

Labour supply and taxation with restricted choices

IFS Working Paper W14/04

Magali Beffy
Richard Blundell
Antoine Bozio
Guy Laroque

The Institute for Fiscal Studies (IFS) is an independent research institute whose remit is to carry out rigorous economic research into public policy and to disseminate the findings of this research. IFS receives generous support from the Economic and Social Research Council, in particular via the ESRC Centre for the Microeconomic Analysis of Public Policy (CPP). The content of our working papers is the work of their authors and does not necessarily represent the views of IFS research staff or affiliates.

Labour Supply and Taxation with Restricted Choices*

Magali Beffy, Richard Blundell, Antoine Bozio and Guy Laroque[†]

September 2012, this draft February 2014

Abstract

A model of labour supply is developed in which individuals face restrictions on hours choices. Observed hours reflect both the distribution of preferences and the distribution of offers. In this framework the choice set is limited and observed hours may not appear to satisfy the revealed preference conditions for ‘rational’ choice. We show first that when the offer distribution is known, preferences can be identified. We then show that, where preferences are known, the offer distribution can be fully recovered. We also develop conditions for identification of both preferences and the offer distribution. We illustrate this approach in a labour supply setting with non-linear budget constraints. The occurrence of non-linearities in the budget constraint can directly reveal restrictions on choices. This framework is then used to study the labour supply choices 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.

*We thank seminar participants at Sciences-Po for their comments and the data archive UK and IPUMS for data access. This research is funded by the ERC (WSCWTBDS) 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

Observed hours of work among working age adults display considerable variation both over time and in the cross-section, especially among women with children (see, for example, Blundell, Bozio and Laroque, 2011). In this paper we ask under what conditions this variation can be used to identify preferences for labour supply. We start with the premise that it is unlikely that all workers are free to choose their working hours (see, for example, Blundell, Brewer and Francesconi, 2008). 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 welfare 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. There are dominant alternatives in the budget set. We may observe a lack of bunching at kink points, inconsistent with the pure neoclassical model of unrestricted hours choices (Saez, 2010; and Chetty et al, 2011).

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. We are not sure of their choice set. This is similar to the idea of a ‘consideration set’ in the modern literature on bounded rationality, see Kfir and Spiegel (2011), for example. 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 is one where rational choices are made from a set of hours restricted by the offers made by employers. Nonetheless, there are many similarities between the two frameworks. In this paper we develop and estimate a structural model of labour supply choices that embeds restrictions on the set of available hours.

We are not the first to examine hours restrictions in labour supply models. There is a long history of incorporating such restrictions into labour supply models, including, Aaberge (1999, 2009), Altonji and Paxson (1992), Bloemen (2000, 2008), Dickens and Lundberg (1993), Van Soest et al (1990), Ham and Reilly (2002) and Dagsvik and Strom (2006); see also the recent discussion in Chetty (2012). The ideas we develop here place these models of 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 do not make their choices over the whole set, 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)), and with 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 approaches 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. 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 provide direct evidence of hours restrictions by recording individuals working at hours of work that would be strictly dominated by other choices were a full range of hours choices to be available.

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). We use data from the UK Family Expenditure Survey over this period which 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. 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 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 rejections of the neoclassical 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 in the standard multinomial choice and random utility models. 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 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 and a higher level of employment 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 unrestricted hours choices at the extensive and intensive margins. In this specification there are non-linear taxes and fixed costs of work but otherwise workers are free to choose their hours of work.

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

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

subject to an intertemporal budget constraint

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

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 income after taxes are deducted and benefits received for someone who works h hours at wage w . This function will also depend on other characteristics that change tax rates and eligibility. 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 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 $u_t(c_t, h_t) + \lambda_t R(w_t, h_t)$. 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 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 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

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

where L ($= 100$, for example) is a physiological upper bound on the number of hours worked weekly, γ and ϕ are non-negative parameters, and the positive factor a , which

governs the substitution between consumption and leisure, has the form

$$\ln(a) = Z^a \beta^a + \sigma^a \varepsilon^a, \quad (2)$$

where Z^a contains observable characteristics, while ε^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 ε^b reflects unobservable heterogeneity in work costs across individuals.

The specification of log market wages is given by

$$\ln(w) = Z^w \beta^w + \sigma^w \varepsilon^w, \quad (4)$$

and for log consumption

$$\ln(c) = Z^c \beta^c + \sigma^c \varepsilon^c. \quad (5)$$

The residuals $(\varepsilon^a, \varepsilon^b, \varepsilon^c, \varepsilon^w)$ are assumed to be standard centered jointly normal.

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

$$\frac{(L - h)^{1-\phi}}{1 - \phi} a(\alpha) + c^{-\gamma} R(w, h), \quad (6)$$

and the household chooses to stay out of the labour market whenever the inequality

$$\frac{(L - h^e)^{1-\phi}}{1 - \phi} a(\alpha) + c^{-\gamma} [R(w, h^e) - b] < \frac{L^{1-\phi}}{1 - \phi} a(\alpha) + c^{-\gamma} R(w, 0).$$

is satisfied. It will be useful when deriving analytic expressions for the likelihood function to let

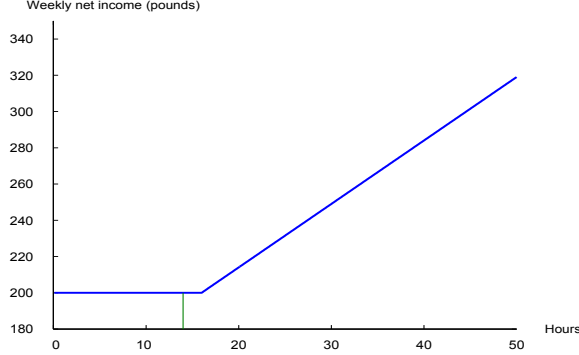
$$v(h) = \frac{(L - h)^{1-\phi}}{1 - \phi},$$

and

$$V(c, h, a, b) = av(h) + c^{-\gamma} [R(w, h) - b], \quad (7)$$

where we leave the wage and other exogenous variables as mute arguments of V .

Figure 1: A standard example of budget constraint



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. In 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 write the revealed preference inequality as

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

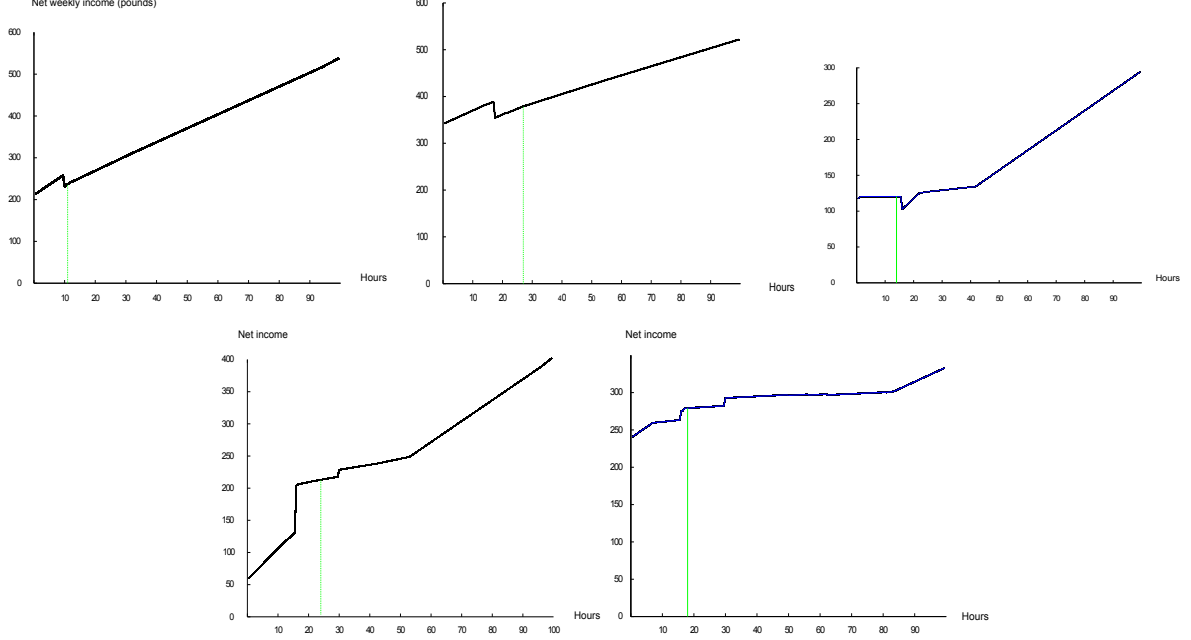
where h^e is the observed choice and h is any other possible length of the workweek. Using the specification for a , we can separate the cases where h is smaller than h^e from those where h is larger than h^e . That is

$$c^\gamma a \leq \left\{ \frac{R(w, h) - R(w, h^e)}{\frac{(L-h^e)^{1-\phi}}{1-\phi} - \frac{(L-h)^{1-\phi}}{1-\phi}} \right\}, \quad (8)$$

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

$$c^\gamma a \geq \left\{ \frac{R(w, h) - R(w, h^e)}{\frac{(L-h^e)^{1-\phi}}{1-\phi} - \frac{(L-h)^{1-\phi}}{1-\phi}} \right\} \quad (9)$$

Figure 2: Some examples of budget constraints for in work people



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.

$$\min_{h \leq h^e} (1 - \phi) \frac{R(w, h) - R(w, h^e)}{(L - h^e)^{1-\phi} - (L - h)^{1-\phi}} \geq \max \left[0, \max_{h \geq h^e} (1 - \phi) \frac{R(w, h) - R(w, h^e)}{(L - h^e)^{1-\phi} - (L - h)^{1-\phi}} \right]. \quad (10)$$

Inequality (10) can easily be checked on real data, since the only parameter that appears in it is ϕ . In fact there are two ways of violating the condition: the positivity of the left hand side of (10) 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 graphs in the upper panel of Figure 2); the second inequality on the other hand does depend on ϕ and is illustrated in the two graphs in the lower panel of Figure 2. These are budget constraints for specific individual women drawn from the sample we use in estimation. We return to this discussion when we examine the details of the data and estimation below. First we 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 β , observed exogenous characteristics Z and unobserved characteristics ε , chooses h that maximizes

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

The possible choices h belong to a finite set \mathcal{H} made of I elements $\{h_1, \dots, h_I\}$. 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 U , 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 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 assumptions under which it still allows to identify the parameters β 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 individuals draw independently two offers from g and choose the one that yields the highest utility.¹ The distribution

¹Note 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 though that the distribution of offers is independent of the past hours worked. This will be an important extension for future work.

of the observed choices $\ell_{2i}(Z, \beta)$ (the first index ‘2’ serves to mark that there are two offers) then takes the form

$$\ell_{2i}(Z, \beta) = g_i^2 + 2g_i \sum_{j \neq i} g_j p_i(\{i, j\}, Z, \beta), \quad (11)$$

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.

4.2 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 neoclassical unrestricted 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 (IIA), 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 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 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 p in the interior of the simplex of \mathcal{R}^I that satisfies the system of equations*

$$\ell_i = g_i^2 + 2g_i p_i \sum_{j \neq i} \frac{g_j}{p_i + p_j} \text{ for } i = 1, \dots, I. \quad (12)$$

Proof: For all i , denote

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

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

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

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

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

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

$$\ell_i = P_i(p_1, \dots, p_{I-1}, \bar{p}_I) \text{ for } i = 1, \dots, I - 1$$

with the unknowns (p_1, \dots, 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

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

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

$$\left| \frac{\partial P_i}{\partial p_i} \right| > \sum_{j=1, j \neq i}^{I-1} \left| \frac{\partial P_j}{\partial p_i} \right|,$$

that is

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

or

$$\sum_{j=1}^{I-1} g_j p_j \left(\frac{1}{p_i + p_j} - \frac{1}{(p_i + p_j)^2} \right) + g_I p_I \left(\frac{1}{p_i + p_I} \right) > 0.$$

Since $p_i + p_j$ is larger than 1, the inequality is satisfied, and the right hand side mapping is univalent on $[1, \infty)^{I-1}$, which completes the proof. \square

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 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. We do not know whether there can be multiple candidates.²

5.2 The random utility model

Consider now the random utility model where the agent has utility $a_i - \varepsilon_i$ for alternative i , $i = 1, \dots, I$, and under full optimization, knowing the value of her utilities, chooses the alternative which gives the highest utility. This is close to our labour supply model with discrete hours. We are able to show that our identification results extend to this model.

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 a_i . We denote F_{ij} the (assumed to be differentiable) cumulative distribution function of $\varepsilon_i - \varepsilon_j$ so that

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

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

Lemma 2. *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 a_i with $a_I = 0$ that satisfies the system of equations*

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

²There 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.

Proof: For all $i = 1, \dots, I - 1$, denote

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

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 (13), 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

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

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 ℓ_i of agents getting two independent offers are constrained by the model, and whether the observation of ℓ allows to recover the probability of offers g . From (11), we have by definition

$$\ell_i = g_i^2 + 2g_i \sum_{j \neq i} g_j p_{ij} \tag{14}$$

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

$$p_{ij} + p_{ji} = 1. \tag{15}$$

Lemma 3. *Given the choice probabilities p_{ij} , $p_{ij} \geq 0$ satisfying (15), 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 (14).*

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

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

By construction, for g in the simplex of \mathcal{R}^I , under (15), $\bar{Q}(g)$ also belongs to the simplex of \mathcal{R}^I . Indeed

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

Consider the mapping

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

First note that Γ is well defined: since g , \bar{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 $\bar{Q}_i(g^*) = 0$, so that

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

It follows that at the fixed point

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

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

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

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

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 $j \neq i$

$$p_{ij} = 0.$$

In this case the two-offer model allows us to rationalize the data by letting for $i = 1$ to k

$$\ell_i = g_i^2 + 2g_i \sum_{j \neq i, j \leq k} g_j p_{ij}.$$

This nonlinear system 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, without further assumptions, be separately identified.

Consider the 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 ζ . 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 $i = 1, \dots, I-1$ is a smooth function of a finite parameter vector γ and, in a similar fashion, the offer probability g_i is a smooth function of a finite parameter vector β , where

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

and where I is the number of possible choices.

From (13) we can write the system of equations

$$Q_i = -\ell_i + g_i(\beta)^2 + 2g_i(\beta) \sum_{j \neq i} g_j(\beta) F_{ij}(a_i(\gamma) - a_j(\gamma)) \text{ for } i = 1, \dots, I - 1. \quad (16)$$

For identification we require full column rank of the matrix

$$\Pi = \begin{bmatrix} \frac{\partial Q}{\partial a} & \frac{\partial Q}{\partial \gamma} \\ \frac{\partial Q}{\partial g} & \frac{\partial Q}{\partial \beta} \end{bmatrix}. \quad (17)$$

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

$$\frac{\partial Q_i}{\partial a_i} = 2g_i \sum_{j \neq i} g_j f(a_i - a_j) \quad (18)$$

$$\frac{\partial Q_i}{\partial a_j} = -2g_i g_j f(a_i - a_j) \quad (19)$$

$$\frac{\partial Q_i}{\partial g_i} = 2g_i + 2 \sum_{j \neq i} g_j F(a_i - a_j) \quad (20)$$

$$\frac{\partial Q_i}{\partial a_j} = 2g_i g_j F(a_i - a_j) \quad (21)$$

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 Π , (18) .. (21), suggests no natural linear dependence and, in general, the rank condition should be satisfied. Additionally, as we have illustrated in section 3, there will be some workers facing non-convex budget constraints which will further help with identification.

8 Data and Sample Likelihood

8.1 The Data

The sample 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 the UK, see Adam, Browne and Heady (2010). The data provide detailed diary and face to face interview information on consumption expenditures, usual hours worked, gross wage earnings, education

qualifications and household demographics. Tables 1 and 2 provide some basic descriptive statistics.

The overall sample contains some 11,458 women spread fairly evenly across the six years under study. A large group of women in this sample have minimal education qualifications, meaning that they left formal schooling at the minimum school leaving age of 16. The majority of the rest have completed secondary school with less than 20% having a college or university degree. The modal number of children 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). Almost 80% of the women in our sample are married or cohabiting (we label all these as ‘cohabiting’), leaving just over 20% 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 sample likelihood

In our sample we observe the employment status and the consumption expenditure of the women in the survey, their earnings and weekly hours when employed. We use the IFS TaxBen tax simulation model to recover income, the net of tax and benefits, for each household.

To construct the likelihood for this sample we derive the distribution of employment, hours, consumption and wages from the model, given the parameters and the distribution of the unobservables $(\varepsilon^b, \varepsilon^a, \varepsilon^c, \varepsilon^w)$.

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 :

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

8.2.1 The neoclassical model

When the individual is in work, 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 then 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 , the inequality above can be rewritten equivalently as

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

Let

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

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

Gathering up these terms, the likelihood function associated with a worker with h hours at work, conditional on hourly wage w and consumption c is given by

$$\int_{\bar{\varepsilon}^a(c, h, w)}^{\underline{\varepsilon}^a(c, h, w)} F^b(\varepsilon, c, h, w) \phi(\varepsilon) d\varepsilon.$$

For an individual who is not in employment, the computation is different because neither the wage nor potential hours of work are observed. For this individual 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 not working is

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

The likelihood function is the integral of this expression with respect to ε^a and w .

8.2.2 The two-offer model

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 therefore³

$$G(h|c, w) = g(h) \left\{ \sum_{h' < h} 2g(h') \int_{-\infty}^{\alpha(c, h, h', w)} F^b(\varepsilon, c, h, w) \phi(\varepsilon) d\varepsilon \right. \\ \left. + g(h) + \sum_{h' > h} 2g(h') \int_{\alpha(c, h, h', w)}^{+\infty} F^b(\varepsilon, c, h, w) \phi(\varepsilon) d\varepsilon \right\}.$$

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

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

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

³The probability of getting a couple of offers (h, h') , $h \neq h'$, is $2g(h)g(h')$, while that of getting (h, h) is $g(h)^2$.

9 Empirical results

9.1 Failure of the neoclassical model

Turning first to the results for the neoclassical model, we find that 7.5% of sampled women are observed working at hours that do not comply with the left hand side of neoclassical revealed preference inequality (10). For this group we can reject the neoclassical 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 estimated unknown parameters. The actual budget constraints for some of the individuals in this rejection group were used in the upper panel of Figure 2 above.

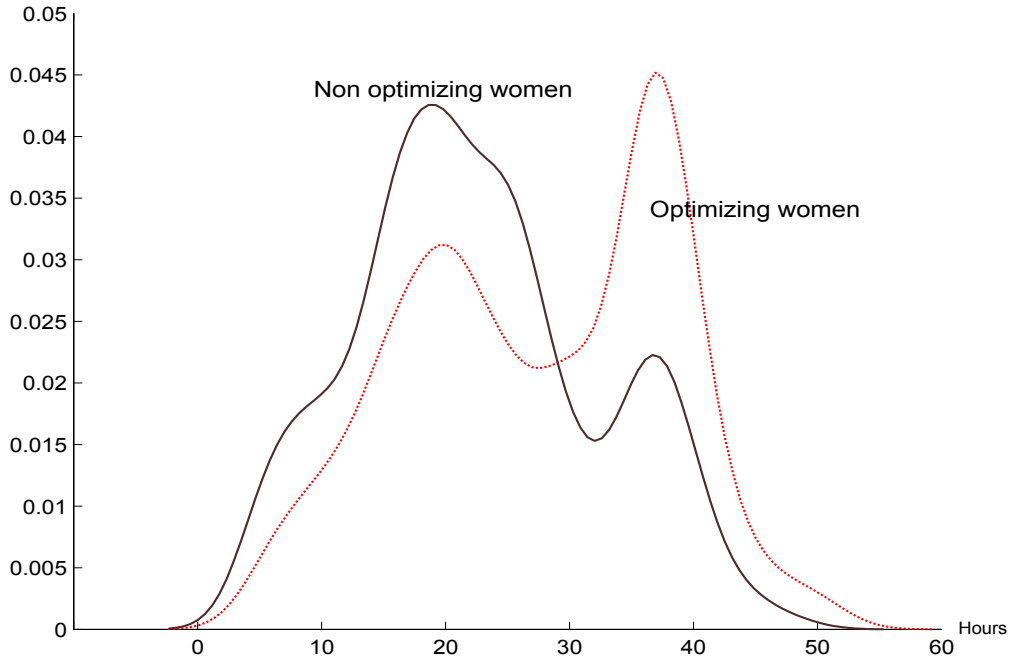
In practice, to incorporate individual observations that fail to satisfy the revealed preference inequality, we put a lower bound of 10^{-311} on the likelihood function of any observation, with a smooth approximation of the function between 10^{-310} and the lower bound. Using this bound, and the estimated model parameters from the neoclassical model, we find an additional 4.5% of failures. Overall we find a total of 12% of the sample that fail the revealed preference inequality.

In Table 3 we contrast the characteristics of these (very low likelihood) observations with the remaining sample. The sample of women whose observed hours of work fails the neoclassical model are more often lone mothers than married ones, their average wage is lower than the other women's average wage and, as Figure 3 shows, their distribution of hours worked is shifted to the left.

9.2 Parameter Estimates

Table 4 presents the estimation results for the parameters of preferences, fixed costs and the offer distribution, respectively. The first column of results are those for the neoclassical model and in the second column are the estimates for the two-offer specification. The ϕ and γ 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 the specification of $\ln(a)$ in equation (2). Following these are the parameters of fixed costs (3). The only remaining parameters in the neoclassical model are the variance and covariances parameters that link the unobserved heterogeneity terms in preferences, wages (4) and consumption (5).

Figure 3: Hours distributions for optimising and non optimising individuals



The unobserved heterogeneity terms are assumed to have a joint normal distribution as described in Section 3.

For the two-offer specification of the restricted choice model, described in (11) of section 4.1 above, offers are modelled as a mixture of two independent normals. The associated parameter estimates are presented in the final column of Table 4. The two-offer model is estimated on the whole sample as the framework is designed to incorporate potential failures of the revealed preference condition. It is noticeable that the preference parameters ϕ and γ are both smaller in this two-offer model, suggesting a more responsive underlying preference structure. These estimates of the parameters of the offer distribution suggest offers concentrated at full-time (around 36) hours and part-time (around 17) hours.

9.3 Frisch Elasticities with Linear Budget Constraints

To further understand the underlying ‘shape’ of preferences underlying these two estimated models, we compute the distribution of Frisch elasticities at the intensive margin

assuming no hours restrictions. That is we simply use the estimated utility parameters. In this exercise we also assume a linear budget constraint.

Frisch elasticities hold the marginal utility of consumption constant and, in our additive utility specification (1), the labour supply elasticity just depends on ϕ , L and a_t . The estimated elasticities are displayed in Table 5 and refer to the two columns of estimates in Table 4. The estimated labour supply elasticities are positive across the distribution and are moderately sized. They show a moderately higher underlying responsiveness for the preferences estimated in the two-offer model. This is as we might expect. Since, if we do not account for restrictions on the offer set, estimated preference parameters could reflect behaviour that appears less responsive.

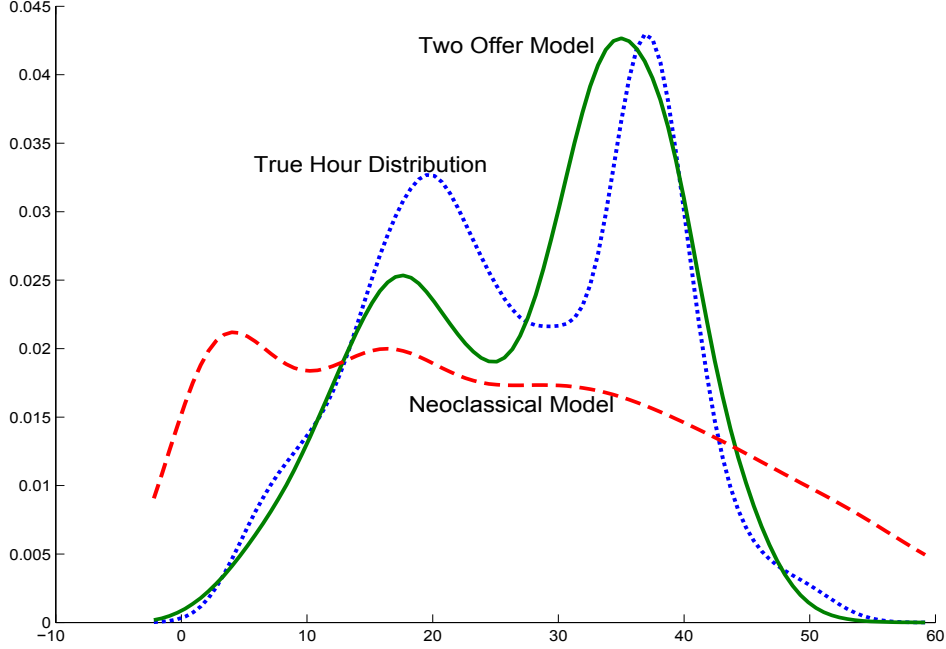
9.4 Model Fit

The contrast between the two model specifications is perhaps best displayed in Figure 4 which plots the simulated hours distributions against the actual hours distribution. The neoclassical model does quite poorly in replicating the twin peaks of the actual hours distribution, predicting hours choices that tend to be too low. The two-offer model provides a much better fit.

Examining the differences between the model specifications in Table 4 more closely we notice that the γ parameter changes dramatically across specifications. This is the marginal utility of consumption parameter in utility specification (1), determining the size of wealth and income effects. Notice that the Frisch elasticities in Table 5, which hold the marginal utility of consumption fixed, do not depend on the γ parameter. It does not enter the expression for the marginal utility of hours in this utility specification. The same is true for the fixed cost parameters. Fixed costs enter the work decision but not the hours of work decision once in work. Nonetheless these parameters will enter the simulation of the hours and employment in both models.

When using the parameters from the neoclassical model to simulate employment, we find 60.25 percent of the women in the sample would choose to be in employment. As would be expected, this closely accords with the actual percentage in work in the sample. However, if we use the preference and fixed cost parameters from the two-offer model to simulate employment choices, we find a larger employment rate of 77.84 percent. It appears that the restrictions in the two-offer model significantly reduce the number in employment relative to those who would, in the ‘long-run’, choose to work.

Figure 4: Hours distributions: observed, neoclassical and two offers model



9.5 Elasticities with Non-linear Budget Constraints

As an alternative to the elasticities presented in Table 5 which assume a linear budget constraint, we can compute elasticities that take explicit account for the nonlinearities in the budget constraint. For each individual these nonlinearities transform the impact of a gross wage change. Given the complex nature of the budget constraints facing the individuals in our sample this can have a sizable impact on the resulting estimated elasticities.

To derive these elasticities we follow three steps: First, we jointly simulate the endogenous variables (log of wages, log of consumption, preference for leisure and cost of working) given the estimated parameters and the exogenous variables. Second, using the IFS tax and benefit simulator (TaxBen), we assess the new budget constraints $(R(w^S, h), h = 0 \text{ or } h \in \mathcal{H})$ and $(R(w^S(1 + x\%), h), h = 0 \text{ or } h \in \mathcal{H})$ associated to the simulated wages for each individual. Finally, we optimise the value function over the whole

set of possible hours:

$$\begin{aligned} V(h, w_i^S, \beta_i^S, c_i^S, \varepsilon_i^S) &= a(\varepsilon_i^S)v(h) + c_i^{S,-\gamma} (R(w_i, h) - \beta_i^S) \quad \forall h > 0 \\ V(0, w_i^S, \beta_i^S, c_i^S, \varepsilon_i^S) &= a(\varepsilon_i^S)v(h) + c_i^{S,-\gamma} R(w_i, 0) \end{aligned}$$

Each individual chooses

$$h_i^o = \arg \max_{h=0, h \in \mathcal{H}} V(h, w_i^S, \beta_i, c_i^S, \varepsilon_i^S)$$

We repeat the optimisation with $(w_i^S(1 + x\%), c_i^S, \varepsilon_i^S, \beta_i^S)$, and each individual chooses

$$h_i^E = \arg \max_{h=0, h \in \mathcal{H}} V(h, w_i^S(1 + x\%), \beta_i^S, c_i^S, \varepsilon_i^S)$$

To compare elasticities at the intensive margin, Table 6 presents the elasticities for both sets of preference parameter estimates using the same sample of women employed under both model specifications. The intensive elasticity computation is as follows: among those who initially work ($h_i^o > 0$),

$$\varepsilon_{\text{Intensive}} = \frac{1}{\#\{h_i^o > 0\}} \sum_{i/h_i^o > 0} \frac{(h_i^E - h_i^o)}{h_i^o} \cdot \frac{1}{x\%}.$$

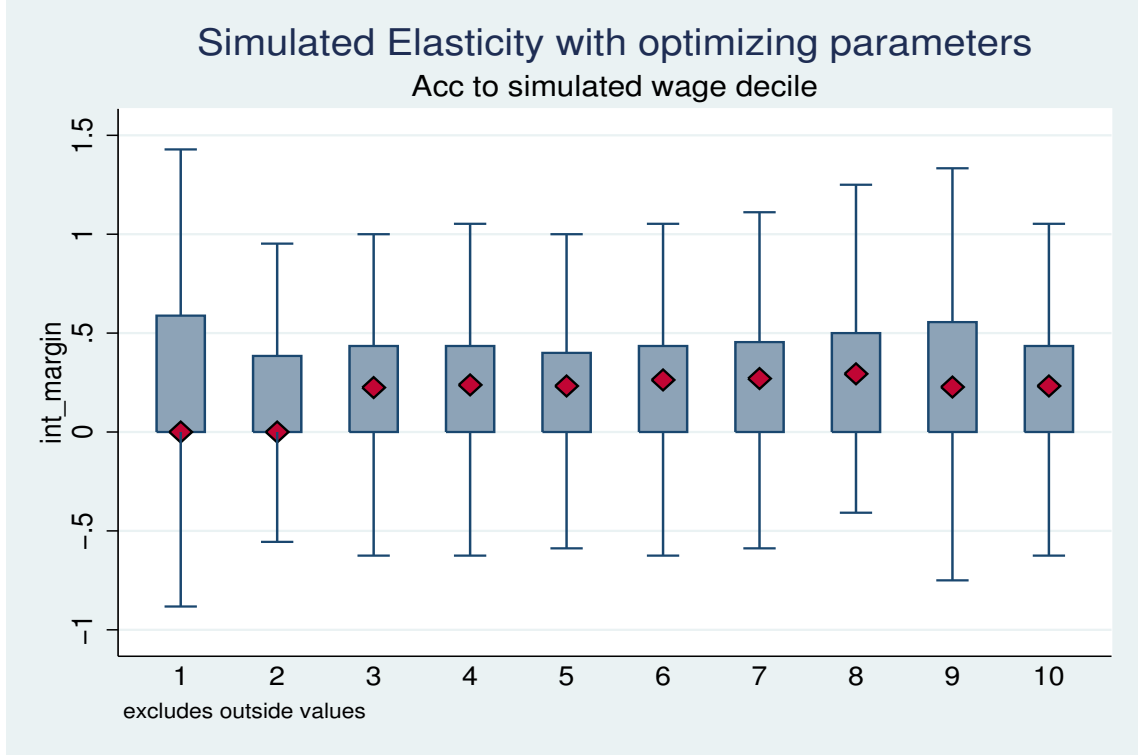
For the neoclassical model, the median value for the distribution of estimated intensive elasticities is .24 with a P75 to P25 range of 0.45 to 0.0.⁴ For the preference parameters under the restricted model assumptions the elasticities are very similar. These estimates lie in the range of estimates from various studies of female labour supply in the UK and in North America, see Blundell and MaCurdy (1999). Figures 5 and 6 present the distribution of these estimated elasticities according to the percentiles of the wage distribution.

9.6 Short-run elasticities in the two-offer model

For the two-offer model we can compute two types of elasticities. In Tables 5 and 6 the elasticities are those implied by the preference and fixed cost parameters alone, as if the restrictions on hours were lifted. We think of these as ‘long-run’ elasticities, directly comparable to those reported for the neoclassical model. Alternatively we can account

⁴They are estimated with a 1% wage increase, and for hours ranging from 1 to 90, every hour. The tails of the intensive elasticity distribution are sensitive to the step choice, but its median is quite robust.

Figure 5: Simulated Elasticities across the Wage Distribution: Neoclassical Model



for the estimated offer distribution. In this case choices are highly restricted. A single alternative is on offer. We think of these as ‘short-run’ elasticities.

To calculate the distribution of these elasticities, we follow the same overall procedure as above, but we additionally draw two offers from the offer distribution: $(h_i^{S,1}, h_i^{S,2}) \sim g$. We thus have two sets of comparison and actual hours will take the form

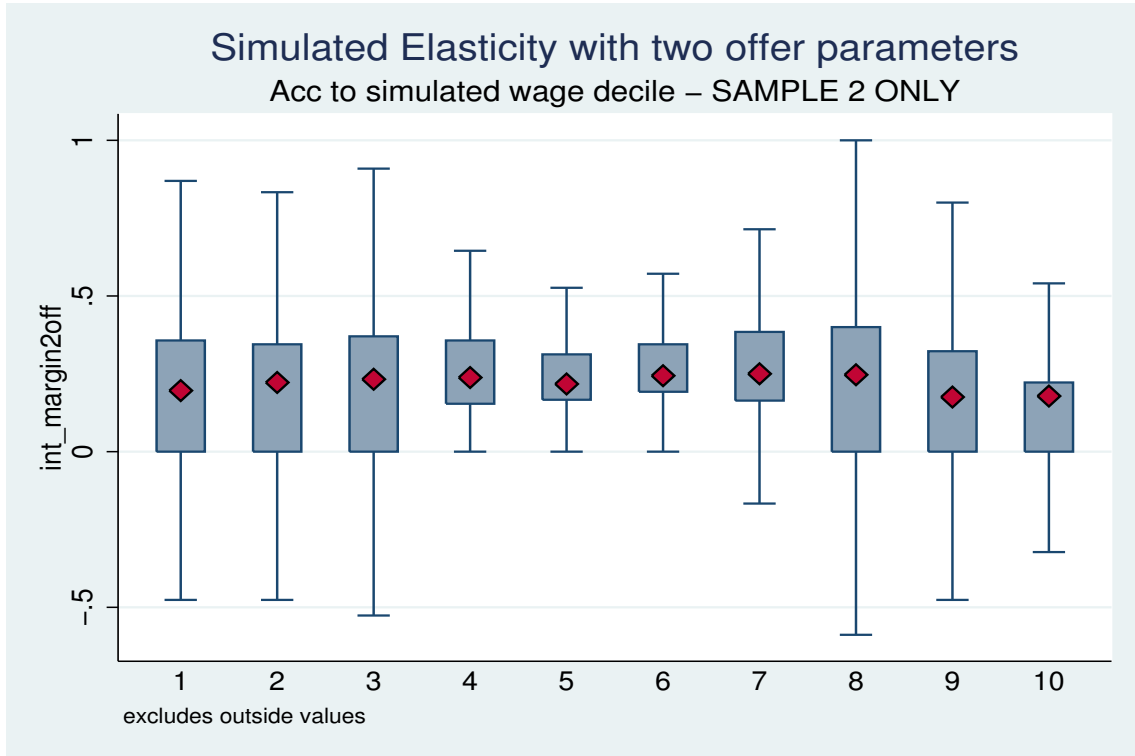
$$h_i^{obs} = \arg \max_{0, h_i^{S,1}, h_i^{S,2}} V(h^S, w_i^S, \beta_i^S, c_i^S, \varepsilon_i^S)$$

Similarly, with the same couple of simulated hours, the choice problem under the new budget set is given by

$$h_i^E = \arg \max_{0, h_i^{S,1}, h_i^{S,2}} V(h, w_i^S(1 + x\%), \beta_i^S, c_i^S, \varepsilon_i^S).$$

For the majority of individuals we find the change in wage is not sufficient to generate changes in hours given the (restricted) choices on offer. Those that do move tend to move

Figure 6: Simulated Elasticities across the Wage Distribution: Two-Offer Model



in large steps. The implied mean intensive elasticity is .40. This number is the average of no change for most of the population and large switches in hours for a few individuals. At the employment margin the restriction on offers is reflected in a short-run extensive elasticity of .21.

10 Conclusions

In this paper we have developed a model of employment and hours 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 revealed preference conditions for optimal choice.

We illustrated this framework in a model of hours and employment with nonlinear budget constraints where observed labor supply may not be reconciled with standard neo-classical optimisation 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.

The new framework was then 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. The results point to a small but important group of workers who fail the neoclassical choice model. For the remainder of the sample the neoclassical model suggests a distribution of estimated Frisch elasticities at the intensive margin with a median of .39 and interdecile range between 0.24 and 0.78. In contrast, the results from the two-offer model reveal an estimated offer distribution with twin peaks centered around full-time and part-time hours. The estimated preference parameters from the two-offer models imply larger elasticities with a median of .49 and an interdecile range between 0.27 and 0.96.

Accounting for restrictions on the choice set changes the underlying pattern of preference parameters. Individuals appear more responsive once restrictions are accounted for and the model simulations predict a higher level of employment were hours restrictions to be completely lifted. Actual responses though are quite different. Only a few individuals find it worth adjusting in the short-run to small changes in wages. But those that do adjust appear to make large changes.

The two-offer specification we adopt in the application in this paper is restrictive. In future work we intend to develop the n-offer case, allowing a much 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.

References

- [1] Aaberge, R., Colombino, U., and Strom, S. (1999), ‘Labour Supply in Italy: An Empirical Analysis of Joint Household Decisions, with Taxes and Quantity Constraints’, *Journal of Applied Econometrics*, 14, 403-22.
- [2] Aaberge, R., U. Colombino and T. Wennemo (2009), ”Evaluating Alternative Representations Of The Choice Sets In Models Of Labor Supply,” *Journal of Economic Surveys*, Wiley Blackwell, vol. 23(3), pages 586-612, 07.
- [3] Adam, Stuart., Browne, James. and Heady, Chris. (2010), ‘Taxation in the UK’, in Mirrlees, James., Adam, Stuart., Besley, Tim., Blundell, Richard., Bond, Steve.,

- Chote, Robert., Gammie, Malcolm., Johnson, Paul., Myles, Gareth. and Poterba, James. (eds), *Dimensions of Tax Design: The Mirrlees Review*, Oxford University Press for Institute for Fiscal Studies, (Oxford).
- [4] Altonji, J. G. (1986), 'Intertemporal Substitution in Labour Supply: Evidence from Micro Data', *Journal of Political Economy*, 94, S176-S215.
- [5] Altonji, J. G. and Paxson, C. (1992), 'Labor Supply, Hours Constraints, and Job Mobility', *Journal of Human Resources*, 27, 256-278.
- [6] Arrufat, J. L., and Zabalza, A. (1986), 'Female Labour Supply with Taxation, Random Preferences and Optimization Errors', *Econometrica*, 54, 47-63.
- [7] Bloemen, H. (2000) 'A Model of Labor Supply with Job Offer Restrictions', *Labour Economics*, 7,297–312.
- [8] Bloemen, H. (2008) 'Job Search, Hours Restrictions, and Desired Hours of Work', *Journal of Labor Economics*, 26, 137-179.
- [9] Blomquist, S and Newey, W. (2002), 'Nonparametric Estimation with Nonlinear Budget Constraints', *Econometrica*, 70, 2455-80.
- [10] Blundell, R., Bozio, A. and Laroque, G. (2011), 'Labour Supply and the Extensive Margin', *American Economic Review* 101(3), 482-486.
- [11] Blundell, R., Bozio, A. and Laroque, G. (2013), 'Extensive and Intensive Margins of Labour Supply: Working Hours in the US, UK and France' *Fiscal Studies*, 34 (1), 1-29, March.
- [12] Blundell, R., Brewer, M. and Francesconi, M. (2008), 'Job changes, hours changes and the path of labour supply adjustment', *Journal of Labor Economics* 26, 3, 421-453.
- [13] Blundell, R., Duncan, A., and Meghir, C. (1992), 'Taxation in Empirical Labour Supply Models: Lone Mothers in the UK', *Economic Journal*, 102, 265-78.
- [14] Blundell, R., Duncan, A., and Meghir, C. (1998), 'Estimating Labor Supply Responses Using Tax Reforms', *Econometrica*, 66, 827-61.
- [15] Blundell, R., Ham, J., and Meghir, C. (1987), 'Unemployment and Female Labour Supply', *Economic Journal*, 97, 44-64.

- [16] Blundell, R. and MaCurdy, T. (1999), ‘Labor Supply: A Review of Alternative Approaches’, in *Handbook of Labor Economics*, Ashenfelter and Card (eds), 3A.
- [17] Blundell, R. and Walker, I. (1986), ‘A Life Cycle Consistent Empirical Model of Family Labour Supply Using Cross Section Data’, *Review of Economic Studies*, 53, 539-58.
- [18] Bourguignon, F., and Magnac, T. (1990), ‘Labor Supply and Taxation in France’, *Journal of Human Resources*, 25, 358-89.
- [19] Chetty, Raj (2012) ‘Bounds on elasticities with optimization frictions: a synthesis of micro and macro evidence on labor supply’, *Econometrica* 80.3, pp. 968–1018.
- [20] Chetty, Raj, John Friedman, Tore Olsen, and Luigi Pistaferri (2011a) ‘Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records’, *Quarterly Journal of Economics*, 126.2, pp. 749-804.
- [21] Cogan, J. (1981), ‘Fixed Costs and Labor Supply’, *Econometrica*, 49, 945-63.
- [22] Dagsvik, J.K. and Z. Jia (2006) ‘Labor Supply as a Choice among Latent Job Opportunities: a Practical Empirical Approach’, Discussion Papers 481, Research Department of Statistics Norway.
- [23] Dagsvik, J.K. and S. Strom (2006) ‘Sectoral Labor Supply, Choice Restrictions and Functional Form’ *Journal of Applied Econometrics*, 21, 803-826.
- [24] Dickens, W. and S. Lundberg (1993) ‘Hours Restrictions and Labor Supply’, *International Economic Review*, 34, 169-191
- [25] Kfir, E. and S. Spiegel (2011) ‘Consideration Sets and Competitive Marketing’, *Review of Economic Studies*, 78, 235-262.
- [26] Ham, J. and K.T. Reilly (2002) ‘Testing Intertemporal Substitution, Implicit Contracts, and Hours Restriction Models of the Labor Market Using Micro Data’, *American Economic Review*, 92, 905-927.
- [27] Hausman, J. (1985), ‘The Econometrics of Nonlinear Budget Sets’, *Econometrica*, 53, 1255-82.
- [28] Heckman, J. (1974a), ‘Shadow Prices, Market Wages, and Labor Supply’, *Econometrica* 42(4), 679-694.

- [29] Heckman, J. (1974b), ‘Effects of Childcare Programs on Women’s Work Effort’, *Journal of Political Economy*, 82, S136-S163.
- [30] Heckman, J. (1979), ‘Sample Selection Bias as a Specification Error’, *Econometrica* 47(1), 153-161.
- [31] Heckman, J. (1993), ‘What Has Been Learned About Labor Supply in the Past Twenty Years?’, *American Economic Review* 83(2), 116-121.
- [32] Heckman, J. and MaCurdy, T. (1980), ‘A Lifecycle Model of Female Labor Supply’, *Review of Economic Studies*, 47, 47-74.
- [33] Ilmakunnas, S. and S. Pudney (1990) ‘A Model of Female Labour Supply in the Presence of Hours Restrictions’, *Journal of Public Economics*, 41, 183-210.
- [34] Meghir, Costas and David Phillips (2010). “Labour supply and taxes”. In: Dimensions of Tax Design: the *Mirrlees Review*. Ed. by James Mirrlees, S Adam, Timothy Besley, Richard Blundell, S Bond, R Chote, M Gammie, P Johnson, G Myles, and James Poterba. Oxford University Press.
- [35] Moffitt, R. (1984), ‘The Estimation of a Joint Wage-Hours Labor Supply Model’, *Journal of Labor Economics*, 2, 550-66.
- [36] Saez, Emmanuel (2010). “Do Taxpayers Bunch at Kink Points?” *American Economic Journal: Economic Policy*, August, 180-212.
- [37] Stewart, M.B. and J.K. Swaffield (1997) ‘Constraints on the Desired Hours of Work of British Men’ *Economic Journal*, 107, 520-535.
- [38] Tummers, M. P. and I. Woittiez (1990) ‘A Simultaneous Wage and Labor Supply Model with Hours Restrictions’ *Journal of Human Resources*, 26, 393-423.
- [39] van Soest, A., I. Woittiez and A. Kapteyn (1990) ‘Labour Supply, Income Taxes and Hours Restrictions in the Netherlands’, *Journal of Human Resources*, 25, 517-558.

Table 1: Some Descriptive Statistics

	Sample	1997	1998	1999	2000	2001	2002
Size	11,458	1,794	1,858	1,935	1,850	2,116	1,905
Education Level							
1	56.41	58.86	57.91	55.71	55.46	56.81	53.86
2	25.82	24.92	25.30	27.29	26.49	24.86	26.09
3	17.77	16.22	16.79	17.00	18.05	18.34	20.05
Number of children							
1	38.53	40.41	36.44	40.05	38.54	37.33	38.58
2	41.55	41.64	42.84	39.12	39.78	42.58	43.25
3	14.51	13.32	14.69	15.35	16.38	13.89	13.49
4	3.98	3.40	4.63	3.77	4.16	4.40	3.46
5+	1.42	1.23	1.40	1.71	1.14	1.80	1.21
Youngest kid between 0 and 4	29.25	30.99	29.98	28.63	26.81	29.68	29.40
Youngest kid between 5 and 10	30.30	28.76	27.93	28.53	28.16	33.51	34.38
London	9.68	10.54	8.93	10.13	8.97	10.07	9.40
Cohabitant	76.94	77.81	78.20	75.71	78.22	76.04	75.91
Spouse Inwork Status							
If inwork	65.92	66.27	67.24	66.75	66.84	65.56	63.22
If out of work	59.20	57.12	63.89	60.51	59.02	57.86	56.24

Table 2: Consumption, Wages and Hours of Work

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

Table 3: Non optimizing and optimizing agents

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

Table 4: Estimation results

	Optimizing	Two Offer
ϕ	6.27 (0.00)	5.49 (0.10)
γ	1.30 (0.06)	0.20 (0.02)
a : Constant	23.84 (0.31)	23.89 (0.46)
a : Cohabitant	-1.90 (0.07)	-0.47 (0.05)
a : Youngest kid between 0 and 4	1.20 (0.04)	1.23 (0.03)
a : Youngest kid between 5 and 9	0.07 (0.02)	0.25 (0.02)
b : Constant	0.97 (1.34)	0.73 (0.43)
b : Cohabitant	0.86 (1.23)	0.68 (0.25)
b : Number of kids	0.83 (0.37)	0.44 (0.20)
σ^b	4.42 (0.24)	
$\rho(\varepsilon^a, \varepsilon^c)$	-0.34 (0.02)	
$\rho(\varepsilon^a, \varepsilon^w)$	0.06 (0.01)	
σ^a	2.05 (0.02)	
m_1		16.71 (0.12)
σ_1		6.35 (0.11)
m_2		36.47 (0.15)
σ_2		5.27 (0.06)
p_1		0.47 (0.01)

Table 5: Frisch (Intensive) Elasticity Estimates

	Neoclassical Model	Two-Offer Model
<i>Mean</i>	0.57 (0.47)	0.68 (0.56)
p10	0.24	0.27
p20	0.27	0.31
p30	0.28	0.34
p40	0.32	0.41
p50	0.39	0.49
p60	0.51	0.61
p70	0.64	0.73
p80	0.78	0.96
p90	1.07	1.33

Notes: see the text for a discussion of the computation of these elasticities.

Table 6: Intensive Elasticity Estimates Accounting for Nonlinear Taxes

Intensive Elasticity	Neoclassical Model	Two-Offer Model
mean	1.37	1.34
p5	-0.77	-.36
p10	0.00	0.00
p25	0.00	0.00
p50	0.24	0.21
p75	0.45	0.33
p90	1.11	0.65
p95	2.00	1.07

Notes: see the text for a discussion of the computation of these elasticities.