Labour Supply and Taxation with Restricted Choices

Magali Beffy, Richard Blundell, Antoine Bozio and Guy Laroque

June 2013
Preliminary Draft

Abstract
A model of labour supply and consumption is developed in which individuals face restrictions on the offer distribution of possible hours choices. Observed hours reflect both the distribution of preferences and the distribution of offers. The choice set is limited and observed hours will not necessarily satisfy the conditions for optimal choice. The leading example we consider is of individuals selecting from at most two offers. We illustrate this in a labour supply model with nonlinear budget constraints. We show first that when the offer distribution is known, preferences can be identified. We are also able to 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. This model is used to study the labour supply and consumption choices of a large sample of women in the UK, accounting for nonlinear budget constraints and fixed costs of work.

*We thank the data archive UK and IPUMS for data access. This research is funded by the ERC and also by the ESRC through the Center for the Microeconomic Analysis of Public Policy at IFS.
†Beffy: Institute for Fiscal Studies (IFS), 7 Ridgmount Street WC1E 7AE London United Kingdom, magali_b@ifs.org.uk. Blundell: University College London (UCL) and IFS, r.blundell@ucl.ac.uk. Bozio: Paris School of Economics (PSE) and IFS, a.bozio@ifs.org.uk. Laroque: UCL, IFS and Sciences-Po, g.laroque@ucl.ac.uk.
1 Introduction

There is considerable variation in observed hours worked over time and in the cross-section among working age adults, especially among women with children (see, for example, Blundell, Bozio and Laroque, 2011). We ask under what conditions can this variation be used to identify preferences for labour supply. It is unlikely that all workers are free to choose their working hours. Different individuals at different times of their career may well face different sets of offered hours from which they can choose. Some may have no choice at all. This becomes particularly noticeable when individuals face nonlinear tax and benefit systems (Hausman, 1985). The resulting budget constraints can give rise to ranges of hours that would never be rationally chosen if, that is, other hours choices had been available. They can also give rise to a lack of bunching at kink points and notches that would normally occur with completely flexible choices (Saez, 2010). Consequently, we may observe individuals with working hours that are inconsistent with optimising behaviour and unrestricted hours choices. The observed hours may however be consistent with optimal choice given a restricted choice set. As empirical economists we typically do not know the complete set of alternatives available to individuals in our data. We are not sure of their choice set. Moreover different individuals may well face different constraint sets. In this paper we develop and estimate a structural model of labour supply that embeds restrictions on the distribution of available hours offers.

We are not the first to examine hours restrictions in labour supply models. There is a long history of incorporating hours restrictions into labour supply models, including, Aaberge (1999, 2009), Bloemen (2000, 2008), Dickens and Lundberg (1993), Van Soest et al (1990), Ham and Reilly (2002) and Dagsvik and Strom (2006). The ideas we develop here extend these models to 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 the agents do not make their choices over the whole set of possible choices, but on a random subset of it. We analyze 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. We first consider the case where the econometrician knows the probability distribution of offered choices. In the more complete model we generalise this to make the distribution of offers unknown but restrict it to be a function of a finite set of unknown parameters.

The labour supply model we consider 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 (1974)), with fixed costs of work (Heckman (1974, 1979), Cogan (1981)), intertemporal choices (Heckman and MaCurdy (1981)). 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 is shown to approach the neoclassical 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 and the nature of their decisions in the face of non-linear budget constraints generated through the working of the tax and welfare system. We provide direct evidence of hours restrictions by recording individuals working at hours of work that would be strictly dominated by other choices were they to be available.

We use data from the UK Family Expenditure Survey which records hours worked and a detailed measure of consumption expenditure. 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, Browne and Heady (2010). 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. Finally we use the consumption measure in the FES to ensure the hours of work decision is consistent with a life-cycle model (see Blundell and Walker, 1986).

The layout of the remainder of the paper is as follows. The next section lays out the intertemporal labour supply model with nonlinear budget constraints and fixed costs. In section 3, we then consider the interpretation of rejects of the unrestricted choice model. Section 4 develops a model of labour supply in which individuals face a two-offer distribution over possible hours choices. In section 5 we show that when the offer distribution is known preferences can be identified. We are also able to show in section 6 that, where preferences are known, the offer distribution can be fully recovered. Section 7 develops
conditions for identification of both the parameters of preferences and of the offer distribution. In section 8 this model is used to study the labour supply and consumption choices of a large sample of women in the UK, accounting for nonlinear budget constraints and fixed costs of work. We find a small but significant group of women whose observed choices cannot be rationalized by the standard neoclassical choice model. We then estimate a parametric specification of the two offer model. The estimated offer distributions are presented in Section 9. These point to larger underlying elasticities were restrictions to be absent. Together with the estimated preferences for hours and employment we argue that the framework provide a compelling empirical framework for understanding observed hours and employment. Section 10 concludes.

2 A model of hours, employment and consumption

We begin by laying out a neoclassical labour supply model in an intertemporal setting with choices at the extensive and intensive margins. In this specification there are nonlinear taxes and fixed costs of work but otherwise workers are free to choose their optimal hours of work.

At date $t$, the typical household chooses her consumption $c_t$ and labour supply $h_t$, maximizing

$$E_t \int_t^T u_t(c, h) \, \tau$$

subject to an intertemporal budget constraint

$$\int_t^T \exp[-r(\tau - t)] \{c - R(w, h) + b_{h>0}\} \, \tau \leq S_t.$$  

Here $u_t$ is the instantaneous utility index, a concave twice differentiable function of the vector $(c, h)$ of consumption and hours of work. It is increasing in consumption, decreasing in hours. The consumption good is the numeraire.

The function $R(w, h)$ denotes the after tax and benefit income of someone who works $h$ hours at wage $w$. 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 $\lambda_t$ 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 satisfies the first-order
condition
\[ \frac{\partial u}{\partial c}(c_t, h_t) = \lambda_t. \]

Also, if the household works, the optimal hours maximize \( u_t(c_t, h_t) + \lambda_t R(w_t, h_t) \). Let \((c^e, h^e)\), the optimal choice of the working household, \( c^o \) the consumption of the household with the worker out of the labour market. The household will be observed out of the labour market whenever

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

In this 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.

### 3 Rejections of the rational choice model

To interpret inconsistencies of observed behaviour with rational choice within the framework developed above it is useful to place some structure on preferences and individual heterogeneity.

Consider the following utility specification, separable in consumption and leisure

\[ u_t(c, h) = \frac{c^{1-\gamma}}{1-\gamma} + \frac{(L-h)^{1-\phi}}{1-\phi}a_t, \quad (1) \]

where \( L (= 100, \text{for example}) \) is a physiological upper bound on the number of hours worked weekly, \( \gamma \) and \( \phi \) are non-negative parameters, and the positive factor \( a_t \), which governs the substitution between consumption and leisure, has the form

\[ \ln(a_t) = Z^a_{\beta^a} + \sigma^a \varepsilon^a, \quad (2) \]

where \( Z^a \) contains observable characteristics, while \( \varepsilon^a \) stands for unobservable preference heterogeneity. We also posit the following stochastic specification for the fixed cost of being employed

\[ b = Z^b_{\beta^b} + \sigma^b \varepsilon^b \quad (3) \]

where \( \varepsilon^b \) reflects unobservable heterogeneity in work costs across individuals.
The specification of log market wages is given by
\[ \ln(w) = Z^w \beta^w + \sigma^w \varepsilon^w, \] (4)
and for log consumption
\[ \ln(c) = Z^c \beta^c + \sigma^c \varepsilon^c. \] (5)

The residuals \((\varepsilon^a, \varepsilon^b, \varepsilon^c, \varepsilon^w)\) are assumed to be standard centered jointly normal. We begin by assuming \(\varepsilon^b\) strongly exogenous independent of \((\varepsilon^a, \varepsilon^c, \varepsilon^w)\).

From the analysis of the previous section of the optimizing household, we have the marginal utility of wealth given by
\[ \lambda = c^{-\gamma}, \]
the optimal hours \(h^e\) when working maximize
\[ \frac{(L - h)^{1-\phi}}{1 - \phi} a(\alpha) + c^{-\gamma} R(w, h), \] (6)
and the household chooses to stay out of the labour market whenever
\[ \frac{(L - h^e)^{1-\phi}}{1 - \phi} a(\alpha) + c^{-\gamma} R(w, h^e) - b < \frac{L^{1-\phi}}{1 - \phi} a(\alpha) + c^{-\gamma} R(w, 0). \]

It will be useful when deriving analytic expressions for the likelihood function to let
\[ v(h) = \frac{(L - h)^{1-\phi}}{1 - \phi}, \]
\[ V(c, h, a, b) = a v(h) + c^{-\gamma} [R(w, h) - b], \] (7)
where we leave the wage and other exogenous variables as mute arguments of \(V\).

For a disposable income function \(R\) which is not concave in \(h\), some values of hours may never be chosen by an optimizing agent who behaves according to the above model. Figure 1, with weekly hours on the x axis and disposable income on the y axis, illustrates this point. On this Figure, the \(R\) function has a flat horizontal portion between 0 and 16 hours per week, so that a choice of 14 hours, shown with a vertical line on the graph, can never be optimal, since it is strictly dominated by shorter positive hours, whatever the unobservable characteristics.

To check whether this phenomenon is important on real data, we rewrite the optimiza-
tion condition (6) as
\[
\frac{(L - h^e)^{1-\phi}}{1-\phi}a + c^{-\gamma}R(w, h^e) - \frac{(L - h)^{1-\phi}}{1-\phi}a - c^{-\gamma}R(w, h) \geq 0,
\]
where $h^e$ is the observed choice and $h$ is any other possible length of the workweek. Using the specification for $a$, separating the cases where $h$ is smaller than $h^e$ from those where $h$ is larger than $h^e$

\[
c^\ast a \leq \left\{ \frac{R(w, h) - R(w, h^e)}{(L-h^e)^{1-\phi}} - \frac{R(w, h^e) - R(w, h)}{(L-h)^{1-\phi}} \right\},
\]
for all $h$ smaller than $h^e$, with the inequality in the other direction

\[
c^\ast a \geq \left\{ \frac{R(w, h) - R(w, h^e)}{(L-h^e)^{1-\phi}} - \frac{R(w, h^e) - R(w, h)}{(L-h)^{1-\phi}} \right\}
\]
for all $h$ larger than $h^e$. The choice $h^e$ is compatible with optimization under our specification if and only if there is an $a$ satisfying the two above inequalities, i.e.

\[
\min_{h \leq h^e} (1 - \phi) \frac{R(w, h) - R(w, h^e)}{(L-h^e)^{1-\phi} - (L-h)^{1-\phi}} \geq \max_{h \geq h^e} (1 - \phi) \frac{R(w, h) - R(w, h^e)}{(L-h^e)^{1-\phi} - (L-h)^{1-\phi}}.
\]

This inequality can easily be checked on real data, since the only parameter that appears in it is $\phi$. In fact there are two ways of violating the condition: the positivity of the left hand side does not depend at all on the parametric specification, but only on the shape of the function $R$ and on the value of hours $h^e$ (see three top graphs on Figure 2); the second
inequality on the other hand does depend on $\phi$ and is illustrated in the two bottom graphs. These are examples from the sample we use in estimation and we return to this discussion when we examine the details of the data and estimation below. First we need to develop a coherent model of restricted choice that can account for such observations.

4 A model with restrictions on offered hours

To introduce our extension of the standard model, we consider choices over discrete hours. In this framework the typical worker characterized by a parameter $\beta$, observed exogenous characteristics $Z$ and unobserved characteristics $\varepsilon$ chooses $h$ that maximizes

$$U(h, Z, \beta, \varepsilon).$$

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 $\beta$ and $Z$, any distribution of $\varepsilon$ yields a probability distribution on $H$. We shall denote the probability of choosing $h_i$ in $H$ as...
We assume that given \( U \), the observation of the family of probabilities \( p_i(H,Z,\beta) \) identifies the parameter \( \beta \), 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, for some \( h_j \) alternative we have

\[
p_j(H,Z,\beta) = 0,
\]

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

To tackle this issue, we suppose that the agents do not make their choices over the whole set \( \mathcal{H} \), but on a random subset of it. We analyze how this modified model works, and in particular the sets of assumptions under which it still allows to identify the parameters \( \beta \) of the underlying structural model.

### 4.1 The two-offer model

Suppose that there is a distribution of offers, the probability of being offered \( h_i \) being equal to \( g_i \), \( g_i > 0 \), \( \sum_{i=1}^{I} g_i = 1 \). First consider the case where the individuals draw independently two offers from \( g \) and choose the one that yields the highest utility. \(^{1}\) The distribution of the observed choices \( \ell_2(Z,\beta) \) (the first index 2 serves to mark that there are two offers) then takes the form

\[
\ell_2(Z,\beta) = g_i^2 + 2g_i \sum_{j \neq i} g_j p_i(i \cup \{j\},Z,\beta),
\]

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

\(^{1}\)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 will imply though that the distribution of offers is independent of the past hours worked.
\( p_i(\{i, j\}, Z, \beta) + p_j(\{i, j\}, Z, \beta) = 1 \) 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 of them from the mere observation of the choice distribution \( \ell \). We explore how well chosen restrictions allow to identify the remaining parameters.

4.2 Increasing the number of offers

When the probability \( g \) has full support, in the two offer case 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 choice model as \( n \) goes to infinity.

5 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.

5.1 Independence of irrelevant alternatives

As a first step, consider the case of independence of irrelevant alternatives, where for all \( i, j \)

\[
p_i(\{i, j\}, X, \beta) = \frac{p_i(\mathcal{H}, X, \beta)}{p_i(\mathcal{H}, X, \beta) + p_j(\mathcal{H}, X, \beta)},
\]

or

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

where to alleviate notation we drop the arguments \( Z \) and \( \beta \), 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 1. Let \( \ell \) 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

\[
\ell_i = g_i^2 + 2g_ip_i \sum_{j \neq i} \frac{g_j}{p_i + p_j} \quad \text{for } i = 1, \ldots, I.
\]

(9)

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

\[
\bar{p}_I \geq \frac{p^0_i}{\min_i p^0_i} \quad \text{and} \quad \bar{p}_I \geq \frac{p^1_i}{\min_i p^1_i},
\]

and define \( \lambda^0 \) and \( \lambda^1 \) through

\[
\lambda^0 p^0_I = \lambda^1 p^1_I = \bar{p}_I.
\]

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

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

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

\[
\ell_i = P_i(p_1, \ldots, p_{I-1}, \bar{p}_I) \quad \text{for } i = 1, \ldots, I - 1
\]

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

\[
\frac{\partial P_i}{\partial p_i} = 2g_i \sum_{j=1}^{I} \frac{g_j p_j}{p_i + p_j},
\]
and for \( j \) different from \( i \)

\[
\frac{\partial P_i}{\partial p_j} = -2g_ip_i \frac{g_j}{(p_i + p_j)^2}.
\]

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

\[
\left| \frac{\partial P_i}{\partial p_i} \right| > \sum_{j=1, j\neq i}^{l-1} \left| \frac{\partial P_i}{\partial p_j} \right|,
\]

that is

\[
2g_i \sum_{j=1}^{l} \frac{g_j p_j}{p_i + p_j} > \sum_{j=1, j\neq i}^{l-1} 2g_j p_j \frac{g_j}{(p_i + p_j)^2}
\]
or

\[
\sum_{j=1}^{l-1} g_j p_j \left( \frac{1}{p_i + p_j} - \frac{1}{(p_i + p_j)^2} \right) + g_i p_i \left( \frac{1}{p_i + p_i} \right) > 0.
\]

Since \( p_i + p_j \) is larger than 1, the inequality is satisfied, and the right hand side mapping is univalent on \([1, \infty)^{l-1}\), 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 \( J \)

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

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 1 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. I do not know whether there can be multiple candidates.\(^2\)

\(^2\)There cannot be two solutions with two disjoint sets \( J_1 \) and \( J_2 \). Indeed one would need to have

\[
\ell_{J_1} = g_{J_1}^2 \quad \ell_{J_2} = g_{J_2}^2,
\]
5.2 The random utility model

Consider now another example where the agent has utility \( a_i - \varepsilon_i \) 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. This is closest to the labour supply model with discrete hours.

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

\[
p_i(\{i, j\}) = F_{ij}(a_i - a_j).
\]

Since only the differences \( a_i - a_j \) can be identified, we normalize \( a_I \) to zero. As in the IIA case, the number of unknowns is equal to the number of equations, and we may hope for exact identification. Indeed

**Lemma 2.** Let \( \ell \) 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 \( a_i \) with \( a_I = 0 \) that satisfies the system of equations

\[
\ell_i = g_i^2 + 2g_i \sum_{j \neq i} g_j F_{ij}(a_i - a_j) \quad \text{for } i = 1, \ldots, I.
\]  

**Proof:** For all \( i = 1, \ldots, I - 1 \), denote

\[
Q_i(a_i) = -\ell_i + g_i^2 + 2g_i \sum_{j \neq i} g_j F_{ij}(a_i - a_j),
\]

and \( Q(a) \) the \( I - 1 \) vector obtained by stacking up the \( Q_i \)s. By construction, any \( a \) such that \( Q(a) = 0 \) satisfies (10), since the \( I \)th 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

\[
\frac{\partial Q_i}{\partial a_i} = 2g_i \sum_{j \neq i} g_j f_{ij}(a_i - a_j),
\]

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.
while for $j \neq i, j \neq I$, 
\[ \frac{\partial Q_i}{\partial a_j} = -2g_i g_j f_{ij}(a_i - a_j). \]

The diagonal terms are positive and the off-diagonal negative. The sum of the elements on line $i$ is positive equal to 
\[ 2g_i g_I f_{ii}(a_i). \]

6 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 for all $i$ different from $j$. We study whether the choices $\ell_i$ of agents getting two independent offers are constrained by the model, and whether the observation of $\ell$ allows to recover the probability of offers $g$. From (8), we have by definition

\[ \ell_i = g_i^2 + 2g_i \sum_{j \neq i} g_j p_{ij} \]  

(11)

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

\[ p_{ij} + p_{ji} = 1. \]  

(12)

Lemma 3. Given the choice probabilities $p_{ij}, p_{ij} \geq 0$ satisfying (12), for any observed probability $\ell_i$ in the simplex of $R^I$, there exists a unique offer probability $g_i$ in the simplex of $R^I$ which satisfies (11).

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

\[ Q_i(g) = g_i^2 + 2g_i \sum_{j \neq i} g_j p_{ij}. \]

By construction, for $g$ in the simplex of $R^I$, under (12), $Q(g)$ also belongs to the simplex of $R^I$. Indeed

\[ \sum_{i=1}^I Q_i(g) = \sum_{i=1}^I \left[ g_i^2 + 2g_i \sum_{j<i} g_j \right] = 1. \]
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 \( \Gamma \) is well defined: since \( g, \overline{Q} \) and \( \ell \) all belong to the simplex, the denominator is larger than 1. Therefore \( \Gamma \) 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
\[ \max(0, g_i^* - \overline{Q}_i(g^*) + \ell_i) = g_i^* - \overline{Q}(g^*) + \ell, \]
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
\[ g_i \left[ 2g_i + 2 \sum_{j \neq i} g_j p_{ij} \right] > \sum_{j \neq i} 2g_j g_i p_{ij}. \]

As we have seen in section 3, in the non-linear budget constraint cases that we are interested in, the choice probabilities can exclude some alternatives, say \( i = 1 \) to \( k \). In the current setup, this means that, for all \( i \leq k \), for all \( j \neq i \)
\[ p_{ij} = 0. \]

Then the two offers model allows to rationalize the data by letting for \( i = 1 \) to \( k \)
\[ \ell_i = g_i^2 + 2g_i \sum_{j \neq i, j \leq k} g_j p_{ij}, \]
a nonlinear system which can be shown by an adaptation of the proof of the above lemma to have a unique solution, satisfying $\sum_1^k g_i = \sqrt{\sum_1^k \ell_i}$.

7 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 $g_i$. In the absence of non-convexities in the budget constraint the choice probabilities and the offer probabilities will not be separately identified.

Consider the labour supply setting for our empirical application. The utility from hours choice $h_i$ is given by $V(c, h_i, a, b)$ in (7). Preference and fixed cost heterogeneity enter through $a$ and $b$ respectively and depend on a set of exogenous observed characteristics $z$ and unobserved heterogeneity $\zeta$. Although consumption and wages are also both treated as endogenous they are determined outside the within period hours choice and so we condition on them in the arguments here. The utility from choice $h_i$ is then given by $U(h_i, z, \zeta)$ and the probability that hours $h_i$ are chosen when the pair $(h_i, h_j)$ are available is given by

$$p_{ij} = \Pr[U(h_i, z, \zeta) - U(h_j, z, \zeta) > 0].$$

To make progress with identification in this case we assume that each $a_i$ for $i = 1, \ldots, I-1$ is a smooth function of a finite parameter vector $\gamma$ and, in a similar fashion, the offer probability $g_i$ is a smooth function of finite parameter vector $\beta$, where

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

and where $I$ is the number of possible choices.

From (10) we can write the system of equations

$$Q_i = -\ell_i + g_i(\beta)^2 + 2g_i(\beta) \sum_{j \neq i} g_j(\beta) F_{ij}(a_i(\gamma) - a_j(\gamma)) \text{ for } i = 1, \ldots, I - 1. \quad (13)$$

For identification we require full column rank of the matrix

$$\Pi = \begin{bmatrix} \partial Q / \partial a & \partial Q / \partial \gamma \\ \partial a / \partial \gamma & \partial g / \partial \beta \end{bmatrix}. \quad (14)$$

where the matrix of derivatives relating to the $Q_i$ has elements of the form.
\[
\frac{\partial Q_i}{\partial a_i} = 2g_i \sum_{j \neq i} g_j f(a_i - a_j) \quad (15)
\]
\[
\frac{\partial Q_i}{\partial a_j} = -2g_i g_j f(a_i - a_j) \quad (16)
\]
\[
\frac{\partial Q_i}{\partial g_i} = 2g_i + 2 \sum_{j \neq i} g_j F(a_i - a_j) \quad (17)
\]
\[
\frac{\partial Q_i}{\partial a_j} = 2g_i g_j F(a_i - a_j) \quad (18)
\]

We note that \( \frac{\partial Q_i}{\partial a_i} > 0 \) and \( \frac{\partial Q_i}{\partial a_j} < 0 \) where the row sum is also positive.

Inspection of the elements of \( \Pi \), (15),.. (18), suggests no natural linear dependence and, in general, the rank condition should be satisfied. In the application there will be some workers facing non-convex budget constraints which will further help with identification.

8 Data and Sample Likelihood

8.1 The UK FES Data

The sample we use is made up of women with children, either lone or married mothers. We use years 1997 to 2002 of the UK Family Expenditure Survey (FES) and hence have information about consumption, usual hours worked, gross wage earnings, household information. Tables 1 and 2 provide some basic descriptive statistics.

The overall sample contains some 11,448 women spread family evenly across the six years under study. The majority of women in this sample have just basic education qualifications, meaning that they left formal schooling at the minimum school leaving age of 16. The majority of the rest stay on to complete secondary school with less than 20% having a college or university degree. The modal family size is two and a little under 30% of the sample have a youngest child aged less than 5 (the formal school entry age in the UK). A little under 80% of the women in our sample are married or cohabiting (we label all these as ‘cohabiting’), leaving between 20% and 25% of the mothers in the sample as single parents.

The median hours of work for this sample is between 25 and 30 hours per week with a wide distribution.
8.2 The likelihood function

We observe the employment status and the consumption of the workers in the survey, their wage rates and weekly hours when employed. We want to derive the distribution of these quantities from the model, given the parameters and the distribution of the disturbances \((\varepsilon^a, \varepsilon^b, \varepsilon^c, \varepsilon^w)\).

We first consider the employment status. Knowing \((\varepsilon^a, \varepsilon^c, \varepsilon^w)\), i.e. consumption, wage and the parameter \(a\), at weekly hours \(h\) one is 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 of \(b\), the probability of this event knowing \(a\) is easily computed from the cumulative distribution of \(\varepsilon^b\):

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

8.2.1 The model with unrestricted hours choices

When the household works, the best choice of hours \(h\) is such that

\[
V(c, h, a, b) \geq V(c, h', a, b) \text{ for all } h' \text{ in } \mathcal{H}.
\]

By linearity the \(b\) term drops from these inequalities, which become

\[
av(h) + c^{-\gamma}R(w, h) \geq av(h') + c^{-\gamma}R(w, h')
\]

for all \(h'\) in the choice set. From the monotonicity of \(v\), this can be rewritten equivalently

\[
\max_{h' > h} c^{-\gamma} \frac{R(w, h') - R(w, h)}{v(h) - v(h')} \leq a \leq \min_{h' < h} c^{-\gamma} \frac{R(w, h) - R(w, h')}{v(h') - v(h)}.
\]

Let

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

\[
\varepsilon^a(c, h, w) = \frac{1}{\sigma^a} \left\{-\gamma \ln c + \ln \left[\min_{h' < h} \frac{R(w, h) - R(w, h')}{v(h') - v(h)}\right] - Z^a \beta^a\right\}.
\]
The likelihood function associated with a worker spending \( h \) hours at work, conditional on hourly wage \( w \) and consumption \( c \) is

\[
\int_{\pi^a(c,h,w)} \int F^b(\varepsilon, c, h, w) \phi(\varepsilon) d\varepsilon.
\]

For an individual who is not in employment, the computation is different because we do not observe her wage, nor her potential hours. The cost of going to work \( b \) is larger than the benefit, whatever the hours worked, so that for a wage rate \( w \) the probability of non working is

\[
1 - \max_{h>0} F^b(\varepsilon^a, c, h, w).
\]

The likelihood function is the integral of the above expression with respect to \( \varepsilon^a \) and \( w \).

### 8.2.2 The two-offer model

When the household works and receive two offers \( h \) and \( h' \), \( h \) is preferred to \( h' \) 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
\[
G(h|c, w) = g(h) \left\{ \sum_{h' < h} 2g(h') \int_{-\infty}^{\alpha(c,h',w)} F^h(\varepsilon, c, h, w) \phi(\varepsilon) d\varepsilon 
+ \sum_{h' > h} 2g(h') \int_{\alpha(c,h',w)}^{+\infty} F^h(\varepsilon, c, h, w) \phi(\varepsilon) d\varepsilon \right\}.
\]

Finally the probability of being unemployed 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
\[
\sum_h \sum_{h'} g(h)g(h') \int_{-\infty}^{\infty} \Phi\left[ \frac{1}{\sigma_b} \left( R(w, 0) + c^\gamma v(0) \exp(Z^a \beta^a + \sigma^a \varepsilon) + Z^b \beta^b \right.ight.
- \max(R(w, h) + c^\gamma v(h) \exp(Z^a \beta^a + \sigma^a \varepsilon), R(w, h') + c^\gamma v(h') \exp(Z^a \beta^a + \sigma^a \varepsilon)) \left.\right) \right] \phi(\varepsilon) d\varepsilon.
\]

Since the wage of the unemployed agents is not known, $w$ has to be integrated out in the above expression.

9 Estimation results

Returning first to the results for the optimising model. We find that 7.5% of sampled women do not comply with the optimisation inequality presented in section 3. For this group we reject the optimising with unrestricted hours offer nonparametrically as there are clearly hours of work that strictly dominate the observed choices. The budget constraints associated with some of this group were used in Figure 2 above. In Table 3 we present some of their characteristics. Note that these ‘non optimizing’ women are more often lone mothers than married ones. Their average wage is lower than the optimising women’s average wage. We notice too that their hours distribution is on the left hand side of the optimising hour distribution.

Table 4 presents the estimation results for the parameters of preferences, fixed costs and the offer distribution respectively. In the first column are the results for the ‘optimising’ model that assumes unrestricted hours choices and in the second panel are the estimates for the two-offer specification. The $\phi$ and $\gamma$ parameters refer to the exponents on hours (non-market work) and on consumption as described in the utility specification (1) of section 3. The next panel refers to the parameters that influence the marginal utility of hours through

\[\text{The probability of getting a couple of offers } (h, h'), h \neq h', \text{ is } 2g(h)g(h'), \text{ while that of getting } (h, h) \text{ is } g(h)^2.\]
Figure 3: Hours distributions for optimising and non optimising individuals
the specification of \( \ln(a) \) in equation (2). Following these are the parameters of fixed costs (3). The only remaining parameters in the unrestricted choice model are the variance and covariances parameters that link the unobserved heterogeneity terms in preferences, wages (4) and consumption (5). The unobserved heterogeneity terms are assumed to have a joint normal distribution as described in Section 3.

The estimated labour supply elasticities at the intensive and extensive margins for the optimising model specification are presented in Table 5. The median value for the distribution of estimated intensive elasticities is .39 with a p90 to P10 range of 1.07 to .24. These estimates lie in the range of estimates from various studies of female labour supply in the UK and in North America, see Blundell and MaCurdy (1999). At the extensive margin the corresponding estimates are **.

The model with restricted choices can be estimated on the whole sample and the estimated parameters for this specification are presented in the second column of Table 4. It is noticeable that the exponent parameters \( \phi \) and \( \gamma \) are both smaller in this specification which allows for restricted choices, suggesting a more responsive underlying preference structure. This can be seen from the distribution of implied elasticities which would occur in the absence of hours restrictions. The final column in Table 5 presents these elasticities. The median is larger at .49. Note too that this is the distribution for the full sample inclusive of the 7.5% of individuals that had to be excluded from the optimising model estimation. The elasticities are also larger across the distribution as can be seen from P90-P10 range which is now 1.34 to .27. (Note that we need to add more on the simulated elasticities and the extensive elasticities too for the two-offer model).

In addition to the preference and fixed cost parameters there are the parameters of the offer distribution as described in (8) of section 4.1 above. Offers are modelled as a mixture of two independent normals and the parameter estimates are in the final panel of Table 4. They suggest offers concentrated around full time at 36 hours and low part-time at 17 hours.

10 Conclusions

In this paper we have developed a model of labour supply and consumption in which individuals face an offer distribution over possible hours choices. Observed hours reflect both the distribution of preferences and the distribution of offers. The leading example is of individuals selecting from two offers. Their choice set is limited and their observed
hours will not necessarily satisfy the conditions for optimal choice.

We illustrated this framework in a labour supply model with nonlinear budget constraints where observed behaviour cannot necessarily be reconciled with rational choice theory. We showed first that when the offer distribution is known preferences can be identified. We were also able to show that, where preferences are known, the offer distribution can be fully recovered. We then developed conditions for identification of both the parameters of preferences and of the offer distribution.

This model has been used to study the labour supply and consumption choices of a large sample of women in the UK, accounting for nonlinear budget constraints and fixed costs of work. The results point to a small but important group of workers who fail the unrestricted choice optimising model. For the remainder of the sample the optimising model suggests a distribution of estimated intensive elasticities with a median of .39 and with a p90 to P10 range of 1.07 to .24. The model with restricted choices can be estimated on the whole sample and suggest an underlying P90- P10 range for the intensive elasticities of 1.34 to .27 with a median around .5 were restrictions on offers to be removed. Additionally we recover a distribution of hours offers with twin peaks centered around full-time and part-time hours.

References


Table 1: Some Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>Sample</th>
<th>1997</th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>11,458</td>
<td>1,794</td>
<td>1,858</td>
<td>1,935</td>
<td>1,850</td>
<td>2,116</td>
<td>1,905</td>
</tr>
<tr>
<td>Education Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>56.41</td>
<td>58.86</td>
<td>57.91</td>
<td>55.71</td>
<td>55.46</td>
<td>56.81</td>
<td>53.86</td>
</tr>
<tr>
<td>2</td>
<td>25.82</td>
<td>24.92</td>
<td>25.30</td>
<td>27.29</td>
<td>26.49</td>
<td>24.86</td>
<td>26.09</td>
</tr>
<tr>
<td>3</td>
<td>17.77</td>
<td>16.22</td>
<td>16.79</td>
<td>17.00</td>
<td>18.05</td>
<td>18.34</td>
<td>20.05</td>
</tr>
<tr>
<td>Number of children</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>38.53</td>
<td>40.41</td>
<td>36.44</td>
<td>40.05</td>
<td>38.54</td>
<td>37.33</td>
<td>38.58</td>
</tr>
<tr>
<td>2</td>
<td>41.55</td>
<td>41.64</td>
<td>42.84</td>
<td>39.12</td>
<td>39.78</td>
<td>42.58</td>
<td>43.25</td>
</tr>
<tr>
<td>4</td>
<td>3.98</td>
<td>3.40</td>
<td>4.63</td>
<td>3.77</td>
<td>4.16</td>
<td>4.40</td>
<td>3.46</td>
</tr>
<tr>
<td>5+</td>
<td>1.42</td>
<td>1.23</td>
<td>1.40</td>
<td>1.71</td>
<td>1.14</td>
<td>1.80</td>
<td>1.21</td>
</tr>
<tr>
<td>Youngest kid between 0 and 4</td>
<td>29.25</td>
<td>30.99</td>
<td>29.98</td>
<td>28.63</td>
<td>26.81</td>
<td>29.68</td>
<td>29.40</td>
</tr>
<tr>
<td>Youngest kid between 5 and 10</td>
<td>30.30</td>
<td>28.76</td>
<td>27.93</td>
<td>28.53</td>
<td>28.16</td>
<td>33.51</td>
<td>34.38</td>
</tr>
<tr>
<td>Cohabitant</td>
<td>76.94</td>
<td>77.81</td>
<td>78.20</td>
<td>75.71</td>
<td>78.22</td>
<td>76.04</td>
<td>75.91</td>
</tr>
<tr>
<td>Spouse Inwork Status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>If in work</td>
<td>65.92</td>
<td>66.27</td>
<td>67.24</td>
<td>66.75</td>
<td>66.84</td>
<td>65.56</td>
<td>63.22</td>
</tr>
<tr>
<td>If out of work</td>
<td>59.20</td>
<td>57.12</td>
<td>63.89</td>
<td>60.51</td>
<td>59.02</td>
<td>57.86</td>
<td>56.24</td>
</tr>
</tbody>
</table>
Table 2: Consumption, Wages and Hours of Work

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>St Error</th>
<th>p1</th>
<th>p5</th>
<th>p10</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>p90</th>
<th>p95</th>
<th>p99</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Log of consumption</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>5.30</td>
<td>0.59</td>
<td>3.92</td>
<td>4.26</td>
<td>4.49</td>
<td>4.95</td>
<td>5.34</td>
<td>5.67</td>
<td>6.00</td>
<td>6.21</td>
<td>6.67</td>
</tr>
<tr>
<td>1998</td>
<td>5.28</td>
<td>0.59</td>
<td>3.88</td>
<td>4.24</td>
<td>4.48</td>
<td>4.92</td>
<td>5.30</td>
<td>5.67</td>
<td>6.01</td>
<td>6.20</td>
<td>6.63</td>
</tr>
<tr>
<td>1999</td>
<td>5.33</td>
<td>0.61</td>
<td>3.84</td>
<td>4.23</td>
<td>4.49</td>
<td>4.96</td>
<td>5.37</td>
<td>5.72</td>
<td>6.06</td>
<td>6.25</td>
<td>6.77</td>
</tr>
<tr>
<td>2000</td>
<td>5.37</td>
<td>0.60</td>
<td>3.93</td>
<td>4.33</td>
<td>4.59</td>
<td>5.01</td>
<td>5.40</td>
<td>5.75</td>
<td>6.09</td>
<td>6.30</td>
<td>6.83</td>
</tr>
<tr>
<td>2001</td>
<td>5.35</td>
<td>0.61</td>
<td>3.92</td>
<td>4.33</td>
<td>4.54</td>
<td>4.98</td>
<td>5.37</td>
<td>5.75</td>
<td>6.08</td>
<td>6.32</td>
<td>6.89</td>
</tr>
<tr>
<td>2002</td>
<td>5.36</td>
<td>0.61</td>
<td>3.87</td>
<td>4.32</td>
<td>4.58</td>
<td>4.98</td>
<td>5.38</td>
<td>5.75</td>
<td>6.12</td>
<td>6.29</td>
<td>6.80</td>
</tr>
<tr>
<td><strong>Hourly wage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>0.31</td>
<td>0.31</td>
<td>0.08</td>
<td>0.10</td>
<td>0.12</td>
<td>0.16</td>
<td>0.23</td>
<td>0.35</td>
<td>0.55</td>
<td>0.72</td>
<td>1.42</td>
</tr>
<tr>
<td>1998</td>
<td>0.31</td>
<td>0.27</td>
<td>0.07</td>
<td>0.10</td>
<td>0.12</td>
<td>0.16</td>
<td>0.24</td>
<td>0.37</td>
<td>0.56</td>
<td>0.72</td>
<td>1.26</td>
</tr>
<tr>
<td>1999</td>
<td>0.32</td>
<td>0.33</td>
<td>0.07</td>
<td>0.10</td>
<td>0.12</td>
<td>0.16</td>
<td>0.23</td>
<td>0.36</td>
<td>0.61</td>
<td>0.79</td>
<td>1.71</td>
</tr>
<tr>
<td>2000</td>
<td>0.31</td>
<td>0.31</td>
<td>0.07</td>
<td>0.10</td>
<td>0.12</td>
<td>0.16</td>
<td>0.23</td>
<td>0.35</td>
<td>0.57</td>
<td>0.75</td>
<td>1.44</td>
</tr>
<tr>
<td>2001</td>
<td>0.34</td>
<td>0.45</td>
<td>0.08</td>
<td>0.11</td>
<td>0.13</td>
<td>0.18</td>
<td>0.25</td>
<td>0.39</td>
<td>0.62</td>
<td>0.81</td>
<td>1.57</td>
</tr>
<tr>
<td>2002</td>
<td>0.34</td>
<td>0.48</td>
<td>0.07</td>
<td>0.11</td>
<td>0.13</td>
<td>0.18</td>
<td>0.25</td>
<td>0.37</td>
<td>0.56</td>
<td>0.79</td>
<td>1.64</td>
</tr>
<tr>
<td><strong>Usual hours of work</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>26.23</td>
<td>10.95</td>
<td>5</td>
<td>8</td>
<td>12</td>
<td>18</td>
<td>25</td>
<td>36</td>
<td>40</td>
<td>42</td>
<td>50</td>
</tr>
<tr>
<td>1998</td>
<td>26.05</td>
<td>10.81</td>
<td>5</td>
<td>8</td>
<td>11</td>
<td>18</td>
<td>25</td>
<td>36</td>
<td>39</td>
<td>40</td>
<td>48</td>
</tr>
<tr>
<td>1999</td>
<td>26.75</td>
<td>10.81</td>
<td>5</td>
<td>9</td>
<td>12</td>
<td>18</td>
<td>27</td>
<td>37</td>
<td>40</td>
<td>42</td>
<td>49</td>
</tr>
<tr>
<td>2000</td>
<td>27.74</td>
<td>10.68</td>
<td>6</td>
<td>10</td>
<td>14</td>
<td>20</td>
<td>30</td>
<td>37</td>
<td>40</td>
<td>43</td>
<td>50</td>
</tr>
<tr>
<td>2001</td>
<td>27.28</td>
<td>10.61</td>
<td>6</td>
<td>10</td>
<td>13</td>
<td>19</td>
<td>28</td>
<td>37</td>
<td>40</td>
<td>42</td>
<td>48</td>
</tr>
<tr>
<td>2002</td>
<td>27.41</td>
<td>10.23</td>
<td>6</td>
<td>10</td>
<td>14</td>
<td>20</td>
<td>28</td>
<td>37</td>
<td>40</td>
<td>41</td>
<td>48</td>
</tr>
</tbody>
</table>
Table 3: Non optimizing and optimizing agents

<table>
<thead>
<tr>
<th></th>
<th>Optimizing agents</th>
<th>Non optimizing agents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion among ‘in work’ women</td>
<td>0.88</td>
<td>0.12</td>
</tr>
<tr>
<td>Age at end of studies</td>
<td>17.5</td>
<td>16.7</td>
</tr>
<tr>
<td>Age</td>
<td>38.0</td>
<td>37.2</td>
</tr>
<tr>
<td>Hourly wage</td>
<td>7.4</td>
<td>5.5</td>
</tr>
<tr>
<td>Marginal tax rate</td>
<td>0.30</td>
<td>0.60</td>
</tr>
<tr>
<td>Usual weekly hours</td>
<td>27.5</td>
<td>22.6</td>
</tr>
<tr>
<td>Log of consumption</td>
<td>5.5</td>
<td>5.1</td>
</tr>
<tr>
<td>Number of kids</td>
<td>1.7</td>
<td>1.8</td>
</tr>
<tr>
<td>A kid younger than 4</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>The youngest kid between 5 and 10</td>
<td>0.3</td>
<td>0.4</td>
</tr>
<tr>
<td>London</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Cohabitant</td>
<td>0.88</td>
<td>0.46</td>
</tr>
<tr>
<td>Spouse inwork</td>
<td>0.72</td>
<td>0.22</td>
</tr>
<tr>
<td>Out of work income</td>
<td>355</td>
<td>199</td>
</tr>
<tr>
<td>In work income</td>
<td>513</td>
<td>255</td>
</tr>
</tbody>
</table>
Table 4: Estimation results

<table>
<thead>
<tr>
<th></th>
<th>Optimizing</th>
<th>Two Offer</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>6.27</td>
<td>5.49</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>1.30</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>$a$: Constant</td>
<td>23.84</td>
<td>23.89</td>
</tr>
<tr>
<td></td>
<td>(0.31)</td>
<td>(0.46)</td>
</tr>
<tr>
<td>$a$: Cohabitant</td>
<td>-1.90</td>
<td>-0.47</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>$a$: Youngest kid between 0 and 4</td>
<td>1.20</td>
<td>1.23</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>$a$: Youngest kid between 5 and 9</td>
<td>0.07</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>$b$: Constant</td>
<td>0.97</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>(1.34)</td>
<td>(0.43)</td>
</tr>
<tr>
<td>$b$: Cohabitant</td>
<td>0.86</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>(1.23)</td>
<td>(0.25)</td>
</tr>
<tr>
<td>$b$: Number of kids</td>
<td>0.83</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>(0.37)</td>
<td>(0.20)</td>
</tr>
<tr>
<td>$\sigma^b$</td>
<td>4.42</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.24)</td>
<td></td>
</tr>
<tr>
<td>$\rho(\varepsilon^a, \varepsilon^c)$</td>
<td>-0.34</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td></td>
</tr>
<tr>
<td>$\rho(\varepsilon^a, \varepsilon^w)$</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td>$\sigma^a$</td>
<td>2.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td></td>
</tr>
<tr>
<td>$m_1$</td>
<td>16.71</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td></td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>6.35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td></td>
</tr>
<tr>
<td>$m_2$</td>
<td>36.47</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td></td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>5.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td></td>
</tr>
<tr>
<td>$p_1$</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Optimizing Model</td>
<td>Two-Offer Model</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Mean</td>
<td>0.57 (0.47)</td>
<td>0.68 (0.56)</td>
</tr>
<tr>
<td>p10</td>
<td>0.24</td>
<td>0.27</td>
</tr>
<tr>
<td>p20</td>
<td>0.27</td>
<td>0.31</td>
</tr>
<tr>
<td>p30</td>
<td>0.28</td>
<td>0.34</td>
</tr>
<tr>
<td>p40</td>
<td>0.32</td>
<td>0.41</td>
</tr>
<tr>
<td>p50</td>
<td>0.39</td>
<td>0.49</td>
</tr>
<tr>
<td>p60</td>
<td>0.51</td>
<td>0.61</td>
</tr>
<tr>
<td>p70</td>
<td>0.64</td>
<td>0.73</td>
</tr>
<tr>
<td>p80</td>
<td>0.78</td>
<td>0.96</td>
</tr>
<tr>
<td>p90</td>
<td>1.07</td>
<td>1.33</td>
</tr>
</tbody>
</table>

Elasticity distributions are assessed assuming the absence of hour restrictions.