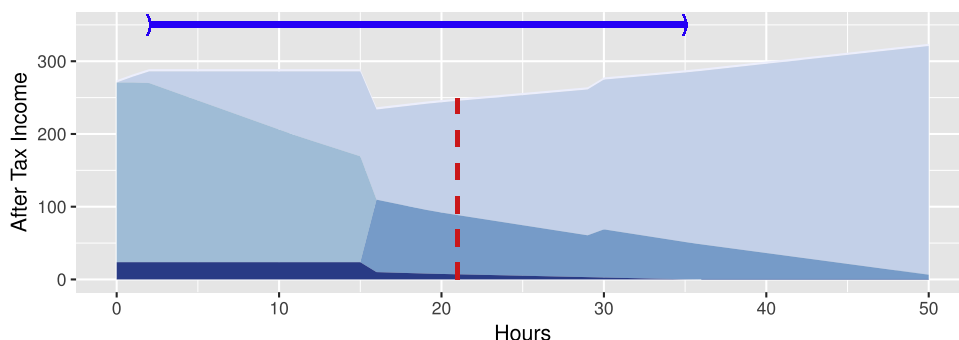
(c) Women with unemployed husband in 2002

Fig. 1. Some examples of budget constraints for in work people.
NOTES: The vertical red dotted lines correspond to the hours of work observed in the data, the horizontal thick blue lines at the top of the graphs correspond to the set of hours strictly dominated.
Source: Family expenditure survey 1997 and 2002; IFS-Taxben.

with the first term on the right hand side corresponding to identical offers (leaving no choice to the decision maker), and the second reflecting choices among all possible couples of offers.

There are I equations, of which only $I - 1$ are independent: the sum of all the equations is identically equal to 1 (on the right hand side, this follows from the observation that $p_i(\{i, j\}, Z, \beta) + p_j(\{i, j\}, Z, \beta) = 1$ for all i, j). On the right hand side, there are potentially $I(I-1)/2 + I - 1$ unknowns: the choice probabilities p and the distribution of offers g. There is no possibility to identify all these unknown parameters from the mere observation of the choice distribution ℓ. Below we explore alternative restrictions that deliver identification.

2.3.2. Increasing the number of offers

In the two-offer case, when the probability g has full support, the choice sets are all the pairs made from elements of $\mathcal{H}$, allowing repetitions. More generally the number of offers n determines the cardinality of the choice sets. If the draws are independent, for any finite n, there is a positive probability that there is no real choice: all the elements in the choice set are identical. However when n increases, this probability goes to zero and more importantly the probability that the choice set contains all the elements of $\mathcal{H}$ goes to one. The n-offer model converges towards the standard unrestricted choice model as n goes to infinity.

3. Identification

3.1. Recovering choices, knowing the offer distribution

Even if the offer distribution g is given, the number of unrestricted choice probabilities among pairs a priori is $I(I-1)/2$, larger than $I-1$ for I greater than 2. We have to restrict the number of structural unknowns, imposing consistency requirements across pairs. We consider a random utility model as this is close to our labour supply model with discrete hours. In Appendix B we also derive our identification results for the independence of irrelevant alternatives (IIA) model.

The agent has utility $V_i - \varepsilon_i$ for alternative i, $i = 1, \ldots, I$, and under full optimization, knowing the value of her utilities, chooses the alternative which gives the highest utility.

The econometrician is supposed to know the joint distribution of the continuous variables ε_i in the economy, and wants to infer from observed hours choices the values of the parameters V_i. We denote F_{ij} the (assumed to be differentiable) cumulative distribution function of $\varepsilon_i - \varepsilon_j$ so that

$$p_i(\{i, j\}) = F_{ij}(V_i - V_j).$$

Since only the differences $V_i - V_j$ can be identified, we normalize V_I to zero. The number of unknowns is equal to the number of equations, and we may hope for exact identification. Indeed

Lemma 1. *Let ℓ and g be two probability vectors in the simplex of $\mathcal{R}^I$, whose components are all positive. There exists at most a unique vector V_i with $V_I = 0$ that satisfies the system of equations*

$$\ell_i = g_i^2 + 2g_i \sum_{j \neq i} g_j F_{ij}(V_i - V_j) \text{ for } i = 1, \ldots, I. \tag{3}$$

Proof. For all $i = 1, \ldots, I - 1$, denote

$$Q_i(V) = -\ell_i + g_i^2 + 2g_i \sum_{j \neq i} g_j F_{ij}(V_i - V_j),$$

and $Q(V)$ the $I - 1$ vector obtained by stacking up the Q_is. By construction, any V such that $Q(V) = 0$ satisfies (3), since the Ith equation follows from summing up the $I - 1$ first ones.

The result then follows from Gale Nikaido since the Jacobian of Q is everywhere a dominant diagonal matrix. Indeed

$$\frac{\partial Q_i}{\partial V_i} = 2g_i \sum_{j \neq i} g_j f_{ij}(V_i - V_j),$$

while for $j \neq i, j \neq I$,

$$\frac{\partial Q_i}{\partial V_j} = -2g_i g_j f_{ij}(V_i - V_j).$$

The diagonal terms are positive and the off-diagonal negative. The sum of the elements on line i is positive equal to

$$2g_i g_I f_{iI}(V_i). \quad \square$$

3.2. Recovering the offer distribution, knowing choice probabilities

In contrast to the previous section, assume that we know the theoretical choice probabilities over all pairs of alternatives: p_{ij} denotes the probability of choosing i when both i and j are available. We study whether the choices ℓ_i of agents getting two independent offers are constrained by the model, and whether the observation of ℓ allows recovering the probability of offers g. From (2), we have by definition

$$\ell_i = g_i^2 + 2g_i \sum_{j \neq i} g_j p_{ij} \tag{4}$$

where for all couples (i, j), $i \neq j$,

$$p_{ij} + p_{ji} = 1. \tag{5}$$

Lemma 2. *Given the choice probabilities p_{ij}, $p_{ij} \geq 0$ satisfying* (5), *for any observed probability ℓ_i in the simplex of $\mathcal{R}^I$, there exists a unique offer probability g_i in the simplex of $\mathcal{R}^I$ which satisfies* (4).

Proof. We first prove the existence of g, then its uniqueness. For all i, define

$$\overline{Q}_i(g) = g_i^2 + 2g_i \sum_{j \neq i} g_j p_{ij}.$$

By construction, for g in the simplex of $\mathcal{R}^I$, under (5), $\overline{Q}(g)$ also belongs to the simplex of $\mathcal{R}^I$. Indeed

$$\sum_{i=1}^{I} \overline{Q}_i(g) = \sum_{i=1}^{I} \left[g_i^2 + 2g_i \sum_{j < i} g_j \right] = 1.$$

Consider the mapping

$$\Gamma_i(g) = \frac{\max(0, g_i - \overline{Q}_i(g) + \ell_i)}{\sum_{j=1}^{I} \max(0, g_j - \overline{Q}_j(g) + \ell_j)}.$$

First note that Γ is well defined: since g, $\overline{Q}$ and ℓ all belong to the simplex, the denominator is larger than 1. Therefore Γ maps continuously the simplex into itself and it has a fixed point, say g^*. If $g_i^* = 0$, by definition $\overline{Q}_i(g^*) = 0$, so that

$$g_i^* - \overline{Q}_i(g^*) + \ell_i = \ell_i.$$

It follows that at the fixed point

$$\max(0, g_i^* - \overline{Q}_i(g^*) + \ell_i) = g_i^* - \overline{Q}_i(g^*) + \ell_i,$$

the denominator is equal to 1, and $\ell = \overline{Q}(g^*)$ as desired.

Uniqueness follows from the univalence of $\overline{Q}$. This is a consequence of the fact that the Jacobian of $\overline{Q}$ is a dominant diagonal matrix, with weights (g_i): for all i

$$g_i \frac{\partial \overline{Q}_i}{\partial g_i} > \sum_{j \neq i} g_j \frac{\partial \overline{Q}_i}{\partial g_j}.$$

Indeed

$$g_i \left[2g_i + 2 \sum_{j \neq i} g_j p_{ij} \right] > \sum_{j \neq i} 2g_j g_i p_{ij}. \quad \square$$

3.3. Recovery of choice and offer probabilities

In general we will neither have prior knowledge of the theoretical choice probabilities p_{ij} nor of the offer probabilities g_i. Without assumptions it follows from Lemmas 1 and 2, that the choice probabilities and the offer probabilities will not be separately identified.

In order to establish identification, we first restrict the choice and offer probabilities to be functions of a finite set of parameters. In a second step, we use exclusion restrictions from the budget constraint to deliver semiparametric identification of the offer probabilities. Finally we investigate the role played by dominated regions of the constraint. We illustrate the finite sample properties of our estimators by Monte-Carlo simulations in Appendix C.

3.3.1. Parametric identification

In this first case we assume the utility from choice h_i is given by a function of a finite set of preference parameters β and individual unobserved heterogeneity ε: $V_i(\beta, \varepsilon) = V(h_i, \beta, \varepsilon)$, and the probability that hours h_i are chosen when the pair (h_i, h_j) is available is given by

$$p_{ij}(\beta) = \Pr[V_i(\beta, \varepsilon) - V_j(\beta, \varepsilon) > 0].$$

To make progress with identification in this case we assume that the offer probability g_i is a smooth function of a finite parameter vector γ, and that

$$\dim[\beta : \gamma] \leq I - 1,$$

where I is the number of possible choices.

From (3) we can write the system of equations

$$Q_i = -\ell_i + g_i(\gamma)^2 + 2g_i(\gamma) \sum_{j \neq i} g_j(\gamma) p_{ij}(\beta) \text{ for } i = 1, \ldots, I - 1. \tag{6}$$

Table 1
Some descriptives statistics.
Source: Family expenditure survey 1997–2002. The sample comprises women with children, either single or married mothers. The first line gives the sample size, and the remainder is expressed in percentage of the sample.

	Sample	1997	1998	1999	2000	2001	2002
Size	10 575	1704	1734	1812	1671	1918	1736
Education							
Level 1	54.82	56.92	56.29	53.81	54.28	55.58	52.02
Level 2	26.55	25.65	26.12	28.42	26.75	25.34	27.07
Level 3	18.63	17.43	17.59	17.77	18.97	19.08	20.91
Number of children							
1	35.88	38.38	32.93	37.2	36.03	35.35	35.43
2	43.45	43.54	44.58	40.56	42.01	44.58	45.39
3	15.22	13.67	16.55	16.11	16.94	13.97	14.23
4	3.96	3.17	4.44	4.08	3.95	4.28	3.8
5+	1.48	1.23	1.5	2.04	1.08	1.82	1.15
Age of the youngest child							
Between 0 and 4	40.99	43.31	43.25	40.73	39.44	39.47	39.92
Between 5 and 10	33.82	32.81	32.41	33.33	34.05	34.62	35.6
London	9.37	10.09	9.34	9.6	8.74	9.7	8.7
Cohabitant	77.86	78.46	79.07	76.21	78.7	77.95	76.9
Spouses inwork status							
Women inwork	63.23	62.09	60.67	61.81	63.73	65.02	65.96
Spouse in work	69.51	70.01	69.38	67.77	70.14	70.13	69.7
Both inwork	50.01	49.71	49.37	48.84	50.27	50.52	51.38
Both out of work	17.27	17.61	19.32	19.26	16.4	15.38	15.73

For identification we require full column rank of the matrix

$$\Pi = \left[\frac{\partial Q}{\partial \beta}, \frac{\partial Q}{\partial \gamma} \right] = \left[\frac{\partial Q}{\partial p} \frac{\partial p}{\partial \beta}, \frac{\partial Q}{\partial g} \frac{\partial g}{\partial \gamma} \right]. \tag{7}$$

where the matrix of derivatives relating to the Q_i has elements of the form

$$\frac{\partial Q_i}{\partial p_{ij}} = 2g_i g_j \tag{8}$$

$$\frac{\partial Q_i}{\partial g_i} = 2g_i + 2\sum_{j\neq i} g_j p_{ij}(\beta) \tag{9}$$

$$\frac{\partial Q_i}{\partial g_j} = 2g_i p_{ij}(\beta) \tag{10}$$

So we have:

$$\frac{\partial Q_i}{\partial \beta} = 2\sum_{j\neq i} g_i g_j \frac{\partial p_{ij}}{\partial \beta} \tag{11}$$

$$\frac{\partial Q_i}{\partial \gamma} = 2\left(g_i + \sum_{j\neq i} g_j p_{ij}(\beta) \right) \frac{\partial g_i}{\partial \gamma} + 2\sum_{j\neq i} g_j p_{ij}(\beta)\frac{\partial g_j}{\partial \gamma} \tag{12}$$

Inspection of the elements of Π, (11) and (12), suggests no natural linear dependence and, in general, the rank condition is satisfied. In addition to that, note that any flat budget constraint would result in choice probabilities that do not depend on parameters, which would imply $\frac{\partial p_{ij}}{\partial \beta} = 0$ for some couples (i, j).

In any finite sample application we would require sufficient variation in the points of support. In Appendix C.2, we consider a simulation analysis with parametric choices for preferences and the offer distribution. The results (see Table 13) show that the parameter estimates are, in general, estimated with precision and with little bias. They also show that the shape of the budget constraint, combined with shape restrictions in preferences imposed by theoretical aspects of the model, plays an important role in the mean squared error of the estimates. As noted in Section 2.2, flat or decreasing budget constraints lead to dominated choices that allow to estimate the offer distribution g more precisely, but the loss of information on preference parameters leads to less precision of these estimates.

Finally, we can also see from the last panel of this illustration that the simultaneous presence of individuals facing different budget constraints helps to recover precise estimates of the offer distribution and the preference parameters. We now turn to exclusion restrictions that can deliver semi-parametric identification of the model.

Table 2
Descriptive statistics by level of labour supply.
Source: Family expenditure survey 1997–2002. The sample comprises women with children, either single or married mothers. The first line gives the sample size, and the remainder is expressed in percentage of the sample.

	Not working	Working 1 to 20 h	Working 21 to 34 h	Working 35 h and more
Size	3888	2424	1909	2354
Education				
Level 1	63.3	56.64	50.81	42.18
Level 2	21.89	26.98	30.54	30.59
Level 3	14.81	16.38	18.65	27.23
Number of children				
1	29.48	30.36	37.19	51.06
2	41.1	49.5	47.67	37.68
3	19.55	16.38	12.26	9.3
4	6.66	3.18	2.2	1.74
5+	3.22	0.58	0.68	0.21
Age of the youngest child				
Between 0 and 4	51.67	39.77	33.47	30.71
Between 5 and 10	30.56	38	37.14	32.2
London	11.75	6.89	7.23	9.73
Cohabitant	68.26	83.95	84.91	81.73
Spouses inwork status				
Women inwork	0	100	100	100
Spouse in work	53.03	80.07	80.83	76.68
Both inwork	0	80.07	80.83	76.68
Both out of work	46.97	0	0	0

Table 3
Consumption, Wages and hours of Work.
Source: Family expenditure survey 1997–2002. The sample comprises women with children, either single or married mothers. Each panel respectively provides descriptive statistics for consumption, wages, and hours. For each variable, the columns of the table show sample statistics: mean, standard deviation, and percentiles of the distribution. The rows describe the full sample and the six years from 1997 to 2002.

	Mean	St. Dev.	p1	p5	p10	p25	p50	p75	p90	p95	p99
Log of consumption											
All	5.3	0.6	3.85	4.25	4.49	4.93	5.33	5.68	6.02	6.23	6.74
1997	5.26	0.58	3.86	4.25	4.45	4.9	5.31	5.64	5.97	6.18	6.66
1998	5.24	0.58	3.88	4.21	4.44	4.88	5.26	5.63	5.95	6.14	6.56
1999	5.29	0.6	3.84	4.21	4.46	4.92	5.33	5.68	6	6.21	6.71
2000	5.33	0.6	3.87	4.3	4.53	4.96	5.36	5.71	6.05	6.28	6.8
2001	5.33	0.6	3.89	4.31	4.52	4.94	5.35	5.72	6.05	6.26	6.85
2002	5.33	0.61	3.82	4.3	4.54	4.96	5.35	5.72	6.1	6.29	6.78
Hourly wage											
All	7.24	4.75	1.84	2.89	3.31	4.14	5.85	9	12.71	15.5	25.39
1997	6.67	4.59	1.83	2.78	3.12	3.84	5.22	8.03	11.59	14.55	26.18
1998	6.61	4.33	1.6	2.59	3.02	3.78	5.34	8.14	12.02	14.51	22.33
1999	7.01	4.43	1.65	2.79	3.32	4	5.64	8.75	12.52	15.16	23.29
2000	7.27	4.17	2.23	3.02	3.36	4.27	6.12	9.25	12.52	15.15	21.03
2001	7.87	5.16	2.33	3.16	3.56	4.44	6.35	10.02	13.49	16.88	27.55
2002	7.87	5.41	1.91	3.21	3.65	4.55	6.43	9.54	13.51	16.33	31.08
Usual hours of work											
All	26	11	3	7	11	18	26	37	40	42	50
1997	25	12	3	6	10	16	25	36	40	42	50
1998	25	11	3	7	10	16	25	36	39	40	50
1999	26	11	4	7	11	17	26	37	40	42	50
2000	27	12	3	7	12	18	28	37	40	44	53
2001	27	11	4	8	12	18	27	37	40	42	51
2002	27	11	4	9	13	19	26	37	40	42	50

3.3.2. Semi-parametric identification using exclusion restrictions

We assume preferences depend on a finite parameter vector β and also that there are exclusion restrictions from the budget constraint. We place no restrictions on the offer probabilities g_i. We write utility from hours choice h as a function of an exogenous observed characteristic Z. In this framework, budget constraint heterogeneity is the result of such exclusion restrictions that have an impact on the budget constraint without altering the distribution of offers. Thus, any exogenous wage or other income variation would be a good candidate for Z. For any realization z of Z, we can write the likelihood

Table 4
Estimation Results for 1997–2002 years: Preferences.
Source: Family expenditure survey 1997–2002. The sample comprises women with children, either single or married mothers. Estimates were obtained by maximum-likelihood. Model 1 is the baseline model. Model 2 allows for correlation between wages, consumption and preferences. In addition to this, Model 3 includes exclusion variables in the offer distribution. Standard-errors are given in parentheses.

	Model 1	Model 2	Model 3
ϕ: yob > 1963	7.66 (0.661)	7.766 (0.822)	15.067 (1.9541)
ϕ: yob $\leq$ 1963	7.62 (0.6193)	7.298 (0.777)	14.402 (1.8481)
γ: yob > 1963	0.001 (0.0882)	0.002 (0.234)	0.003 (0.3945)
γ: yob $\leq$ 1963	0.001 (0.0798)	0.656 (0.2015)	1.112 (0.3284)
β^a: Const	31.671 (2.8479)	32.279 (3.9346)	62.139 (9.0152)
β^a: Cohab	0.5 (0.1077)	0.053 (0.1768)	−0.133 (0.2836)
β^a: Youngest Kid age 0–4	1.51 (0.1291)	1.721 (0.1751)	2.997 (0.4237)
β^a: Youngest Kid age 5–9	0.758 (0.1077)	0.898 (0.1282)	1.394 (0.2407)
β^a: yob $\leq$ 1963	0.154 (1.873)	−5.277 (2.1096)	−8.473 (3.2075)
β^a: 2 kids	0.849 (0.0944)	0.918 (0.1262)	1.446 (0.2542)
β^a: 3 kids	0.815 (0.1194)	0.847 (0.1513)	1.549 (0.31)
β^a: 4 kids or more	0.917 (0.1847)	1.008 (0.2263)	1.973 (0.4534)
$\rho(\varepsilon^a, \varepsilon^c)$	0 (.)	−0.339 (0.0554)	−0.342 (0.0505)
$\rho(\varepsilon^a, \varepsilon^w)$	0 (.)	−0.249 (0.0397)	−0.323 (0.0396)
σ_a	1.376 (0.0838)	1.542 (0.1308)	2.881 (0.3824)
β^b: Const	36.356 (4.2809)	35.155 (4.8589)	38.335 (3.7302)
β^b: Cohab	−18.013 (3.371)	−18.956 (3.8128)	−10.576 (2.7693)
β^b: Youngest Kid age 0–4	17.837 (4.3898)	18.274 (4.8321)	10.564 (3.7806)
β^b: Youngest Kid age 5–9	−5.498 (3.6711)	−6.976 (4.0496)	−5.424 (3.159)
β^b: yob $\leq$ 1963	−2.101 (3.5497)	−1.746 (3.9668)	−1.487 (3.0314)
β^b: 2 kids	−11.332 (3.0743)	−12.615 (3.4264)	−10.609 (2.6814)
β^b: 3 kids	−1.448 (4.0398)	0.426 (4.4764)	−3.592 (3.4936)
β^b: 4 kids or more	19.936 (6.1746)	24.471 (6.9246)	12.735 (5.183)
β^b: London	33.304 (3.9481)	35.815 (4.3995)	25.743 (3.5809)
σ_b	67.142 (2.9846)	74.749 (3.4361)	53.036 (2.1681)
Lik	−3.26528	−3.26031	−3.25496
N	10 575	10 575	10 575

function:

$$\ell_i(z) = g_i^2 + 2g_i \sum_{j \neq i} g_j p_{ij}(z, \beta)$$

A necessary condition for joint identification of choice and offer probabilities is that there is enough variation in choice probabilities due to Z. It can easily be illustrated in the case of a discrete Z taking K distinct values. In such case, the number of distinct likelihood contributions is equal to $(I - 1) \times K$, and the number of unknowns is equal to $I - 1 + \dim(\beta)$. Thus, low variation in Z would fail identification.

The second set of Monte-Carlo simulations presented in Appendix C.3 illustrates this case. In these simulations, the offer distribution is unrestricted, but the utility of hours depends on a finite set of parameters. We observe that as soon as the data generating process includes more than a unique linear budget constraint, the parameters estimates are reasonably distributed around the true value (see Figs. 13 and 14). From Table 14, we see that the inclusion of an additional linear budget constraint dramatically reduces the mean-squared-error, and the decomposition of the mean-squared error shows that this reduction is mainly driven by a reduction in the bias, although some reduction in the variance is also observed.

3.3.3. Nonparametric identification using dominated regions of the budget constraint

Finally, we show that the distribution of offers is identified where the data contains cases of flat or decreasing budget constraints. For a particular z, shape restrictions would reveal preference probabilities regardless of the value of the

parameters:

$$\forall \beta, \ i < j : \ p_{ij}(z, \beta) = 1$$

If we are able to isolate households with this property, Lemma 2 applied to this subpopulation identifies the distribution of offers. Then with this distribution of offers, Lemma 1 allows to identify the random utilities from women facing increasing budget constraints.

4. Data, model specification and sample likelihood

4.1. The data

The sample of families we use comprises women with children, either single or married mothers. We use years 1997 to 2002 of the UK Family Expenditure Survey (FES) as this covers the period of key reforms to the welfare and tax-credit system in Britain, see Adam et al. (2010). The data provide detailed diary and face to face interview information on consumption expenditures, usual hours worked, gross wage earnings, education and household demographics. Tables 1 and 3 provide some basic descriptive statistics.

The final sample we use contains some 10,575 women spread fairly evenly across the six years under study. As documented in Table 1, a large group of women in this sample have relatively low education, meaning that they left formal schooling at the minimum school leaving age of 16 (Level 1). The majority of the rest have completed secondary school (Level 2) with less than 20% having a college or university degree (Level 3). The modal number of children is two and a little more than 40% of the sample have a youngest child aged less than 5 (the formal school entry age in the UK). Almost 78% of the women in our sample are married or cohabiting (we label all these as 'cohabiting'), leaving just over 22% of the mothers in the sample as single parents.

Table 2 describes the relation between observed labour supply and different covariates. First, we can see that, compared to the distribution of the whole sample presented in Table 1, women who are not working are less educated (63.3% of non-working women have the lowest level of education vs. 54.8% in the whole sample). Women with 3 children or more are also over-represented among non-working women, as well as women with a very young child, women living in London, and non-cohabiting women.

The last three columns of Table 2 compare working women with different levels of hours. If hours were independent of the characteristics the column entries would be identical. Women with more children, and women with a child aged 0 to 4 tend to work fewer hours. In contrast, despite the negative association with the extensive margin, living in London has a positive impact on hours of work. This pattern accords with fixed costs of working in London, incorporated in the structural model estimates below. Cohabiting is also associated with different intensive and extensive patterns since 83.95% of women working less than 20 h are cohabiting, whereas only 81.73% of women working more than 35 h are cohabiting. The same type of pattern appears for women whose youngest child is aged between 5 and 10 years old: they tend to be more often working than the rest of women, but work less when working.

Table 3 presents the distribution of log real consumption, real hourly wages and usual hours of work. For example, the median wage is £5.85 per hour and the average wage is £7.24 per hour (all in 1997 prices). The median usual hours of work is 26 h per week with a wide distribution.

4.2. Model specification

Preferences over consumption and hours of work are given by the unitary utility function:

$$u(c, h) = \frac{c^{1-\gamma}}{1-\gamma} + \frac{(L-h)^{1-\phi}}{1-\phi} a, \tag{13}$$

where $L (= 100)$ is a physiological upper bound on the number of hours worked weekly, γ and ϕ are non-negative parameters, and the strictly positive factor a governs the substitution between consumption and leisure. The disutility of labour is

$$v(h) = \frac{(L-h)^{1-\phi}}{1-\phi}.$$

With this specification, the Lagrangian of Section 2.1 is

$$\mathcal{L}(c, h, w, \lambda) = u(c, h) + \lambda [R(w, h) - b1\{h > 0\} - c] \tag{14}$$

The first order condition in c gives

$$\lambda = c^{-\gamma}.$$

The a factor has the form:

$$\ln(a) = Z^a \beta^a + \sigma^a \varepsilon^a, \tag{15}$$

Table 5
Estimation Results for 1997–2002 years: Offer Distribution.
Source: Family expenditure survey 1997–2002. The sample comprises women with children, either single or married mothers. Estimates were obtained by maximum-likelihood. Model 1 is the baseline model. Model 2 allows for correlation between wages, consumption and preferences. In addition to this, Model 3 includes exclusion variables in the offer distribution. Standard-errors are given in parentheses.

	Model 1	Model 2	Model 3
m_1: Const	14.673 (1.0425)	15.19 (0.8551)	13.85 (1.1455)
σ_1: Const	17.564 (0.7718)	16.591 (0.6333)	18.811 (0.8301)
m_2	37.869 (0.066)	37.839 (0.0656)	37.815 (0.0644)
σ_2	1.67 (0.0599)	1.671 (0.0595)	1.661 (0.0584)
p_1: Const	0.834 (0.007)	0.836 (0.0067)	0.891 (0.0098)
p_1: Edu 2	1 (.)	1 (.)	0.543 (0.0527)
p_1: Edu 3	1 (.)	1 (.)	0.344 (0.0346)
p_1: London	1 (.)	1 (.)	0.946 (0.1284)
p_1: year 99–00	1 (.)	1 (.)	0.915 (0.0901)
p_1: year 01–02	1 (.)	1 (.)	0.953 (0.0937)

where Z^a contains observable characteristics, while ε^a stands for unobserved preference heterogeneity. We also posit the following stochastic specification for the fixed cost of being employed

$$b = Z^b \beta^b + \sigma^b \varepsilon^b \tag{16}$$

where ε^b reflects unobserved heterogeneity in work costs across individuals.

Log market wages are specified as a function of covariates reflecting human capital accumulation Z^w:

$$\ln(w) = Z^w \beta^w + \sigma^w \varepsilon^w. \tag{17}$$

Since the consumption process is external to hours and employment choices we specify the reduced form for log consumption

$$\ln(c) = Z^c \beta^c + \sigma^c \varepsilon^c. \tag{18}$$

The residuals $(\varepsilon^a, \varepsilon^b, \varepsilon^c, \varepsilon^w)$ are assumed to be joint normally distributed and independent from Z^a, Z^b, Z^c and Z^w. We work under additional assumptions and exclusion restrictions.

First, baseline covariates that drive preference for leisure, Z^a, include a dummy for cohabiting, the age of the youngest kid, the number of children, and a cohort effect. Second, we assume that in addition to that, the fixed cost of working depends on whether the household is located in London, and that the unobserved component, ε^b, is independent from the other dimensions of unobserved heterogeneity, $\varepsilon^a, \varepsilon^c$, and ε^w. Third, covariates in the wage equation are chosen to reflect human capital accumulation: they include age and education in a flexible way, and the number of kids as it is likely to be correlated with work interruptions. We also include year dummies as well as a dummy for living in London to control for geographic and temporary variations in wages. Finally, the consumption equation includes all variables used in the wage equation with additional controls for family composition likely to influence household total consumption: whether the woman is cohabiting or not, and a set of dummy variables for the age of the youngest child. Thus, identification of the correlation between ε^w and ε^a is mainly driven by variables related to human capital, and time and spatial variations of wages, the correlation between ε^w and ε^c by household structure variables, and the correlation between ε^c and ε^a by a combination of both types of exclusion restrictions.

Offers are modelled as a mixture of two independent normal distributions of mean m_k and standard deviation σ_k. We truncate these distributions at 0 and 66 h and denote their cdf by $\tilde{\Phi}(., m_k, \sigma_k)$.[3] The mixture probability p_1 depends on observable exogenous covariates (See Table 5). The cdf of the offer distribution is:

$$p(h|Z^o) = \tilde{\Phi}(h, m_1, \sigma_1)p_1(Z^o) + \tilde{\Phi}(h, m_2, \sigma_2)(1 - p_1(Z^o)).$$

We use a discretized version of this distribution:

$$g(h|Z^o) = p(h + 1|Z^o) - p(h|Z^o).$$

Identification would probably be achieved by the joint normality of the error terms, but is also secured by the various exclusion restrictions.

---

[3] The formula for the truncated normal is

$$\tilde{\Phi}(x, m_k, \sigma_k) = \frac{\Phi(x, m_k, \sigma_k)}{\Phi(66, m_k, \sigma_k) - \Phi(0, m_k, \sigma_k)} 1\{x \in [0, 66]\}$$

4.3. Sample likelihood for the two-offer model

We derive the distribution of employment, hours, consumption and wages from the model assumptions. To allow for the endogeneity of consumption in the determination of hours, employment and wages, we need to account for the dependence between the reduced form error in (18) and the other error terms of the model in (15), (16) and (17). We adopt a two-step control function approach (see Blundell and Powell (2003)) in which the reduced form parameters for the consumption equation are estimated in an initial step.[4] At the second step the estimated error term ε^c is added as additional regressor in (15), (16), and (17). The estimated parameters in the conditional likelihood for wage, hours and employment are then adjusted to account for this initial estimation.

To formulate the likelihood we require an expression for the probability of being employed and choosing working hours h conditional on (c, w). Recall, the probability of being offered a job with working hours h is given by $g(h|Z^o)$. The probability of receiving a couple of offers (h, h'), $h \neq h'$, is $2g(h|Z^o)g(h'|Z^o)$, while that of receiving (h, h) is $g(h|Z^o)^2$.

First, consider the employment status. Assume $(\varepsilon^a, \varepsilon^c, \varepsilon^w)$ are known, i.e. consumption, wage and the parameter a. At weekly hours h an individual is observed employed when

$$av(h) + c^{-\gamma}[R(w, h) - b] > av(0) + c^{-\gamma}R(w, 0),$$

or

$$b < R(w, h) - R(w, 0) + ac^\gamma[v(h) - v(0)].$$

From the expression for fixed costs of work b, the probability of this event knowing a is easily computed from the cumulative distribution of ε^b:

$$F^b(\varepsilon^a, c, h, w) = \Phi\left[\frac{R(w, h) - R(w, 0) + c^\gamma[v(h) - v(0)]\exp(Z^a\beta^a + \sigma^a\varepsilon^a) - Z^b\beta^b}{\sigma^b}\right].$$

When the individual would like to work she can choose from two offers h and h'. Offer h is preferred to offer h' when

- either h is larger than h' and

$$a \leq c^{-\gamma}\frac{R(w, h) - R(w, h')}{v(h') - v(h)}$$

    which can be written equivalently

$$\varepsilon^a \leq \alpha(c, h, h', w) = \frac{1}{\sigma^a}\left\{-\gamma \ln c + \ln\left[\frac{R(w, h) - R(w, h')}{v(h') - v(h)}\right] - Z^a\beta^a\right\},$$

- or h is smaller than h' and

$$c^{-\gamma}\frac{R(w, h') - R(w, h)}{v(h) - v(h')} \leq a,$$

    which is also

$$\frac{1}{\sigma^a}\left\{-\gamma \ln c + \ln\left[\frac{R(w, h') - R(w, h)}{v(h) - v(h')}\right] - Z^a\beta^a\right\} = \alpha(c, h, h', w) \leq \varepsilon^a.$$

The probability of being employed and choosing h, conditional on (c, w), is therefore

$$\ell(h|c, w) = g(h|Z^o)^2 \int_{-\infty}^{+\infty} F^b(\varepsilon, c, h, w)\phi(\varepsilon)d\varepsilon$$

$$+ g(h|Z^o)\sum_{h' < h} 2g(h')\int_{-\infty}^{\alpha(c, h, h', w)} F^b(\varepsilon, c, h, w)\phi(\varepsilon)d\varepsilon$$

$$+ g(h|Z^o)\sum_{h' > h} 2g(h')\int_{\alpha(c, h, h', w)}^{+\infty} F^b(\varepsilon, c, h, w)\phi(\varepsilon)d\varepsilon$$

To get the complete likelihood, we multiply the previous expression by the density of the wage conditional on consumption. Given the joint normality assumption we have:

$$\ell(w|c) = \frac{1}{\sqrt{\sigma_w^2(1 - \rho_{wc}^2)}}\varphi\left(\frac{\log w - Z^w\beta^w - \sigma^w\rho_{wc}\varepsilon^c/\sigma^c}{\sqrt{\sigma_w^2(1 - \rho_{wc}^2)}}\right)$$

---

[4] These parameters were obtained from a simple linear regression of log consumption on Z^c. The results of this first step are given in Table 12 in Appendix A.

Table 6
Estimation Results for 1997–2002 years: Wage Equation.
Source: Family expenditure survey 1997–2002. The sample comprises women with children, either single or married mothers. Estimates were obtained by maximum-likelihood. Model 1 is the baseline model. Model 2 allows for correlation between wages, consumption and preferences. In addition to this, Model 3 includes exclusion variables in the offer distribution. Standard-errors are given in parentheses.

	Model 1	Model 2	Model 3
β^w: Const	−4.356 (0.8335)	−4.403 (0.8365)	−4.79 (0.8403)
β^w: year 98	−0.031 (0.0185)	−0.031 (0.0185)	−0.033 (0.0186)
β^w: year 99	0.015 (0.0183)	0.014 (0.0183)	0.013 (0.0184)
β^w: year 00	0.064 (0.0188)	0.064 (0.0188)	0.062 (0.0189)
β^w: year 01	0.126 (0.0176)	0.127 (0.0176)	0.126 (0.0178)
β^w: year 02	0.107 (0.0183)	0.106 (0.0183)	0.103 (0.0185)
β^w: Age	1.106 (0.6056)	1.118 (0.6076)	1.277 (0.6097)
β^w: Age^2	−0.234 (0.1586)	−0.234 (0.1591)	−0.267 (0.1597)
β^w: Age^3	0.012 (0.0135)	0.012 (0.0136)	0.015 (0.0136)
β^w: Edu	0.355 (0.0342)	0.358 (0.0342)	0.372 (0.0345)
β^w: Edu^2	−0.793 (0.0808)	−0.784 (0.0807)	−0.803 (0.0813)
β^w: Edu $\times$ Age	0.136 (0.0337)	0.121 (0.0338)	0.111 (0.034)
β^w: London	0.145 (0.0176)	0.148 (0.0177)	0.147 (0.018)
β^w: 2 kids	−0.045 (0.0121)	−0.045 (0.0121)	−0.046 (0.0121)
β^w: 3 kids	−0.127 (0.0163)	−0.128 (0.0164)	−0.132 (0.0164)
β^w: 4 kids or more	−0.204 (0.031)	−0.204 (0.0311)	−0.214 (0.0311)
$\rho(\varepsilon^c, \varepsilon^w)$	0.634 (0.0198)	0.62 (0.0219)	0.632 (0.0286)
σ_w	0.571 (0.0076)	0.563 (0.0076)	0.573 (0.0078)

Finally the probability of being out of employment at a given wage w in the two offer model is obtained by summing over all the couples (h, h'), the probability of preferring not to work. Let us define the probability of not working conditional on offers h and h', and unobserved heterogeneity ε:

$$P(0|c, w, h, h', \varepsilon) = \Phi\left[\frac{1}{\sigma^b}\left(R(w, 0) + c^\gamma v(0)\exp(Z^a\beta^a + \sigma^a\varepsilon) + Z^b\beta^b\right.\right.$$

$$\left.\left. - \max\left\{R(w, h) + c^\gamma v(h)\exp(Z^a\beta^a + \sigma^a\varepsilon), R(w, h') + c^\gamma v(h')\exp(Z^a\beta^a + \sigma^a\varepsilon)\right\}\right)\right]$$

Then we have the contribution of non workers conditional on wage and consumption:

$$\ell(0|c, w) = \sum_h \sum_{h'} g(h|Z^o)g(h'|Z^o)\int_{-\infty}^{\infty} P(0|c, w, h, h', \varepsilon)\phi(\varepsilon)d\varepsilon.$$

Since the wage of non-workers is not known, w has to be integrated out in the above expression.

5. Empirical results

We first present the estimates of the model parameters.

5.1. Parameter estimates for the two-offer model

Table 4 presents the estimation results for the parameters of preferences and fixed costs. The offer distribution parameters are presented in Table 5. Table 6 presents the estimates of the wage equation. Three different models were estimated. The first column presents the estimates of the baseline model (Model 1) in which we treat wages and consumption as exogenous in the determination of hours and employment. That is we set the correlation between the reduced form errors and the unobservable preference errors to zero. The baseline specification also excludes covariates from the specification of the offer distribution.

The ϕ and γ parameters refer to the exponents on hours (non-market work) and on consumption as described in the utility specification (13) of Section 4.2. We let these parameters vary with the cohort of birth of women.

Table 7
Model fit: hours, employment and wages.
Source: Family expenditure survey 1997–2002. The sample comprises women with children, either single or married mothers. Estimates were obtained by maximum-likelihood. Model 1 is the baseline model. Model 2 allows for correlation between wages, consumption and preferences. In addition to this, Model 3 includes exclusion variables in the offer distribution. Simulation is based on maximum-likelihood estimates.

	Data	Model 1	Model 2	Model 3
$\bar{\bar{E}}$	0.63	0.63	0.63	0.62
Hours				
$\bar{h}\|E = 1$	26.39	26.24	26.21	26.39
$sd(h\|E = 1)$	11.35	11.42	11.51	11.43
$h\|E = 1$, p 25%	18.00	17.00	17.00	18.00
$h\|E = 1$, p 50%	26.00	27.00	27.00	27.00
$h\|E = 1$, p 75%	37.00	36.00	36.00	36.00
Wages				
$\overline{\log(w)}\|E = 1$	1.82	1.86	1.87	1.88
$sd(\log(w)\|E = 1)$	0.54	0.63	0.62	0.62
$\log(w)\|E = 1$, p 25%	1.42	1.44	1.45	1.46
$\log(w)\|E = 1$, p 50%	1.77	1.84	1.85	1.85
$\log(w)\|E = 1$, p 75%	2.20	2.28	2.28	2.29
Joint distribution				
$\bar{h}\|\log(w) \leq 1.5$	21.63	23.32	22.82	23.56
$sd(h\|\log(w) \leq 1.5)$	11.01	11.55	11.62	11.41
$\bar{h}\|\log(w) \in]1.5, 2]$	27.60	26.16	26.07	26.11
$sd(h\|\log(w) \in]1.5, 2])$	10.88	11.31	11.38	11.33
$\bar{h}\|\log(w) > 2$	29.40	28.32	28.62	28.48
$sd(h\|\log(w) > 2)$	10.72	10.96	10.93	11.09
$corr(\log(w), h)$	0.23	0.19	0.22	0.19

The next panel refers to the parameters that influence the marginal utility of hours through the specification of $\ln(a)$ in Eq. (15). We find that cohabiting women have a higher preference for leisure, and that this preference is also higher when the youngest kid is younger. Following these are the fixed costs parameters of Eq. (16). The fixed cost for the reference category of lone mothers who has one kid aged more than 10 years old, and who lives out of the London region is about £36 a week. Cohabitation lowers this cost by £18, and the cost increases with the number of children. Living in London also increases the cost of working by more than £33 per week.

The comparison of the impact of covariates on intensive and extensive margins can be related to the description of the data in Section 4.1. Cohabiting women face a lower fixed cost but have a higher utility from leisure, which is compatible with the fact that they are participating more but provide shorter hours. The same applies for women whose youngest kid is aged between 5 and 9 years old.

For the two-offer specification of the restricted choice model, described in Section 4.3, offers are modelled as a mixture of two independent normal distributions. The associated parameter estimates are presented in Table 5. These estimates suggest offers concentrated at full-time (around 38) hours and having a mode at part-time (around 15) hours.

The second column (Model 2) presents the results for a model in which we allow correlation between unobserved heterogeneity terms in preferences, wages (Eq. (17)) and consumption (Eq. (18)). The correlation between consumption and preferences, $\rho(\varepsilon^a, \varepsilon^c)$ is significantly different from zero. Contrary to what was found in model 1, we find a significant increase in the γ parameter for the elder cohort of women born before 1963. Other parameters are qualitatively similar to those of the baseline model.

The last column of Tables 4 and 5 shows the results for a model in which we let the distribution of offers depend on three additional covariates: education, living in London and year dummies. These variables enter the mixture parameter of the offer distribution. The last panel of Table 5 shows the odd-ratios of the mixture parameter with respect to each of the variables. The more educated, the more likely women are to receive an offer from the higher (full-time) mixture, but we find no statistically significant difference along location and years. Accounting for this heterogeneity also affects the estimates of preferences and fixed cost parameters. From that specification, we see a higher income effect in particular for the older cohort. Moreover, the results show a stronger negative correlation between unobserved heterogeneity terms, and a stronger preference for leisure for women who have young kids.

5.2. Model fit

Table 7 summarizes the employment and distribution of hours obtained from simulation of the two-offer model. All three models predict employment with accuracy and do particularly well in replicating the twin peaks of the actual hours distribution. Fig. 2 plots the simulated hours distributions against the actual hours distribution. As expected, we see that
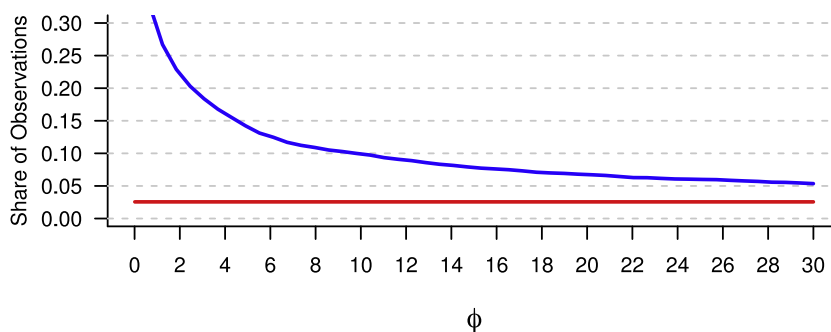
(a) Model 1   (b) Model 2   (c) Model 3

Fig. 2. Hours distributions: observed and two offers model.
Source: Family expenditure survey 1997–2002. The sample comprises women with children, either single or married mothers. Estimates were obtained by maximum-likelihood. Model 1 is the baseline model. Model 2 allows for correlation between wages, consumption and preferences. In addition to this, Model 3 includes exclusion variables in the offer distribution. Simulation is based on maximum-likelihood estimates.

Table 8
Estimation Results for 1997–2002 years: Rejection of the Unconstrained Model.
Source: Family expenditure survey 1997–2002. The sample comprises women with children, either single or married mothers. Estimates were obtained by maximum-likelihood. Model 1 is the baseline model. Model 2 allows for correlation between wages, consumption and preferences. In addition to this, Model 3 includes exclusion variables in the offer distribution.

	Model 1	Model 2	Model 3
NP-Reject (%)	2.56	2.56	2.56
ϕ-Reject (%)	11.02	11.14	7.93

Model 3, allowing for observable heterogeneity in the distribution of offers, fits the distribution of hours better than the first two specifications.

Table 7 also compares the mean and variance of the log-wage distribution of the data to the one simulated by the model. The lower panel of Table 7 gives a description of the joint distribution of hours and wages. The fit is less precise than the one of unconditional moments but it shows that our model is able to reproduce the positive correlation between hours and wages that is observed in the data.

5.3. Observations rejecting the unrestricted model

From the data we find that about 2.6% (see Table 8) of working women are observed working at hours that belong to the set of irrational hours H^{ir}. For this group we can reject the unrestricted choice model as there are alternative hours of work that strictly dominate the observed choices. This is a nonparametric rejection of the unrestricted choice model in the sense that the rejection does not depend on the specification, provided the utility function is increasing in consumption and leisure. The actual budget constraints for some of the individuals in this rejection group were used in Fig. 1.

In addition to these 2.6%, we observe 0.4% of working women who would earn more by staying out of employment. Again these observations reject the model whenever the utility function has the usual monotonicity properties and the fixed cost of work is non-negative.

From the specification of the model, we can also quantify the share of individuals rejecting the unconstrained model at the particular value of the parameter estimates. To do so, we need to determine women for whom the likelihood of observed hours is zero under the unconstrained model.

We start from the revealed preference inequality in the unrestricted case:

$$\frac{(L - h^e)^{1-\phi}}{1-\phi} a + c^{-\gamma} R(w, h^e) - \frac{(L - h)^{1-\phi}}{1-\phi} a - c^{-\gamma} R(w, h) \geq 0,$$

where h^e is the observed choice and h is any other possible length of the workweek. Using the specification for a, we can separate the cases where h is smaller than h^e from those where h is larger than h^e. That is

$$c^\gamma a \leq \left\{ \frac{R(w, h) - R(w, h^e)}{\frac{(L-h^e)^{1-\phi}}{1-\phi} - \frac{(L-h)^{1-\phi}}{1-\phi}} \right\}, \tag{19}$$

Fig. 3. Rejection of the unrestricted model according to ϕ.

Table 9
Observations rejecting the unrestricted choice model.
Source: Family expenditure survey 1997–2002. The sample we use comprises women with children, either single or married mothers. Computation based on the estimates of Model 3.

	Observations …		Type of rejection	
	…not rejecting the unrestricted model	…rejecting the unrestricted model	Parametric	Non parametric
Proportion among 'in work' women	0.92	0.08	0.05	0.03
Age at end of studies	17.52	16.64	16.80	16.29
Age	37.42	36.42	36.61	36.02
Hourly wage	7.40	5.33	5.86	4.21
Marginal tax rate	2.33	−1.16	0.55	−4.75
Usual weekly hours	26.93	20.18	21.87	16.63
Log of consumption	5.48	4.99	5.04	4.89
Number of kids	1.78	1.92	1.87	2.02
A kid younger than 4	0.35	0.29	0.28	0.30
The youngest kid between 5 and 10	0.35	0.41	0.41	0.41
London	0.08	0.06	0.05	0.08
Cohabitant	0.87	0.45	0.47	0.41
Spouse inwork	0.84	0.28	0.36	0.10
Out of work income	357.06	205.89	210.48	196.26
In work income	509.73	245.11	263.24	207.06

for all h smaller than h^e, with the inequality in the other direction

$$c^\gamma a \geq \left\{ \frac{R(w, h) - R(w, h^e)}{\frac{(L-h^e)^{1-\phi}}{1-\phi} - \frac{(L-h)^{1-\phi}}{1-\phi}} \right\} \tag{20}$$

for all h larger than h^e.

The choice of observed hours h^e is compatible with unconstrained optimization under our specification if and only if there is an a satisfying the two above inequalities, i.e.

$$\min_{h \leq h^e} (1-\phi) \frac{R(w, h) - R(w, h^e)}{(L-h^e)^{1-\phi} - (L-h)^{1-\phi}} \geq \max \left[0, \max_{h \geq h^e} (1-\phi) \frac{R(w, h) - R(w, h^e)}{(L-h^e)^{1-\phi} - (L-h)^{1-\phi}} \right]. \tag{21}$$

The only parameter that appears in this inequality is ϕ. In fact there are two ways of violating the condition: the positivity of the left hand side of (21) only depends on the shape of the function R, and corresponds to the non-parametric rejection; the second inequality on the other hand does depend on ϕ.

For each value of ϕ, Fig. 3 gives the proportion of the observations that would fall into the parametric rejection set. This proportion decreases with ϕ. For low values of ϕ, the share of individuals who work a number of hours incompatible with the model is above 30%. This number decreases to around 5% for very high values of ϕ.

Given the value of ϕ obtained from the estimation of the third model, 7.9% of working women violate the revealed preference inequality. Table 9 contrasts the characteristics of these observations with the rest of the sample. Individuals in this set are more often lone mothers than married ones, their wage is lower than average and, as Fig. 4 shows, their distribution of hours worked is shifted to the left.
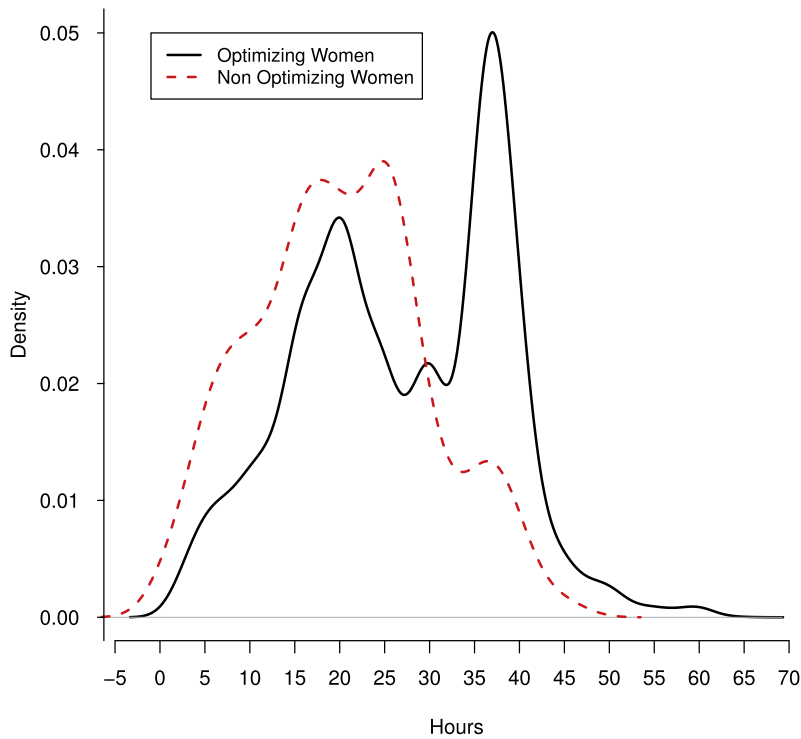
Fig. 4. Hours distributions and rejection of the unrestricted model (Model 3).
Source: Family expenditure survey 1997–2002. The sample comprises women with children, either single or married mothers. Estimates were obtained by maximum-likelihood. Model 1 is the baseline model. Model 2 allows for correlation between wages, consumption and preferences. In addition to this, Model 3 includes exclusion variables in the offer distribution. Simulation is based on maximum-likelihood estimates.

5.4. Elasticities with linear budget constraints

To further describe the preferences underlying the model, we compute the distribution of Frisch and Marshallian elasticities at the intensive margin assuming no hours restrictions and a linear budget constraint.

Frisch elasticities hold the marginal utility of consumption constant and, in our additive utility specification (13), the labour supply elasticity just depends on ϕ and L. The Marshallian elasticities account for the change in consumption that is induced by the within period change in labour earnings, holding non-labour income constant. Table 10 shows the elasticities obtained from the three different models. Note that in model 1 we find no differences between Frisch and Marshallian elasticities, since the income effect is close to 0 (see estimates of γ in Table 4), but this model does not account for the correlation between wages, consumption and preferences. Fig. 5 displays more details of estimated elasticities from model 3. The estimated Frisch labour supply elasticities are positive across the distribution and are moderately sized. The Marshallian elasticities account for the income effect. Results in Fig. 5 show that these are smaller, and can even be negative.

5.5. Model simulations

In that section, we provide simulations from the model to describe the importance of hours restrictions and how would women react to a wage increase. To simulate women decisions, we use observed consumption and covariates from the data. We do not allow consumption to change in these simulations. Conditional on these variables, for each woman i, we draw wages w_i^s, preferences a_i^s and fixed costs b_i^s from the estimated distributions (Model 3). We evaluate the Lagrangian (14):

$$\mathcal{L}(c_i, h, w_i^s) = a_i^s v(h) + c_i^{-\gamma} \left(R(w_i^s, h) - b_i^s 1\{h > 0\} \right)$$

We constrain women to choose between two different offers $h_i^{s,1}\ h_i^{s,2}$ that are drawn from the estimated offer distribution g. For each individual, the choice of hours worked is then:

$$h_i^c = \arg \max_{0, h_i^{s,1}, h_i^{s,2}} \mathcal{L}(c_i, h, w_i^s)$$

Maximization over the whole possible set of hours yields the chosen unconstrained supply of hours h_i^u.

Table 10
Estimation Results for 1997–2002 years: Elasticity Assuming Linear Budget Constraint.
Source: Family expenditure survey 1997–2002. The sample comprises women with children, either single or married mothers. Estimates were obtained by maximum-likelihood. Model 1 is the baseline model. Model 2 allows for correlation between wages, consumption and preferences. In addition to this, Model 3 includes exclusion variables in the offer distribution.

	Model 1	Model 2	Model 3
Frisch elasticity			
Mean	0.58	0.59	0.30
q5	0.18	0.18	0.10
q25	0.22	0.23	0.12
q50	0.37	0.39	0.20
q75	0.60	0.62	0.32
q95	1.73	1.71	0.88
Marshallian elasticity			
Mean	0.58	0.48	0.20
q5	0.18	−0.04	−0.12
q25	0.22	0.15	0.03
q50	0.37	0.29	0.12
q75	0.60	0.52	0.26
q95	1.73	1.59	0.76
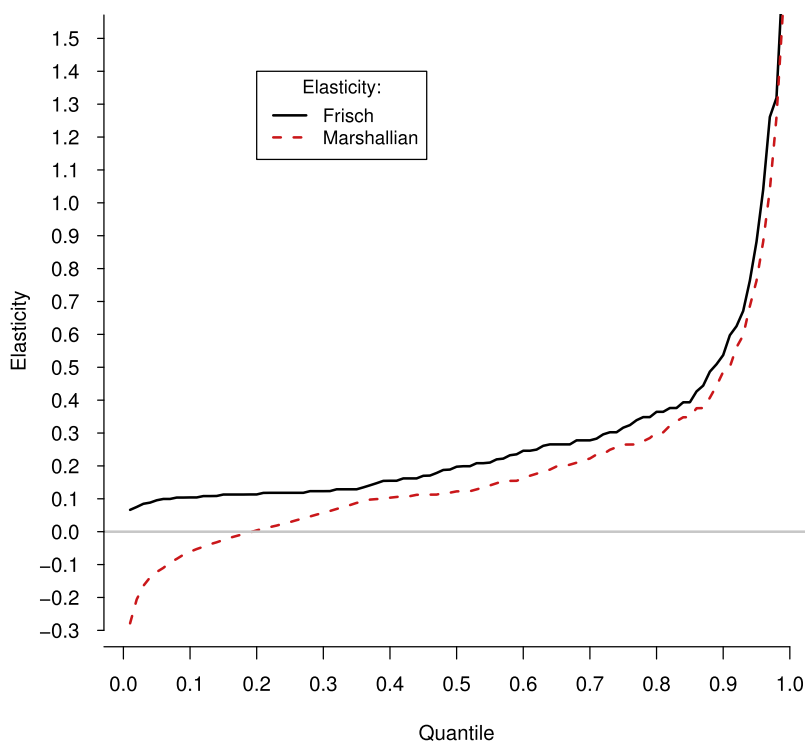
Fig. 5. Intensive margin elasticities with linear budget constraints.
Source: Family expenditure survey 1997–2002. The sample we use comprises women with children, either single or married mothers. Computation obtained from the estimates of Model 3.

5.5.1. Importance of hours restrictions

Comparing the two simulated distributions of hours, we find an employment rate of 71 percent in the unconstrained case, larger than the 62.5 percent obtained in the constrained case. It appears that the restrictions in the two-offer model significantly reduce the number in employment relative to those who would choose to work if they were not constrained. Fig. 6 shows the prediction of the hours distribution using the estimated preference parameters assuming that women are not constrained at all. The resulting distribution of hours is of course very different, reflecting the importance of the specification of the distribution of offers. In addition to the large difference in employment, the modes of the hours distribution move downwards when one goes from the unconstrained to the two-offer case, as well as the average (35.5 vs. 26.2 h).
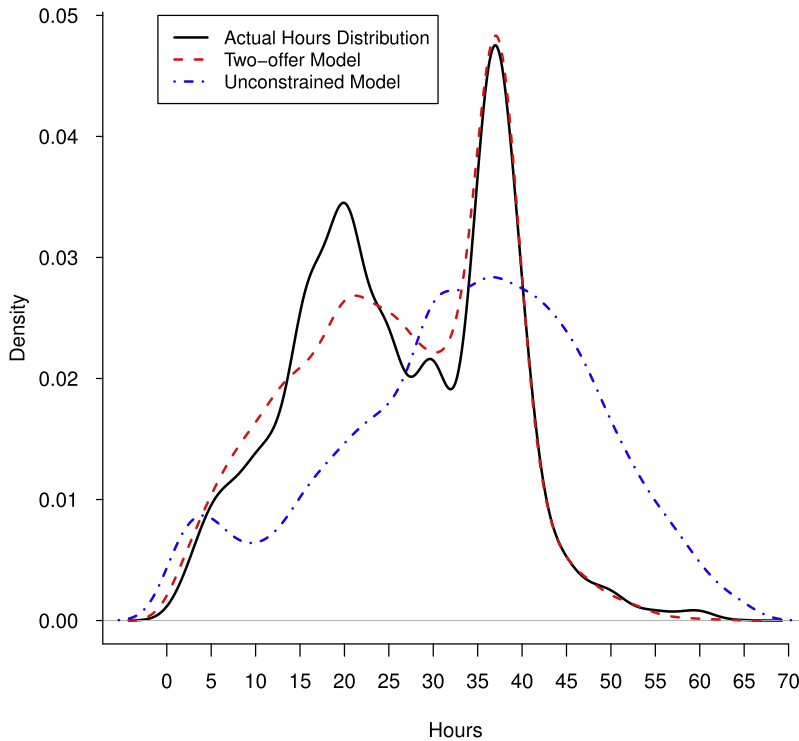
Fig. 6. Predicted hours with and without restrictions.
Source: Family expenditure survey 1997–2002. The sample we use comprises women with children, either single or married mothers. Computation obtained from the estimates of Model 3.

5.5.2. Labour supply responses to a wage increase

We now focus on the impact of an increase of before-tax-income on labour supply decisions. We start from the baseline cases previously detailed. From the baseline simulated wage, we consider an increase of $x = 10\%$. To do so, we compute the corresponding budget constraint functions $R(w_i^s(1 + x), h)$.

Using these new budget constraints, we derive new labour supply decisions. In the case of the unrestricted model, we have:

$$h_i^{(u,x)} = \arg \max_{h=0, h\in\{1,\ldots,H\}} \mathcal{L}(c_i, h, w_i^s(1+x))$$

In the case of the restricted model, we keep the offer fixed. The new chosen hours are obtained from:

$$h_i^{(c,x)} = \arg \max_{0, h_i^{s,1}, h_i^{s,2}} \mathcal{L}(c_i, h, w_i^s(1+x)).$$

In both cases, the average intensive margin response is obtained as:

$$\varepsilon_{\text{Intensive}} = \frac{1}{\sharp(h_i^k > 0)} \sum_{i/h_i^k > 0} \frac{(h_i^{(k,x)} - h_i^k)}{h_i^k} \cdot \frac{1}{x}.$$

and the average extensive margin response as:

$$\varepsilon_{\text{Extensive}} = \frac{(\overline{E_i^{(k,x)}} - \overline{E_i^k})}{\overline{E_i^k}} \cdot \frac{1}{x}.$$

where $\overline{E_i^{(k,x)}}$ and $\overline{E_i^k}$ are equal to one if $h_i^{(k,x)}$ and h_i^k are respectively positive.

These (average) labour supply responses are obtained by keeping c_i constant.

Table 11 presents the mean of these responses that were obtained for a 10% wage increase with and without hours constraints. Accounting for non-linearities in the budget constraint leads to higher responses (.35 vs. .30 in the case of the Frisch elasticity). These estimates are relatively modest in size but lie in the range of estimates of intensive labour supply elasticities found in the literature, for example see Blundell and MaCurdy (1999).

Table 11
Average Elasticities Model 3.
Source: Family expenditure survey 1997–2002. The sample comprises women with children, either single or married mothers. Computation based on estimates of Model 3.

	Unconstrained	2 Offers
Extensive margin	0.25	0.27
Average intensive margin	0.35	0.16
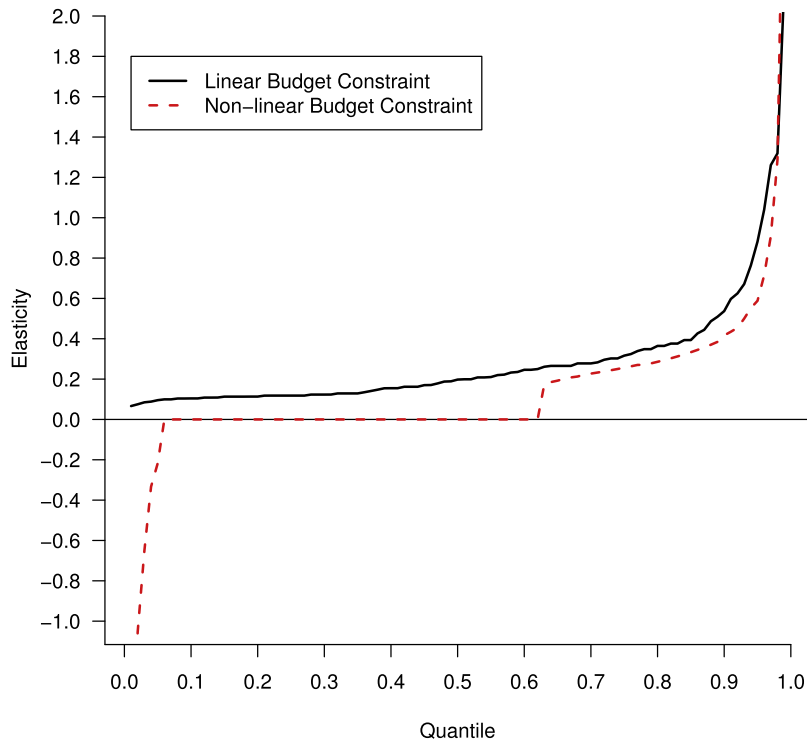
Fig. 7. Labour supply responses with Linear and Non-linear Budget Constraints.
Source: Family expenditure survey 1997–2002. The sample we use comprises women with children, either single or married mothers. Computation obtained from the estimates of Model 3.

We also note from Fig. 7 that in the non-linear budget constraint case many women do not react to a change in wages leading to a large range of no change in hours.

Table 11 also shows the differences in response if we account for hours restrictions. In that case, intensive margin responses are much lower (.16), but extensive margin responses are slightly higher in the constrained than in the unconstrained case: 0.27 vs. 0.25.

6. Summary and conclusions

In this paper we have developed a model of employment and hours in which individuals face restrictions over possible hours choices. Hours choices are made on a random subset of possible hours, and observed hours reflect both the distribution of preferences and the limited choice set. Consequently observed choices do not necessarily satisfy the revealed preference conditions of the standard labour supply model with unrestricted hours choices. Instead, rational choices are made from a set of job packages restricted by the hours offered by employers.

The example we explore in detail in this paper is of individuals selecting from two offers. We show first that, when the offer distribution is known, preferences can be identified. Second, we show that, where preferences are known, the offer distribution can be fully recovered. We then develop conditions for identification of both the parameters of preferences and of the offer distribution. We show that the existence of individuals facing flat budget constraints in addition to other individuals facing more regular budget constraints allows for semi-parametric identification of the model (non-parametric identification of the offer distribution, and parametric identification of choice probabilities). A Monte-Carlo analysis suggests that heterogeneity in budget constraints should be sufficient to provide precise estimates of the parametric model we detail in the next section.

Table 12

First stage reduced form estimates.

Source: Family expenditure survey 1997–2002. The sample comprises women with children, either single or married mothers. Estimates were obtained by OLS.

	Estimate	Std. Error	t-value
β^c : Const	0.18	0.69	0.26
β^c : year 98	−0.04	0.02	−2.48
β^c : year 99	0.01	0.02	0.84
β^c : year 00	0.03	0.02	2.15
β^c : year 01	0.03	0.01	1.84
β^c : year 02	0.04	0.02	2.20
β^c : Age	0.77	0.51	1.51
β^c : Age^2	−0.17	0.14	−1.29
β^c : Age^3	0.01	0.01	0.87
β^c : Edu	0.28	0.03	9.13
β^c : Edu^2	−0.76	0.07	−10.43
β^c : Edu $\times$ Age	0.18	0.03	5.87
β^c : London	0.03	0.02	2.16
β^c : 2 kids	0.05	0.01	5.04
β^c : 3 kids	0.05	0.01	3.49
β^c : 4 kids or more	0.03	0.02	1.60
β^c : Cohab	0.68	0.01	60.78
β^c : Youngest Kid age 0–4	−0.02	0.02	−1.42
β^c : Youngest Kid age 5–9	−0.05	0.01	−3.71
σ^c	0.46		

The new framework is used to study the labour supply choices of a large sample of women in the UK, accounting for nonlinear budget constraints and fixed costs of work. With nonlinear budget sets observed labour supply may not be reconciled with standard optimization theory. The results point to a small but important group of workers who fail the standard choice model with unrestricted choices. This motivates the estimation of a two-offer model, which provides a satisfactory fit of the data. We specify a mixture of normals for the offer distribution which is allowed to depend on education region and calendar time. The estimated offer distribution features the observed twin peaks centred around full-time and part-time hours.

Accounting for restrictions on the choice set changes the estimated pattern of preference parameters. Individuals appear more responsive once restrictions are accounted for and the model simulations predict a higher level of employment were restrictions to be removed.

The two-offer specification we adopt in the application in this paper is nevertheless restrictive. In future work we intend to develop the *n*-offer case, allowing a more flexible specification of the effective choice set. In particular, we could allow the number of alternative choices to vary by location, age, education and point in the business cycle. It would also be natural to extend the framework to allow for offers over wage–hours packages.

Appendix A. First stage estimates: Consumption equation

See Table 12.

Appendix B. Independence of irrelevant alternatives

Consider the case of independence of irrelevant alternatives (IIA), where for all i, j

$$p_i(\{i, j\}, Z, \beta) = \frac{p_i(\mathcal{H}, Z, \beta)}{p_i(\mathcal{H}, Z, \beta) + p_j(\mathcal{H}, Z, \beta)},$$

or

$$p_i(\{i, j\}) = \frac{p_i}{p_i + p_j},$$

where to alleviate notation we drop the arguments Z and β, and denote by p_i the probability of choosing i among the whole set of alternatives. In this circumstance the number of unknowns is equal to the number of equations, and we may hope for exact identification. Indeed

Lemma 3. *Let ℓ and g be two probability vectors in the simplex of $\mathcal{R}^I$, whose components are all positive. There exists at most a unique vector p in the interior of the simplex of $\mathcal{R}^I$ that satisfies the system of equations*

$$\ell_i = g_i^2 + 2g_i p_i \sum_{j \neq i} \frac{g_j}{p_i + p_j} \text{ for } i = 1, \ldots, I. \tag{22}$$
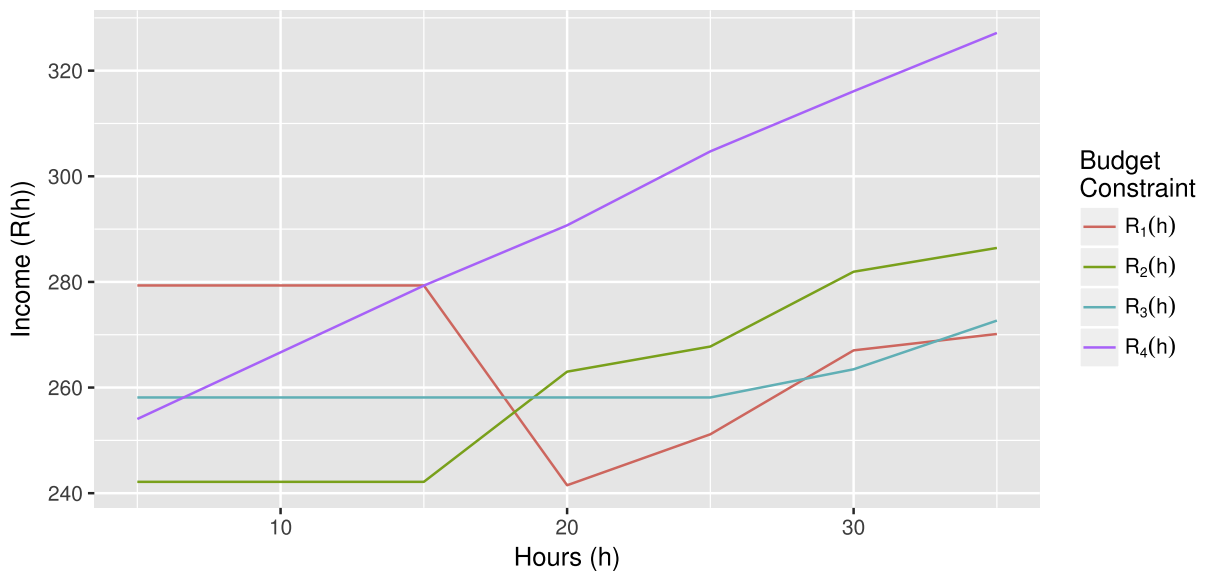
Fig. 8. Four budget constraints from the data.

Table 13
Monte-Carlo estimates, parametric specification.

	Parameter	True value	Mean	Median	MSE	Var	Bias^2
	$\log((1-p)/p)$	0.1	0.1	0.1	0.001	0.001	$< 1e-4$
	$\log(\alpha)$	-1.609	-1.546	-1.609	0.262	0.258	0.004
R_1	β	31	30.991	31.001	0.162	0.162	$< 1e-4$
	$\log(\sigma)$	0.693	0.787	0.693	0.786	0.777	0.009
	$\log(\phi)$	1.792	1.775	1.79	0.024	0.024	$< 1e-4$
	$\log((1-p)/p)$	0.1	0.102	0.1	0.003	0.003	$< 1e-4$
	$\log(\alpha)$	-1.609	-1.599	-1.609	0.031	0.031	$< 1e-4$
R_2	β	31	30.986	31	0.16	0.16	$< 1e-4$
	$\log(\sigma)$	0.693	0.719	0.693	0.067	0.067	0.001
	$\log(\phi)$	1.792	1.79	1.792	0.001	0.001	$< 1e-4$
	$\log((1-p)/p)$	0.1	0.103	0.103	0.001	0.001	$< 1e-4$
	$\log(\alpha)$	-1.609	-1.609	-1.609	0.027	0.027	$< 1e-4$
R_3	β	31	31.014	31.014	0.288	0.288	$< 1e-4$
	$\log(\sigma)$	0.693	0.462	0.693	1.381	1.328	0.053
	$\log(\phi)$	1.792	1.752	1.8	0.019	0.017	0.002
	$\log((1-p)/p)$	0.1	0.079	0.1	0.029	0.029	$< 1e-4$
	$\log(\alpha)$	-1.609	-1.637	-1.626	0.002	0.002	0.001
R_4	β	31	30.974	30.989	0.049	0.048	0.001
	$\log(\sigma)$	0.693	1.252	1.068	0.864	0.552	0.312
	$\log(\phi)$	1.792	1.742	1.791	0.043	0.04	0.002
	$\log((1-p)/p)$	0.1	0.1	0.101	0.001	0.001	$< 1e-4$
	$\log(\alpha)$	-1.609	-1.511	-1.609	0.513	0.504	0.01
All	β	31	30.99	31	0.168	0.168	$< 1e-4$
	$\log(\sigma)$	0.693	0.712	0.693	0.078	0.077	$< 1e-4$
	$\log(\phi)$	1.792	1.813	1.792	0.013	0.012	$< 1e-4$

Proof. For all i, denote

$$P_i(p) = g_i^2 + 2g_ip_i \sum_{j \neq i} \frac{g_j}{p_i + p_j}$$

for p in $\mathcal{R}_+^I$. For any $\lambda \neq 0$, observe that $P_i(\lambda p) = P_i(p)$. Suppose by contradiction that there are two solutions p^0 and p^1 to the system of equations both belonging to the interior of $\mathcal{R}_+^I$. Choose $\bar{p}_I$ such that

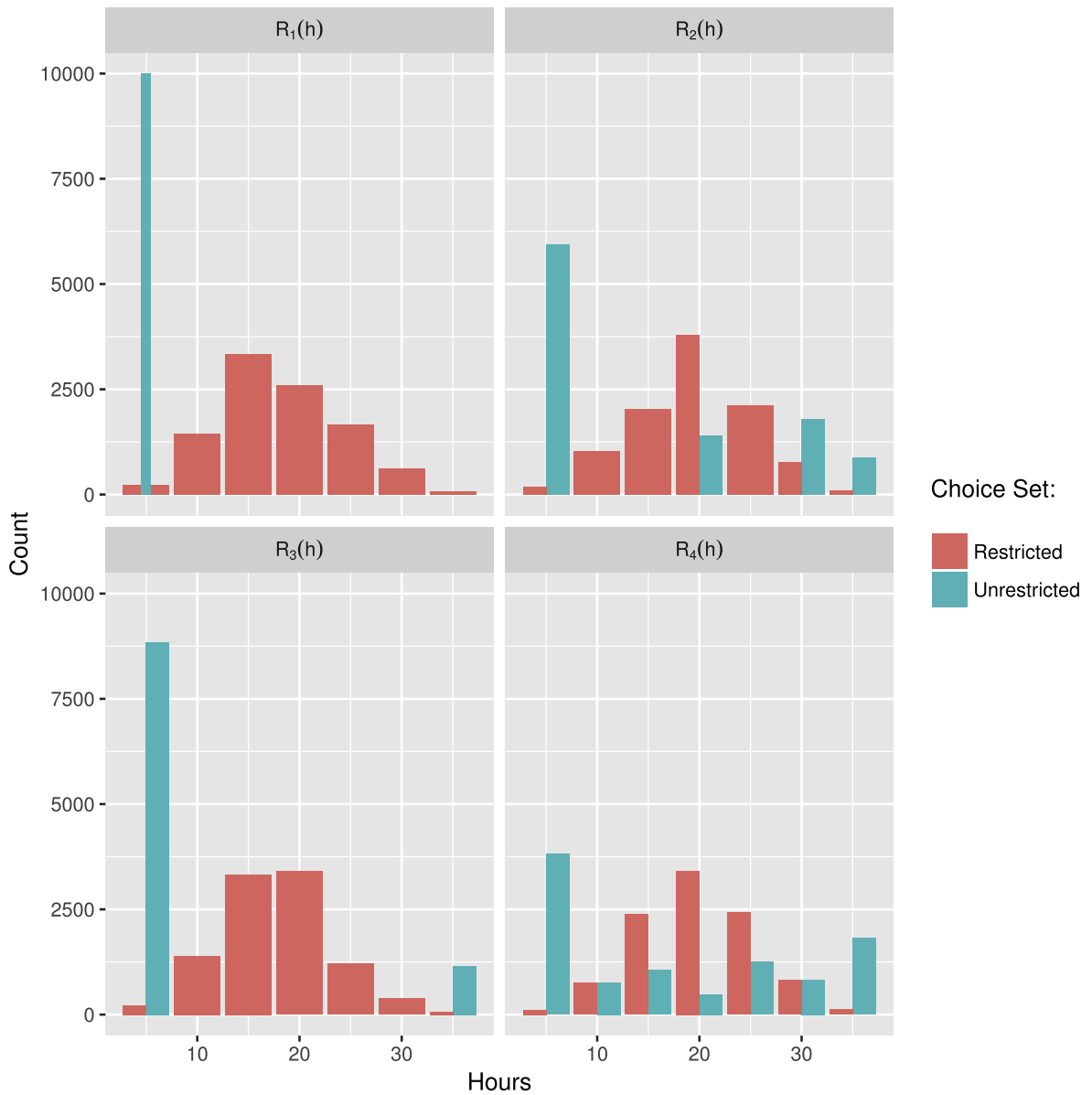$$\bar{p}_I \geq \frac{p_I^0}{\min_i p_i^0} \text{ and } \bar{p}_I \geq \frac{p_I^1}{\min_i p_i^1},$$

Fig. 9. Labour supply decisions for the four budget constraints from the data.

and define λ^0 and λ^1 through

$$\lambda^0 p_I^0 = \lambda^1 p_I^1 = \bar{p}^I.$$

This construction implies that the two vectors $\lambda^0 p^0$ and $\lambda^1 p^1$ are both solutions of

$$\ell_i = P_i(p) \text{ for } i = 1, \ldots, I - 1,$$

have all their coordinates larger than 1, with n'th coordinate normalized at $\bar{p}_I$. We therefore study the reduced system of $I - 1$ equations

$$\ell_i = P_i(p_1, \ldots, p_{I-1}, \bar{p}_I) \text{ for } i = 1, \ldots, I - 1$$
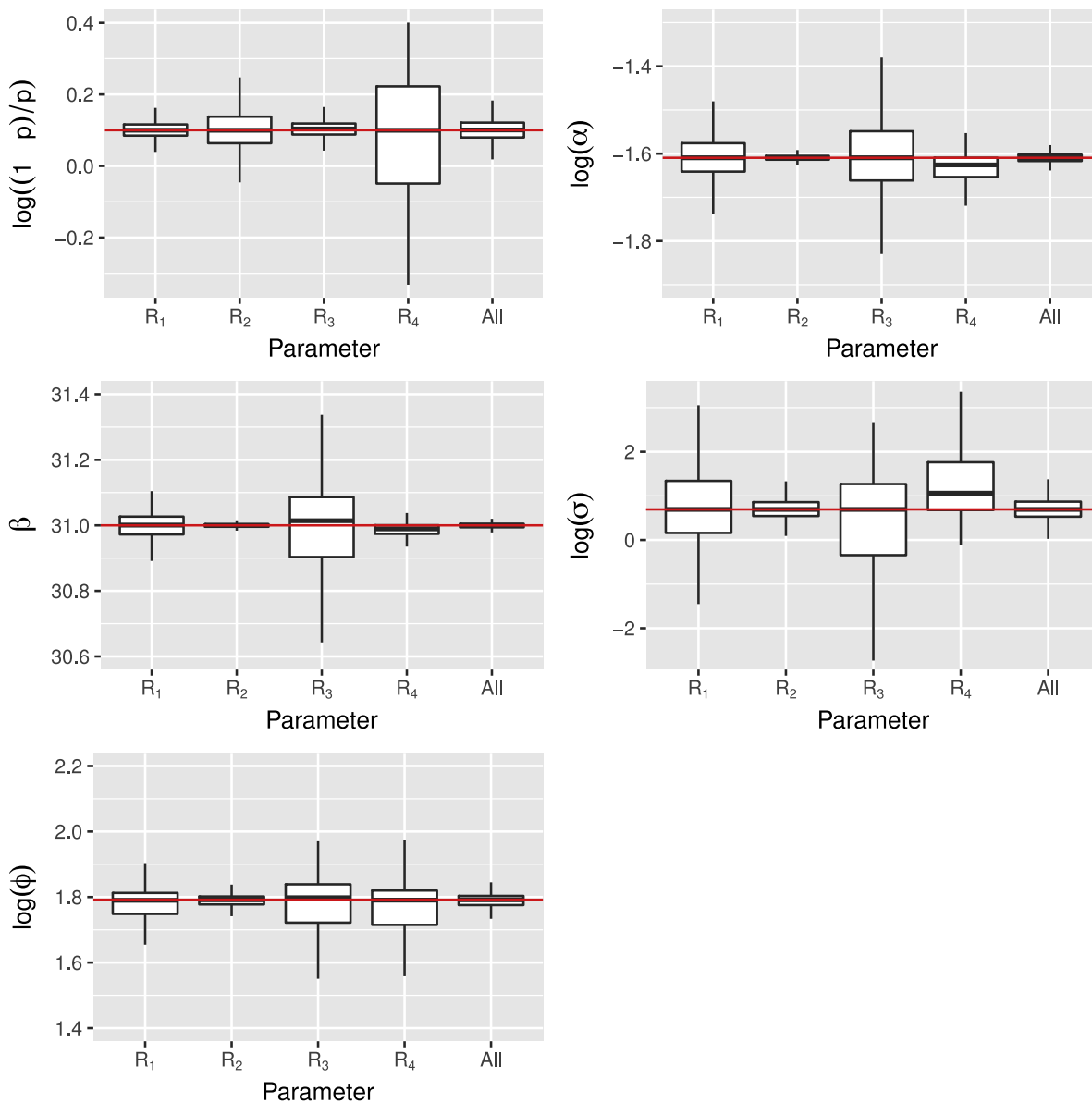
Fig. 10. Monte Carlo result for parametric models.

with the unknowns $(p_1, \ldots, p_{I-1})$ in $[1, \infty)^{I-1}$. The fact that it has at most a unique root follows from Gale Nikaido, once it is shown that the Jacobian of P is everywhere a dominant diagonal matrix. We have

$$\frac{\partial P_i}{\partial p_i} = 2g_i \sum_{j=1, j \neq i}^{I} \frac{g_j p_j}{(p_i + p_j)^2},$$

and for j different from i

$$\frac{\partial P_i}{\partial p_j} = -2g_i p_i \frac{g_j}{(p_i + p_j)^2}.$$

The property of diagonal dominance with equal weights to all terms is equivalent to

$$\left| \frac{\partial P_i}{\partial p_i} \right| > \sum_{j=1, j \neq i}^{I-1} \left| \frac{\partial P_j}{\partial p_i} \right|,$$
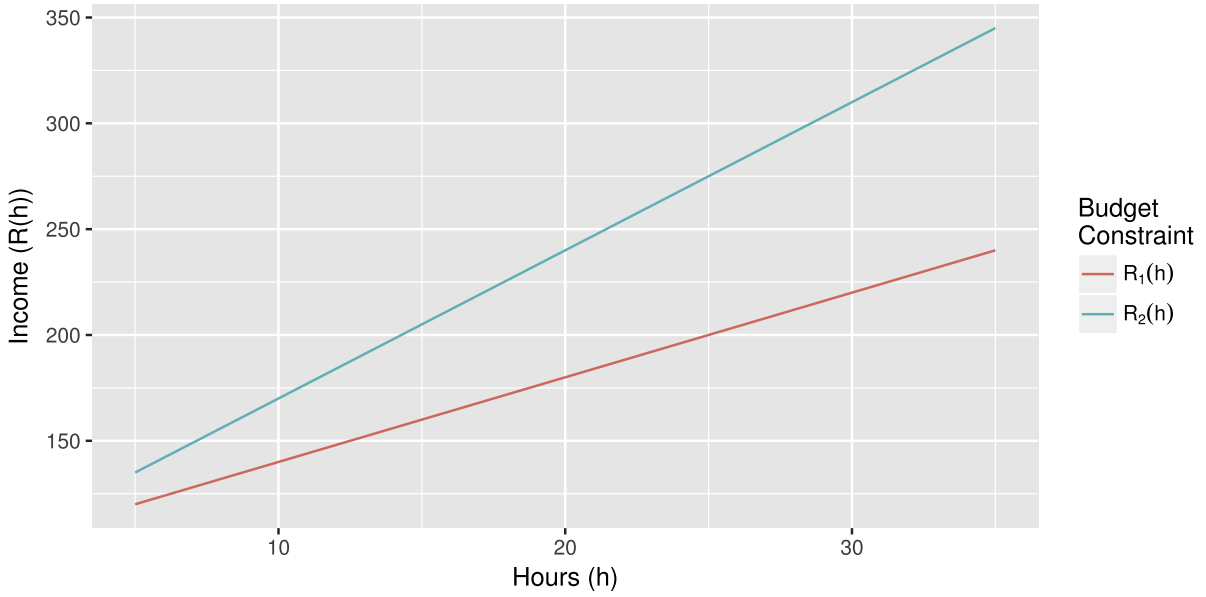
Fig. 11. Two linear budget constraints.

that is

$$2g_i \sum_{j=1,j\neq i}^{I} \frac{g_j p_j}{(p_i + p_j)^2} > \sum_{j=1,j\neq i}^{I-1} 2g_i p_j \frac{g_j}{(p_i + p_j)^2}$$

or

$$\frac{g_I p_I}{(p_i + p_I)^2} > 0.$$

The inequality is satisfied, and the right hand side mapping is univalent on $[1, \infty)^{I-1}$, which completes the proof. □

As we noted in Section 4, there may be cases which would never be rationally chosen. In these situations we can put zero weights on some of the decisions, that is $p_j = 0$ for some subset J of the alternatives. A simple manipulation of the system of equations, using the equality $p_i(\{i, j\}) + p_j(\{i, j\}) = 1$ even when the marginal probabilities are zero, yield

$$\ell_J = \sum_{j \in J} \ell_j = \left(\sum_{j \in J} g_j \right)^2 = g_J^2,$$

and for all i not in J

$$\ell_i = g_i(1 + 2g_J) + 2g_i p_i \sum_{k \notin J, k \neq i} \frac{g_k}{p_i + p_k},$$

where the notation p_J denotes the sum of the components of the vector p with indices in J. A minor adaptation of the proof of Lemma 3 then shows that the vector p is uniquely determined. Using the first equation, a natural procedure is to compute the non-negative difference $\ell_J - g_J^2$ for all subsets J of indices. The candidates J for the solution are the ones for which the difference is zero. We do not know whether there can be multiple candidates.[5]

---

[5] There cannot be two solutions with two disjoint sets J_1 and J_2. Indeed one would need to have

$$\ell_{J_1} = g_{J_1}^2 \quad \ell_{J_2} = g_{J_2}^2,$$

which implies

$$\ell_{J_1 \cup J_2} = g_{J_1}^2 + g_{J_2}^2 < g_{J_1 \cup J_2}^2,$$
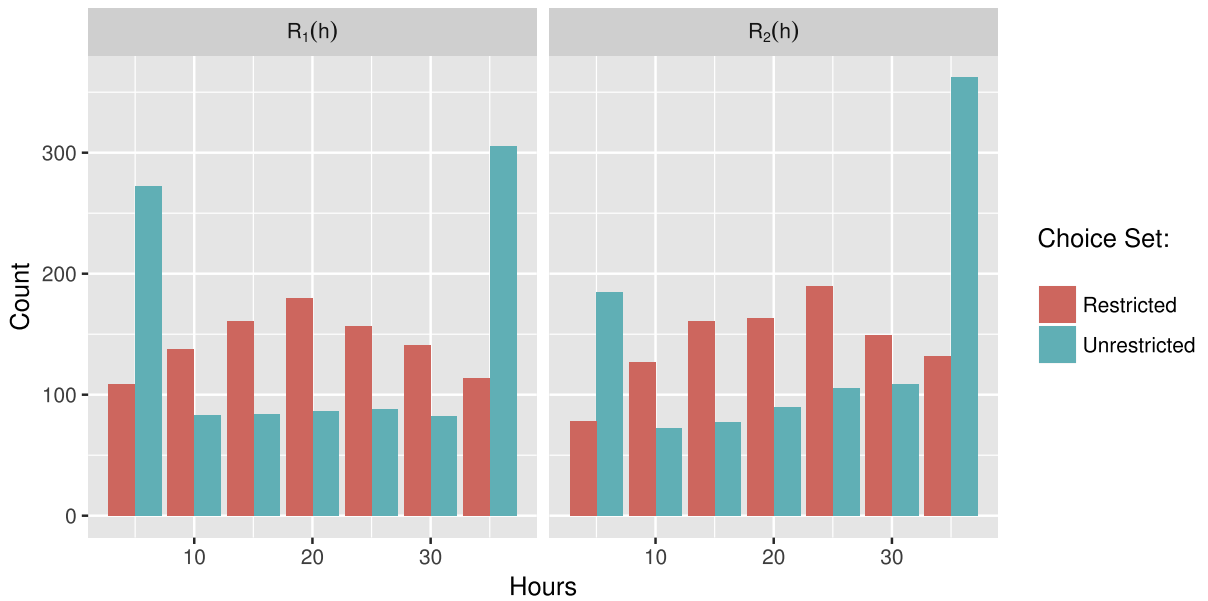
which is impossible.

Fig. 12. Restricted and unrestricted labour supply decisions with linear budget constraints.

Appendix C. Monte-Carlo analysis

In this section we set a Monte-Carlo analysis in order to illustrate the identification results from Section 3.3, and the finite sample properties of the estimates.

In all Monte-Carlo experiments preferences are fixed and the choices are made among a number of offers. Each offer consists in a number of hours h, which is associated to income through a budget constraint $R(h)$.

The utility function associated to h is specified as:

$$U(h, \varepsilon) = \frac{R(h)^{1-\alpha}}{1-\alpha} + \exp(\beta + \sigma\varepsilon)\frac{(100-h)^{1-\phi}}{1-\phi} = v_1(R(h), \alpha) + \exp(\beta + \sigma\varepsilon)v_2(h, \phi)$$

$(\alpha, \beta, \phi, \sigma)$ is the set of preference parameters that we want to estimate. Unobserved preferences for leisure are given by ε, which is supposed to follow a standardized normal distribution.

The offer distribution takes $I = 7$ different discrete values in $\mathcal{S} = \{5, 10, 15, 20, 25, 30, 35\}$. The probability that one of the two independent offers is equal to h is equal to $g(h)$.

C.1. Likelihood of the model

From the distributions of preferences and offers, we can derive the individual contribution to the likelihood. For an individual working h hours, the contribution is:

$$\ell(h) = g(h)^2 + 2g(h)\sum_{h' \neq h} g(h')P\left[U(h, \varepsilon) > U(h', \varepsilon)\right]$$

The probability to prefer the actual number of hours h, compared to the alternative offer h' can be detailed:

$$P\left[U(h, \varepsilon) > U(h', \varepsilon)\right] = P(v_1(R(h), \alpha) + e^{\beta + \sigma\varepsilon}v_2(h, \phi) > v_1(R(h'), \alpha) + e^{\beta + \sigma\varepsilon}v_2(h', \phi))$$
$$= P(e^{\beta + \sigma\varepsilon}[v_2(h, \phi) - v_2(h', \phi)] > v_1(R(h'), \alpha) - v_1(R(h), \alpha))$$
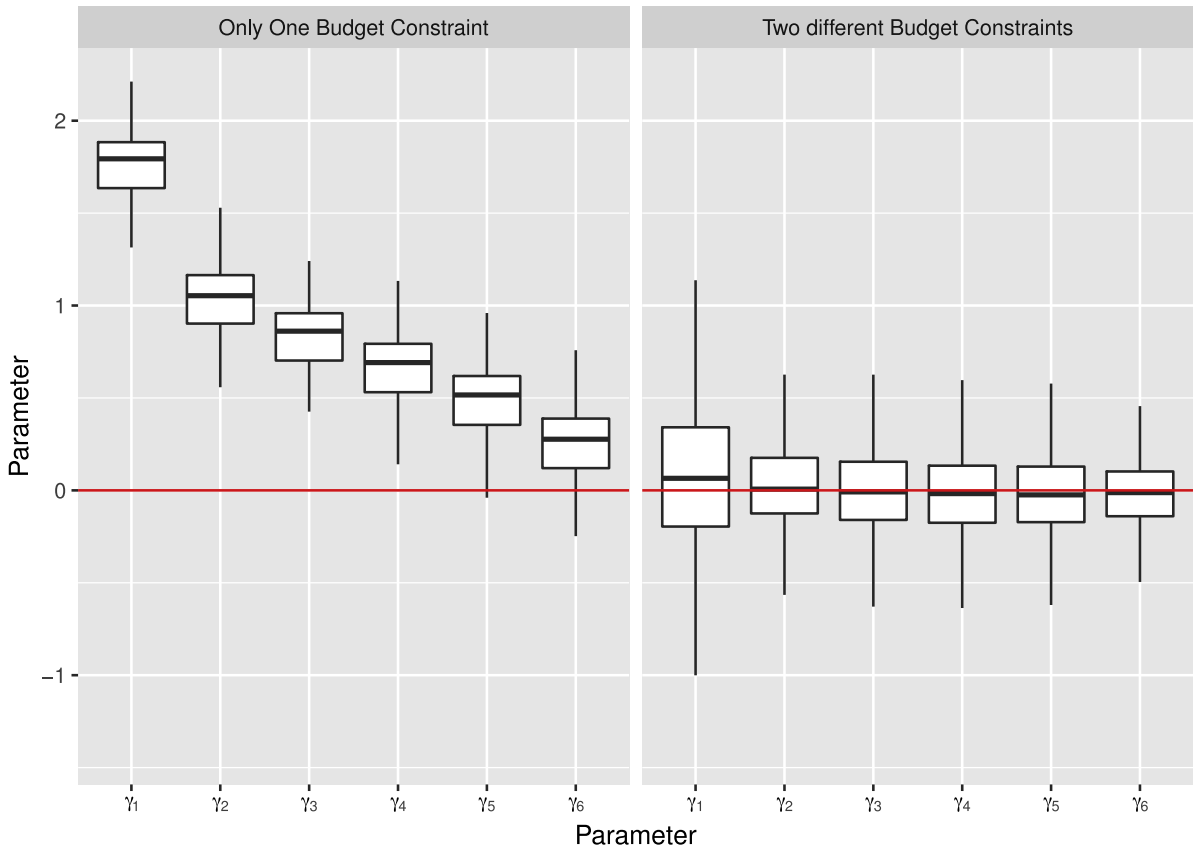
Fig. 13. Distribution of Monte-Carlo estimates (offer distribution parameters).

v_2 is decreasing, so if $h' < h$, $v_2(h, \phi) - v_2(h', \phi) < 0$, so we have:

$$P\left[U(h, \varepsilon) > U(h', \varepsilon)\right] = P\left[\beta + \sigma\varepsilon < \log\left(\frac{v_1(R(h'), \alpha) - v_1(R(h), \alpha)}{v_2(h, \phi) - v_2(h', \phi)}\right)\right]$$

$$= \Phi\left[\frac{\log\left(\frac{v_1(R(h'),\alpha)-v_1(R(h),\alpha)}{v_2(h,\phi)-v_2(h',\phi)}\right) - \beta}{\sigma}\right]$$

else if $h' > h$:

$$P\left[U(h, \varepsilon) > U(h', \varepsilon)\right] = 1 - \Phi\left[\frac{\log\left(\frac{v_1(R(h'),\alpha)-v_1(R(h),\alpha)}{v_2(h,\phi)-v_2(h',\phi)}\right) - \beta}{\sigma}\right]$$

Note that in both cases, the ratio $\frac{v_1(R(h'),\alpha)-v_1(R(h),\alpha)}{v_2(h,\phi)-v_2(h',\phi)}$ may be negative. This corresponds to cases where h will be preferred to h' (if $h < h'$) with probability 1, or where h will never be preferred to h' (if $h > h'$).

C.2. Parametric offer distribution

We first start by illustrating the importance of shape restrictions of the budget constraint in the parametric case. In this case, we suppose that the offer distribution is binomial, with parameter p. Then the probability of drawing h_i is:

$$g(h_i) = \binom{I-1}{i-1} p^{i-1}(1-p)^{I-i}.$$

Each Monte-Carlo sample is made of a population of 1,000 women. We use four different budget constraints, and implement also a fifth experiment which uses all four budget constraints together. The budget constraints are taken from the actual data.
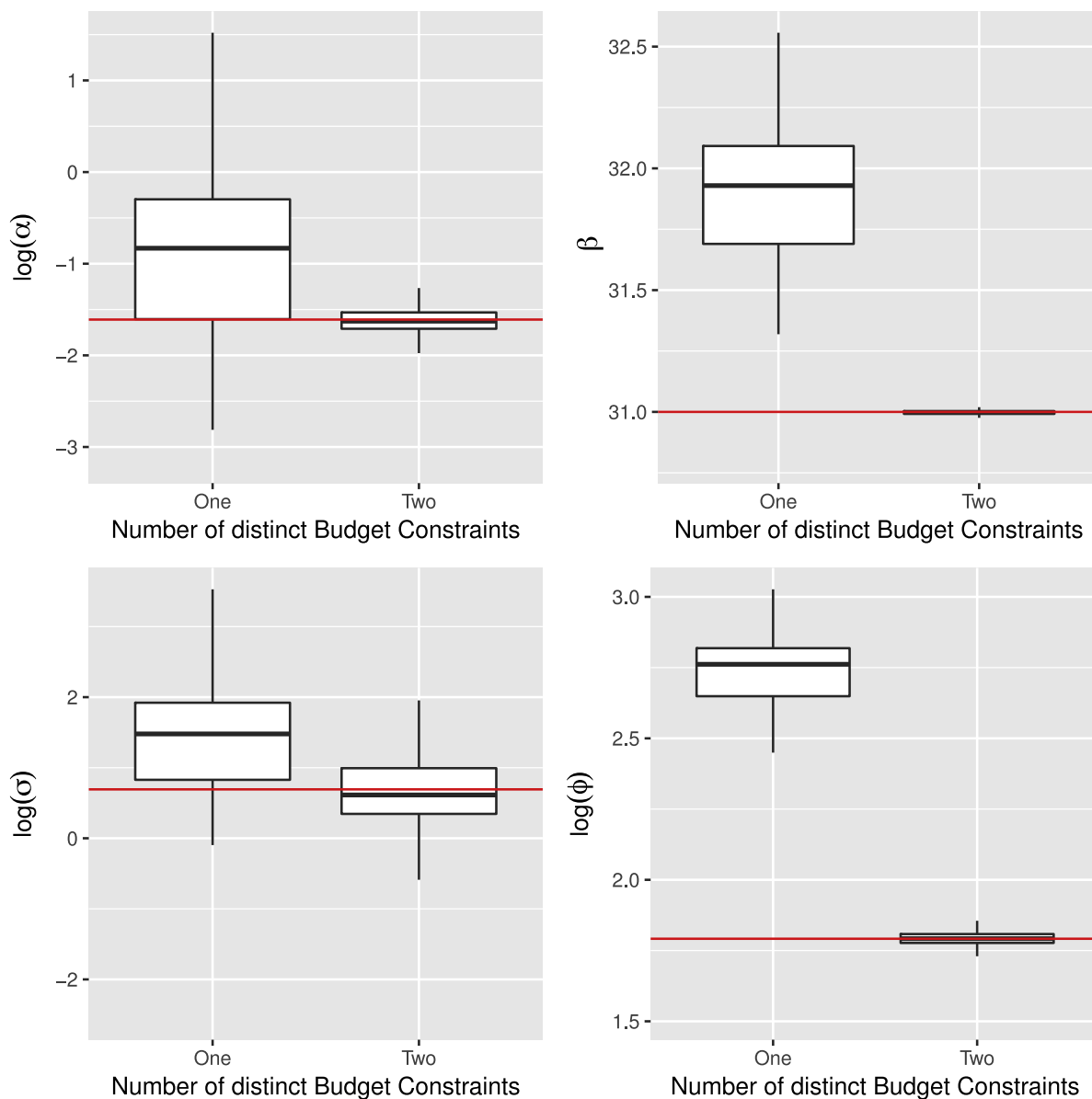
Fig. 14. Distribution of Monte-Carlo estimates (Preference parameters).

We can see from Fig. 8 that these budget constraints have very different patterns. Budget constraints R_1, R_2 and R_3 admit large dominated regions, whereas the fourth one is strictly increasing in hours.

The consequences of these budget constraints in terms of labour supply decisions are shown in Fig. 9 where we display the frequency of each choice in two different situations: when women choices are restricted, among two random offers, and when they can freely choose in the support of hours. We see that unrestricted individuals facing budget constraint 1, would always choose the lowest point of support. Budget constraints 2, and 3 also lead to a partial coverage of the support of hours in the unrestricted case, whereas in the last case heterogeneity in preferences leads to a full coverage of the support. In the two-offer case, the support is fully covered whatever the budget constraint, but each budget constraint leads to a slightly different distribution of observed hours.

We now study whether we can recover the parameters used to generate the data. We estimate the model by maximum likelihood using several starting values. We repeat the experiment 1,500 times.

Fig. 10 (and Table 13) show the results of these estimations. Each graph corresponds to one parameter, and each box plot gives the distribution of Monte-Carlo estimates obtained for each budget constraint.

We observe that the budget constraints with larger sets of dominated hours lead to more precise estimates of the offer distribution parameter, but on the contrary less precise estimates of the preference parameters. This result is consistent

Table 14
Monte-Carlo estimates, semi-parametric specification.

Number of BC	Parameter	True value	Mean	Median	MSE	Var	Bias^2
1	γ_1	0	1.4351	1.7938	2.6356	0.5765	2.0595
	γ_2	0	0.8677	1.0532	0.97	0.2173	0.7528
	γ_3	0	0.7093	0.8613	0.6537	0.1507	0.5031
	γ_4	0	0.5754	0.6914	0.4352	0.1042	0.331
	γ_5	0	0.4406	0.5161	0.2608	0.0666	0.1942
	γ_6	0	0.2537	0.2767	0.0997	0.0353	0.0644
	$\log(\alpha)$	−1.6094	−0.8239	−0.8311	1.145	0.5283	0.617
	β	31	31.7846	31.9291	0.8125	0.1971	0.6156
	$\log(\sigma)$	0.6931	1.4699	1.4791	0.9925	0.3893	0.6034
	$\log(\phi)$	1.7918	2.57	2.7617	0.7767	0.1711	0.6057
2	γ_1	0	0.0688	0.0651	0.3711	0.3666	0.0047
	γ_2	0	−0.0042	0.0068	0.1799	0.18	< 1e − 4
	γ_3	0	−0.02	−0.0083	0.126	0.1256	4e-04
	γ_4	0	−0.0302	−0.0175	0.0914	0.0906	9e-04
	γ_5	0	−0.0282	−0.0245	0.0656	0.0648	8e-04
	γ_6	0	−0.0251	−0.0122	0.0374	0.0368	6e-04
	$\log(\alpha)$	−1.6094	−1.6196	−1.6268	0.0777	0.0777	1e-04
	β	31	31.0614	30.9978	0.0688	0.065	0.0038
	$\log(\sigma)$	0.6931	0.7116	0.6138	0.3293	0.3292	3e-04
	$\log(\phi)$	1.7918	1.8195	1.7921	0.0313	0.0305	8e-04

with the fact that dominated hours of the budget constraint provide information about the offer distribution g, since the probability of individual choosing this number of hours is equal to the probability of getting this offer twice. On the other hand, the probability of choosing these dominated hours compared to any other is equal to 0 regardless of the value of the preference parameters, so that we then have less information allowing to pin down preference parameters.

Note also that the identification from the functional form of the offer distribution and preferences may be weak. This may be the case with the fourth budget constraint, which is close to linear, where the distribution of estimates of preference parameters is not always centred on the true value.

Finally, the last set of estimates combines all four different budget constraints for the simulation of the data. We see from the results of this set of estimates that it combines the precision of both cases where some regions are dominated, and cases where the monotonicity of the budget constraint leads to precise estimates of the preference parameters. The inclusion of several budget constraints can be viewed as the use of an exclusion variable that would impact the shape of the budget constraint without affecting the offer distribution. We show below that this even allows to recover the non-parametric distribution of the (discrete) offer distribution.

C.3. Non parametric offer distribution and linear budget constraints

In the second Monte-Carlo experiment, we let the offer distribution unrestricted. So we exactly have $I − 1$ parameters to recover: $\{\gamma_1, \ldots, \gamma_{I-1}\}$, and the probability of getting offer h_i is:

$$g(h_i) = \frac{\exp(\gamma_i)}{1 + \sum_{j=1}^{I-1} \exp(\gamma_j)}$$

and

$$g(h_I) = \frac{1}{1 + \sum_{j=1}^{I-1} \exp(\gamma_j)}$$

The model leads to $I − 1$ different likelihood contributions, and we have $I − 1$ parameters to identify for the offer distribution in addition to the four preference parameters. In order to have identification, we need at least to have one exclusion restriction. To do so, we allow for two different linear budget constraints, R_1 and R_2, leading to $2 \times (I-1)$ likelihood contributions, without increasing the number of parameters to estimate.

The two budget constraints are illustrated by Fig. 11. Fig. 12 shows the distribution of women labour supply decision when they face restrictions or not. In the first case, they optimize over the whole set $\mathcal{S}$, whereas in the second case they make their choice between two offers drawn from a uniform distribution.

Each Monte-Carlo sample is made of a population of 1,000 women. We simulate two types of populations. In the first type, all women face the same budget constraint (R_1), whereas in the second type, half of the population faces R_1, and the other half faces R_2. For each sample, we estimate the model by maximum likelihood, using different sets starting value among the true vector of parameters and random starting values. We repeat the experiment 1,500 times.

The distribution of estimates is summarized in the boxplot presented in Figs. 13 and 14 (more detailed results are provided in Table 14). Each box gives the first and ninth deciles, and quartiles of the distribution of estimates, and the red horizontal lines give the true value of the parameters used in the simulation.

The estimation procedure seems to be biased when the population is only made of women facing the R_1 budget constraint. On the contrary, when we simulate populations facing two different budget constraints, the estimates of the model seem to be well behaved.

References

Aaberge, R., Colombino, U., Strom, S., 1999. Labour supply in italy: An empirical analysis of joint household decisions, with taxes and quantity constraints. J. Appl. Econometrics 14, 403–422.

Aaberge, R., Colombino, U., Wennemo, T., 2009. Evaluating alternative representations of the choice sets in models of labor supply. J. Econ. Surv. 23 (3), 586–612, Wiley Blackwell.

Adam, Stuart, Browne, James, Heady, Chris, 2010. Taxation in the UK. In: Mirrlees, James, Adam, Stuart, Besley, Tim, Blundell, Richard, Bond, Steve, Chote, Robert, Gammie, Paul, Myles, Gareth, Poterba, James (Eds.), Dimensions of Tax Design: The Mirrlees Review. Oxford University Press for Institute for Fiscal Studies, Oxford.

Altonji, J.G., Paxson, C., 1992. Labor supply, hours constraints, and job mobility. J. Hum. Resour. 27, 256–278.

Bloemen, H., 2000. A model of labor supply with job offer restrictions. Lab. Econ. 7, 297–312.

Bloemen, H., 2008. Job search, hours restrictions, and desired hours of work. J. Lab. Econ. 26, 137–179.

Blundell, R., Bozio, A., Laroque, G., 2011. Labour supply and the extensive margin. Amer. Econ. Rev. 101 (3), 482–486.

Blundell, R., Brewer, M., Francesconi, M., 2008. Job changes, hours changes and the path of labour supply adjustment. J. Lab. Econ. 26 (3), 421–453.

Blundell, R., MaCurdy, T., 1999. Labor supply: A review of alternative approaches. In: Ashenfelter, Card (Eds.), Handbook of Labor Economics. 3A.

Blundell, R., Powell, J.L., 2003. Endogeneity in nonparametric and semiparametric regression models. In: Dewatripont, M., Hansen, L.P., Turnovsky, S.J. (Eds.), Advances in Economics and Econometrics, Theory and Applications, Eighth World Congress, Vol. II. Cambridge University Press, Cambridge, pp. 272–273.

Blundell, R., Walker, I., 1986. A life cycle consistent empirical model of family labour supply using cross section data. Rev. Econom. Stud. 53, 539–558.

Chetty, Raj, 2012. Bounds on elasticities with optimization frictions: a synthesis of micro and macro evidence on labor supply. Econometrica 80 (3), 968–1018.

Chetty, Raj, John, Friedman, Tore, Olsen, Luigi, Pistaferri, 2011. Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from danish tax records. Q. J. Econ. 126 (2), 749–804.

Cogan, J., 1981. Fixed costs and labor supply. Econometrica 49, 945–963.

Dagsvik, J.K., Strom, S., 2006. Sectoral labor supply, choice restrictions and functional form. J. Appl. Econometrics 21, 803–826.

Dickens, W., Lundberg, S., 1993. Hours restrictions and labor supply. Internat. Econom. Rev. 34, 169–191.

Ham, J., Reilly, K.T., 2002. Testing intertemporal substitution, implicit contracts, and hours restriction models of the labor market using micro data. Amer. Econ. Rev. 92, 905–927.

Hausman, J., 1985. The econometrics of nonlinear budget sets. Econometrica 53, 1255–1282.

Heckman, J., 1974a. Shadow prices, market wages, and labor supply. Econometrica 42 (4), 679–694.

Heckman, J., 1974b. Effects of childcare programs on womens work effort. J. Polit. Econ. 82, S136–S163.

Heckman, J., 1979. Sample selection bias as a specification error. Econometrica 47 (1), 153–161.

Heckman, J., MaCurdy, T., 1980. A lifecycle model of female labor supply. Rev. Econom. Stud. 47, 47–74.

Kfir, E., Spiegler, R., 2011. Consideration sets and competitive marketing. Rev. Econom. Stud. 78, 235–262.

Kline, P., Tartari, M., 2016. Bounding the labor supply responses to a randomised welfare experiment: A revealed preference approach. Amer. Econ. Rev. 106 (4), 972–1014.

Saez, Emmanuel, 2010. Do taxpayers bunch at kink points? Amer. Econ. J. Econ. Policy August, 180–212.

van Soest, A., Woittiez, I., Kapteyn, A., 1990. Labour supply, income taxes and hours restrictions in the netherlands. J. Hum. Resour. 25, 517–558.