Employment, Hours of Work and the Optimal Taxation of Low Income Families

Richard Blundell — UCL & IFS

Andrew Shephard — Princeton University

First version August 2008;
This version November 2010

Abstract

The design of an optimal tax schedule is examined using a structural labour supply model. The model incorporates unobserved heterogeneity, fixed costs of work, childcare costs and the detailed non-convexities of the tax and transfer system. The analysis concerns optimal design under social welfare functions with different degrees of inequality aversion. It also considers purely Pareto improving reforms. We explore the gains from tagging according to child age and also examine the case for the use of hours-contingent payments. Using the UK tax treatment of lone parents as our policy environment, the results point to a reformed nonlinear tax schedule with tax credits only optimal for low earners. The results also suggest a welfare improving role for tagging according to child age and also for hours-contingent payments, although the case for the latter is mitigated when hours cannot be monitored or recorded accurately by the tax authorities.

Acknowledgements: This study is part of the research program of the ESRC Centre for the Microeconomic Analysis of Public Policy at the IFS. We thank Stuart Adam, Mike Brewer, Guy Laroque, Ian Preston, Emmanuel Saez, Florian Scheuer, seminar participants at MIT, Jerusalem, NIESR, SOLE and participants in the Mirrlees Review, the editor and the referees for helpful comments. We are alone responsible for all errors and interpretations.

1 Introduction

This paper examines the optimal design of earnings taxation using a structural labour supply model. The analysis concerns the optimal choice of the tax rate schedule in a Mirrlees (1971) framework extended to allow for unobserved heterogeneity, fixed costs of work, childcare costs and the detailed non-convexities of the tax and transfer system. We consider the implications for the optimal tax schedules of allowing for different degrees of inequality aversion. We also considers purely Pareto improving reforms.

The contribution of this paper is fourfold. First, we take the structural model of employment and hours of work seriously in designing the schedule of taxes and benefits. Second, we assess the role of tagging taxes by the age of children. Third, we consider the case where hours of work are partially observable to the tax authorities. Fourth, we identify and quantify inefficiencies in the actual tax and transfer system.

In the empirical literature on labour supply certain common and robust features of estimated labour supply responses of the low paid have emerged. Specifically, the importance of distinguishing between the intensive margin of hours of work and the extensive margin where the work decision is made. Labour supply elasticities for certain groups of working age individuals appear to be much larger at the extensive margin, see Blundell and Macurdy (1999), for example. As Saez (2002) and Laroque (2005) have shown, empirical results on the responsiveness of different types of individuals at difference margins of labour supply have strong implications for the design of earnings taxation.

The UK tax treatment of lone parents is used as the empirical environment for our policy reform analysis.[1] As in North America this group has been the subject of a number of tax and benefit reforms, see Blundell and Hoynes (2004), for example. These reforms can provide useful variation for assessing the reliability

[1]Taxation design in the UK was also recently explored by Brewer et al. (2010) as part of the Mirrlees Review.

of structural models. In particular, we use the 1999 Working Families' Tax Reform (WFTC) in the UK which considerably increased the generosity of in-work benefits/tax credits for lone parents, see Adam and Browne (2009).

The microeconometric analysis here is based on a stochastic discrete choice labour supply model (Hoynes, 1996; Keane and Moffitt, 1998; Blundell et al., 2000; van Soest et al., 2002). We find strong differences in the responsiveness of labour supply at the extensive and the intensive margin. We also find that these responses vary according to the age of children. We use this variation to explore the welfare gains from tagging according to child age. Our results suggest a welfare improving role for such tagging and also suggest pure tax credits at low earnings may be optimal, but only for mothers with school aged children.

The WFTC system uses hours-contingent payments.[2] Eligibility requires parents with children to be working in a job that involves at least 16 hours of work per week. There is a further supplement if the parent works 30 hours or more. We explore the optimality of such eligibility rules. Given the likely difficulties in recording and monitoring hours of work, we also consider the optimal tax schedule when declared weekly hours can be, in part, manipulated by the individual and also when the hours can only be recorded with measurement error. Our results point to welfare gains from hours-contingent payments, especially at full-time work. However, the case is substantially mitigated when hours cannot be monitored or recorded accurately by the tax authorities.

The rest of the paper proceeds as follows. In the next section we develop the analytical framework for optimal design within a stochastic structural labour supply model. In section 3 we outline the WFTC reform in the UK and its impact on work incentives. Section 4 outlines the structural microeconometric model, while in section 5 we describe the data and model estimates. Section 6 uses

[2]Hours conditions are also used in the tax credit systems in Ireland and New Zealand. They are also used in the design of work conditioned earnings supplements, for example in the Canadian Self-Sufficiency Project (Card and Robins, 1998) and in the TANF programme of welfare payment in the US (Moffitt, 2003). It has also been proposed as a mechanism for improving tax design, see Keane (1995), although not within an optimal tax framework.

these model estimates to derive optimal tax schedules. We provide evidence for lowering the marginal rates at lower incomes and also document the importance of allowing the tax schedule to depend on the age of children. We also discuss how introducing hours rules affects tax design, and how important these are likely to be in terms of social welfare. In section 7 we examine reforms that are strictly Pareto improving. We quantify the inefficiency under the existing system, and the extent to which the requirement that no individual is made worse off would act as a constraint on the welfare maximisation problem. Finally, section 8 concludes.

2 The Optimal Design Problem

The policy analysis here concerns the choice of a tax schedule in which the government is attempting to allocate a fixed amount of revenue R to a specific demographic group – single mothers – in a way which will maximise the social welfare for this group. Such a schedule balances redistributive objectives with efficiency considerations. Redistributive preferences are represented through the social welfare function defined as the sum of transformed individual utilities, where the choice of transformation reflects the desire for equality.[3]

In this section we develop an analytical framework for the design of tax and transfer policy that allows for two scenarios. In the first only earnings are observable by the tax authority, in the second we allow for partial observability of hours of work. Rather than assuming that individuals are unconstrained in their choice of hours, we suppose that only a finite number of hours choices are available, with hours of work h chosen from the finite set $\mathcal{H} = \{h_0, \ldots h_J\}$.[4]

The formulation of the optimal tax design problem will depend upon what

[3]The framework developed here contrasts with our later exploration in section 7. In that analysis we do not adopt a social welfare function, but rather seek to identify Pareto improving reforms to the actual UK tax and transfer system.

[4]An alternative model which incorporates constraints on the labour supply choices in an optimal design problem is developed in Aaberge and Colombino (2008).

information is observable to the tax authorities. We always assume that the government can observe earnings wh and worker characteristics X, and we shall also allow for the possibility of observing some hours of work information. In much of our analysis we will assume that rather than necessarily observing the actual hours h that are chosen, the tax authorities is assumed to only be able to observe that they belong to some closed interval $\mathbf{h} = [\underline{h}, \overline{h}] \in \mathcal{H}$ with $\underline{h} \leq h \leq \overline{h}$. For example, the tax authorities may be able to observe whether individuals are working at least h_B hours per week, but conditional on this, not how many. Depending on the size of the interval, this framework nests two important special cases; (i) when hours are perfectly observable $h = \underline{h} = \overline{h}$ for all $h \in \mathcal{H}$; (ii) only earnings information is observed $\mathbf{h} = \mathcal{H}_{++}$ for all $h > 0$. In general this is viewed as a problem of partial observability since actual hours h are always contained in the interval $\mathbf{h}$. In our later analysis in section 6.4 we will explore the effect that both random hours measurement error, and possible direct hours misreporting have upon the optimal design problem.

Work decisions by individuals are determined by their preferences over consumption c and labour hours h, as well as possible childcare requirements, fixed costs of work, and the tax and transfer system. Preferences are indexed by observable characteristics X, including the number and age of her children, and vectors of unobservable (to the econometrician) characteristics ϵ and ε. The vector ε corresponds to the additive hours (or state) specific errors in the utility function, and we let $U(c, h; X, \epsilon, \varepsilon) = u(c, h; X, \epsilon) + \varepsilon_h$ represent the utility of a single mother who consumes c and works h hours. We will assume that she consumes her net income which comprises the product of hours of work h and the gross hourly wage w plus non-labour income and transfer payments, less taxes paid, childcare expenditure, and fixed costs of work. In what follows we let F denote the distribution of state specific errors ε, and G denote the joint distribution of (X, ϵ).[5]

In our empirical analysis individual utilities $U(c, h; X, \epsilon, \varepsilon)$ will be described

[5]Throughout our analysis we assume that ε is independent of both ϵ and X.

by a parametric utility function and a parametric distribution of unobserved heterogeneity (ϵ, ε). Similarly, a parametric form will be assumed for the stochastic process determining fixed costs of work and childcare expenditure. To maintain focus on the optimal design problem, we delay this discussion regarding the econometric modelling until section 4; for now it suffices to write consumption c at hours h as $c(h; T, X, \epsilon)$,[6] where $T(wh, \mathbf{h}; X)$ represents the tax and transfer system. Non-labour income, such as child maintenance payments, enter the tax and transfer schedule T through the set of demographics X, and for notational simplicity we abstract from the potential dependence of the tax and transfer system on childcare expenditure. Taking the schedule T as given, each single mother is assumed to choose her hours of work $h^* \in \mathcal{H}$ to maximise her utility. That is:

$$h^* = \underset{h \in \mathcal{H}}{\arg\max}\, U(c(h; T, X, \epsilon), h; X, \epsilon, \varepsilon). \tag{1}$$

We assume that the government chooses the tax schedule T to maximise a social welfare function W that is represented by the sum of transformed utilities:

$$W(T) = \int_{X,\epsilon} \int_{\varepsilon} \Upsilon(U(c(h^*; T, X, \epsilon), h^*; X, \epsilon, \varepsilon)) dF(\varepsilon) dG(X, \epsilon) \tag{2}$$

where for a given cardinal representation of U, the utility transformation function Υ determines the governments relative preference for the equality of utilities.[7] This maximization is subject to the incentive compatibility constraint which states that lone mothers choose their hours of work optimally given T as in (1) and the government resource constraint:

$$\int_{X,\epsilon} \int_{\varepsilon} T(wh^*, \mathbf{h}^*; X) dF(\varepsilon) dG(X, \epsilon) \geq \overline{T} (\equiv -R). \tag{3}$$

In our empirical application we will restrict T to belong to a particular parametric

[6] Conditional on work hours h, consumption will not depend on ε given our assumption that ε enters the utility function additively and is independent of (X, ϵ).

[7] Given the presence of preference heterogeneity, a more general formulation would allow the utility transformation function Υ to vary with individual characteristics.

class of tax functions. This is discussed in section 6 when we empirically examine the optimal design of the UK tax and transfer schedule.

3 Tax Credit Reform

The increasing reliance on tax-credit policies during the 1980s and 1990s, especially in the UK and the US, reflected the secular decline in the relative wages of low skilled workers with low labour market attachment together with the growth in single-parent households (see Blundell, 2002, and references therein). The specific policy context for this paper is the Working Families' Tax Credit (WFTC) reform which took place in the UK at the end of 1999. A novel feature of the British tax credit system is that it makes use of minimum hours conditions in addition to an employment condition. Specifically, WFTC eligibility required a working parent to record at least 16 hours of work per week. Moreover there was a further hours contingent bonus for working 30 hours or more.

As in the US, the UK has a long history of in-work benefits, starting with the introduction of Family Income Supplement (FIS) in 1971. In 1988 FIS became Family Credit (FC), and in October 1999, Working Families' Tax Credit was introduced. While these programmes have maintained a similar structure, the reforms have been associated with notable increases in their generosity. As described above, an important feature of British programmes of in-work support since their inception – and in contrast with programmes such as the US Earned Income Tax Credit – is that awards depend not only on earned and unearned income and family characteristics, but also on a minimum weekly hours of work requirement. While under FIS this minimum requirement was always 24 hours per week, the April 1992 reform that occurred during the life of FC reduced this to 16 hours per week, where it has stayed since.[8] The impact of this reform to FC on single parents' labour supply is ambiguous: those working more than 16

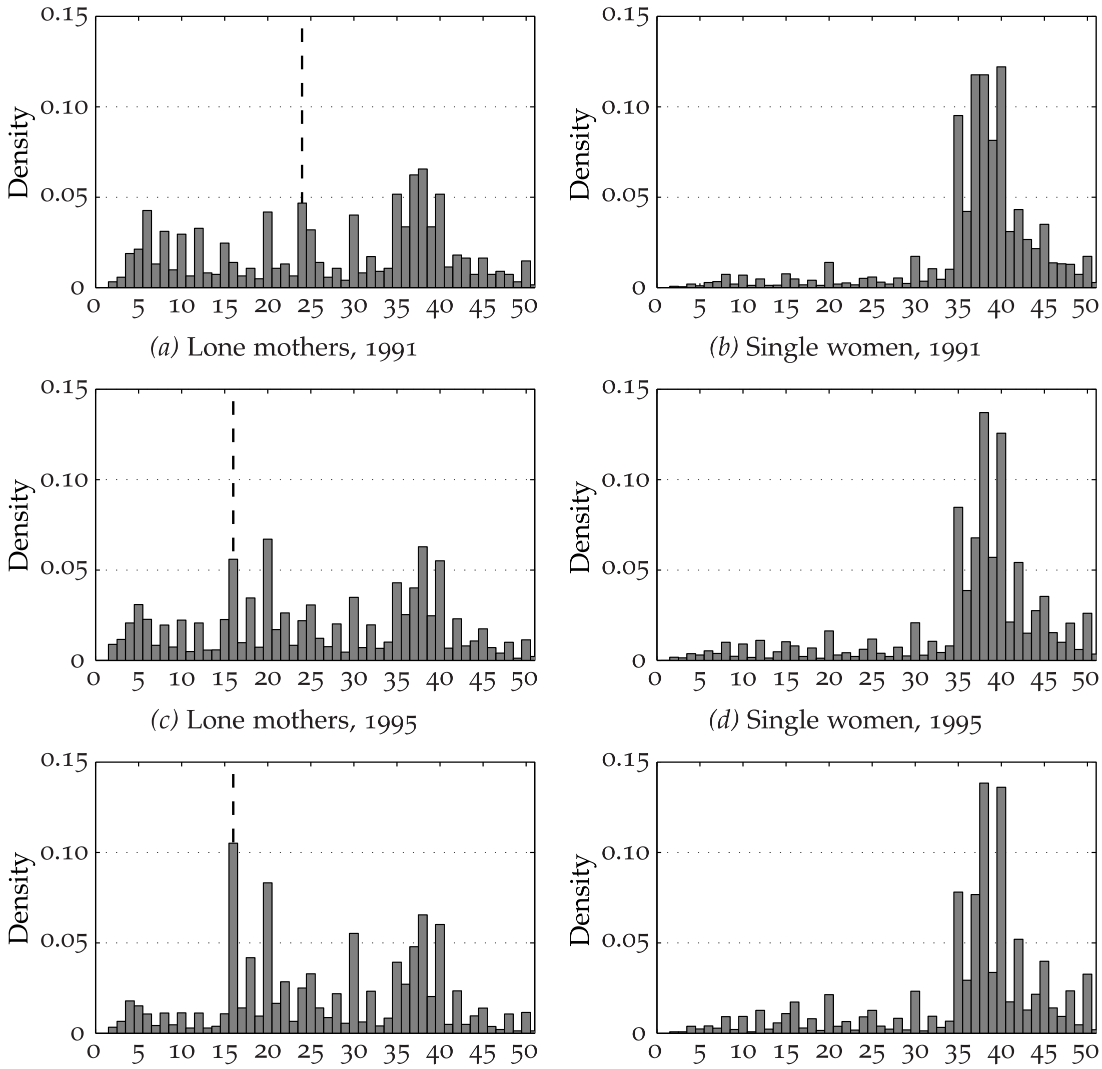
[8] In 1995, there was another reform to Family Credit in the form of an additional (smaller) credit for those adults working full time (defined as 30 or more hours a week).

hours a week had an incentive to reduce their weekly hours to (no less than) 16, while those previously working fewer than 16 hours had an incentive to increase their labour supply to (at least) the new cut-off. Figure 1 shows that the pattern of observed hours of work over this period strongly reflects these incentives. Single women without children were ineligible.

The tax design problem we discuss here relates directly to the features of the WFTC. Indeed we assess the reliability of our labour supply model in terms of its ability to explain behaviour before and after the reform. There were essentially five ways in which WFTC increased the level of in-work support relative to the previous FC system: (i) it offered higher credits, especially for families with younger children; (ii) the increase in the threshold meant that families could earn more before it was phased out; (iii) the tax credit withdrawal rate was reduced from 70% to 55%; (iv) it provided more support for formal childcare costs through a new childcare credit; (v) all child maintenance payments were disregarded from income when calculating tax credit entitlement. The main parameters of FC and WFTC are presented in Table 1.

The WFTC reform increased the attractiveness of working 16 or more hours a week compared to working fewer hours, and the largest potential beneficiaries of WFTC were those families who were just at the end of the FC benefit withdrawal taper. Conditional on working 16 or more hours, the theoretical impact of WFTC is as follows: (i) people receiving the maximum FC award will face an income effect away from work, but not below 16 hours a week; (ii) people working more than 16 hours and not on maximum FC will face an income effect away from work (but not below 16 hours a week), and a substitution effect towards work; (iii) people working more than 16 hours and earning too much to be entitled to FC but not WFTC will face income and substitution effects away from work if they claim WFTC (see Blundell and Hoynes, 2004).

When analyzing the effect of the WFTC programme it is necessary to take an



(a) Lone mothers, 1991

(b) Single women, 1991

(c) Lone mothers, 1995

(d) Single women, 1995

(e) Lone mothers, 2002

(f) Single women, 2002

Figure 1: Female hours of work by survey year. Figure shows the distribution of usual hours of work for women by year and presence of children. Sample is restricted to women aged 18–45. Calculated using UK Labour Force Survey data (for 1991) and UK Quarterly Labour Force Survey data (1995 and 2002). Horizontal axes measure weekly hours of work; the vertical line indicates the minimum hours eligibility.

Table 1: Parameters of FC/WFTC

	April 1999 (FC)	October 1999 (WFTC)	June 2000 (WFTC)	June 2002 (WFTC)
Basic Credit	49.80	52.30	53.15	62.50
Child Credit				
under 11	15.15	19.85	25.60	26.45
11 to 16	20.90	20.90	25.60	26.45
over 16	25.95	25.95	26.35	27.20
30 hour credit	11.05	11.05	11.25	11.65
Threshold	80.65	90.00	91.45	94.50
Taper rate	70% after income tax and National Insurance	55% after income tax and National Insurance	55% after income tax and National Insurance	55% after income tax and National Insurance
Childcare	Expenses up to £60 (£100) for 1 (more than 1) child under 12 disregarded when calculating income	70% of total expenses up to £100 (£150) for 1 (more than 1) child under 15	70% of total expenses up to £100 (£150) for 1 (more than 1) child under 15	70% of total expenses up to £135 (£200) for 1 (more than 1) child under 15

Notes: All monetary amounts are in pounds per week and expressed in nominal terms. Minimum FC/WFTC award is 50p per week in all years above.

integrated view of the tax system. This is because tax credit awards in the UK are counted as income when calculating entitlements to other benefits, such as Housing Benefit and Council Tax Benefit. Families in receipt of such benefits would gain less from the WFTC reform than otherwise equivalent families not receiving these benefits; Figure 2 illustrates how the various policies impact on the budget constraint for a low wage lone parent. Moreover, there were other important changes to the tax system affecting families with children that coincided with the expansion of tax credits, and which make the potential labour supply responses considerably more complex. In particular, there were increases in the generosity of Child Benefit (a cash benefit available to all families with children regardless of income), as well as notable increases in the child additions in Income Support (a welfare benefit for low income families working less than 16 hours a week).[9]

4 A Structural Labour Supply Model

The labour supply specification develops from earlier studies of structural labour supply that use discrete choice techniques and incorporate non-participation in transfer programmes, specifically Hoynes (1996) and Keane and Moffitt (1998). Our aim is to construct a credible model of labour supply behaviour that adequately allows for individual heterogeneity in preferences and can well describe observed labour market outcomes. As initially discussed in section 2, lone mothers have preferences defined over consumption c and hours of work h. Hours of work h are chosen from some finite set $\mathcal{H}$, which in our main empirical results will correspond to the discrete weekly hours points $\mathcal{H} = \{0, 10, 19, 26, 33, 40\}$.[10] In

[9] For many families with children, these increases in out-of-work income meant that, despite the increased generosity of in-work tax credits, replacement rates remained relatively stable. There were also changes to the tax system that affected families both with and without dependent children during the lifetime of WFTC: a new 10% starting rate of income tax was introduced; the basic rate of income tax was reduced from 23% to 22%; there was a real rise in the point at which National Insurance (payroll tax) becomes payable.

[10] These hours points correspond to the empirical hours ranges 0, 1–15, 16–22, 23–29, 30–36 and 37+ respectively.

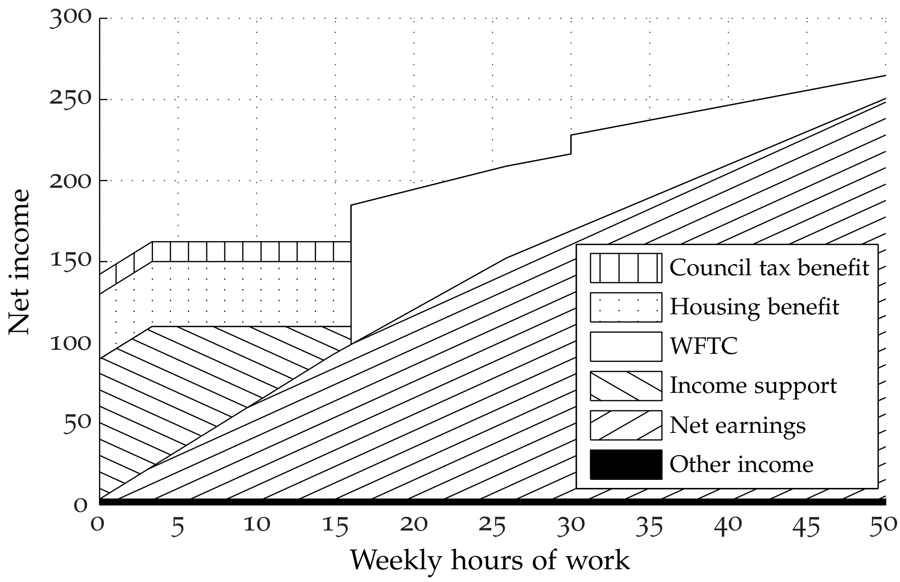


Figure 2: Tax and transfer system interactions. Figure shows interaction of tax and transfer system under April 2002 system for a lone parent with a single child aged 5, average band C council tax, £40 per week housing costs, £6 gross hourly wage rate, and no childcare costs. All incomes expressed in April 2002 prices. Calculated using FORTAX.

section 6 we also present results which allow for a finer discretization of weekly hours.

We augment the framework presented in section 2 to allow the take-up of tax-credits to have a direct impact on preferences through the presence of some stigma or hassle cost (discussed further below), and we use P (equal to one if tax credits are received, zero otherwise) to denote the endogenous programme participation decision.[11] These preferences may vary with observable demographic characteristics X (such as age, region, the number and age of children), and vectors of unobservable (to the econometrician) characteristics ϵ and ε. As described in section 2, ε is used specifically to denote the additive state specific errors which are attached to each discrete hours point. We shall assume that these follow a standard Type-I extreme value distribution.

All the estimation and simulation results presented here assume preferences

[11] All other transfer programmes are assumed to have complete take-up. This could be generalised in future work.

of the form:

$$u(c,h,P;X,\epsilon) = \alpha_y(X,\epsilon)\frac{c^{\theta_y}-1}{\theta_y} + \alpha_l(X)\frac{(1-h/H)^{\theta_l}-1}{\theta_l} - P\eta(X,\epsilon) \tag{4}$$

where $H = 168$ denotes the total weekly time endowment, and where the set of functions $\alpha_y(X,\epsilon)$, $\alpha_l(X)$ and $\eta(X,\epsilon)$ capture observed and unobserved preference heterogeneity.[12] The function $\eta(X,\epsilon)$ is included to reflect the possible disutility associated with claiming in-work tax credits ($P = 1$), and its presence allows us to rationalise less then complete take-up of tax credit programmes. In each case we allow observed and unobserved heterogeneity to influence the preference shifter functions through appropriate index restrictions. We assume that $\log\alpha_y(X,\epsilon) = X_y'\beta_y + \epsilon_y$ and $\log\alpha_l(X) = X_l'\beta_l$; programme participation costs are assumed to be linear in parameters, $\eta(X,\epsilon) = X_\eta'\beta_\eta + \epsilon_\eta$.

The choice of hours of work h affects consumption c through two main channels: firstly, through its direct effect on labour market earnings and its interactions with the tax and transfer system; secondly, working mothers may be required to purchase childcare for their children which varies with maternal hours of employment. Given the rather limited information that our data contains on the types of childcare use, we take a similarly limited approach to modelling, whereby hours of childcare use h_c is essentially viewed as a constraint: working mothers are required to purchase a minimum level of childcare $h_c \geq \alpha_c(h,X,\epsilon)$ which varies stochastically with hours of work and demographic characteristics. Since we observe a mass of working mothers across the hours of work distribution who do not use any childcare, a linear relationship (as in Blundell et al., 2000) is unlikely to be appropriate. Instead, we assume the presence of some underlying latent variable that governs both the selection mechanism and the value of required childcare itself. More specifically, we assume that the total childcare hours

[12]In the empirical application we assess the sensitivity of our results to these parametric assumptions.

constraint is given by:

$$\alpha_c(h, X, \epsilon) = \mathbf{1}(h > 0) \times \mathbf{1}(\epsilon_{c_X} > -\beta_{c_X} h - \gamma_{c_X}) \times (\gamma_{c_X} + \beta_{c_X} h + \epsilon_{c_X}) \tag{5}$$

where $\mathbf{1}(\cdot)$ is the indicator function, and where the explicit conditioning of the parameters and the unobservables on demographic characteristics X reflects the specification we adopt in our estimation, where we allow the parameters of this stochastic relationship to vary with a subset of observable characteristics X_c (specifically, the number and age composition of children). Total weekly childcare expenditure is then given by $p_c h_c$ with p_c denoting the hourly price of childcare. Empirically, we observe a large amount of dispersion in childcare prices, with this distribution varying systematically with the age composition of children. This is modelled by assuming that p_c follows some distribution $p_c \sim F_c(\cdot; X_c)$ which again varies with demographic characteristics.

Individuals are assumed to face a budget constraint, determined by a fixed gross hourly wage rate (assumed to be generated by a log-linear relationship of the form $\log w = X_w' \beta_w + \epsilon_w$) and the tax and transfer system. We arrive at our measure of consumption by subtracting both childcare expenditure $p_c h_c$ (which also interacts with the tax and transfer system) and fixed work-related costs from net-income. These fixed work-related costs help provide a potentially important wedge that separates the intensive and extensive margin. They reflect the actual and psychological costs that an individual has to pay to get to work. We model work-related costs as a fixed, one-off, weekly cost subtracted from net income at positive values of working time: $f = \alpha_f(h; X) = \mathbf{1}(h > 0) \times X_f' \beta_f$. It then follows that consumption at a given hours and programme participation choice is given by:

$$c(h, P; T, X, \epsilon) = wh - T(wh, \mathbf{h}, P; X) - p_c h_c - f \tag{6}$$

where non-labour income, such as child maintenance payments, enter the tax and transfer schedule T through the set of demographic characteristics X, and with the explicit conditioning of T on childcare expenditure suppressed for notational simplicity.

In order to fully describe the utility maximization problem of lone mothers, we denote $P^*(h) \in \{0, E(h; X, \epsilon)\}$ as the optimal choice of programme participation for given hours of work h, where $E(h; X, \epsilon) = 1$ if the individual is eligible to receive tax credits at hours h, and zero otherwise. Assuming eligibility, it then follows that $P^*(h) = 1$ if and only if the following condition holds:

$$u(c(h, P = 1; T, X, \epsilon), h, P = 1; X, \epsilon) \geq u(c(h, P = 0; T, X, \epsilon), h, P = 0; X, \epsilon) \quad (7)$$

where $c(h, P; X, \epsilon)$ is as defined in equation 6. It then follows that the optimal choice of hours $h^* \in \mathcal{H}$ maximises $U(c(h, P^*(h); T, X, \epsilon), h, P^*(h); X, \epsilon, \varepsilon)$ subject to the constraints as detailed above.

5 Data and Estimation

5.1 Data

We use six repeated cross-sections from the Family Resources Survey (FRS), from the financial year 1997/8 through to 2002/3, which covers the introduction and subsequent expansion of WFTC. The FRS is a cross-section household-based survey drawn from postcode records across Great Britain: around 30,000 families with and without children each year are asked detailed questions about earnings, other forms of income and receipt of state benefits.

Our sample is restricted to lone mothers who are aged between 18 and 45 at the interview date, not residing in a multiple tax unit household, and not in receipt of any disability related benefits. Dropping families with missing observations of crucial variables, and those observed during the WFTC phase-in period of October 1999 to March 2000 inclusive, restricts our estimation sample to 7,090 lone mothers.

5.2 Estimation

The full model (preferences, wages, and childcare) is estimated simultaneously by maximum likelihood; the likelihood function is presented in Appendix A.[13] We incorporate highly detailed representations of the tax and transfer system using FORTAX (Shephard, 2009). The budget constraints vary with individual circumstances, and reflect the complex interactions between the many components of the tax and transfer system. To facilitate the estimation procedure, the actual tax and transfer schedules are modified slightly to ensure that there are no discontinuities in net-income as either the gross wage or childcare expenditure vary for given hours of work. We do not attempt to describe the full UK system here, but the interested reader may consult Adam and Browne (2009) and O'Dea et al. (2007) for recent surveys; see Shephard (2009) for a discussion of the implementation of the UK system in FORTAX.

For the purpose of modelling childcare, we define six groups by the age of youngest child (0–4, 5–10, and 11–18) and by the number of children (1 and 2 or more). The stochastic relationship determining hours of required childcare $\alpha_c(h, X, \epsilon)$ varies within each of these groups, as does the childcare price distribution $F_c(\cdot; X_c)$. Using data from the entire sample period, the childcare price distribution is discretised into either four price points (if the youngest child is aged 0–4 or 5–10) or 2 points (if the youngest child is aged 11-18). In each case, the zero price point is included. The positive price points p_c are fixed prior to estimation and correspond to the mid-points in equally sized groups amongst those using paid childcare (these values are presented alongside the estimation results in Table 2). The probability that lone mothers face each of these discrete price points is estimated together with the full model.

[13]This simultaneous estimation procedure contrasts with labour supply studies in the UK that have used discrete choice techniques. Perhaps largely owing to the complexity of the UK transfer system, these existing studies (such as Blundell et al., 2000) typically pre-estimate wages which allows net-incomes to be computed prior to the main preference estimation. In addition to the usual efficiency arguments, the simultaneous estimation here imposes internal coherency with regards to the various selection mechanisms.

We impose concavity on the utility function by restricting the power terms θ_l and θ_y to be between 0 and 1 (see equation 4). The unobserved wage component ϵ_w and the random preference heterogeneity terms $(\epsilon_y, \epsilon_\eta, \epsilon_{c_X})$ are assumed to be normally distributed. Given the difficulty in identifying flexible correlation structures from observed outcomes (see Keane, 1992), we allow ϵ_y to be correlated with ϵ_w, but otherwise assume that the errors are independent. The integrals over ϵ in the log-likelihood function are approximated using Gaussian quadrature with 11 nodes in each integration dimension. See Appendix A for further details.

5.3 Specification and Structural Parameter Estimates

The estimates of the parameters of our structural model are presented in Table 2. The age of the youngest child has a significant impact on the estimated fixed costs of work α_f; fixed work related costs are higher by around £16 per week if the youngest child is of pre-school age. The presence of young children also has a significant effect on the linear preference terms α_y (negatively) and α_l (positively). Parents with more children are also estimated to have a higher valuation for leisure, as well as higher fixed costs of work.

Lone mothers who are older are estimated to have a lower preference for both consumption and leisure, but higher costs of claiming in-work support. Meanwhile, the main impact of education comes primarily on the preference for leisure α_l; mothers who have completed compulsory schooling have a lower preference for leisure. Ethnicity enters the model through both fixed costs of work and programme participation costs η; we find that programme participation costs are significantly higher for non-white lone mothers. Programme participation costs are found to fall significantly following the introduction of WFTC, although the reduction in the first year is small (as captured by the inclusion of a zero-one dummy variable in the first year of WFTC). In contrast to many theoretical optimal tax studies which assume that preferences are quasi-linear in consumption, our estimate of θ_y places significant curvature on consumption. The estimate of

Table 2: Maximum likelihood estimation results

Preference parameters

	constant	youngest child 0–4	youngest child 5–10	number of children -1	age	compulsory schooling	non-white	London	WFTC period	year 2000
α_y	1.570	-0.441	-0.171	0.018	-0.021	-0.091	–	–	–	–
	(0.128)	(0.119)	(0.096)	(0.039)	(0.007)	(0.094)				
α_l	2.673	0.251	0.203	0.132	-0.035	-0.341	–	–	–	–
	(0.117)	(0.125)	(0.113)	(0.033)	(0.006)	(0.070)				
θ_y	0.301	–	–	–	–	–	–	–	–	–
	(0.085)									
θ_l	1.000	–	–	–	–	–	–	–	–	–
	(–)									
α_f	0.295	0.164	0.029	0.057	0.005	0.072	-0.078	0.261	–	–
	(0.076)	(0.089)	(0.068)	(0.033)	(0.005)	(0.063)	(0.049)	(0.044)		
η	0.982	–	–	–	0.017	-0.116	0.544	–	-0.438	0.388
	(0.208)				(0.009)	(0.161)	(0.181)		(0.117)	(0.134)
σ_y	0.668	–	–	–	–	–	–	–	–	–
	(0.050)									
σ_η	2.182	–	–	–	–	–	–	–	–	–
	(0.195)									
ρ_{yw}	0.241	–	–	–	–	–	–	–	–	–
	(0.042)									

Continued . . .

Table 2: (continued)

Childcare parameters

	1 child youngest age 0–4	1 child youngest age 5–10	1 child youngest age 11–1	2+ children youngest age 0–4	2+ children youngest age 5–10	2+ children youngest age 11–1
γ_c	5.697	-6.371	-26.633	7.237	-22.996	-57.585
	(1.917)	(1.371)	(4.966)	(3.435)	(3.041)	(10.100)
β_c	0.694	0.654	0.283	1.180	1.270	0.640
	(0.064)	(0.047)	(0.150)	(0.131)	(0.109)	(0.301)
σ_c	13.234	11.779	24.528	27.206	27.428	42.603
	(0.474)	(0.314)	(2.246)	(0.941)	(0.872)	(3.751)
$\Pr(p_{cc}^1)$	0.179	0.173	0.145	0.152	0.133	0.175
	(0.019)	(0.018)	(0.036)	(0.019)	(0.016)	(0.048)
$\Pr(p_{cc}^2)$	0.206	0.181	–	0.192	0.147	–
	(0.021)	(0.019)	–	(0.023)	(0.018)	–
$\Pr(p_{cc}^3)$	0.244	0.191	–	0.289	0.162	–
	(0.024)	(0.020)	–	(0.030)	(0.020)	–
p_{cc}^1	0.000	0.000	0.000	0.000	0.000	0.000
p_{cc}^2	0.937	0.804	1.887	0.516	0.570	1.658
p_{cc}^3	2.172	1.594	–	1.547	1.474	–
p_{cc}^4	3.440	2.579	–	2.949	2.474	–

Wage equation

constant	education	age	age squared	non-white	London	1998	1999	2000	2001	2002	σ_w
-0.010	0.097	0.050	-0.051	-0.046	0.192	-0.005	0.025	0.129	0.146	0.144	0.404
(0.067)	(0.004)	(0.012)	(0.017)	(0.026)	(0.024)	(0.025)	(0.029)	(0.025)	(0.023)	(0.023)	(0.005)

Notes: All parameters estimated simultaneously by maximum likelihood, using FRS data and with sample selection as detailed in section 5.1. Standard errors calculated using the outer product of gradients method. Incomes are expressed in hundreds of pounds per week in April 2002 prices. Age and age squared are defined in terms of deviations from the median value; age squared is divided by one hundred. Compulsory schooling is equal to 1 if the individual completed school at age 16 or above. Education measures age that education was completed. London is equal to one if resident in the Greater London area. WFTC period is equal to one if individual is interviewed post-October 1999. Standard errors are presented in parentheses.

θ_l is equal to the upper bound imposed so that estimated preferences are linear in leisure.

Both the intercept γ_c and the slope coefficient β_c in the childcare equation are typically lower for those with older children. This reflects the fact that lone mothers with older children use childcare less, and that the total childcare required varies less with maternal hours of work. To rationalise the observed distributions, we require that the standard deviation σ_c is also larger for those with older children. As noted in section 5.2, the price distribution of childcare for each group was discretised in such a way that amongst those mothers using paid childcare, there are equal numbers in each discrete price group. Our estimates attach greater probability on the relatively high childcare prices (and less on zero price) than in our raw data. Individuals who do not work are therefore more likely to face relatively expensive childcare were they to work.

The hourly log-wage equation includes the age at which full-time education was completed (which enters positively), and both age and age squared (potential wages are increasing in age, but at a diminishing rate). Lone mothers who reside in the Greater London area have significantly higher wages, and the inclusion of time dummies track the general increase in real wages over time. There is considerable dispersion in the unobserved component of log-wages.

The within sample fit of the model is presented in Tables 3 and 4. The estimated model matches the observed employment states and the take-up rate over the entire sample period very well (see the first two columns of Table 3). We slightly under predict the number of lone mothers working 19 hours per week, and slightly over predict the number working either 26 or 33 hours per week, but the difference is not quantitatively large. Similarly, we obtain very good fit by age of youngest child. The fit to the employment rate is also encouraging, and the difference between predicted and empirical hours frequencies never differs by more than around three percentage points and is typically smaller. Furthermore, despite the relatively simple stochastic specification for childcare, our model per-forms reasonably well in matching both the use of childcare by maternal employment hours (both overall and by age of youngest child), and conditional hours of childcare. Full results are presented in the Supplementary Material.

The fit of the model over time is presented in Table 4. Fitting the model over time is more challenging given that time only enters our specification in a very limited manner - through the wage equation and via the change in the stigma costs of the accessing the tax credit. Despite this we are able to replicate the 9 percentage point increase in employment between 1997/98 and 2002/03 reasonably well with our model, although we do slightly under predict the growth in part-time employment over this period.

To understand what our parameter estimates mean for labour supply behaviour we simulate labour supply elasticities under the actual 2002 tax systems across a range of household types. All elasticities are calculated by simulating a 1% increase in consumption at all positive hours points.[14] The results of this exercise are presented in Table 5. Across our sample of single mothers, we obtain an overall participation elasticity of 0.77, with our estimates implying a lower participation elasticity for single mothers whose youngest child is under 4 (an elasticity of 0.66), while they are significantly higher for mothers with school aged children (0.90 if youngest child is aged 5-10; 0.75 if the youngest child is aged 11-18).

Intensive elasticities, which here measure the responsiveness of hours worked amongst employed single mothers to changes in in-work consumption, are small and are also increasing for parents with older children. Since mothers with older children also work longer hours on average (see Table 3), these intensive elasticities also reflect larger increases in absolute hours for these groups. Compensated intensive elasticities are slightly higher. Finally, the total hours elasticities reported in the table combine these intensive and extensive responses.[15] Here,

[14]In the Supplementary Material we also present elasticity measures which are calculated by increasing the gross wage by 1%. The tax and transfer system introduces a substantial wedge between these alternative elasticity measures.

[15]The total hours elasticity η_{total} is related to the intensive and extensive elasticities (respectively

Table 3: Predicted and empirical frequencies by age of youngest child

	All		0–4		5–10		11–18	
	Predicted	Empirical	Predicted	Empirical	Predicted	Empirical	Predicted	Empirical
0 hours	0.549	0.550	0.704	0.708	0.490	0.489	0.319	0.320
	(0.005)	(0.006)	(0.007)	(0.008)	(0.009)	(0.010)	(0.012)	(0.013)
10 hours	0.078	0.068	0.063	0.049	0.090	0.083	0.086	0.081
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)	(0.007)
19 hours	0.105	0.134	0.089	0.108	0.117	0.156	0.117	0.147
	(0.002)	(0.004)	(0.003)	(0.006)	(0.003)	(0.007)	(0.004)	(0.010)
26 hours	0.079	0.057	0.054	0.035	0.090	0.068	0.112	0.082
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.005)	(0.003)	(0.007)
33 hours	0.087	0.077	0.048	0.042	0.099	0.086	0.152	0.136
	(0.002)	(0.003)	(0.002)	(0.004)	(0.003)	(0.005)	(0.004)	(0.009)
40 hours	0.103	0.115	0.044	0.058	0.114	0.120	0.214	0.234
	(0.003)	(0.004)	(0.003)	(0.004)	(0.005)	(0.006)	(0.010)	(0.012)
Take-up	0.769	0.764	0.840	0.788	0.768	0.781	0.702	0.715
rate	(0.010)	(0.009)	(0.010)	(0.017)	(0.011)	(0.013)	(0.016)	(0.018)

Notes: Empirical frequencies calculated using FRS data with sample selection as detailed in Section 5.1. The discrete points 0, 10, 19, 26, 33 and 40 correspond to the hours ranges 0, 1–15, 16–22, 23–29, 30–36 and 37+ respectively. Empirical take-up rates calculated using reported receipt of FC/WFTC with entitlement simulated using FORTAX. Predicted frequencies are calculated using FRS data and the maximum likelihood estimates from Table 2. Standard errors are in parentheses, and calculated for the predicted frequencies by sampling 500 times from the distribution of parameter estimates and conditional on the sample distribution of observables.

Table 4: Predicted and empirical frequencies: 1997–2002

	1997		2002	
	Predicted	Empirical	Predicted	Empirical
0 hours	0.595	0.600	0.493	0.507
	(0.007)	(0.014)	(0.007)	(0.013)
10 hours	0.079	0.080	0.079	0.062
	(0.003)	(0.008)	(0.003)	(0.006)
19 hours	0.098	0.110	0.116	0.155
	(0.003)	(0.009)	(0.003)	(0.010)
26 hours	0.069	0.043	0.090	0.063
	(0.002)	(0.006)	(0.002)	(0.007)
33 hours	0.072	0.063	0.104	0.093
	(0.002)	(0.007)	(0.002)	(0.008)
40 hours	0.086	0.104	0.119	0.120
	(0.004)	(0.009)	(0.003)	(0.009)
Take-up	0.736	0.684	0.808	0.838
rate	(0.013)	(0.029)	(0.014)	(0.016)

Notes: Empirical frequencies calculated using FRS data with sample selection as detailed in Section 5.1. The discrete points 0, 10, 19, 26, 33 and 40 correspond to the hours ranges 0, 1–15, 16–22, 23–29, 30–36 and 37+ respectively. Empirical take-up rates calculated using reported receipt of FC/WFTC with entitlement simulated using FORTAX. Predicted frequencies are calculated using FRS data and the maximum likelihood estimates from Table 2. Standard errors are in parentheses, and calculated for the predicted frequencies by sampling 500 times from the distribution of parameter estimates and conditional on the sample distribution of observables.

Table 5: Simulated elasticities

	All		0–4		5–10		11–18	
	Uncomp.	Comp.	Uncomp.	Comp.	Uncomp.	Comp.	Uncomp.	Comp
Participation	0.770	0.770	0.663	0.663	0.897	0.897	0.745	0.745
Intensive	0.042	0.123	0.032	0.094	0.043	0.128	0.047	0.136
Total Hours	1.534	1.616	2.253	2.317	1.590	1.676	1.007	1.097

Notes: All elasticities simulated under actual 2002 tax systems with complete take-up of WFTC. Elasticities are calculated by increasing consumption by 1% at all positive hours choices. Participation elasticities measure the percentage point increase in the employment rate; intensive elasticities measure the percentage increase in hours of work amongst workers in the base system; total hours elasticities measure the percentage increase in total hours

the lower employment rates for single mothers with younger children produces somewhat higher total hours elasticities for these groups.

5.4 Simulating the WFTC Reform

Before we proceed to consider optimal design problems using our structural model, we first provide an evaluation of the impact of the WFTC reform discussed in section 3 above on single mothers. This exercise considers the impact of replacing the actual 2002 tax systems with the April 1997 tax system on the 2002 population. This exercise is slightly different to simply examining the change in predicted states over this time period as it removes the influence of changing demographic characteristics.

The results of this policy reform simulation are presented in Table 6. Overall we predict that employment increased by 5 percentage points as a result of these reforms, with the increase due to movements into both part-time and full-time employment. Comparing with Table 4 we find the reform explains around a half of the rise in employment over this period. The predicted increase in take-up

$\eta_{extensive}$ and $\eta_{intensive}$) according to $\eta_{total} = \eta_{intensive} + (Q/P) \times \eta_{extensive}$. Here, P denotes the employment rate, and Q is the ratio of average hours of new workers, relative to the initial average hours of existing workers.

Table 6: Impact of reforms: 1997-2002

	1997 system	2002 system	change
0 hours	0.546	0.493	-0.053
	(0.007)	(0.007)	(0.007)
10 hours	0.079	0.079	0.000
	(0.003)	(0.003)	(0.001)
19 hours	0.105	0.116	0.010
	(0.003)	(0.003)	(0.004)
26 hours	0.076	0.090	0.014
	(0.002)	(0.002)	(0.002)
33 hours	0.082	0.104	0.022
	(0.002)	(0.002)	(0.002)
40 hours	0.112	0.119	0.007
	(0.004)	(0.004)	(0.001)
Take-up	0.697	0.808	0.111
rate	(0.013)	(0.014)	(0.017)

Notes: impact of tax and transfer system reforms on hours of work and take-up simulated using FRS 2002 data by replacing actual 2002 tax systems with the April 1997 tax system. Standard errors are in parentheses and are calculated by sampling 500 times from the distribution of parameter estimates and conditional on the sample distribution of observables.

of tax credits is also substantial, with this increase driven both by the changing entitlement and the estimated reduction in programme participation costs.

6 The Optimal Design of the Tax and Transfer Schedule

In this section we use our structural model to examine the design of the tax and transfer schedule. This shows the key importance of the differences in labour supply responses at the extensive and intensive margin. We also examine the welfare cost from moving to an administratively simpler linear tax system. The variation in response elasticities noted in our discussion of the estimated model above points to potential gains from allowing the optimal schedule to vary with children's age. We investigate such a design.

Given the use of a minimum hours condition for eligibility in the British tax credit system, we also consider the design in the case of a minimum hours rule. We show that if hours of work are partially (but otherwise accurately) observable, then there can be modest welfare gains from introducing an hours rule for lone mothers. However, accurately observing hours of work is crucial for this result. Our results suggest that if hours of work are subject to measurement error – whether this be random or due to direct misreporting – then the welfare gains that can be realised may be much reduced. Our analysis here therefore supports the informal discussion regarding the inclusion of hours in the tax base in Banks and Diamond (2010). Before detailing these results, we first turn to the choice of social welfare transformation and the parameterisation of the tax and transfer schedule.

6.1 Optimal Tax Specification

We have shown that using parameter estimates from a structural model of labour supply, the behaviour of individuals can be simulated as the tax and transfer system is varied. With these heterogeneous labour supply responses allowed for, the structural model provides all the necessary information to maximise an arbitrary social welfare function, subject to a government budget constraint. Note that our analysis here integrates that tax and transfer system.

To implement the optimal design analysis we approximate the underlying non-parametric optimal schedule by a piecewise linear tax schedule that is characterised by a level of out-of-work income (income support), and nine different marginal tax rates. These marginal tax rates, which are restricted to lie between -100% and 100%, apply to weekly earnings from £0 to £400 in increments of £50, and then all weekly earnings above £400. We do not tax any non-labour sources of income, and do not allow childcare usage to interact with tax and transfer schedule unless explicitly stated. When we later allow for partial observability of hours we introduce additional payments that are received only if the individual fulfills the relevant hours criteria.

In all of these illustrations we condition upon the presence of a single child, and we set the value of government expenditure equal to the predicted expenditure on this group within our sample. Conditioning upon this level of expenditure we numerically solve for the tax and transfer schedule that maximises social welfare. In this section we adopt the following utility transformation in the social welfare function:

$$\Upsilon(U;\theta) = \frac{(\exp U)^{\theta} - 1}{\theta} \tag{8}$$

which controls the preference for equality by the one dimensional parameter θ and also permits negative utilities which is important in our analysis given that the state specific errors ε can span the entire real line. When θ is negative, the function in equation 8 favours the equality of utilities; when θ is positive the reverse is true. By L'Hôpital's rule $\theta = 0$ corresponds to the linear case. Note that $-\theta = -\Upsilon''(U;\theta)/\Upsilon'(U;\theta)$ so that $-\theta$ can be interpreted as the coefficient of absolute inequality aversion.

We solve the schedule for a set of parameter values $\theta = \{-0.4, -0.2, 0.0\}$ and then derive the social weights that characterise these redistributive preferences. We do not consider cases where $\theta > 0$. The presence of state specific Type-I extreme value errors, together with our above choice of utility transformation has some particularly convenient properties, as the follow Proposition now demonstrates.

Proposition 1. *Suppose that the utility transformation function is as specified in equation (8). If $\theta = 0$ then conditional on X and ϵ the integral over (Type-I extreme value) state specific errors ε in equation (2) is given by:*

$$\log\left(\sum_{h\in\mathcal{H}} \exp(u(c(h;T,X,\epsilon),h;X,\epsilon))\right) + \gamma$$

where $\gamma \approx 0.57721$ is the Euler-Mascheroni constant. If $\theta < 0$ then conditional on X

and ϵ the integral over state specific errors is given by:

$$\frac{1}{\theta}\left[\Gamma(1-\theta)\times\left(\sum_{h\in\mathcal{H}}\exp(u(c(h;T,X,\epsilon),h;X,\epsilon))\right)^{\theta}-1\right]$$

where Γ is the gamma function.

Proof. The result for $\theta = 0$ follows directly from an application of L'Hôpital's rule, and the well known result for expected utility in the presence of Type-I extreme value errors (see McFadden, 1978). See Appendix B for a proof in the case where $\theta < 0$. □

This proposition, which essentially generalizes the result of McFadden (1978), facilitates the numerical analysis as the integral over state specific errors does not require simulating. Moreover, the relationship between the utilities in each state, and the contribution to social welfare for given (X, ϵ) is made explicit and transparent.

6.2 Implications for the Tax Schedule

The underlying properties from the labour supply model, together with the choice of social welfare weights, are the key ingredients in the empirical design problem. We have seen from Table 5 that the intensive and extensive labour supply responses differ substantially. As expected this is reflected in the optimal tax results. For the choice of utility transformation function in equation (8) we examine the impact of alternative θ values. In Table 7 we present the underlying social welfare weights evaluated at the optimal schedule (discussed below) according to these alternative θ values. For all three values of θ considered here the weights are broadly downward sloping. For the most part we focus our discussion here on the -0.2 value, although we do provide a sensitivity of our results to the choice of θ and find the broad conclusions are robust to this choice.

Table 7: Social welfare weights under optimal system

Weekly	$\theta = -0.4$		$\theta = -0.2$		$\theta = 0.0$	
Earnings	Density	Weight	Density	Weight	Density	Weight
0	0.398	1.378	0.367	1.305	0.281	1.073
0–50	0.055	1.340	0.051	1.218	0.039	0.968
50–100	0.109	1.088	0.104	1.071	0.088	0.935
100–150	0.101	0.907	0.110	0.987	0.123	1.015
150–200	0.100	0.718	0.111	0.855	0.136	1.024
200–250	0.078	0.563	0.087	0.721	0.115	1.021
250–300	0.049	0.457	0.054	0.615	0.071	0.959
300–350	0.043	0.347	0.046	0.504	0.060	0.945
350–400	0.021	0.307	0.023	0.454	0.029	0.880
400+	0.046	0.184	0.047	0.305	0.058	0.806

Notes: Table presents social welfare weights under optimal structure of marginal tax rates and out-of-work income under range of distributional taste parameters θ as presented in Table 8. All incomes are in pounds per week and are expressed in April 2002 prices. Welfare weights are obtained by increasing consumption uniformly in the respective earnings range and calculating a numerical derivative; weights are normalized so that the earnings-density-weighted sum under optimal system is equal to unity.

In the first three columns of Table 8 we present the optimal tax and transfer schedules across the alternative θ values (also see Figure 3). We also present standard errors for the parameters of the optimal tax schedule. We obtain these by sampling 500 times from the distribution of parameter estimates and re-solving for the optimal schedule conditional on the sample distribution of covariates. In all the simulations performed here, the structure of marginal tax rates suggests lower rates at the lowest positive earnings levels. In particular, marginal rates are typically much lower in the first tax bracket (earnings up to £50 per-week) than at higher earnings. Apart from the $\theta = 0.0$ case, marginal tax rates are much higher in the second bracket (weekly earnings between £50 and £100), but then fall before proceeding to generally increase with labour earnings. As we increase the value of θ (corresponding to less redistributive concern), we obtain reductions in the value of out-of-work income. This is accompanied by broad decreases in marginal tax rates, except in the first tax bracket where marginal tax rates are

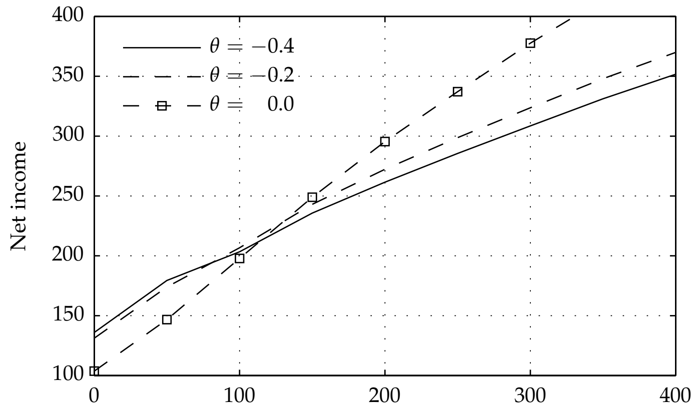


Figure 3: Optimal tax schedules with alternative values of θ. All incomes are measured in April 2002 prices and are expressed in pounds per week.

largely unchanged. The social welfare weights presented in Table 7 reflect these changes.

The results presented in Table 8 point towards a non-linear tax schedule over a large range of earnings. For each value of θ considered we quantify the welfare gains from allowing for such non-linearity by calculating the increase in government expenditure required such that the value of social welfare under the optimal linear tax system is the same as under the non-linear systems above. This produces optimal constant marginal tax rates of 43.5%, 37.6% and 11.3% (for $\theta = -0.4$, $\theta = -0.2$ and $\theta = 0.0$ respectively). The welfare gains from non-linearity are modest; in the illustrations when $\theta = -0.2$, government expenditure would need to increase by 1.5% to achieve the same level of social welfare.

6.3 Tagging by age of child

Before exploring the use of hours contingent payments in the tax schedule we consider how the optimal schedule varies by age of children, should the government decide to condition (or tag) the tax and transfer schedule upon this information.[16]

[16] The nature of the optimal income tax schedule in the presence of tagging was theoretically explored by Akerlof (1978).

Table 8: Optimal tax schedules

Weekly	No hours			19 hours			Optimal hours		
Earnings	$\theta = -0.4$	$\theta = -0.2$	$\theta = 0.0$	$\theta = -0.4$	$\theta = -0.2$	$\theta = 0.0$	$\theta = -0.4$	$\theta = -0.2$	$\theta = 0.0$
0–50	0.132	0.144	0.139	0.266	0.280	0.252	0.053	0.056	0.072
	(0.028)	(0.025)	(0.029)	(0.029)	(0.034)	(0.037)	(0.031)	(0.028)	(0.028)
50–100	0.520	0.344	-0.022	0.995	0.899	0.328	0.778	0.646	0.295
	(0.030)	(0.030)	(0.044)	(0.006)	(0.034)	(0.062)	(0.030)	(0.032)	(0.044)
100–150	0.354	0.275	-0.022	0.466	0.355	-0.013	0.535	0.481	0.267
	(0.019)	(0.020)	(0.037)	(0.027)	(0.019)	(0.039)	(0.021)	(0.022)	(0.030)
150–200	0.483	0.414	0.069	0.503	0.440	0.090	0.698	0.650	0.321
	(0.014)	(0.017)	(0.033)	(0.014)	(0.017)	(0.035)	(0.028)	(0.030)	(0.050)
200–250	0.520	0.471	0.167	0.535	0.484	0.173	0.672	0.638	0.338
	(0.015)	(0.017)	(0.038)	(0.015)	(0.017)	(0.039)	(0.030)	(0.032)	(0.051)
250–300	0.540	0.501	0.189	0.551	0.512	0.197	0.659	0.632	0.338
	(0.020)	(0.021)	(0.040)	(0.020)	(0.022)	(0.042)	(0.043)	(0.045)	(0.060)
300–350	0.546	0.514	0.266	0.554	0.521	0.270	0.644	0.618	0.365
	(0.023)	(0.025)	(0.053)	(0.024)	(0.026)	(0.053)	(0.038)	(0.040)	(0.064)
350–400	0.590	0.561	0.285	0.604	0.575	0.293	0.728	0.715	0.458
	(0.019)	(0.020)	(0.040)	(0.019)	(0.021)	(0.042)	(0.029)	(0.031)	(0.054)
400+	0.616	0.599	0.401	0.623	0.607	0.403	0.687	0.676	0.477
	(0.008)	(0.009)	(0.023)	(0.008)	(0.009)	(0.024)	(0.008)	(0.009)	(0.029)
Out-of-work	135.975	131.170	103.651	136.226	131.361	104.407	137.262	132.204	106.153
Income	(s1.672)	(s1.680)	(s3.308)	(1.704)	(1.686)	(3.348)	(1.740)	(1.736)	(3.300)
Hours bonus	–	–	–	36.290	38.698	23.231	44.056	48.632	47.995
				(1.670)	(1.357)	(2.944)	(2.037)	(1.540)	(5.140)
Hours point	–	–	–	19	19	19	33	33	33

Notes: Table presents optimal structure of marginal tax rates and out-of-work income under range of distributional taste parameters θ. All incomes are in pounds per week and are expressed in April 2002 prices. Standard errors are in parentheses and are calculated by sampling 500 times from the distribution of parameter estimates and conditional on the sample distribution of observables.

Note that WFTC awards depended upon on the age of children (see the different rates in Table 1) as do other parts of the UK tax and transfer system (including Income Support, the main transfer available to low income families working less than 16 hours per week).

Since our model is static this exercise ignores the dynamics that are introduced by the child ageing process. Clearly, such considerations could be important for the optimal design problem. Nonetheless, this remains an important benchmark case and is likely to still yield important insights, particularly if the population of interest have a sufficiently low discount factor, or are liquidity constrained.

We proceed to solve the optimal tax schedules for three different groups on the basis of the age of youngest child: under 4, aged 5 to 10 and 11 to 18. Since the childcare requirements of mothers with young children are considerably higher (see the estimates in Table 2), we also allow for a childcare expenditure subsidy of 70% (which corresponds to the formal childcare subsidy rate under WFTC) to facilitate the comparison of marginal tax rates across these groups. We first solve for these schedules separately when we condition on the predicted expenditure on each of these groups in our sample; we then solve for these schedules jointly allowing the division of overall expenditure to be re-optimised. Results are presented in Tables 9a and 9b. Figure 4 illustrates how the optimal schedules vary with the age of children, in the case with fixed group expenditure and where $\theta = -0.2$.

While the overall structure of the schedules (firstly, when we condition on within group expenditure – see Table 9a) retain many of the features present in our earlier simulations, our optimal tax simulations here reveal some important differences by the age of children. In particular, marginal tax rates tend to be higher at low earnings for lone mothers with younger children: in the first tax bracket marginal tax rates for the youngest group are around 40 percentage points higher than for the oldest group. Amongst women with children from the oldest

Table 9: Optimal tax system by age of child with childcare subsidy (conditional on group expenditure)

(a) Fixed expenditure division

Weekly	0–4			5–10			11–18		
Earnings	$\theta = -0.4$	$\theta = -0.2$	$\theta = 0.0$	$\theta = -0.4$	$\theta = -0.2$	$\theta = 0.0$	$\theta = -0.4$	$\theta = -0.2$	$\theta = 0.0$
0–50	0.198	0.287	0.432	-0.003	0.006	0.085	-0.107	-0.111	-0.009
50–100	0.503	0.344	0.043	0.545	0.370	0.013	0.478	0.279	-0.013
100–150	0.309	0.232	-0.033	0.395	0.320	0.038	0.445	0.343	-0.004
150–200	0.478	0.415	0.151	0.517	0.444	0.085	0.552	0.472	0.086
200–250	0.490	0.442	0.149	0.579	0.537	0.265	0.577	0.510	0.154
250–300	0.557	0.526	0.348	0.532	0.480	0.101	0.674	0.629	0.222
300–350	0.530	0.496	0.220	0.640	0.614	0.449	0.488	0.441	0.160
350–400	0.592	0.563	0.384	0.583	0.540	0.168	0.771	0.734	0.383
400+	0.607	0.590	0.431	0.640	0.622	0.420	0.654	0.631	0.377
Out-of-work income	140.950	139.152	126.405	131.855	125.374	95.572	118.382	106.947	66.850

(b) Optimal expenditure division

Weekly	0–4			5–10			11–18		
Earnings	$\theta = -0.4$	$\theta = -0.2$	$\theta = 0.0$	$\theta = -0.4$	$\theta = -0.2$	$\theta = 0.0$	$\theta = -0.4$	$\theta = -0.2$	$\theta = 0.0$
0–50	0.167	0.265	0.429	-0.002	0.008	0.085	-0.121	-0.115	-0.009
50–100	0.535	0.368	0.047	0.536	0.362	0.016	0.441	0.254	-0.024
100–150	0.316	0.238	-0.028	0.398	0.323	0.041	0.458	0.353	-0.015
150–200	0.473	0.406	0.156	0.519	0.447	0.088	0.564	0.483	0.073
200–250	0.482	0.433	0.153	0.584	0.541	0.268	0.585	0.517	0.146
250–300	0.544	0.513	0.351	0.533	0.482	0.104	0.685	0.640	0.209
300–350	0.523	0.490	0.223	0.643	0.618	0.450	0.495	0.447	0.154
350–400	0.581	0.551	0.387	0.585	0.543	0.171	0.780	0.742	0.372
400+	0.602	0.584	0.433	0.642	0.623	0.422	0.660	0.636	0.370
Out-of-work income	156.618	154.340	123.959	127.071	120.336	93.975	100.615	90.768	71.954

Notes: Table presents optimal structure of marginal tax rates and out-of-work income by age of youngest child under range of distributional taste parameters θ. All schedules calculated with an uncapped childcare subsidy equal to 70%. All incomes are in pounds per week and are expressed in April 2002 prices.

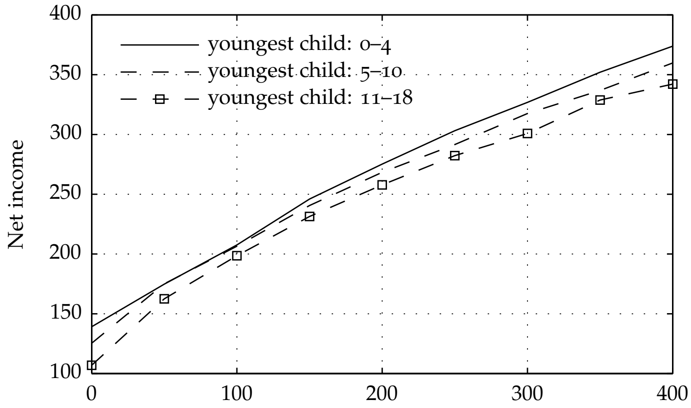


Figure 4: Optimal tax schedules by age of child. All schedules are calculated with fixed expenditure division and with $\theta = -0.2$. All incomes are measured in April 2002 prices and are expressed in pounds per week.

child age group we also obtain pure tax credits (negative marginal tax rates). The higher marginal tax rates at low earnings for parents with younger children are also accompanied by higher levels of out-of-work support for these groups.

Conditioning upon within group expenditure levels makes an implicit assumption on the weight that the government attaches on the welfare of parents with children of different ages. Under the assumption that the government places equal valuation on the welfare of individuals in each of these groups we solve for the three optimal schedules jointly (see Table 9b). Relative to the previous simulations, this makes the differences across groups more pronounced. In particular, there are notable increases in expenditure (and out-of-work income levels) for lone mothers with younger children. While there are some changes in the structure of marginal tax rates (due to income effects) these changes are somewhat smaller in magnitude.

The welfare gains from tagging on the basis of age of children can be calculated in much the same way as when comparing a non-linear schedule to one which is linear. The potential welfare gains appear reasonably large: relative to a system where tagging by the age of youngest child is not possible, government expenditure would have to increase by 2.6% (when $\theta = -0.2$) to obtain the same level of social welfare as that achieved when such tagging is possible. These gains are even larger when more redistributive preferences are considered.

6.4 Introducing an Hours Rule

For several decades the UK's tax credits and welfare benefits have made use of rules related to weekly hours of work. As discussed in section 3, individuals must work at least 16 hours a week to be eligible for in-work tax credits, and receive a further smaller credit when working 30 or more hours. While many theoretical models rule out the observability of any hours information, this design feature motivates us to explore the optimal structure of the tax and transfer system when hours can be partially observed as set out in section 2. We begin by assuming that the tax authority is able to observe whether individuals are working 19 hours or more, which roughly corresponds to the placement of the main 16 hours condition in the British tax-credit system, and for now we do not allow for any form of measurement error. In this case the tax authority is able to condition an additional payment on individuals working such hours. When the tax authority is only able to observe earnings, it is unable to infer whether an individual with a given level of earnings is low wage-high hours, or high wage-low hours. Since the government may value redistribution more highly in the former case, it may be able to better achieve its goals by introducing an hours rule into the system.

The results of this exercise are presented in columns 4–6 in Table 8, and the $\theta = -0.2$ case is also presented in Figure 5a.[17] Relative to the optimal system when such a rule is not implementable, the hours bonus increases marginal rates in the part of the earnings distribution where this hours rule would roughly come into effect (particularly in the £50 to £100 earnings bracket) while marginal rates further up the distribution, as well as the level of out-of-work support, are essentially unchanged. As a result of this, some non-workers with low potential wages may be induced to work part-time, while some low hours individuals will either

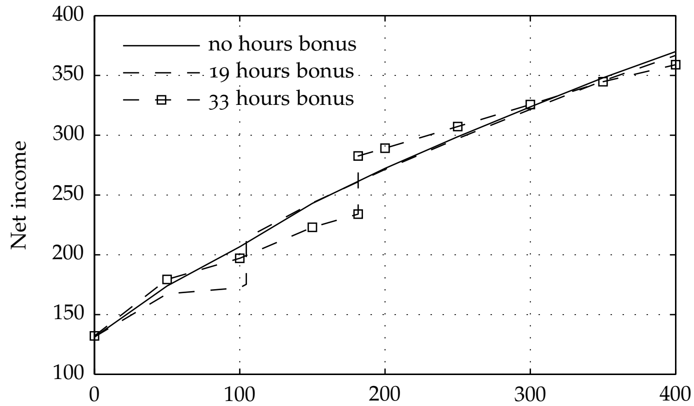
[17] The figure assumes a constant hourly wage rate of £5.50.

not work or increase their hours. Similarly, some high earnings individuals will reduce their hours to that required for the bonus. The hours bonus is sufficiently large for lone mothers such that the participation participation tax rate at 19 hours when earning the minimum wage rate is effectively zero.

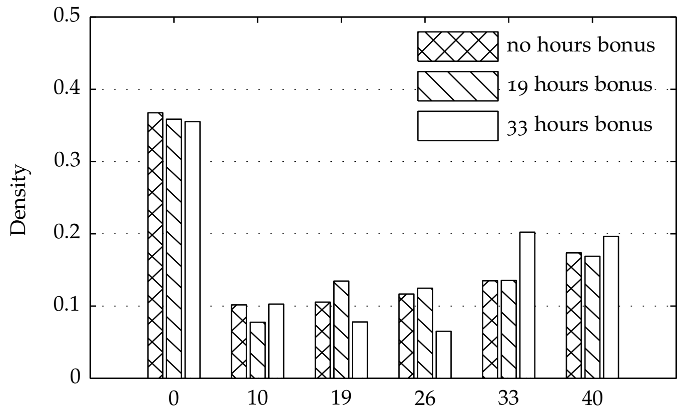
Although there are some notable changes in the structure of the constraint when hours information is partially observable, it does not follow that it necessarily leads to a large improvement in social welfare. Indeed, in the absence of the hours conditioning, there are only few individuals working less than 19 hours (see Figure 5b when $\theta = -0.2$) so the potential that it offers to improve social welfare appear limited. We now attempt to provide some guidance concerning the size of the welfare gain from introducing hours rules. The exact experiment we perform is as follows: we calculate the level of social welfare under the optimal schedule with hours contingent payments, and then determine the increase in expenditure that is required to obtain the same level of social welfare in the absence of such hours conditioning. In conducting this experiment we allow all the parameters of the (earnings) tax schedule to vary so this is obtained at least cost.

Perhaps unsurprisingly, these welfare gains are found to be relatively small; in both the $\theta = -0.4$ and $\theta = -0.2$ cases the expenditure increase required to achieve the level of social welfare obtained under the 19 hours rule is a little under 1% of the current level. When the least redistributive preferences are considered, this falls to just 0.2%. Even without allowing for any form of measurement error, it follows that unless the costs of partial hours observability is sufficiently low, it would appear difficult to advocate the use of a 19 hour rule based upon this analysis. This has very important policy implications given that the UK tax credit system makes heavy use of very similar hours conditions.[18]

[18]This finding contrasts with Keane and Moffitt (1998) which considered introducing a work subsidy in a model with three employment states (non-workers, part-time and full-time work) and multiple benefit take-up. Even small subsidies were found to increase labour supply and to reduce dependence on welfare benefits. In contrast to our application (where we are moving from a base with marginal rates well below 100% at low earnings), their simulations considered introducing the subsidy in an environment where many workers faced marginal effective tax rates



(a) Optimal tax schedules



(b) Distribution of work hours

Figure 5: Optimal tax schedules with hours bonuses and associated hours distribution. All schedules are calculated with $\theta = -0.2$ and assuming a gross hourly wage of £5.50. All incomes are measured in April 2002 prices and are expressed in pounds per week.

6.4.1 An Optimal Hours Rule?

The social welfare gains from introducing a 19 hours rule appear to be only very modest in size at best. In this section we explore whether there are potentially larger gains by allowing the choice of the point at which the hours rule becomes effective to be part of the optimal design problem. The parameters of the optimal tax schedules for all θ are presented in columns 7–9 of Table 8, while the optimal schedule when $\theta = -0.2$ is also shown in Figure 5a. Apart from when considering the least redistributive government preferences, we obtain an optimal hours rule at the fifth (out of six) discrete hours point, which corresponds to 33 hours per week.[19] We also note that the size of the optimally placed hours bonus always exceeds that calculated when the hours rule became effective at 19 hours per week.

Introducing an hours rule further up the hours distribution allows the government to become more effective in distinguishing between high wage/low effort and high effort/low wage individuals than at 19 hours to the extent that few higher wage individuals would choose to work very few hours. Relative to the schedule when the hours rule is set at around 19 hours, this alternative placement tends to make people with low and high earnings better off, while people in the middle range lose. While we again obtain very small adjustments to the level of out-of-work income, there are much more pronounced changes to the overall structure of marginal rates. In particular, there are large reductions in the marginal tax rate in the first tax bracket, while marginal rates now become higher at higher earnings. Figure 5b shows the resulting impact on the hours distribution when $\theta = -0.2$.

As before, we attempt to quantify the benefits from allowing for hours conditioning. Performing the same experiment as we conducted under the 19 hours rule we find that the required increase in expenditure is considerably larger than

which often exceeded 100%.

[19]When $\theta = 0.0$ the optimal placement shifts to 40 hours per-week.

that obtained previously. We find that a 2.5% increase in expenditure would be required to achieve the same level of social welfare when $\theta = -0.2$ (with very similar increases for the alternative θ values). While this is clearly not a "huge" amount, we believe that if hours can be accurately observed (as this analysis so far assumes), then this still represents a non-trivial welfare gain. In any case, if the government wishes to maintain the use of hours conditional eligibility, the analysis here suggests that it may be able to improve design by shifting towards a system that primarily rewards full-time rather than part-time work.[20]

6.5 Discrete hours sensitivity analysis

Before considering how our view regarding hours rules is affected by the presence of measurement error and hours misreporting, we first explore the sensitivity of our results with respect to the number of hours points available. The results reported here double the number of positive hours points (so a total of 11 discrete hours points) and re-estimate the structural model using these.[21] With the new set of parameter estimates, we again simulate a set of optimal tax schedules. The pure earnings schedules are very similar to those obtained with 6 discrete hours points; there are very similar levels of out-of-work income, and marginal rates from moderate earnings levels. The only notable difference is that the marginal tax rates in the first bracket are now slightly higher, while those in the second bracket are slightly lower.

The same general findings are true in the simulations with hours of work bonuses (both fixed, and with optimal hours bonus placement). Moreover, both the size and placement of these hours contingent payments are essentially the same as before. Full results from this exercise are presented in the Supplementary

[20] The welfare gains from a part-time hours rule are also typically small if we condition the tax system by the age of children as described in section 6.3. And while the welfare gains from an optimally placed (full-time) hours rule are also small for mothers with pre-school aged children, these gains are found to be much more substantial for parents with school age children. Full results are available upon request.

[21] The discrete points are now placed at 0, 5, 10, 14.5, 19, 22.5, 26, 29.5, 33, 36.5, and 40.

Material.

6.6 Measurement error and hours misreporting

The results presented so far have not allowed for any form of measurement error. While earnings may not always be perfectly measured, it seems likely that there is more scope for mismeasurement of hours as they are conceivably harder to monitor and verify. Indeed, the presence of hours rules in the tax and transfer system presents individuals with an incentive to not truthfully declare whether they satisfy the relevant hours criteria. Relative to when hours are always accurately reported, this would seem to weaken the case for introducing a measure of hours in the tax base. In this section we quantify the importance of such measurement error by considering two alternative scenarios: firstly, we consider the case where hours are imperfectly observed due to random measurement error; secondly, we allow individuals to directly misreport their hours of work to the tax authorities.

In performing this analysis it is necessary to modify our analytical framework from section 2 to distinguish between actual hours of work h, and reported hours of work h_R. While actual hours continue to determine both leisure and earnings, reported hours of work directly affect consumption through the tax schedule, with $T = T(wh, \mathbf{h_R}; X)$. They will also have a direct impact on utility when we allow for individual hours misreporting (discussed below).

6.6.1 Measurement error

We allow for random measurement error by adding an independent and normally distributed error term ν to work hours h to form a pseudo reported hours measure, $\tilde{h}_R = h + \nu$. Actual reported hours h_R are then given by the nearest discrete hours point in the set of hours $\mathcal{H}_{++}$. We assume that ν has zero mean, and in Table 10 we show how the size of the hours bonus and the associated welfare gain, vary as the standard deviation of the measurement error term σ_ν increases

Table 10: The effect of random measurement error on the optimal hours bonus

Standard	$\theta = -0.4$			$\theta = -0.2$			$\theta = 0.0$		
Deviation	bonus	hours	welfare	bonus	hours	welfare	bonus	hours	welfare
0	44.06	33	2.24%	48.63	33	2.46%	51.70	40	2.44%
2	42.08	33	2.10%	46.48	33	2.30%	50.85	40	2.38%
4	38.28	33	1.82%	42.28	33	1.99%	43.53	40	1.82%
6	34.38	33	1.58%	37.82	33	1.71%	38.28	33	1.30%
8	28.26	33	1.22%	31.09	33	1.32%	31.49	33	1.02%
10	23.58	33	0.96%	25.73	33	1.03%	26.10	33	0.80%
12	21.55	26	0.77%	23.69	26	0.82%	22.88	33	0.68%
14	17.75	26	0.59%	18.33	33	0.63%	19.00	33	0.51%

Notes: Table shows how the optimal placement and size of hours contingent payments varies with random hours measurement error. Standard Deviation refers to the standard deviation of the additive independent normally distributed hours measurement error term. The columns "welfare" refer to the percentage increase in required expenditure to achieve the same level of social welfare compared to when no hours conditioning is performed. All incomes are in pounds per week and are expressed in April 2002 prices.

in value. A clear pattern emerges. Across all values of θ, the size of the optimal hours bonus declines as reported hours become less informative. Furthermore, the placement of the optimal hours rule is reduced by a single discrete hours category for relatively high values of σ_ν (although a non-monotonic relationship is obtained in the case that $\theta = -0.2$). In the simulations where the standard deviation of the error term is between 4 and 8 (so that a single standard deviation results in reported hours differing from actual hours by a single category), the welfare gain from using hours information falls by between around 20% and 40%. The presence of random measurement error clearly reduces the desirability of conditioning upon hours, and if it is modest or large in size, then the welfare gains that are achievable are only small.

6.6.2 Hours misreporting

We have shown that random measurement error reduces the extent to which the government may wish to condition upon hours of work, and it also diminishes the welfare gains that are achievable. In the case of hours conditioning, it is plau-

Table 11: The effect of hours misreporting on the optimal hours bonus

Misreporting	$\theta = -0.4$			$\theta = -0.2$			$\theta = 0.0$		
Cost	bonus	hours	welfare	bonus	hours	welfare	bonus	hours	welfare
∞	46.53	33	2.31%	51.46	33	2.52%	54.80	40	2.57%
0.64	46.52	33	2.31%	51.45	33	2.52%	54.79	40	2.57%
0.32	45.25	33	2.28%	49.89	33	2.50%	53.76	40	2.56%
0.16	33.73	33	1.95%	37.74	33	2.12%	41.71	40	2.16%
0.08	24.24	33	1.36%	26.54	33	1.52%	29.26	40	1.63%
0.04	14.46	33	0.89%	15.89	33	1.00%	17.41	40	1.13%
0.02	9.24	33	0.58%	10.72	33	0.67%	12.44	40	0.83%
0.01	7.21	33	0.43%	8.12	33	0.52%	9.17	40	0.72%

Notes: Table shows how the optimal placement and size of hours contingent payments varies with the utility cost of hours misreporting. "Misreporting Cost" refers to the additive utility cost associated with misreporting, and is measured per-hour overstated and relative to standard deviation of the state specific error ε. The columns "welfare" refer to the percentage increase in required expenditure to achieve the same level of social welfare compared to when no hours conditioning is performed. All incomes are in pounds per week and are expressed in April 2002 prices.

sible that the form of misreporting is likely to be more systematic than random measurement error. Here we modify our setup to allow individuals to directly misreport their reported hours of work. We let h_B be the required hours of work to receive a bonus (received if $h \geq h_B$), and we continue to let h_R denote reported hours of work. Misreporting is only possible if $h > 0$, so that the tax authorities can always accurately observe employment status. If individuals misreport their hours of work then they must incur a utility cost, which is assumed to depend upon the distance $h_R - h$. Since misreporting hours is costly, it is only necessary to consider the cases when hours are truthfully revealed $h_R = h$, or when $h_R = h_B > h$.

We therefore modify the individual utility function by including $h_R - h$ as an explicit argument, so that $U = u(c, h, h_R - h; X, \epsilon) + \varepsilon_h$. This modified utility function is as in equation 4 but now with the additional cost term $b \times (h_R - h)$ subtracted from u whenever $h_R > h$.[22] If misreporting is not possible, then this

[22]In practice misreporting costs are likely to vary with both observed and unobserved worker characteristics. While it is sufficient to model this as a single cost for the purpose of our discussion

is equivalent to $b = \infty$. We do not allow individuals to manipulate their earnings wh. At a given actual hours of work $h < h_B$ individuals will report their hours as $h_R = h_B$ if and only if the utility gain exceeds the cost. That is:

$$u(c(h, T(wh, \mathbf{h_B}; X), X, \epsilon), h, h_B - h; X, \epsilon) > u(c(h, T(wh, \mathbf{h}; X), X, \epsilon), h, 0; X, \epsilon).$$

We refer to the parameter b as the misreporting cost, and in the results presented in Table 11 this is measured relative to the standard deviation of the state specific error ε. With an hours bonus payable at 33 hours per week (for example), a value of $b = 0.16$ would mean that the utility cost of reporting 33 hours when actual hours are 26 is equivalent to a $0.16 \times (33 - 26) = 1.12$ standard deviation change in the realisation of the state specific error. The table illustrates that as the utility cost of misreporting becomes very low, the welfare gain from using reported hours of work effectively disappears (but the optimal placement remains unchanged for all values considered). Again, this suggests that the welfare gains from using hours of work information may be small unless the scope for misreporting hours of work is limited.

7 Pareto Improving Reforms

The analysis of the previous section delivered some strong results. In particular, it pointed to marginal rates which are somewhat lower than under the actual UK system, particularly at low levels of earnings, and also suggested a welfare enhancing role of tagging taxes by the age of children. The analysis also had some important implications concerning the use of hours conditions in the tax schedule; the welfare gains from using a part time hours rule - a prominent feature of the UK system - appears limited. Larger gains may be realised by primarily rewarding full-time, but even these gains are mitigated by the presence of misreporting and measurement error.

and simulations here, our framework can easily be extended to incorporate such heterogeneity.

All these results were, however, derived under the assumption of a specific class of social welfare function with varying degrees of inequality aversion. In this section we are concerned with the extent to which these features are also implied solely by efficiency. To that end, we wish to identify a set of reforms that result in *Pareto improvements*. This exercise is closely related to Werning (2007), who characterized the set of Pareto efficient tax systems within the Mirrlees (1971) model, and proposed a test for efficiency through the lens of that model.

7.1 Conceptual framework

The exact experiment that we conduct here is as follows. We take the actual 2002 tax and transfer systems T with complete take up of tax credits and calculate the maximised value of utility for all X and all (ϵ, ε) subject to the individual incentive compatibility constraint (equation 1) and individual budget constraint (equation 6). With slight abuse of our earlier notation, we denote these maximised utility levels as $U(T, X, \epsilon, \varepsilon)$. We consider reforms to the tax and transfer system T by constructing a new tax and transfer system T^*, where $T^* = T + T'$. While T accurately reflects the full heterogeneity in the actual tax system (a function of demographics X, earnings wh, hours h and childcare expenditure $p_c h_c$), we will restrict ourselves to reforms where T' is a function only of earnings wh and later will also allow it to be a function of partially observed hours of work. Maximised utility levels as a function of T^* and individual heterogeneity $(X, \epsilon, \varepsilon)$ are denoted by $U(T^*, X, \epsilon, \varepsilon)$.

As in section 6, we parametrically specify T' and then proceed to search for the parameters of this schedule which maximise the revenue of the government, subject to the requirement that each individual is at least as well off as under the actual tax and transfer systems T. That is, we require that $U(T^*, X, \epsilon, \varepsilon) \geq U(T, X, \epsilon, \varepsilon)$ for all $(X, \epsilon, \varepsilon)$. If revenue is not maximised under the existing system then it can not be Pareto efficient, since it would be possible to reform the system in a direction which, by raising revenue, allows the welfare of some individuals to be improved without harming others. Note that Pareto improvements in this setting require *reductions* in tax schedules.

7.2 Implications of efficiency for the tax schedule

The results of this exercise are presented in column 2 of Table 12. We again restrict ourselves to a piecewise linear schedule, but allow for an increased number of tax brackets to help identify regions where Pareto improvements are obtainable. Reductions in the tax schedule are found for weekly earnings between 225 and 400 pounds per week. This is precisely the range where the density of earnings is falling most quickly (see column 1 in the same table). The table also quantifies the inefficiency under the existing system by comparing the actual and maximised revenue levels from this exercise. The same metric was proposed by Werning (2007) but was not quantitatively explored. As a result of this reform, we find that the government expenditure on single mothers is reduced by less that 0.1%. Thus, the increase in tax revenue that this particular reform delivers is clearly very small. Of course, this metric does not quantify any gains that accrue to single mothers as a result of the reductions in the tax schedules that they face.

Before we explore incorporating partial hours observability into T', we first consider a somewhat more relaxed criterion where we integrate over some dimensions of the unobserved heterogeneity and require that individuals are made no worse off for all (X, ϵ_w). The inequality constraints are then replaced by:

$$\int_{\epsilon_{-w}} \int_{\varepsilon} U(T^*, X, \epsilon, \varepsilon) dF(\varepsilon) dG(X, \epsilon_{-w}|\epsilon_w) \geq \int_{\epsilon_{-w}} \int_{\varepsilon} U(T, X, \epsilon, \varepsilon) dF(\varepsilon) dG(X, \epsilon_{-w}|\epsilon_w)$$

for all (X, ϵ_w). This may be viewed as an appropriate criterion if we think of social welfare conditional on characteristics X and idiosyncratic productive capacity ϵ_w. Note that this relaxed criterion does not necessarily require reductions in the tax schedule everywhere. The results are shown in column 4 of Table 12, and are extremely similar to those obtained in our initial exercise.

Table 12: Pareto improving changes to the tax schedule

Weekly	Base	Conditional on $(X, \epsilon, \varepsilon)$		Conditional on (X, ϵ_w)	
Earnings	Density	No hours rule	Hours rule	No hours rule	Hours rule
0–25	0.005	0.000	0.000	0.000	-0.297
25–50	0.041	0.000	0.000	0.000	0.243
50–75	0.046	0.000	0.000	0.000	0.194
75–100	0.044	0.000	0.000	0.000	-0.119
100–125	0.054	0.000	0.000	0.000	0.025
125–150	0.048	0.000	0.000	0.000	0.192
150–175	0.049	0.000	0.000	0.000	-0.231
175–200	0.042	0.000	0.000	0.000	-0.075
200–225	0.034	-0.076	-0.076	-0.083	0.167
225–250	0.032	0.077	0.077	0.088	-0.048
250–275	0.021	-0.435	-0.435	-0.456	-0.092
275–300	0.020	0.064	0.064	0.074	-0.107
300–325	0.016	-0.073	-0.073	-0.052	0.072
325–350	0.018	0.273	0.273	0.167	0.074
350–375	0.010	0.170	0.170	0.253	0.193
375–400	0.008	0.000	0.000	0.059	0.224
400–425	0.008	0.000	0.000	0.026	0.107
425–450	0.007	0.000	0.000	-0.030	-0.354
450–475	0.007	0.000	0.000	-0.038	0.178
475–500	0.006	0.000	0.000	-0.001	-0.001
500+	0.027	0.000	0.000	0.000	-0.269
Out-of-work Income		0.000	0.000	0.000	0.269
Bonus at 16 hours		–	0.000	–	-1.370
Bonus at 30 hours		–	0.000	–	18.616
Change in employment		0.002	0.003	0.003	0.006
Change in expenditure		-0.090%	-0.090%	-0.095%	-0.692%

Notes: Table presents changes to the structure of marginal tax rates, out-of-work income, and hours contingent payments that yield Pareto improvements conditional on $(X, \epsilon, \varepsilon)$ and (X, ϵ_w) respectively. The base system refers to the actual 2002 tax and transfer system with complete take-up of tax credits. All incomes are in pounds per week and are expressed in April 2002 prices.

7.3 Incorporating hours information

We now consider the use of hours information to improve efficiency. The hours rules in T' are restricted to operate at the same location as under the actual systems (that is, further payments are received if working at the discrete points corresponding to more than 16 and more than 30 hours per-week). Here we abstract from any form of hours mismeasurement. Note that if we condition on all the observed and unobserved heterogeneity in the tax system, then Pareto improvements do not permit any reductions in these hours contingent payments since it would make individuals with a particularly high attachment to a given hours state worse off. This severely limits the potential for reforms to the hours rules to yield Pareto improvements. Indeed, the revenue maximizing tax schedules (column 3) does not alter the hours bonuses, with the reformed schedule the same as reported in column 2 of the same table.

Unsurprisingly, the more relaxed criterion produces quite different results as we are integrating over the unobserved heterogeneity ε that is responsible for this hours attachment. The results from this exercise (see column 5) point to a small increase in out-of-work income, together with a reduction in the size of the part-time hours bonus and a large increase in the full-time hours bonus. There are also pronounced changes to marginal tax rates over the entire distribution of labour earnings. This reform produces larger reductions in government expenditure relative to when we did not adjust the size of the hours bonuses (around 1%). Moreover, the direction of this reform is consistent with our earlier results in section 6 when we adopted a social welfare function with varying degrees of inequality aversion.

7.4 The constraints on social welfare maximisation

The requirement that no individual is made worse off following a tax reform is a demanding criterion, particularly in the presence of preference heterogeneity.

In this subsection we seek to quantify the extent to which imposing this requirement may restrict the potential for social welfare improving reforms. To do this, we first perform a similar exercise to that in section 6 by calculating the social welfare maximizing tax schedule (under the same set of redistributional parameters that was previously considered) subject to the usual incentive compatibility constraints and government revenue constraint. We then proceed to calculate the increase in revenue that is required such that this same level of social welfare is achieved, but subject to the additional requirement that no individual is made worse off relative to the actual tax and transfer.[23]

For each value of the redistributional taste parameter θ we conduct four sets of simulations; when individuals are made no worse off conditional on the full set of observable and unobservable characteristics $(X, \epsilon, \varepsilon)$ both with and without possible hours rules, and also when we only condition on demographics and productive capacity (X, ϵ_w) (again, with and without possible hours rules). The results of this exercise are presented in Table 13, which shows the proportional increase in required government expenditure. The table shows that the constraint that no individual is made worse off would impose a significant constraint on the welfare maximisation problem.

8 Summary and Conclusions

The aim of this paper has been to examine the optimal design of the tax schedule using a stochastic structural labour supply model. The application focussed on the design of the tax schedule for parents with children, in particular single mothers in the UK. The structural labour supply model was shown to be reliable and found to match closely the changes in observed behaviour that followed a large reform to the tax credit system in the UK.

[23]The "unconstrained" maximisation problem that we consider here differs slightly from that considered in section 6; the tax schedule is now constructed in the same way as when we were examining the Pareto improving tax reforms.

Table 13: Increases in expenditure to make no individual worse off

	Conditional on $(X, \epsilon, \varepsilon)$		Conditional on (X, ϵ_w)	
	No hours rule	Hours rule	No hours rule	Hours rule
$\theta = -0.4$			1.92%	2.03%
$\theta = -0.2$			2.50%	2.50%
$\theta = \quad 0.0$			6.67%	5.95%

Notes: Table presents the increase in government expenditure required such that the value of maximised social welfare (under the additional requirement that no individual is made worse off conditional on $(X, \epsilon, \varepsilon)$ and (X, ϵ_w) respectively) is the same relative to when this constraint is not imposed.

The optimal design problem has been developed within an extended Mirrlees framework which incorporates unobserved heterogeneity, the non-convexities of the tax and welfare system as well as allowing for childcare costs and fixed costs of work. We considered social welfare improving designs for a variety of social welfare functions that display inequality aversion and we have also examined purely Pareto improving reforms.

To mirror the hours contingent nature of the British tax credit system we developed an analytical framework that explicitly allowed for the tax authorities to have partial observability of hours of work. We contrasted this to the standard case in which only earnings (and employment) are revealed to the tax authority. Reflecting the variation in estimated labour supply responses with the age of children we also considered a design in which there is tagging in the tax schedule according to child age.

When firstly considering social welfare improving designs, our results highlighted a role for conditioning effective tax rates on the age of children. Tax credits being found to be most important for low earning families with school age children. Hours contingent payments, as feature in the British tax credit system, are also found to lead to improvements in the tax design. If the tax authorities are able to choose the lower limit on working hours that trigger eligibility for such families, then we find an empirical case for using a full-time work rule rather than the main part-time rule currently in place for parents in the UK. While this is found to be a more effective instrument, we demonstrate how the welfare gains diminish with both misreporting and measurement error.

We identified inefficiencies in the actual UK tax and transfer system, and characterised purely Pareto improving reforms. Within this framework, and when viewing individual welfare conditional on observable characteristics and productive capacity, we presented a pure efficiency case for moving towards a tax system that places greater emphasis on rewarding full-time rather than part-time work.

Appendix

A Likelihood function

In what follows let $\mathcal{P}_j(X, p_{c_k}, \epsilon) \equiv \Pr(h = h_j | X, p_{c_k}, \epsilon)$ denote the probability of choosing hours $h_j \in \mathcal{H}$ conditional on demographics X, the childcare price p_{c_k}, and the vector of unobserved preference heterogeneity $\epsilon = (\epsilon_w, \epsilon_{c_X}, \epsilon_y, \epsilon_\eta)$. Given the presence of state specific Type-I extreme value errors, this choice probability takes the familiar conditional logit form. We also use $\pi_k(X) \equiv \Pr(p_c = p_{c_k} | X)$ to denote the probability of the lone mother with characteristics X facing childcare price p_{c_k}. In the case of non-workers ($h = h_0$), neither wages nor childcare are observed so that the likelihood contribution is simply given by:

$$\sum_k \pi_k(X) \int_\epsilon \mathcal{P}_0(X, p_{c_k}, \epsilon) dG(\epsilon).$$

Now consider the case for workers when both wages and childcare information is observed so that h_c is not censored at zero. Using $E_h \equiv E(h; X, p_c, \epsilon)$ to denote eligibility for in-work support we define the indicator $\mathcal{D}(e, p) = \mathbf{1}(E_h = e, P = p)$. We also let $\Delta u(h_j | p_{c_k}, X, \epsilon_{|\epsilon_\eta = 0})$ denote the (possibly negative) utility gain from claiming in-work support at hours h_j, conditional on demographics X, the childcare price p_{c_k}, and the vector of unobserved preference heterogeneity ϵ with $\epsilon_\eta = 0$. Suppressing the explicit conditioning for notational simplicity, the likelihood contribution is given by:

$$\prod_k \pi_k(X)^{\mathbf{1}(p_c=p_{c_k})} \int_{\epsilon_y} \left\{ \mathcal{D}(1,1) \int_{\epsilon_\eta<\Delta u} \prod_j \mathcal{P}_j(X, p_{c_k}, \epsilon)^{\mathbf{1}(h=h_j)} \right.$$

$$\left. + \mathcal{D}(1,0) \int_{\epsilon_\eta>\Delta u} \prod_j \mathcal{P}_j(X, p_{c_k}, \epsilon)^{\mathbf{1}(h=h_j)} + \mathcal{D}(0,0) \int_{\epsilon_\eta} \prod_j \mathcal{P}_j(X, p_{c_k}, \epsilon)^{\mathbf{1}(h=h_j)} \right\}$$

$$dG(\epsilon|\epsilon_w = \log w - X_w'\beta_w, \epsilon_c = h_c - \gamma_{c_X} - \beta_{c_X} h)$$

$$g_{w,c}(\log w - X_w'\beta_w, h_c - \gamma_{c_X} - \beta_{c_X} h).$$

If working mothers are not observed using childcare, then h_c is censored at zero and the childcare price also unobserved. If $\bar{\epsilon}_c = -\gamma_{c_X} - \beta_{c_X} h$, then the likelihood contribution is given by:

$$\sum_k \pi_k(X) \iint_{\epsilon_c<\bar{\epsilon}_c, \epsilon_y} \left\{ \mathcal{D}(1,1) \int_{\epsilon_\eta<\Delta u} \prod_j \mathcal{P}_j(X, p_{c_k}, \epsilon)^{\mathbf{1}(h=h_j)} \right.$$

$$\left. + \mathcal{D}(1,0) \int_{\epsilon_\eta>\Delta u} \prod_j \mathcal{P}_j(X, p_{c_k}, \epsilon)^{\mathbf{1}(h=h_j)} + \mathcal{D}(0,0) \int_{\epsilon_\eta} \prod_j \mathcal{P}_j(X, p_{c_k}, \epsilon)^{\mathbf{1}(h=h_j)} \right\}$$

$$dG(\epsilon|\epsilon_w = \log w - X_w'\beta_w) g_w(\log w - X_w'\beta_w).$$

Our estimation also allows for workers with missing wages. This takes a similar form to the above, except that it is now necessary to also integrate over the unobserved component of wages ϵ_w.

All the integration over ϵ is performed using Gaussian Hermite quadrature with 11 nodes in each integration dimension. When it is unnecessary to integrate over the entire real line in a given dimension, a change of variable is conducted so that integration is performed over $[0, +\infty)$, with appropriate semi-Hermite quadrature formulae then applied (see Kahaner et al., 1982).

B Proof of Proposition

For notational simplicity we abstract from the explicit conditioning of utility on observed and unobserved preference heterogeneity and let $u(h) \equiv u(c(h), h; X, \epsilon)$. We then define V as the integral of transformed utility over state specific errors conditional on (X, ϵ):

$$V \equiv \int_{\varepsilon} \Upsilon\left(\max_{h \in \mathcal{H}} \left[u(h) + \varepsilon_h\right]\right) dF(\varepsilon) \tag{A-1}$$

To prove this result we first differentiate V with respect to $u(h)$:

$$\begin{aligned}
\frac{\partial V}{\partial u(h)} &= \int_{\varepsilon} \left(\frac{\partial \Upsilon\left(\max_{h \in \mathcal{H}} \left[u(h) + \varepsilon_h\right]\right)}{\partial u(h)}\right) dF(\varepsilon) \\
&= \int_{\varepsilon} \Upsilon'\left(u(h) + \varepsilon_h\right) \times \mathbf{1}\left(h = \arg\max_{h' \in \mathcal{H}} \left[u(h') + \varepsilon_{h'}\right]\right) dF(\varepsilon)
\end{aligned}$$

Given our choice of utility transformation function in equation 8 and our distributional assumptions concerning ε the above becomes:

$$\begin{aligned}
\frac{\partial V}{\partial u(h)} &= \int_{\varepsilon_h = -\infty}^{\infty} \left\{e^{(u(h) + \varepsilon_h)}\right\}^{\theta} \left(\prod_{h' \neq h} e^{-e^{-\{\varepsilon_h + u(h) - u(h')\}}}\right) \times e^{-\varepsilon_h} e^{-e^{-\varepsilon_h}} d\varepsilon_h \\
&= \left\{e^{u(h)}\right\}^{\theta} \int_{\varepsilon_h = -\infty}^{\infty} \left\{e^{\varepsilon_h}\right\}^{\theta} \times \exp\left(-e^{-\varepsilon_h} \sum_{h' \in \mathcal{H}} e^{-(u(h) - u(h'))}\right) e^{-\varepsilon_h} d\varepsilon_h
\end{aligned}$$

We proceed by using the change of variable $t = \exp(-\varepsilon_h)$ so that the above partial derivative becomes:

$$\frac{\partial V}{\partial u(h)} = \left\{e^{u(h)}\right\}^{\theta} \int_{t=0}^{\infty} t^{-\theta} \times \exp\left(-t \sum_{h' \in \mathcal{H}} e^{-(u(h) - u(h'))}\right) dt$$

By defining $z \equiv t \times \sum_{h' \in \mathcal{H}} e^{-(u(h)-u(h'))}$ we can once again perform a simple change of variable and express the above as:

$$
\begin{aligned}
\frac{\partial V}{\partial u(h)} &= \left\{e^{u(h)}\right\}^{\theta} \left\{\sum\nolimits_{h' \in \mathcal{H}} e^{-(u(h)-u(h'))}\right\}^{\theta-1} \int_{z=0}^{\infty} z^{-\theta} e^{-z} dz \\
&= e^{u(h)} \left\{\sum\nolimits_{h' \in \mathcal{H}} e^{u(h')}\right\}^{\theta-1} \int_{z=0}^{\infty} z^{-\theta} e^{-z} dz \\
&= e^{u(h)} \left\{\sum\nolimits_{h' \in \mathcal{H}} e^{u(h')}\right\}^{\theta-1} \Gamma(1-\theta) \qquad \text{(A-2)}
\end{aligned}
$$

where the third equality follows directly from the definition of the Gamma function $\Gamma(\cdot)$. Note that this integral will always converge given that we are considering cases where $\theta < 0$. Integrating equation A-2 we obtain:

$$
V = \frac{1}{\theta} \left[\Gamma(1-\theta) \times \left(\sum_{h' \in \mathcal{H}} \exp\left\{u(h')\right\} \right)^{\theta} - 1 \right] \qquad \text{(A-3)}
$$

where the constant of integration is easily obtained by considering the case of a degenerate choice set and directly integrating A-1. This completes our proof of the Proposition.

References

Aaberge, R. and U. Colombino (2008): "Designing Optimal Taxes with a Microeconometric Model of Household Labour Supply," CHILD Working Paper 06/08. 4

Adam, S. and J. Browne (2009): "A survey of the UK tax system," Briefing Note 9, Institute for Fiscal Studies, London. 3, 16

Akerlof, G. A. (1978): "The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning," *American Economic Review*, 68, 8–19. 30

Banks, J. and P. Diamond (2010): "The base for direct taxation," in *Dimensions of Tax Design: The Mirrlees Review*, ed. by J. Mirrlees, S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Grammie, P. Johnson, G. Myles, and J. Poterba, Oxford University Press for Institute for Fiscal Studies. 26

Blundell, R. (2002): "Welfare-to-work: Which policies work and why?" in *Proceedings of the British Academy*, vol. 117, 477–524. 7

Blundell, R., A. Duncan, J. McCrae, and C. Meghir (2000): "The Labour Market Impact of the Working Families' Tax Credit," *Fiscal Studies*, 21, 75–104. 3, 13, 16

Blundell, R. and H. W. Hoynes (2004): "Has "In-Work" Benefit Reform Helped the Labor Market?" in *Seeking a Premier Economy: The Economic Effects of British Economic Reforms, 1980-2000*, ed. by R. Blundell, D. Card, and R. B. Freeman, Chicago: University Of Chicago Press, 411–459. 2, 8

Blundell, R. and T. Macurdy (1999): "Labor supply: A review of alternative approaches," in *Handbook of Labor Economics*, ed. by O. C. Ashenfelter and D. Card, Elsevier, vol. 3, Part 1 of *Handbook of Labor Economics*, 1559–1695. 2

Brewer, M., E. Saez, and A. Shephard (2010): "Means testing and tax rates on earnings," in *Dimensions of Tax Design: The Mirrlees Review*, ed. by J. Mirrlees, S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Grammie, P. Johnson, G. Myles, and J. Poterba, Oxford University Press for Institute for Fiscal Studies. 2

Card, D. and P. Robins (1998): "Do financial incentives encourage welfare recipients to work? Evidence from a randomized evaluation of the Self-Sufficiency Project," *Research in Labor Economics*, 17, 1–56. 3

Hoynes, H. W. (1996): "Welfare Transfers in Two-Parent Families: Labor Supply and Welfare Participation Under AFDC-UP," *Econometrica*, 64, 295–332. 3, 11

Kahaner, D., G. Tietjen, and R. Beckman (1982): "Gaussian-quadrature formulas for $\int \exp(-x^2)g(x)dx$," *Journal of Statistical Computation and Simulation*, 15, 155–160. 52

Keane, M. (1995): "A New Idea for Welfare Reform," *Federal Reserve Bank of Minneapolis Quarterly Review*, 19, 2–28. 3

Keane, M. and R. Moffitt (1998): "A Structural Model of Multiple Welfare Program Participation and Labor Supply," *International Economic Review*, 39, 553–589. 3, 11, 36

Keane, M. P. (1992): "A Note on Identification in the Multinomial Probit Model," *Journal of Business & Economic Statistics*, 10, 193–200. 17

Laroque, G. (2005): "Income Maintenance and Labor Force Participation," *Econometrica*, 73, 341–376. 2

McFadden, D. (1978): "Modeling the Choice of Residential Location," in *Spatial Interaction Theory and Planning Models*, ed. by A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull, North Holland, 75–96. 28

Mirrlees, J. A. (1971): "An Exploration in the Theory of Optimum Income Taxation," *The Review of Economic Studies*, 38, 175–208. 2, 44

Moffitt, R. (2003): "The Temporary Assistance for Needy Families Program," in *Means-Tested Transfer Programmes in the US*, ed. by R. Moffitt, University of Chicago Press. 3

O'Dea, C., D. Phillips, and A. Vink (2007): "A survey of the UK benefit system," Briefing Note 13, Institute for Fiscal Studies, London. 16

Saez, E. (2002): "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses," *The Quarterly Journal of Economics*, 117, 1039–1073. 2

Shephard, A. (2009): "FORTAX: Reference Manual," University College London, Manuscript. 16

van Soest, A., M. Das, and X. Gong (2002): "A structural labour supply model with flexible preferences," *Journal of Econometrics*, 107, 345–374. 3

Werning, I. (2007): "Pareto Efficient Income Taxation," Massachusetts Institute of Technology, Manuscript. 44, 45