



Working Papers

www.cesifo.org/wp

Optimal Taxation, Income Inequality and the Household

Patricia Apps
Ray Rees

CESIFO WORKING PAPER NO. 5845
CATEGORY 1: PUBLIC FINANCE
APRIL 2016

An electronic version of the paper may be downloaded

- *from the SSRN website:* www.SSRN.com
- *from the RePEc website:* www.RePEc.org
- *from the CESifo website:* www.CESifo-group.org/wp

ISSN 2364-1428

Optimal Taxation, Income Inequality and the Household

Abstract

This paper analyses optimal piecewise linear tax systems for two-earner households, based on joint and individual incomes respectively. It models the interaction between wage rates and variation in child care prices and productivities as determinants of across-household heterogeneity in second earner labour supply. We find that individual taxation is welfare-superior to joint taxation in models that match the data on empirical wage distributions, on grounds of both efficiency and equity. The main reason for this is the inequity created by income splitting in the presence of a high degree of inequality in the primary earner wage distribution.

JEL-Codes: J220, H210, H240, H310, D130.

Keywords: optimal taxation, labour supply, child care, household production, inequality.

*Patricia Apps**
Faculty of Law
University of Sydney
Australia – Sydney, NSW 2006
patricia.apps@sydney-edu.au

Ray Rees
Professor Emeritus
ray.rees@lrz.uni-muenchen.de

*corresponding author

March 19, 2016

We are very grateful to Vidar Christiansen for many helpful comments on an earlier draft of this paper, as well as to participants in the 2016 CESifo Public Economics Conference, and also to Yuri Andrienko for excellent research assistance. The work was supported under the Australian Research Council's Discovery Project funding scheme (Project ID: DP120104115).

1 Introduction

In the title of a short paper published in 1977, Harvey Rosen posed the question: "Is it time to abandon joint filing?" The case for doing so rested on the observation that joint filing, or income-splitting, equalises the marginal tax rates of a two-earner couple, while the empirical evidence¹ suggested that female workers had significantly higher compensated labour supply elasticities than did prime-age male workers. So, a straightforward application of the Ramsey principle argues for individual income as the tax base with a lower marginal tax rate on female workers at any given income level. Now, almost forty years on, income splitting still seems to be firmly enshrined in the personal income tax systems of three of the world's largest economies, France, Germany and the USA, with no obvious indications of the likelihood of change.²

A strong intuition in support of joint taxation, held by many economists and non-economists alike, is that, after controlling for demographics, a household's standard of living is strictly increasing with its total income from market labour supply. This suggests that a move from joint to individual income as the tax base could have adverse equity effects because, in a tax system with marginal tax rates increasing in income, such a change can result in two households with the same aggregate income paying different amounts of tax, or even one with a higher joint income paying less tax, depending on the relative incomes of the primary and second earner.³ This would seem to violate the principles of both vertical and horizontal equity, and so provides a basis for the argument that joint taxation could be welfare superior to individual taxation.

A starting point for a contrary argument has been succinctly expressed by Feldstein and Feenberg:⁴

The current [US income taxation] system is considered unfair because it imposes the same tax burden on a married couple with one earner as it does on a two-earner couple with the same income. The two-earner couple will in general have more total hours of work and less of the untaxed home services of the second earner.

Thus, a switch to individual taxation could actually yield gains in equity as well as efficiency. Individual taxation at progressive marginal rates allows the value of household production to be taxed indirectly, something not possible under joint taxation. In effect, by eliminating income splitting for high wage couples with low second earner labour supply and therefore high household

¹See Heckman and Killingsworth (1986) for a survey of the literature of that period. For further studies see LaLumia (2008) and Steiner and Wrohlich (2004, 2008).

²Forms of income splitting vary across countries. The US is one of the few countries with full income splitting. Others have partial income splitting or quasi-joint taxation systems due to the withdrawal of family payments or tax credits on joint income. For a comparative analysis of the Australian, German, UK and US systems, see Apps and Rees (2009), Ch. 6.

³We define primary and second earners simply in terms of who has the higher earned income, rather than in terms of gender. The second earner typically also has the lower wage rate, though this need not invariably be the case. However, the average wage rate of second earners is certainly below that of primary earners. In OECD countries, typically around 80% of second earners are female.

⁴From Feldstein and Feenberg (1996).

production, individual taxation partially compensates for the untaxed status of domestic output. In general terms, the issue would seem to be the extent of the gains in equity and efficiency in moving from joint to individual taxation in an economy where household production is a significant form of time use⁵ and households differ widely in their second earner labour supplies, and therefore values of domestic output.

In this paper we explore this issue in depth, by comparing two optimal piecewise linear tax systems for a given population of heterogeneous households, the first with joint, the second with individual incomes as the tax base. We then carry out numerical simulations for the two optimal tax systems and compare the achieved values of social welfare under each. The analysis provides some important insights into why individual taxation is welfare superior, that go beyond the observation made by Feldstein and Feenberg, important though that is.

We find that when we take a realistic view of the empirical wage distributions of primary and second earners, individual taxation is consistently welfare superior to joint taxation and that there is no equity-efficiency trade off in moving between them. The reason for this goes beyond recognition of the untaxed status of household production. The effect of income splitting is essentially to restrict the contribution high wage primary earners make to public expenditure and, under the optimal piecewise linear joint tax, to shift much of the tax burden on to lower wage households (in particular, on to lower wage second earners), whose real living standard is lower, even if comparisons of household earned income suggest the contrary.⁶ In other words, it is placing the tax burden disproportionately on the lower to middle segment of the wage distribution. Then, ending the advantage of income splitting to high income households and decoupling the labour supply elasticities of primary and second earners leads to a more progressive tax system in which equity gains reinforce the efficiency gains of the move to individual taxation.

This analysis of taxing couples under piecewise linear tax systems is new to the tax theory literature,⁷ which up until now has focused either on linear taxation,⁸ tax reform, or on nonlinear taxation⁹ in the tradition of Mirrlees (1971). One reason for our approach here is that in reality almost all tax systems are piecewise linear, and the conditions that characterise the optimal tax rates, as well as their intuitive interpretations, are different to those derived from the mechanism design approach, where incentive compatibility constraints and the implementation of a separating equilibrium across wage types are central to the

⁵For earlier analyses of income taxation with household production see Alesina et al. (2011), Apps and Rees (1988), (1999), Boskin (1995), and Sandmo (1990).

⁶The considerable growth in individual wage and earnings inequality that has taken place since Rosen's paper was published has very much strengthened this point. For more on this in the framework of the present paper see Andrienko et al. (2016).

⁷The literature on optimal piecewise linear income taxation for single person households is also not large. See Sheshinski (1989), Dahlby (1998), (2008), Slemrod et al (1994), and Apps et al (2014).

⁸As in the seminal paper by Boskin and Sheshinski (1983).

⁹See Apps and Rees (2009) for a literature survey.

analysis.

A piecewise linear approach also allows us to use a generalised model of the household in which the specification of a household's type can be far richer than that used in optimal nonlinear taxation models, where the restrictions on the number of dimensions of private information that can be handled severely limits the type specification and general structure of the household model. In this paper we present a new model of the two-earner household in which a household's type depends not only on the wage rates of the two earners but also on the price and quality it faces for the bought in input into household production - represented canonically by child care - and by its own productivity in child care, as determined by its human and physical capital.¹⁰ We define child care broadly, to denote not just physically looking after the child, but also to include all the activities that contribute to the child's welfare and development of human capital. Thus the output of child care can be thought of in terms of child outcomes or, in Gary Becker's terms, of child quality resulting from the investment the household makes in the child.

The paper is set out as follows. In the next section we present the structural household model that provides the analytical basis for the indirect utility and labour supply functions¹¹ used in the tax analysis, and for the later numerical simulations. In section 3 we define the tax systems and carry out the optimal tax analysis for joint and individual taxation respectively. In Section 4 we present the results of the numerical analysis of the optimal tax systems, for two alternative empirical specifications of the household model, and we show that individual taxation is consistently welfare superior under assumptions on productivities and prices that can generate the data on household labour supplies. Section 5 concludes.

2 The Household Model

Households are assumed to consist of couples, all with the same number of children, normalised at one.¹² The primary earner divides his time between market work and leisure, while the second earner allocates her time to market work and to the household production of child care.¹³

¹⁰This is supported by studies that find child outcomes improve with maternal human capital and that parental investment in child care and education rises with family resources. See, for example, Lundberg et al. (2014). For a survey see Almond and Currie (2011).

¹¹Detailed derivations of these are given in Appendix A.

¹²In that case, we are ruling out variations in the number of children as being the main determinant of heterogeneity in second earner labour supply. This is consistent with the empirical evidence: see Apps and Rees (2009). It also implies that we are excluding from the tax analysis single person households and childless couples. This is essentially on the grounds of simplicity and is the subject of further work. For the time being, note that all but a small proportion of the entire optimal tax literature is concerned with singles.

¹³Nothing would be gained by having both parents consume leisure and contribute to household production. Although that would be more realistic, we think the assumption made here captures the salient aspects of reality - the differing margins of substitution facing primary and second earners - while keeping the model simple.

The productivity of the second earner’s time input to child care is assumed to vary randomly across households, with a distribution whose mean value shifts upward with her wage. This reflects differences in human and physical capital. There is in addition a bought-in child care time input. The price of this input *at any given quality* varies randomly across households, while increases in quality shift the mean of this distribution of prices upward. The household chooses its optimal quality level given the market-determined relation between quality and price that it faces. The realisations of the random variables determining productivities and prices are known when decisions are taken and so there is no uncertainty, they are there to generate across-household heterogeneity.

The motivation for this emphasis on variation in the price of child care of a given quality is based upon everyday observation. In media articles and social surveys parents report that the main obstacle to second earner labour supply is the problem of finding child care of an acceptable quality and price. Costs of bought in child care vary not only with its quality or type, or mix of types,¹⁴ but also with location, age of children and other household characteristics. Moreover, households commonly report that net of taxes and other costs of going out to work, child care expenditure can swallow a large part, if not all or even more than all, of the second earner’s income.

An obvious motivation for working even when there is a negative net return is that the second earner is investing in maintaining her work-related human capital over the period in which the demand for child care is strongest, so as to be in a better labour market situation when that demand falls substantially. This cannot be drawn upon formally however in a static model of the type developed here.¹⁵ An alternative is that the productivity/quality of bought in care is sufficiently high that its contribution to child development can offset the negative return. There is no obvious reason for the costs of bought-in contributions to the development of the child’s human capital to be constrained by the income of only the second earner.

We model these observations in the following way. There is a composite market consumption good, x , individuals face given gross wage rates w , representing their productivities in a linear aggregate production technology that produces x , and have earnings $y = wl$ from their labour supply l . In addition to the market consumption good, household utility depends on child care output, z , which is produced using the second earner’s time input, c , and a bought-in child care time input, b , according to a standard linear homogeneous, strictly quasiconcave and increasing production function

$$z = z(kc, qb) \tag{1}$$

¹⁴Child care available to a household can range from grandparents and other family members, neighbours or an *au pair*, through day-care centres and private child-minders, to highly trained tutors. Parents may choose a mix of types to achieve an overall quality at an acceptable price.

¹⁵We need a life cycle-based model of optimal taxation which will show how taxes vary as a household moves over successive phases in its "family life cycle", see Apps and Rees (2009), Ch 5.

where k and q are strictly positive measures of the productivity/quality in child care of one unit of c and b respectively. We assume that k is defined by $k = k(w_2) + \tilde{\kappa}$ with $k'(w_2) > 0$ and $\tilde{\kappa} \in [\kappa_0, \kappa_1] \subset \mathbf{R}$ a zero-mean random variable. The price of bought-in care of a given quality q is given by $p = p(q) + \tilde{\varepsilon} \geq 0$, where $\tilde{\varepsilon} \in [\varepsilon_0, \varepsilon_1] \subset \mathbf{R}$ is another zero-mean random variable and $p'(q) > 0$, $p''(q) \geq 0$.

The "type" of a household depends on its wage pair (w_1, w_2) , and its realisations of home child care productivity $\tilde{\kappa}$ and price of bought in child care $\tilde{\varepsilon}$. A household's type is therefore defined by the 4-tuple $(w_1, w_2, \tilde{\kappa}, \tilde{\varepsilon})$, and we let the index $h \in \mathcal{H} \subset \mathbf{R}^4$ correspond to a particular value of this, with \mathcal{H} the set of 4-tuples. Thus, in this model, at any given primary earner wage rate w_{1h} , across-household heterogeneity is driven by variations in p_h , w_{2h} and k_h .

The household utility function is given by

$$u_h = x_h - \hat{u}_1(l_{1h}) + \hat{u}_2(z(k_h c_h, q_h b_h)) \quad h \in \mathcal{H} \quad (2)$$

The $\hat{u}_1(\cdot)$ function is a strictly increasing and strictly convex function of the primary earner's labour supply l_{1h} , representing the standard trade-off between work and leisure, while $\hat{u}_2(\cdot)$ is a strictly increasing and strictly concave function of z_h and therefore of c_h, b_h .

For the second earner, the time spent in market work and child care must sum to the total time endowment,¹⁶ normalised at 1, and so we have

$$c_h + l_{2h} = 1 \quad h \in \mathcal{H} \quad (3)$$

where l_{2h} is second earner market labour supply.

There is however a further important time constraint: Although second earner time and bought in child care may not be perfect substitutes as inputs in producing child care, realistically it is the case that every hour the second earner spends at work requires an hour of child care, in which case

$$b_h = l_{2h} \quad \forall h \in \mathcal{H} \quad (4)$$

In the absence of taxation, the household budget constraint is then

$$x_h \leq w_{1h} l_{1h} + [w_{2h} - p_h(q_h)] l_{2h} \quad h \in \mathcal{H} \quad (5)$$

As this budget constraint shows, we can view the price of child care as in effect a tax on the second earner's market labour supply¹⁷ and, at a given choice of quality, variations in w_{2h} and p_h have equal but opposite effects.¹⁸

¹⁶For the primary earner, it is sufficient to assume that the convexity of $\hat{u}_1(\cdot)$ is such that $l_{1h} < 1$ at all equilibria.

¹⁷Rather than a fixed cost of working. See also Attanasio et al. (2008) for an empirical study of the effects of child care prices on female labour supply.

¹⁸Though it should be noted that empirically the possible range of variation of w_{2h} - from minimum wage to something typically below the primary earner's wage - is much narrower than that of p_h , which can range from zero to something greater than the primary wage. For that reason, in the numerical analysis in Section 4, we consider the effect of variations in p_h rather than in w_{2h} .

2.1 Household equilibrium

Using the above time constraints, the household's problem can be written as

$$\max_{x_h, l_{1h}, q_h} u_h = x_h - \hat{u}_1(l_{1h}) + \hat{u}_2(z(k_h(1 - l_{2h}), q_h l_{2h})) \quad h \in \mathcal{H} \quad (6)$$

subject to the budget constraint in (5), and $k_h = k(w_{2h}) + \kappa_h > 0$, $p_h = p(q_h) + \varepsilon_h > 0$, with κ_h and ε_h the household's realisations of $\tilde{\kappa}$ and $\tilde{\varepsilon}$ respectively.

The first order condition¹⁹ determining l_{2h}^* is rather more interesting than the conventional expressions determining second earner labour supply. We can write this as

$$MVP_c - MVP_b = w_{2h} - p_h = w_{2h} - (p(q_h^*) + \varepsilon_h) \quad (7)$$

where MVP_b , MVP_c denote the household marginal value products of bought in and domestic child care respectively, in terms of the numeraire consumption, at the household equilibrium. The right hand side of this equation is the household's net marginal return to a unit of market labour supply l_{2h} , and therefore the marginal opportunity cost of domestic child care, given the bought in child care quality q_h^* . The left hand side represents the difference in marginal value products of domestic and bought in care respectively. If we ignore bought in care the condition would be the standard $MVP_c = w_{2h}$. However, since an hour of l_{2h} requires an hour of b_h , the return to market labour supply net of the cost of bought in care will be equated to the difference between marginal value products of the two types of child care.

It is possible to have the case in which the second earner's net return to market labour supply is negative, if in equilibrium the bought in child care input has a higher marginal value product than that of the second earner. This would be true if the bought in child care supplied a sufficiently large element of education or human capital formation.

Given the main concern of this paper, the welfare comparison of alternative tax systems, we interpret the results of the comparative statics analysis of this model²⁰ not as showing how a given "representative household" would respond to a change in the exogenous variables, but rather as suggesting how second earner labour supply and child care quality *vary across households* as we move through the distributions of bought in child care prices, wage rates and child care productivities. These results confirm that we would not expect a clear, positive relationship between household income, on the one hand, and the achieved utility level of the household on the other. This is because the variations in child care prices and productivities generate wide variations in second earner labour supply at any given wage pair, with reductions in this labour supply being associated with increases in the value of the output of child care that may at least compensate for the loss of market income.²¹ Furthermore, given that child care

¹⁹See Appendix A for the full set of first order conditions.

²⁰Given in full in Appendix A.

²¹The model therefore rationalises the statement made by Feldstein and Feenberg quoted in the Introduction.

is a normal good, increases in the primary earner wage have a positive income effect on demand for it, and may therefore induce an increase in second earner child care input, so reducing second earner labour supply. Bought-in child care may then fall but its quality will tend to increase. These comparative statics effects are brought out clearly in the numerical analysis of Section 4 below. First we turn to the optimal tax analysis.

3 Tax Analysis

As Chetty (2009) argues, whatever might be the underlying structural household model, an optimal tax system can at a general level be characterised by a small number of "sufficient statistics", essentially deriving from the joint distribution of the income measure used as the tax base and marginal social utilities of income, and the distribution of the derivatives or elasticities of earnings with respect to tax rates and the lump sum.

In the optimal tax analysis, the key relationships are the social welfare function and the households' indirect utility and earned income functions and their derivatives with respect to the tax parameters. We denote household indirect utility functions by $v(\cdot)$ and individual earnings functions by $y_i(\cdot)$, $i = 1, 2$. In Appendix A we present the details of the derivation of these functions and their properties, based on the household model presented in the previous section. Here we simply assume:

The functions $v(\zeta; w_1, w_2)$, $y_i(\zeta; w_i)$, where ζ , to be specified, denotes a vector of tax variables, are increasing in w_i , $i = 1, 2$, and continuously differentiable in all their arguments.

3.1 Tax systems

The tax system pays households a uniform lump sum²² funded by revenue from taxes on the labour incomes of the two earners across all households, and offers a schedule of marginal tax rates. For any given tax system, households choose their optimal labour supplies and the resulting earnings form the basis for their allocation to a tax bracket - a subset of households with a given marginal tax rate.

As well as the issue of the choice of tax base as between joint and individual incomes, also central is the structure of the rate scale, in particular whether the marginal tax rates applying to successive income brackets should be strictly increasing, or whether over at least some income ranges they should be decreasing. We refer to these as the "convex" and "nonconvex" cases respectively, to describe the types of budget sets in the gross income-net income/consumption plane to which they give rise. For the purposes of this paper we focus on the convex case of a two-bracket piecewise linear system.²³

²²This could be thought of as a standard child benefit.

²³Apps et al. (2014) show that for wage distributions such as those currently prevailing in many OECD countries convex systems are highly likely to be welfare optimal. The numerical

By individual taxation we mean the case in which the two earners' incomes are taxed separately but according to the same tax schedule. This is in contrast to the case in which separate optimal tax schedules are found for primary and second earners respectively, so-called selective taxation. The main reason for constraining the rate schedules to be identical under individual taxation is that in practice, piecewise linear tax systems that are not joint are in fact overwhelmingly of the individual rather than selective kind.²⁴ Moreover, if individual taxation yields higher social welfare than joint taxation under realistic assumptions, this result applies *a fortiori* to selective taxation, since removing the constraint that tax schedules must be identical cannot reduce the maximised value of social welfare and would be expected to increase it. It is not difficult to extend the results of this paper to the selective taxation case, at the cost however of some increase in notational complexity, since there is a range of possibilities concerning how the separate optimal bracket points relate to each other.²⁵

3.2 Joint taxation

There is a two-bracket piecewise linear tax on total household labour earnings, defined by $\zeta_1 = (\alpha, \tau_1, \tau_2, \eta)$, where α is the uniform lump sum paid to every household, τ_1, τ_2 are the marginal tax rates in the lower and upper brackets of the tax schedules, and η is the value of *joint earnings* defining the bracket limit. The household tax function is $T(y_{1h}, y_{2h}) \equiv T(y_h)$, with $y_h = \sum_{i=1}^2 y_{ih}$, where households are indexed by $h \in \mathcal{H}$ and y_{ih} is the labour income of individual $i = 1, 2$ in household h , with by definition $y_{2h} \leq y_{1h}$. This function is given by:

$$T(y_h) = -\alpha + \tau_1 y_h \quad y_h \leq \eta \quad (8)$$

$$T(y_h) = -\alpha + \tau_2 y_h + (\tau_1 - \tau_2)\eta \quad y_h > \eta \quad (9)$$

Given that all households face this identical budget constraint, it is straightforward to show that the optimal income y_h^* for any one household must be in one of three possible subsets,²⁶ which give a partition $\{\mathcal{H}_0, \mathcal{H}_1, \mathcal{H}_2\}$ of the index set \mathcal{H} defined as follows:

$$\mathcal{H}_0 = \{ h \mid 0 \leq y_h^* < \eta \} \quad (10)$$

analysis in the following section finds a global optimum and does not constrain the search to convex systems.

²⁴At the same time, it is possible to find examples of tax systems that contain selective elements. For example in Australia, a portion of family benefits is withdrawn on the basis of the second earner's income. In Germany and the US, contributions to social security, which are effectively part of the tax system, vary with the income of the second earner. See Apps and Rees (2009), Ch 6, and Feldstein and Feenberg (1996). In this paper we focus on the formal tax system, leaving the issue of implicit modifications to it created by social benefit payments and withdrawal rates for future work.

²⁵The same is true of assuming an arbitrary number $m \geq 2$ of tax brackets. For an extension of optimum piecewise linear taxation to this case see Andrienko et al. (2016).

²⁶See Appendix A for details.

$$\mathcal{H}_1 = \{ h \mid y_h^* = \eta \} \quad (11)$$

$$\mathcal{H}_2 = \{ h \mid y_h^* > \eta \} \quad (12)$$

A household's optimum income may be either in the lower tax bracket, at the kink in the budget constraint defined by the bracket limit η , or in the upper tax bracket. In all of what follows we assume that we are dealing with tax systems in which each of these subsets is non-empty. Total household gross and net income are increasing continuously as we move from \mathcal{H}_0 to \mathcal{H}_1 and from \mathcal{H}_1 to \mathcal{H}_2 , while they are both constant in \mathcal{H}_1 . Important points to note are that:

- τ_1 is a marginal tax rate for $h \in \mathcal{H}_0$ but defines an intra-marginal, non-distortionary tax for $h \in \mathcal{H}_1 \cup \mathcal{H}_2$
- A marginal increase in η has no effect for $h \in \mathcal{H}_0$, yields a net welfare gain for almost all $h \in \mathcal{H}_1$, and yields a lump sum income gain proportional to $(\tau_2 - \tau_1)$ for $h \in \mathcal{H}_2$ (recall we assume that $\tau_2 > \tau_1$)
- In effect, for purposes of the tax analysis the household can be treated as a single individual, given that at each level of household income individual earnings are chosen so as to equate marginal effort costs, i.e. to minimise the cost of supplying that level of aggregate income²⁷

3.2.1 Optimal tax analysis

We define dF as the marginal density of household type h . The planner solves

$$\max_{\alpha, \tau_1, \tau_2, \eta} W = \int_{h \in \mathcal{H}} S(v_h) dF \quad (13)$$

subject to the public sector budget constraint²⁸

$$\tau_1 \left[\int_{h \in \mathcal{H}_0} y_h dF + \eta \int_{h \in \mathcal{H}_1} dF \right] + \int_{h \in \mathcal{H}_2} [\tau_2 y_h + (\tau_1 - \tau_2) \eta] dF \geq \alpha \quad (14)$$

where $S(\cdot)$ is a strictly concave and increasing function expressing the planner's distributional preferences over household utilities. From the first order conditions characterising the optimal tax variables²⁹ we can derive:

Proposition 1: The optimal tax system $(\alpha, \tau_1, \tau_2, \eta)$ satisfies the conditions:

$$\int_{\mathcal{H}} (\sigma_h - 1) dF = 0 \quad (15)$$

²⁷Again the details are in Appendix A.

²⁸We assume the aim of taxation is purely redistributive. Adding a non-zero revenue requirement would make no essential qualitative difference to the results.

²⁹Of course, exactly which households will be in which subsets is determined at the optimum, and depends on the values of the tax parameters. The following discussion characterises the optimal solution *given* the allocation of households to subsets that obtains at this optimum. As our later numerical analysis has shown us, it is not a trivial computational task to solve this model.

$$\tau_1^* = \frac{\int_{\mathcal{H}_0} (\sigma_h - 1) y_h^* dF + \eta^* \int_{\mathcal{H}_1 \cup \mathcal{H}_2} (\sigma_h - 1) dF}{\int_{\mathcal{H}_0} (\partial y_h / \partial \tau_1) dF} \quad (16)$$

$$\tau_2^* = \frac{\int_{\mathcal{H}_2} (\sigma_h - 1) (y_h^* - \eta^*) dF}{\int_{\mathcal{H}_2} (\partial y_h / \partial \tau_2) dF} \quad (17)$$

$$\int_{\mathcal{H}_1} \left\{ \sigma_h [(1 - \tau_1) - \frac{\partial \psi}{\partial y_h}] + \tau_1^* \right\} dF = -(\tau_2^* - \tau_1^*) \int_{\mathcal{H}_2} (\sigma_h - 1) dF \quad (18)$$

where y_h^* denotes household income at the optimum and σ_h is the marginal social utility of income to household h .

Condition (15) follows from the quasilinearity of the utility functions and is familiar from linear tax theory³⁰: Denoting the shadow price of the government budget constraint by λ , $\sigma_h \equiv S'(v_h)/\lambda$ is the marginal social utility of income to household h in terms of the numeraire, consumption, and so the optimal lump sum α equalises the average of the marginal social utilities of household income across the population to the marginal cost of the lump sum, which is 1.

The strict concavity of $S(\cdot)$ implies that σ_h is strictly decreasing in v_h . In the standard income tax model, with v_h and y_h co-monotonic, the lower tax bracket would contain not only the lower incomes but also the lower utilities. But because, as shown in the previous section, the household model of this paper does not imply this co-monotonicity, the lower tax bracket may contain households with higher utility than households that are assigned, on the basis of joint income, to the higher bracket. This is of course simply a way of expressing the income-splitting advantage given to households with high primary and low second incomes under a joint taxation system.

In the two conditions corresponding to the tax rates τ_1^* , τ_2^* , the denominators are the frequency-weighted sums of the compensated derivatives of earnings with respect to the tax rates over the relevant subsets, and so give a measure of the marginal deadweight loss of the tax rate at the optimum, the efficiency cost of the tax, for households in the indicated subsets. The numerators give the equity effects.

The two terms in the numerator of (16) correspond to the two ways in which the lower bracket tax rate affects the contributions households make to funding the lump sum payment α . Given their optimal household earnings y_h^* , the first term aggregates the effect of a marginal tax rate change on utility net of its marginal contribution to tax revenue over subset \mathcal{H}_0 . The second term reflects the fact that the lower bracket tax rate is effectively a lump sum tax on income earned by the two higher income brackets, \mathcal{H}_1 and \mathcal{H}_2 , since a change in this tax rate has only an intramarginal effect, changing the tax they pay at a rate given by η , while leaving their (compensated) labour supply unchanged.

Only the first of these two effects is present in the condition (17) corresponding to the higher tax rate. The portion of the income of the households in the higher tax bracket that is taxed at the rate τ_2^* is $(y_h^* - \eta^*)$, and so this weights the effect on social welfare net of the effect on tax revenue. Note that, unlike

³⁰See Sheshinski (1972).

the case of linear income taxation, these numerator terms are not covariances, since the mean of σ_h over each of the subsets is not 1. They are commonly referred to as "distributional characteristics".

Comparing the numerator terms in (16) and (17) shows that each contains the term $\eta^* \int_{\mathcal{H}_2} (\sigma_h - 1) dF$, but with opposite signs. This suggests that the greater the contribution of the lump sum tax on upper income bracket households arising from the tax rate τ_1^* , the smaller is the tax rate τ_2^* , and so the smaller is the distortionary effect on labour supplies in this bracket, other things being equal.³¹ Note also that, other things equal, the more sharply y_h^* increases across households in the upper bracket the greater will be the tax rate τ_2 , implying that tax rates are sensitive to growing inequality in the form of sharp increases in top incomes.³²

Condition (18) corresponding to the optimal bracket limit η^* , has the following interpretation. The left hand side represents the marginal social benefit of a relaxation of the bracket limit. This consists first of all of the gain to all those households that are effectively constrained at η^* , in the sense that they are prepared to increase earnings if these are taxed at τ_1^* but not at τ_2^* - the return to additional labour supply at τ_1^* , but not τ_2^* , exceeds its marginal utility cost.³³ The first term in brackets on the left hand side is the net marginal benefit to these consumers, weighted by their marginal social utilities of income. The second term is the rate at which tax revenue increases given the increase in gross income resulting from the relaxation of the bracket limit.

The right hand side gives the marginal social cost of the relaxation. Since $(\tau_2^* - \tau_1^*) > 0$ by assumption, all households $h \in \mathcal{H}_2$ receive a lump sum income increase at this rate and this is weighted by the deviation of the marginal social utility of income of these households from the average. As long as the sum of these deviations, weighted by the frequencies of the household types, is negative, the marginal cost of the bracket limit increase is a worsening in the equity of the income distribution. The condition then trades off the social value of the gain to households in \mathcal{H}_1 against the social cost of making households in \mathcal{H}_2 better off.

3.3 Individual taxation

There is a two-bracket piecewise linear tax system now applied to individual labour earnings, defined by $\zeta_2 = (a, t_1, t_2, y)$, where a is again a uniform lump sum paid to every household, t_1, t_2 are the marginal tax rates in the lower and upper brackets, and y is the value of *individual earnings* defining the bracket. Thus the individual tax function $\hat{T}(y_{ih})$ is defined by:

$$\hat{T}(y_{ih}) = t_1 y_{ih} \quad y_{ih} \leq y \tag{19}$$

³¹It is this tradeoff which can lead to the nonconvex case in which the upper bracket tax rate is optimally lower than that in the lower bracket. For further discussion see Apps et al. (2014).

³²See Andrienko et al. (2014) for more on this.

³³For the details again see Appendix A.

$$\hat{T}(y_{ih}) = t_2 y_{ih} + (t_1 - t_2)y \quad y_{ih} > y \quad h \in \mathcal{H} \quad (20)$$

and the household tax function is $T(y_{1h}, y_{2h}) \equiv -a + \sum_{i=1}^2 \hat{T}(y_{ih})$. Given that, by definition, $y_{2h}^* \leq y_{1h}^*$ for every household, and that under individual taxation everyone faces the same tax schedule, it is easy to see that there are now six possible subsets of households which form a partition $\{H_0, H_1, \dots, H_5\}$ of the index set \mathcal{H} , defined by

$$H_0 = \{ h \mid 0 \leq y_{ih}^* < y, i = 1, 2 \} \quad (21)$$

$$H_1 = \{ h \mid y_{2h}^* < y = y_{1h}^* \} \quad (22)$$

$$H_2 = \{ h \mid y_{ih}^* = y, i = 1, 2 \} \quad (23)$$

$$H_3 = \{ h \mid y_{2h}^* < y < y_{1h}^* \} \quad (24)$$

$$H_4 = \{ h \mid y_{2h}^* = y < y_{1h}^* \} \quad (25)$$

$$H_5 = \{ h \mid y_{ih}^* > y, i = 1, 2 \} \quad (26)$$

In $H_0 - H_2$ both individuals pay the lower tax rate, in H_3 and H_4 the primary earner alone pays the higher tax rate, and in H_5 both pay the higher tax rate. The difference to the joint taxation case, apart from the obviously finer partition based on individual reactions to the tax system, is that only in subsets H_0 and H_5 , where both the individuals in the household are in the *interior* of the same tax bracket, will the marginal rates of substitution between consumption and labour supply of primary and second earners be equalised. In all other cases they will not in general be the same, as each earner chooses their individually optimal earnings levels.

3.3.1 Optimal tax analysis

To shorten notation denote the subset $H_i \cup H_j$ by H_{ij} , and $H_i \cup H_j \cup H_k$ by H_{ijk} , $i, j, k = 0, \dots, 5$, $i \neq j$, $i, j \neq k$. The planner solves

$$\max_{a, t_1, t_2, y} W = \int_{\mathcal{H}} S(v_h) dF$$

subject now to the public sector budget constraint

$$\int_{H_{012}} t_1 y_h dF + \int_{H_{34}} [t_2 y_{1h} + t_1 y_{2h} + (t_1 - t_2)y] dF + \int_{H_5} [t_2 y_h + 2(t_1 - t_2)y] dF \geq a \quad (27)$$

where again $y_h = \sum_{i=1}^2 y_{ih}$.

In what follows it will be useful to denote by μ_{ih} the value of a relaxation of the bracket limit to an individual at the kink in the budget constraint.³⁴ Also, to shorten notation we denote $\sigma_h - 1$ by δ_h . Then $\delta_h > (<)0$ according as household h is relatively worse (better) off *in utility terms* than the subset of households for which $\sigma_h = 1$.

³⁴See Appendix A.

From the first order conditions for an optimal solution³⁵ we derive:

Proposition 2: The optimal tax system (a, t_1, t_2, y) is characterised by the conditions:

$$\int_{\mathcal{H}} \delta_h dF = 0 \quad (28)$$

$$t_1^* = \frac{\int_{H_0} \delta_h y_h^* dF + \int_{H_{13}} \delta_h (y_{2h}^* + y^*) dF + 2y^* \int_{H_{245}} \delta_h dF}{\int_{H_0} \partial y_{1h} / \partial t_1 dF + \int_{H_{013}} \partial y_{2h} / \partial t_1 dF} \quad (29)$$

$$t_2^* = \frac{\int_{H_{345}} \delta_h (y_{1h}^* - y^*) dF + \int_{H_5} \delta_h (y_{2h}^* - y^*) dF}{\int_{H_{345}} \partial y_{1h} / \partial t_2 dF + \int_{H_5} \partial y_{2h} / \partial t_2 dF} \quad (30)$$

$$\begin{aligned} \int_{H_{12}} (\sigma_h \mu_{1h} + t_1) dF + \int_{H_{24}} (\sigma_h \mu_{2h} + t_1) dF = \\ -(t_2 - t_1) \left[\int_{H_{34}} \delta_h dF + 2 \int_{H_5} \delta_h dF \right] \end{aligned} \quad (31)$$

The first condition, since it involves the entire population, is exactly as for joint taxation. The remaining three conditions have basically the same interpretation as before, but of course the relevant sums are now over subsets of individuals reflecting the partition defined in (21)-(26).

We can use this tax analysis to give us an idea of how the switch from joint to individual taxation could affect welfare. First note that the denominator in the above expression for t_1^* contains more lower wage second earners, while that for t_2^* more high wage primary earners, than the corresponding expressions for τ_1^* and τ_2^* respectively. Other things equal, this would lead us to expect a greater difference between the two tax rates, or higher marginal rate progressivity, in the case of individual taxation, given the stylised facts on the relative elasticities of primary and second earners.³⁶

A similar point can be made in respect of the numerators of the expressions for the upper bracket tax rates in the two cases, which represent the equity terms. In the expression for τ_2^* we have the term $\int_{\mathcal{H}_2} \delta_h (y_h^* - \eta^*) dF$, while for t_2^* we have $\int_{H_{345}} \delta_h (y_{1h}^* - y^*) dF + \int_{H_5} \delta_h (y_{2h}^* - y^*) dF$. The subset \mathcal{H}_2 will contain lower wage two-earner households with close to average welfare weights and therefore δ_h -values close to zero, while the differences $(y_h^* - \eta^*)$ for households with strongly negative δ_h -values will be diminished by the fact that the values of joint income y_h^* will be relatively lower for households with little or no income from the second earner. In contrast, the subsets H_{345} and H_5 contain only the highest earning primary and second earners respectively, with (as shown by the empirical wage distributions in Figure 1 below) very large differences between

³⁵Again, exactly which households will be in which subsets is determined at the optimum, and depends on the values of the tax parameters.

³⁶We should note that the empirical estimates of elasticities are gender-based - female labour supply elasticities are much higher than male - whereas the distinction here between primary and second earners is on the basis of earned income rather than gender. We would argue however that the high female elasticities are based on role rather than gender. As pointed out earlier, it is also still the case that the large majority of second earners are women.

their incomes and the bracket limit. This gives an additional reason to expect that the individual tax system will be very much more marginal rate progressive than the joint system.

These remarks are confirmed by the numerical analysis of the next section, where we highlight this difference in marginal rate progressivity as a major reason for the welfare superiority of individual over joint taxation systems, given the characteristics of the empirical wage distribution.

4 Numerical analysis

A calibrated version of a theoretical model cannot of course be a blueprint for tax reform in any actual economy. However, since Mirrlees (1971), Stern (1976) and Tuomala (1984), there has been a tradition in optimal tax theory of using plausibly calibrated models to clarify the qualitative implications of particular tax models and provide insights which, hopefully, can be followed up in richer empirical models. Following this tradition, in this section we show the sensitivity of optimal tax parameters to the relationship between the second earner's wage rate and price of bought in child care, and to the shapes of the overall distributions of primary and second earner wage rates.

We present simulation results for the optimal parameters of joint and individual taxation for two models, labeled Model 1 and Model 2, that are nested within a general empirical specification of the theoretical model presented in Section 2. The models are calibrated by drawing on household survey data for a sample of two-parent families.

As in most household surveys, each record contains information on hours of work and earnings which allows the computation of a gross wage as a measure of the productivity of market work. However, data on home productivity, k_h , are missing. Some assumption is therefore necessary. Information on quality of bought in child care, q_h , is limited to care type (family relative, child care centre, etc.) with a reported price, p_h , that can vary apparently randomly across households for a given care type.³⁷ Again, an assumption is needed.

In both models we follow the convention of defining the numeraire, consumption x , as a Hicksian composite good priced at one. In Model 1 we make assumptions, detailed below, that ensure that the productivities of the second earner's market labour supply and time spent on child care are symmetric. In addition, the quality/productivity of bought in care, while a different input from own child care time, is assumed to match that of the second earner. As we pointed out in the Introduction, these assumptions are supported by studies that find child outcomes improve with maternal human capital and that parental investment in quality child care and education increases with family resources.

³⁷While household survey data typically report type and price of bought in child care, reliable values for q_h and p_h are difficult to construct because many parents report a mix of care types. The mix frequently includes a grandmother's time, with its reported price typically zero, which of course may not reflect its true opportunity cost.

In Model 2 the productivities of market and domestic work are asymmetric. Consistently with the implicit assumptions of the standard empirical family labour supply model, the productivity of home time is set to a constant across all households.³⁸ With the price of consumption as numeraire set to one, the price of a "quality unit" of bought in care is set to one and quantity is measured by expenditure.³⁹

Section 4.1 discusses the data and Section 4.2 presents a general empirical specification of the theoretical model within which the two models are nested. Sections 4.3 and 4.4 report the results for the optimal values of piecewise linear joint and individual tax variables for each model under a range of parameter values.

4.1 Data

The sample of two-parent families is drawn from the Australian Bureau of Statistics (ABS) 2010 Survey of Income and Housing (SIH).⁴⁰ We first construct a reference wage distribution based on the data for the earnings and hours of work of the primary earner. The wage in each percentile is calculated as average gross hourly earnings, with hours smoothed across the distribution.⁴¹ A second profile representing the average second earner wage at each primary wage percentile is constructed from the data on second earnings and hours.⁴² Both profiles are plotted in Figure 1.

Figure 1 about here

The smoothed profile of primary earner hours is relatively flat, rising only slightly across the primary wage distribution, with an overall mean of 8.7 hours per day for a five day working week. The smoothed profile of second earner hours, while relatively flat, tends to rise across the middle percentiles and then decline towards the top percentiles, with an overall mean close to 4.5 hours per day. However, in contrast to primary hours, there is a high degree of heterogeneity. Over a third of second earners work part time with widely varying hours, around a third work full time and the remainder are not in the workforce.

In the simulation to follow we select parameter values that can generate heterogeneity in second earner hours consistent with the data. For this purpose we split the full sample into subsamples defined with respect to median second

³⁸For an outline of the theoretical framework of the model and survey of empirical applications, see Blundell and MaCurdy (1999), Section 7.

³⁹The limitations of the available data for estimating family labour supply models were clearly set out by Heckman (1974), who points out that these limitations make it impossible to measure the price per unit of quality of bought in care, and that measuring quality by expenditure with a normalised market price of unity is debatable but standard practice for want of anything better. Stern (1975) also notes the potential sensitivity of estimated parameters to the assumed productivity or quality of "leisure".

⁴⁰The sample is selected on the criteria that a child aged from 0 to 9 years is present and the primary earner is aged from 25 to 59 years and works at least 25 hours per week for a wage of at least \$15.00 (the minimum wage in 2010). The sample contains 1860 records.

⁴¹We apply the *lowess* method to obtain a smoothed profile.

⁴²We correct for selectivity bias based on a comparative analysis of predicted wage rates for participant and non-participant sub-samples within quintiles of the primary wage distribution.

earner hours. We label records with second earner hours below the median as "H1" households and those at or above the median, as "H2" households. Overall means of 1.4 hours for the H1 subsample and of 7.5 hours for the H2 subsample indicate the high degree of heterogeneity in second earner labour supplies across the primary wage distribution. Given percentile means of less than 10 hours for both primary and second earners we set the time constraint in the simulations to follow to 10 hours per day

Figure 2 plots the incomes of the H1 and H2 households and the distribution of their implicit expenditures on home child care time, based on the wage profiles in Figure 1 and time spent on home child care, calculated by subtracting the mean of market hours from the 10 hour time constraint.

Figure 2 about here

This figure highlights an important fact. At any given primary wage, differences in household welfare, based on the values of their consumption of both the market good and child care, can be accurately represented by differences in household market income only under some assumption that has the effect of equalising the contributions of H1 and H2 second earners to home child care. This could be achieved either by setting both to zero, or by deflating sufficiently the contribution of those in H1 households relative to those in H2 households. The figure also highlights the fact that the differences between the incomes of the two household types across the middle of the wage distribution are small relative to those between the incomes in the top few percentiles and in all the percentiles below them.

4.2 Empirical specification

We solve for the optimal tax parameters by maximising a social welfare function of the form $[\sum_{h=1}^H u_h^{1-\pi}]^{1/(1-\pi)}$, with π a measure of inequality aversion.

The quasilinear utility function introduced in Section 2 is specified as

$$u_h = x_h - (y_{1h}/w_{1h})^{\alpha_1} + z_h^{\alpha_2} \quad h \in \mathcal{H} \quad (32)$$

with $\alpha_1 = (1 + e_1)/e_1 > 1$, where e_1 is the elasticity of labour supply with respect to the net wage of the primary earner, and $\alpha_2 \in (0, 1)$.⁴³ All households are assumed to have the same preferences and therefore the same parameters in the household utility function, as is usual in optimal tax analysis.

Child care, z , is the output of a CES production function:

$$z_h = \gamma[\beta_h(k_h c_h)^{\rho_h} + (1 - \beta_h)(q_h b_h)^{\rho_h}]^{1/\rho_h} \quad h \in \mathcal{H} \quad (33)$$

where $\gamma > 0$ is a scaling factor, the parameter ρ_h determines the elasticity of substitution between the second earner's home care input and bought in child care, $1/(1 - \rho_h)$, and $\beta_h \in (0, 1)$.

⁴³The numerical analysis was also carried out for the more general case in which the utility function takes the form

$$u_h = x_h^{1-\gamma}/(1-\gamma) - l_{1h}^{1-\delta}/(1-\delta) + z_h^{1-\zeta}/(1-\zeta)$$

The results support the conclusions derived here from the simpler quasi-linear function, and are available from the authors.

We find that $e_1 = 0.1$ generates a primary hours profile across the primary wage distribution that broadly matches the data and we therefore base the simulations on this primary wage elasticity, which is probably at the higher end of empirical estimates of prime age male labour supply elasticities. In the case of the second earner we select parameter values that, for the prices and/or productivities assumed in each model, can generate heterogeneity in second earner labour supplies that broadly reflects the data.

4.3 Model 1: Symmetric market and household productivities

All households are assumed to have the same production technology, in addition to identical preferences. We can therefore drop the h subscript on ρ and β and rewrite the child care production function in (33) as

$$z_h = \gamma[\beta(k_h c_h)^\rho + (1 - \beta)(q_h b_h)^\rho]^{1/\rho} \quad h \in \mathcal{H} \quad (34)$$

where, as set out in Section 2, k and q are strictly positive measures of the productivity/quality in child care of one unit of c and b respectively, k is defined by $k = k(w_2) + \tilde{\kappa}$ and the price of bought-in care of a given quality q is given by $p = p(q) + \tilde{\varepsilon} \geq 0$.⁴⁴ As discussed previously, we assume that the quality of bought in care matches that of the second earner and therefore set $q_h = w_{2h}$.

We define p_z as the implicit price of a quality unit of child care output, z_h , which is equal to its marginal production cost, determined by the net of tax wage rates, prices of bought in care p_h and marginal productivities, the derivatives of (34). This implicit price is of course independent of output z_h given the constant returns assumption.

The first step is to construct a *pre-tax benchmark population of households* giving an optimal second earner labour supply at every wage pair of 5 hours per day (which is marginally above the after-tax overall mean of 4.5 hours) for $\tilde{\kappa} = 0$, $\tilde{\varepsilon} = 0$. For this outcome we set $k_h = w_{2h}$, $p_h = w_{2h}$ and $\beta = 0.5$. We also set $\gamma = 2$ to give $p_z = 1$ at the benchmark equilibrium.

We then perturb the benchmark labour supplies by varying k_h and p_h above and below their benchmark value,⁴⁵ w_{2h} , by the same percentage at each wage pair, solving for the household equilibria in each case. In this way we generate subsamples of households in which second earner labour supplies are respectively above and below the median, corresponding to the households facing higher/lower productivities or lower/higher prices, respectively. This provides us with a population of heterogenous households for which we then compute

⁴⁴Variation in the price of bought in child care may be due to government taxes or subsidies for child care, in addition to stochastic variation in the market price. These may be set by agencies other than the tax authority. An extension of our approach in this paper could of course be used to analyse optimal policies towards child care provision.

⁴⁵As mentioned earlier, theoretically, variations in w_{2h} and p_h have opposite but equivalent effects, but in empirical terms the range of possible variation in p_h is much larger, and so we choose this for purposes of the numerical analysis.

optimal piecewise linear tax systems based on joint and individual incomes respectively. Figure 3 illustrates this procedure for variation in p_h .

Figure 3 about here

Given the above specifications, it is easy to show that an equilibrium of the household at any wage pair that is fully symmetric with respect to bought-in and parental child care and yields an equal division of available time between market and household work must look like that shown in the figure. The marginal value product curves of b_h and c_h are identical and therefore mirror images of each other when c_h is measured from left to right and b_h measured in the reverse direction, and setting $p_h = w_{2h}$ ensures that the curves intersect at $b_h = c_h = 5$, where condition (7) of Section 2 is satisfied. Then perturbing this equilibrium by raising p_h reduces b_h , and therefore market labour supply l_{2h} , and increases c_h , while reducing p_h has the converse effect. The equilibrium division of time use between market labour supply/bought in care and parental child care corresponds to the point at which the vertical distance between the curves is equal to the difference between w_{2h} and p_h . Thus this simple, symmetric benchmark model provides us with a very convenient way of generating whole populations of heterogeneous households for varying distributions of the price of bought-in child care.

We first examine the effects on the benchmark equilibrium of variation in the productivity parameter k . We take the binary productivity distribution $k \in \{k^1, k^2\}$ with 100 H1 households with the higher productivity k^1 and 100 H2 households facing the lower productivity k^2 . The degree of variation is expressed as a proportion of the second earner wage at each percentile.

Table 1 reports results for a productivity variation of $\pm 15\%$. The table compares the optimal marginal tax rates and lump sum, the bracket point and aggregate social welfare, SW, under piecewise linear joint and individual taxation for $\pi = 0.1$ and 0.3 , and for two values of ρ .

Table 1 about here

The results show that individual taxation is superior to joint taxation and becomes increasingly superior as the degree of inequality aversion, π , rises from 0.1 to 0.3. This finding is perhaps to be expected because the variation in k implies that the second earner in the H1 household is more productive at home and faces a lower implicit price of child care, p_z , as shown in the table. The H1 households therefore have a higher level of utility but a lower level of income than the H2 household at any given wage pair. With household income negatively correlated with household utility within each wage percentile, progressive individual taxation, by placing a lower tax burden on the two earner households, will be superior in terms of equity as well as efficiency.

It could be argued however that this case unduly biases our results against joint taxation, and so in the remainder of this section we study the case in which second earner labour supply is generated by variation in the price of bought in child care. Households with higher market labour supply will also have a lower p_z , and so household income and utility will tend to be positively correlated. In that case we have the possibility that a switch from joint to individual taxation could have adverse equity effects that offset the efficiency gains. We solve for

the optimal tax systems first for two degrees and then for four degrees of price variation respectively.

With two degrees of price variation we have the binary price distribution $p \in \{p^1, p^2\}$ with 100 H1 households facing the higher price p^2 and 100 H2 households facing the lower price p^1 . Each degree of variation is again expressed as a proportion of the second earner wage at each percentile.

Tables 2 to 4 report results for, respectively, price variations of $\pm 15\%$, $\pm 25\%$ and $\pm 50\%$. The tables compare the optimal marginal tax rates and lump sum, the bracket point and aggregate social welfare, SW, under piecewise linear joint and individual taxation for $\pi = 0.1$ and 0.3 , and for two values of ρ .

Tables 2 to 4 about here

The values of ρ in the top panel of the tables are selected to generate second earner labour supplies for households H1 and H2 that broadly reflect the overall data means of 1.4 and 7.5 hours, respectively. In the second panel we reduce the value of ρ , and therefore the elasticity of substitution between home and market care, to obtain a smaller gap between the second hours of H1 and H2 households.

To generate hours that broadly match the data as the degree of price variation rises from $\pm 15\%$ in Table 2 to $\pm 50\%$ in Table 4 it is necessary to reduce ρ from 0.85 in Table 2 to 0.75 in Table 3, and then to 0.5 in Table 4. The explanation is straightforward. At a relatively low degree of price variation, such as $\pm 15\%$ in Table 1, the observed heterogeneity in second hours can only be generated by assuming home and bought in care tend to be close substitutes. In contrast, with a high degree of price variation, as in Table 4, matching the data requires a significantly lower value of ρ .

The results in all three tables show that individual taxation is consistently superior to joint taxation. Importantly, individual taxation becomes increasingly superior to joint taxation as the degree of inequality aversion, π , rises from 0.1 to 0.3.

We extend the analysis by combining two degrees of price variation. This gives four rather than two relative prices for child care in each percentile, that is, $p_h \in \{p^1, \dots, p^4\}$, which increases the total number of household records from 200 to 400. Table 5 reports the results for price variations of $\pm 25\%$ and $\pm 50\%$ above and below the benchmark price. Individual taxation remains superior to joint taxation.

Table 5 about here

The well-known disincentive effect of joint taxation resulting from the equalisation of marginal tax rates of primary and second earners is reflected in the lower median second earner labour supplies of both H1 and H2 households as compared to individual taxation.

The tables reflect a further important aspect of the results. When home and bought in care are close substitutes, small differentials in p_h between H1 and H2 households can cause large differences in bought in care and second earner labour supply, and therefore in household income. But with a high elasticity of substitution there is a significantly smaller impact on the implicit prices of child care, p_z , as indicated in the final two columns of the tables, and therefore on

the households' achieved utility levels. Thus large price differentials for bought in care between H1 and H2 households and correspondingly large differences in second earner labour supplies and household labour incomes are associated with smaller differences in household utility levels.

This effect also explains why the SW values rise with the degree of price variation across Tables 2 to 4. The utility gain from the lower price of bought in child care for H2 households is greater than the loss from the higher price for H1 households: utilities are convex and decreasing in the price p_z .

The most striking feature of the results is the change in the structure of marginal tax rates as we switch from joint to individual taxation. The top marginal tax rate under individual taxation is consistently much higher than that under joint taxation, while the bracket limit is somewhat lower. This reflects the shape of the wage profiles in Figure 1. The profiles are relatively flat up to around the 90th percentile and rise steeply thereafter - note that the bracket limit at which tax rates change is almost invariably above the 90th percentile.⁴⁶ To obtain a result in support of joint taxation due to price variation it would be necessary to introduce child care subsidies for H2 households of an order that could offset the utility effects of the dramatic rise in wage rates in the top percentiles. Such subsidies are empirically implausible.

Given the shape of the wage distribution, a progressive income tax redistributes the tax burden from the lower and middle to the top wage percentiles. The differences in top marginal tax rates under the two systems highlight the extent of the income-splitting gain to top primary wage earners under joint taxation, and the extent to which this constrains the distribution of tax burdens. Note also that the lump sums are much lower under joint taxation. Furthermore, as discussed earlier in Section 3, including second earners with relatively higher compensated elasticities in the higher tax bracket under joint taxation significantly reduces the marginal rate progressivity of the tax system. It is counter-productive in terms of both efficiency and equity to use a joint tax system to impose a higher marginal rate on second earners across the lower and middle percentiles of the wage distribution, while holding the marginal tax rate on high primary wage earners at relatively low levels.

4.4 Model 2: Asymmetric market and household productivities

Model 2 incorporates the assumptions on domestic time productivities and quality and price of bought in time adopted in the estimation of two-person family

⁴⁶The effect of the high degree of wage inequality at the top end of the distribution on the structure of piecewise linear tax systems is very significant, and is more comprehensively explored in Andrienko et al. (2016). There, a much simpler formulation of the household model allows the optimal structure of tax systems with up to 4 brackets to be analysed numerically, with results which we regard as important for tax policy. If replicated here, they would imply that the first bracket limit occurs at around the 50th percentile, the second around the 90th and the third at the 99th, with sharply increasing marginal tax rates.

labour supply models underlying much of the literature, as discussed above.⁴⁷ While the productivity of market time is given by the wage, as in the preceding model, the quality/productivity of home child care time, k_h , is implicitly assumed to be constant across all households, and so can be normalised to one. In addition, since this literature defines bought-in child care as expenditure on the market good, we denote expenditure on child care by x_{bh} and rewrite the household's budget constraint as

$$x_h + x_{bh} \leq \sum_{i=1}^2 w_{ih} l_{ih} \quad h \in \mathcal{H} \quad (35)$$

and the child care production function as

$$z_h = \gamma[\beta_h c_h^{\rho_h} + (1 - \beta_h)x_{bh}^{\rho_h}]^{1/\rho_h} \quad h \in \mathcal{H} \quad (36)$$

Thus in this model variation in the *price* of bought in child care across households, an exogenous variable, is necessarily interpreted as a variation in the *quantity* of child care, since this is measured by expenditure, an endogenous variable. In empirical terms, this could be regarded as a misspecification which leads to a form of measurement error. Be that as it may, we take this model at its face value and explore its implications for the comparison of tax systems.

The model also rules out stochastic variation in p_h across households. Heterogeneity in second earner labour supply therefore can only be driven by differences in the elasticity of substitution parameter ρ_h . Again we first construct a pre-tax benchmark population of households giving, for the same value of ρ_h for all households, an optimal second earner labour supply at every wage pair of 5 hours per day. We obtain this benchmark case by setting $\beta_h = 0.5$ and $\gamma = 2$ as in Model 1, and $\rho = 0$ (the Cobb-Douglas case) across all households.

This benchmark case, since it again implies identical marginal value productivities for the inputs c_h and x_{bh} , would if graphed look very similar to that shown in Figure 3 earlier, but now we perturb the equilibrium by varying the determinant of the elasticity of substitution, ρ_h , as just discussed. This generates different pairs of marginal productivity curves for H1 and H2 households which intersect at different levels of labour supplies/domestic child care. We select values of ρ_h above and below $\rho = 0$ that generate labour supplies for H1 and H2 households that broadly match their data means when we solve for the optimal marginal tax rates and lump sum, the bracket point and aggregate social welfare, SW, under piecewise linear joint and individual taxation. The results are reported in Table 6.

Table 6 about here

The first panel of the table presents results for $\rho_{1h} = -0.8$ and $\rho_{2h} = 0.3$, values that generate labour supplies that are reasonably close to the data means of 1.4 and 7.5 hours for household types H1 and H2, for $\pi = 0.1$ and $\pi = 0.3$.

⁴⁷These assumptions also underpin extensions of the unitary model to collective empirical applications with household production. See, for example, Cherchye et al. (2012) and the literature cited there.

The second panel reports results for $\rho_{1h} = -0.4$ and $\rho_{2h} = 0.15$ to illustrate the effect of varying ρ for the same values of π .

The values of ρ_{1h} , ρ_{2h} in the first panel imply that home and bought in care are strong complements for H1, whereas they tend to be substitutes for H2. As a consequence, household H1 specialises more in the use of home time in the production of z , the input with the lower productivity with $w_2 > 1$. H1 households therefore have a lower output of z and face a higher price per unit. For example, with $\pi = 0.1$, under individual taxation H1 faces a price p_z of 9.356 while the price for H2 is only 3.176. As a consequence, since household welfare falls with p_z , it tends to track household income.

However, somewhat surprisingly, this does not lead to a result that supports joint taxation as welfare superior. When we compare aggregate measures of SW we find that individual taxation is again consistently superior to joint taxation, and increasingly so as we increase the degree of inequality aversion. We again find that the rate scale under individual taxation is far more progressive than under joint taxation because a higher marginal rate can, in effect, be applied selectively to the incomes of top primary earners with much higher wage rates and relatively inelastic labour supplies.

As in the preceding Model 1, in which the H1 household is disadvantaged by a higher child care price, the superiority of individual taxation is driven by a gap between the utility levels of those in the top percentiles and those in the lower and middle percentiles that is far greater than the gap between the utility levels of the two household types within a given wage percentile. In other words, together with the efficiency gains, the high degree of inequality across the primary wage distribution drives the result that individual taxation is superior to joint taxation.

5 Conclusions

The choice between joint and individual income as the tax base is an important one because of its implications for the efficiency and equity of the tax system. The majority of households have, actually or potentially, two earners, and the high degree of heterogeneity across households in the second earner's market labour supply and therefore in the production of domestic goods and services presents a challenge in modelling the underlying behavioral unit in the optimal taxation problem. It is necessary to provide an empirically well-founded explanation of this heterogeneity, which at the same time clarifies the relationship between the household's total labour income and its standard of living, since this relationship is, at least *a priori*, an important determinant of the relative merits of joint and individual taxation. It is also essential to take account of the shape of the empirical wage distribution and the way in which the two tax systems define the effects of redistribution of the tax burden in the light of this.

This paper has presented a new model of the household that focuses on child care as the form of household production. The model shows that the nature of the second earner's time constraint and the cost and quality of child care play

a central role in explaining labour supply heterogeneity, as well as yielding the implication that household income is not a reliable indicator of a household's real living standard.⁴⁸

We have argued that in the light of the marked inequality in the primary earner wage distribution, a move from joint to individual taxation results in both a more efficient and a fairer allocation of the aggregate tax burden. In numerical versions of the model calibrated to reflect the empirical distributions of wages and labour supplies, and given standard formulations of society's distributional preferences, it is very difficult to find a case in which the optimal piecewise linear tax system based on joint income yields a higher social welfare than the optimal system based on individual income. Although this result of course depends on a particular parametrisation of the model, the one we adopt seems more plausible *a priori* than the alternatives that would be required to overturn the result.

In carrying out the numerical calculations in Section 4 we have used estimates of labour supply elasticities. This was done despite the recent literature⁴⁹ which argues that, particularly in respect of primary earners at the top of the income distribution, it is the elasticity of taxable earnings (ETI) that really matters, and this is considerably higher than the labour supply elasticities we have used. However, we are convinced by the work of Moffitt and Wilhelm (2000), Saez et al. (2012) and Piketty et al. (2014) that it is the labour supply elasticity rather than the ETI which is the relevant behavioural parameter for normative analysis of the kind carried out in this paper.

The household model we have put forward is based on a "unitary" approach to the household, in which we effectively assume that the welfare weights the household attaches to the wellbeing of the two individuals correspond to those applied by the tax authority or "planner".⁵⁰ However, we can show that the results are strengthened if we extend the model to allow the household's welfare weights to vary positively with net of tax wage rates,⁵¹ and assume that the weight the planner places on the utility of the second earner is at least as high as that she receives in the household. The move from joint to individual taxation in that case yields a form of "double dividend", with a gain in within-household as well as between-household equity. Although we do not believe that the general tax system can be used as an instrument to "fine tune" the allocation of resources within the household, we do believe that individual taxation improves the position of second earners generally.

In summary therefore, we would certainly answer Rosen's question in the affirmative: It *is* time to abandon joint filing.

⁴⁸Thus supporting the point made by Feldstein and Feenberg and quoted in the Introduction.

⁴⁹See Saez et al. (2012) for a survey.

⁵⁰In Apps and Rees (1988) this is referred to as the "non-dissonance" assumption.

⁵¹This can, but need not, be rationalised in terms of a Nash bargaining analysis of household resource allocation. See for example Gugl (2009). It would result from just about any model in which the weight given to an (adult) individual's wellbeing in the household increases with their net of tax wage rate, as in the exchange model of Apps (1982), or indeed earnings, as in the game-theoretic model of Basu (2006).

References

- [1] Alesina, A., Ichino, A., Karabarbounis, L., 2011. Gender based taxation and the division of family chores. *American Economic Journal: Economic Policy* 3(2),1-40.
- [2] Almond, D., Currie, J., 2011. Human capital development before age five. In O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, Vol 4, Part B.
- [3] Andrienko, Y., Apps, P., Rees R., 2016. Optimal taxation and top incomes. *International Tax and Public Finance*. Available at 10.1007/s10797-015-9391-y
- [4] Apps P., 1982. Institutional Inequality and Tax Incidence. *Journal of Public Economics*, 18(2), 217-42.
- [5] Apps, P., Long, N., Rees, R., 2014. Optimal piecewise linear income taxation. *Journal of Public Economic Theory*, 16 (4), 523-545.
- [6] Apps, P., Rees, R., 2009. *Public Economics and the Household*. Cambridge: Cambridge University Press.
- [7] Apps, P., Rees, R., 1999. On the taxation of trade within and between households. *Journal of Public Economics* 73(2), 241-63.
- [8] Apps, P., Rees, R., 1988. Taxation and the household. *Journal of Public Economics* 35(3), 355-69.
- [9] Attanasio, O., Low, H., Sanchez-Marcos, V., 2008. Explaining changes in female labour supply in a life cycle model. *American Economic Review*, 98(4), 1517-52.
- [10] Basu, K., 2006. Gender and say: A model of household behaviour with endogenously determined balance of power. *The Economic Journal*, 116, 558-580.
- [11] Blundell, R., MaCurdy, T., 1999. Labor supply: A review of alternative approaches, in O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, Vol 3. Elsevier: Amsterdam.
- [12] Boskin, M., 1995. Efficiency aspects of the differentiated tax treatment of market and household activity. *Journal of Public Economics* 4, 1-25.
- [13] Boskin, M., Sheshinski, E., 1983. Optimal tax treatment of the family: Married couples. *Journal of Public Economics* 20, 281-297.
- [14] Cherchye, L., De Roch, B., Vermeulen, V., 2012. Married with children: A collective labor supply model with detailed time use and intrahousehold expenditure information, *American Economic Review*, 102 (7), p. 3377-3405.

- [15] Chetty, R., 2009. Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annual Review of Economics*, 1, 451-488.
- [16] Dahlby, B., 1998. Progressive taxation and the marginal social cost of public funds. *Journal of Public Economics* 67, 105-122.
- [17] Dahlby, B., 2008. *The Marginal Cost of Public Funds*. Cambridge: MIT Press.
- [18] Feldstein, M., Feenberg, D., 1996. The taxation of two-earner families, ch. 2 in M Feldstein and J Poterba (eds), *Empirical Foundations of Household Taxation*. University of Chicago Press, Chicago.
- [19] Gugl, E., 2009. Income splitting, specialization and Intrafamily distribution, *Canadian Journal of Economics*, 42(3), 1050-1071.
- [20] Heckman J., 1974, Effects of child-care programs on women's work effort, *Journal of Political Economy*, 82(2), S236-S163.
- [21] Heckman J., Killingsworth, M., 1986. Female labor supply: A survey. In O. Ashenfelter and R. Layard (eds), *Handbook of Labor Economics*, Vol. 1. Amsterdam and New York: North Holland, 103-204.
- [22] LaLumia, S., 2008. The effects of joint taxation of married couples on labor supply and non-wage income, *Journal of Public Economics* 92(7), 1698-1791.
- [23] Lundborg, P., Nilsson, A., Rooth, D., 2014. Parental education and offspring outcomes: Evidence from the Swedish compulsory schooling reform. *American Economic: Applied Economics*, 6(1), 253-278.
- [24] Mirrlees, J., 1971. An exploration in the theory of optimum income taxation, *Review of Economic Studies*, 38, 175-208.
- [25] Moffitt, R., Wilhelm, M.O., 2000. Taxation and the labor supply decisions of the affluent. In *Does Atlas shrug? The economic consequences of taxing the rich*, edited by Joel Slemrod, 193-234. New York: Russell Sage Foundation; Cambridge and London: Harvard University Press.
- [26] Piketty, T., Saez, E., Stantcheva, S., 2014. Optimal taxation of top labor incomes: A tale of three elasticities. *American Economic Journal: Economic Policy*, 6(1), 230-271.
- [27] Rosen, H., 1977. Is it time to abandon joint filing?, *National Tax Journal*, 30, 423-428.
- [28] Saez, E., Slemrod, J., Giertz, S.H., 2012. The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature*, 50(1), 3-50.

- [29] Sandmo, A., 1990. Tax distortions and household production. *Oxford Economic Papers* 42, 78-90.
- [30] Sheshinski, E., 1972. The optimal linear income tax. *Review of Economic Studies* 39, 297-302.
- [31] Sheshinski, E., 1989. Note on the shape of the optimum income tax schedule. *Journal of Public Economics* 40, 201-215.
- [32] Slemrod, J., Yitzhaki, S., Mayshar, J., Lundholm, M., 1994. The optimal two-bracket linear income tax. *Journal of Public Economics*, 53, 269-290.
- [33] Steiner, V., Wrohlich, K., 2008. Introducing family tax splitting in Germany: How would it affect the income distribution, work incentives, and household welfare? *FinanzArchiv* 64(1), 115-142.
- [34] Steiner, V., Wrohlich, K., 2004. Household taxation, income splitting and labor supply incentive: A microsimulation study for Germany, *CESifo Economic Studies* 50, 541-568.
- [35] Stern, N., 1976. On the specification of optimum income taxation." *Journal of Public Economics* 6 (1-2), 123-162.
- [36] Tuomala M., 1984. On the optimal income taxation: some further numerical results, *Journal of Public Economics*, 23, 351-366.

Appendix A

Part 1: Comparative statics analysis of the household model

The comparative statics effects of interest are the signs of the effects on optimal labour supply l_{2h}^* and quality choice q_h^* of changes in the second earner wage rate w_{2h} and the productivity and price parameters $\tilde{\kappa}$ and $\tilde{\varepsilon}$ respectively. The household's problem as stated in the text is

$$\max_{x_h, l_{1h}, q_h} u_h = x_h - \hat{u}_1(l_{1h}) + \hat{u}_2(z(k_h(1 - l_{2h}), q_h l_{2h})) \quad h \in \mathcal{H} \quad (37)$$

$$\text{s.t. } x_h \leq w_{1h}l_{1h} + [w_{2h} - p_h]l_{2h} \quad (38)$$

where:

$$k_h = k(w_{2h}) + \kappa_h > 0, \quad p_h = p(q_h) + \varepsilon_h > 0 \quad (39)$$

and κ_h and ε_h are the household's realisations of $\tilde{\kappa}$ and $\tilde{\varepsilon}$ respectively.

Writing the Lagrange function as

$$\mathcal{L}_h = u_h + \lambda_h(w_{1h}l_{1h} + (w_{2h} - p_h)l_{2h} - x_h) \quad (40)$$

the first order conditions are

$$\frac{\partial \mathcal{L}_h}{\partial x_h} = 1 - \lambda_h = 0 \quad (41)$$

$$\frac{\partial \mathcal{L}_h}{\partial l_{1h}} = \hat{u}'_1(l_{1h}^*) - \lambda_h w_{1h} = 0 \quad (42)$$

$$\frac{\partial \mathcal{L}_h}{\partial l_{2h}} = \left(-\frac{\partial z}{\partial c_h} + \frac{\partial z}{\partial b_h}\right)\hat{u}'_2 + \lambda_h(w_{2h} - p_h) = 0 \quad (43)$$

$$\frac{\partial \mathcal{L}_h}{\partial q_h} = [\hat{u}'_2 z_2 - \lambda_h p'_h(q_h^*)]l_{2h}^* = 0 \quad (44)$$

together with the budget constraint. The household's marginal utility of income is denoted by λ_h . Clearly the additive separability and quasilinearity in the utility function buy a great deal of simplicity in these conditions. The compensated labour supply function for 1 is simply $l_{1h}^*(w_{1h}) \equiv [\hat{u}'_1^{-1}(w_{1h})]$, and the numeraire consumption good absorbs all income effects.

For purposes of the comparative statics analysis it is useful to define

$$\hat{u}_2(z(k_h(1 - l_{2h}), q_h l_{2h})) \equiv \varphi(l_{2h}, q_h, k_h).$$

Moreover, from the first order conditions it is clear that the optimal values l_{2h}^* and q_h^* are fully determined by the two conditions⁵²

$$\varphi_1(l_{2h}^*, q_h^*, k_h) + [w_{2h} - (p(q_h^*) + \varepsilon_h)] = 0$$

⁵²In the standard notation φ_1 denotes $\partial\varphi/\partial l_{2h}$ and so on.

$$\varphi_2(l_{2h}^*, q_h^*, k_h) - p'(q_h^*)l_{2h}^* = 0 \quad (45)$$

In interpreting the comparative statics results of the model, it is useful to begin by assuming that k_h , like w_{2h} and ε_h , is fixed exogenously. Standard comparative statics analysis on these conditions then gives:

Proposition A1:

$$\frac{\partial l_{2h}^*}{\partial w_{2h}} = \frac{-(\varphi_{22} - p''l_{2h}^*)}{D} > 0; \quad (46)$$

$$\frac{\partial q_h^*}{\partial w_{2h}} = \frac{(\varphi_{21} - p')}{D} \begin{matrix} \geq \\ \leq \end{matrix} 0 \quad (47)$$

$$\frac{\partial l_{2h}^*}{\partial \varepsilon_h} = -\frac{\partial l_{2h}^*}{\partial w_{2h}} < 0; \quad \frac{\partial q_h^*}{\partial \varepsilon_h} = -\frac{\partial q_h^*}{\partial w_{2h}} \begin{matrix} \geq \\ \leq \end{matrix} 0 \quad (48)$$

where

$$D \equiv \varphi_{11}\varphi_{22} - (\varphi_{12})^2 \quad (49)$$

and $D > 0$ from the second order condition.

With k_h fixed, second earner labour supply certainly increases with the wage, since

$$\varphi_{22} = \hat{u}_2''(z_2 l_{2h}^*)^2 + \hat{u}_2' z_2 l_{2h}^{*2} < 0 \quad (50)$$

This alone ensures that household income and welfare both increase with the wage. The effect of a wage increase on quality of bought in care is however ambiguous, depending as it does on the sign of φ_{21} , which is the derivative of the marginal value product of q_h (MVP_q) with respect to labour supply, and is given by

$$\varphi_{12} = \varphi_{21} = [(z_{22}q_h^* - z_{12}k_h)l_{2h}^* + z_2]\hat{u}_2' + z_2 l_{2h}^*(z_2 q_h^* - z_1 k_h)\hat{u}_2'' \begin{matrix} \geq \\ \leq \end{matrix} 0 \quad (51)$$

This sign depends on whether parental and bought in child care are Hicksian substitutes or complements, and on which of them has the higher marginal product at the equilibrium, which in turn depends on whether w_{2h} is greater or less than p_h . If the net effect of the increase in l_{2h}^* ($= b_h$) is to reduce MVP_q , then $\varphi_{21} < 0$ and the increase in w_{2h} reduces q_h^* , which is somewhat counterfactual. Intuitively, faced with the need to buy in more child care associated with the increase in labour supply induced by the higher wage rate, the household saves money on child care by reducing its quality. Though not particularly plausible, there are no *a priori* grounds to rule this case out. The alternative case is where the increase in l_{2h}^* increases MVP_q by more than enough to offset the increased price resulting from choosing a higher quality of child care, $(\varphi_{21} - p') > 0$ and so q_h^* increases.

Assuming now that $k_h = k(w_{2h}) + \kappa_h$ we have that in general terms

$$\frac{\partial l_{2h}^*}{\partial w_{2h}} = \frac{\partial l_{2h}^*}{\partial w_{2h}} \Big|_{k_h=const} + \frac{\partial l_{2h}^*}{\partial k_h} k'(w_{2h}) \quad (52)$$

$$\frac{\partial q_h^*}{\partial w_{2h}} = \frac{\partial q_h^*}{\partial w_{2h}} \Big|_{k_h=const} + \frac{\partial q_h^*}{\partial k_h} k'(w_{2h}) \quad (53)$$

implying:

Proposition A2:

$$\frac{\partial l_{2h}^*}{\partial w_{2h}} = \frac{-(\varphi_{22} - p''l_{2h}^*) - [\varphi_{13}(\varphi_{22} - p''l_{2h}^*) - \varphi_{23}(\varphi_{21} - p')]}{D} \begin{matrix} \geq \\ \leq \end{matrix} 0$$

$$\frac{\partial q_h^*}{\partial w_{2h}} = \frac{(\varphi_{21} - p') - [\varphi_{11}\varphi_{23} - \varphi_{13}(\varphi_{21} - p')]}{D} \begin{matrix} \geq \\ \leq \end{matrix} 0 \quad (54)$$

Essentially then, the ambiguity in the effect of the wage rate on labour supply comes about because we may have that $\partial l_{2h}^*/\partial k_h < 0$, or, intuitively, that an increase in the productivity of the second earner's time in child care increases the amount of time she spends on it and therefore reduces her market labour supply, other things being equal. In the case of the effect of the wage rate on bought in quality, the addition of the term $\partial q_h^*/\partial k_h$ suggests the possibility of a "matching effect". If, as seems empirically plausible, increases in the second earner's productivity in child care leads the household to demand increasing quality in bought-in child care, this makes it more likely that the latter increases with the second earner's wage as a measure of her human capital.

In these expressions, we have:

$$\varphi_{11} = (z_{11}k_h^2 - 2z_{12}k_hq_h^* + z_{22}q_h^{*2})\hat{u}'_2 + (z_2q_h^* - z_1k_h)^2\hat{u}''_2 < 0 \quad (55)$$

$$\varphi_{13} = \varphi_{31} = [z_{12}(1-l_{2h}^*)q_h^* - z_{11}(1-l_{2h}^*)k_h - z_1]\hat{u}'_2 + z_1(1-l_{2h}^*)(z_2q_h^* - z_1k_h)\hat{u}''_2 \begin{matrix} \geq \\ \leq \end{matrix} 0 \quad (56)$$

$$\varphi_{23} = \varphi_{32} = z_1z_2(1-l_{2h}^*)l_{2h}^*\hat{u}''_2 + z_{12}l_{2h}^*(1-l_{2h}^*)\hat{u}'_2 \begin{matrix} \geq \\ \leq \end{matrix} 0 \quad (57)$$

Part 2: Household Equilibrium and Indirect Utilities for the Taxation Analysis

To present the results of the analysis⁵³ in the most useful way for purposes of the tax analysis it is useful to reformulate the household model as follows. We rewrite the utility functions as

$$\hat{u}_1(l_{1h}) \equiv \hat{u}_1(y_{1h}/w_{1h}) \equiv \psi_{1h}(y_{1h}; w_{1h})$$

and

$$\hat{u}_2[z(k_h c_h, q_h b_h)] \equiv \hat{u}_2[z(k_h(1 - y_{2h}/w_{2h}), q_h y_{2h}/w_{2h})] \equiv -\psi_{2h}(y_{2h}; w_{2h}) \quad (58)$$

where the $\psi_{ih}(\cdot)$ are strictly increasing and convex in y_{ih} . We retain the assumption of identical preferences $\hat{u}_1(\cdot)$, $\hat{u}_{2h}(\cdot)$ across households, as is usual in optimal tax analysis.⁵⁴

⁵³Throughout this analysis we assume for simplicity that productivities k and q are determined by the wage type of the household as expressed by the gross wage rate w_{2h} rather than by the net of tax wage.

⁵⁴In contrast to this, empirical applications of the standard household model rely on a high degree of preference heterogeneity, expressed in terms of "preference errors", to explain the data. The problem with this is that welfare comparisons across households then become problematic and controversial. This is avoided by having prices and productivities of child care as well as wages be the drivers of across household heterogeneity in labour supplies, since these determine the feasible set rather than preferences.

Joint Taxation

A household $h \in \mathcal{H}$ solves the problem

$$\max_{x_h, y_{ih}} u_h = x_h - \sum_{i=1}^2 \psi_i(y_{ih}; w_{ih}) \quad (59)$$

subject to a budget constraint determined by the tax system. We consider three cases which provide the results we require, the partial derivatives of the household's indirect utility function with respect to the tax parameters. We write below the constraints for each of these cases together with these derivatives.

Case 1. The household is at the optimum in the interior of the lower tax bracket. It therefore faces the budget constraint:

$$x_h = \alpha + (1 - \tau_1) \sum_i y_{ih} \quad (60)$$

and the first order conditions imply:

$$\frac{\partial \psi_i}{\partial y_{ih}} = 1 - \tau_1 \quad i = 1, 2, \quad (61)$$

giving the earnings supply functions $y_{ih}(\tau_1)$. The properties of the functions $\psi_{ih}(\cdot)$ imply

$$\frac{\partial y_{ih}(\tau_1)}{\partial \tau_1} < 0, \quad i = 1, 2, \quad (62)$$

where, note, this is a compensated derivative.

We write the household indirect utility function⁵⁵ as $v_h(\alpha, \tau_1)$, with, by the Envelope Theorem,

$$\frac{\partial v_h}{\partial \alpha} = 1; \quad \frac{\partial v_h}{\partial \tau_1} = -y_h^* = - \sum_i y_{ih}(\tau_1) \quad i = 1, 2, \quad (63)$$

Case 2. The household is effectively constrained at the bracket limit η , in the sense that it chooses $y_h = \eta$, but would prefer to increase its labour supply and earnings if it would be taxed at the rate τ_1 , but not if it would be taxed at the rate τ_2 . We formulate its allocation problem by adding the constraint $y_h \leq \eta$, noting that this will be binding at the optimum.⁵⁶ We can write the first order conditions as

$$(1 - \tau_1) - \frac{\partial \psi_{ih}}{\partial y_{ih}} - \mu_h = 0 \quad i = 1, 2, \quad (64)$$

$$y_h \leq \eta \quad \mu_h \geq 0 \quad \mu_h [y_h - \eta] = 0 \quad (65)$$

where μ_h is the multiplier associated with the constraint $y_h \leq \eta$.

⁵⁵Where no confusion should arise we simplify notation by suppressing the type arguments w_{ih} in the indirect utility functions.

⁵⁶Case 1 can be thought of as the case in which this constraint is non-binding.

We write the indirect utility function as $v_h(\alpha, \tau_1, \eta)$, with, by the Envelope Theorem,

$$\frac{\partial v_h}{\partial \alpha} = 1; \quad \frac{\partial v_h}{\partial \tau_1} = -\eta; \quad \frac{\partial v_h}{\partial \eta} = (1 - \tau_1) - \frac{\partial \psi_i}{\partial y_{ih}} \geq 0 \quad (66)$$

Intuitively, the idea of the expression for $\partial v_h / \partial \eta$ is that a small relaxation of the constraint would increase consumption and utility at the rate $(1 - \tau_1)$, which exceeds for almost every individual the marginal cost of effort $\partial \psi_{ih} / \partial y_{ih}$. In diagrammatic terms, the household is at the kink in its budget constraint at the bracket limit η . The term is zero only if i 's marginal rate of substitution happens to equal $(1 - \tau_1)$ at the kink. Note that condition (68) implies that the individuals' marginal effort costs are equalised also in this type of equilibrium. Given that the household wants to earn the income η , it allocates labour supplies so as to minimise the total utility costs of achieving this.

Case 3. The household is in equilibrium in the interior of the upper income bracket. We therefore replace the previous budget constraint by

$$x_h \leq \alpha + (1 - \tau_2)y_h + (\tau_2 - \tau_1)\eta \quad (67)$$

and the first order conditions imply

$$\frac{\partial \psi_{ih}}{\partial y_{ih}} = 1 - \tau_2 \quad i = 1, 2, \quad (68)$$

giving the earnings supply functions $y_{ih}(\tau_2)$. The properties of the functions $\psi_{ih}(\cdot)$ imply

$$\frac{\partial y_{ih}(\tau_2, w_{ih})}{\partial \tau_2} < 0, \quad \frac{\partial y_{ih}(\tau_2, w_{ih})}{\partial w_{ih}} > 0 \quad i = 1, 2, \quad (69)$$

Writing the indirect utility function as $v_h(\alpha, \tau_1, \tau_2, \eta)$ we now obtain

$$\frac{\partial v_h}{\partial \alpha} = 1; \quad \frac{\partial v_h}{\partial \tau_1} = -\eta; \quad \frac{\partial v_h}{\partial \tau_2} = -(y_h^* - \eta); \quad \frac{\partial v_h}{\partial \eta} = \tau_2 - \tau_1 > 0 \quad (70)$$

In all three cases, it follows from the properties of the function $\psi_{ih}(\cdot)$ that $\partial v_h / \partial w_{ih} > 0$, $i = 1, 2$, $h \in \mathcal{H}$.

We can show that under joint taxation the household can be treated as if it were a single individual. To see this, note that we can solve the household's problem in two steps. First solve $\min_{y_{ih}} \sum_i \psi_i(y_{ih}, w_{ih})$ subject to $\sum_i y_{ih} \leq y_h$ for any given y_h , and define $\psi_h(y_h)$ as the value function of this problem. Then solve $\max_{x_h, y_h} x_h - \psi_h(y_h)$ subject to the relevant budget constraint in each case.

Individual Taxation

With individual income as the tax base, and given that (by definition) the second earner's income is always below that of the primary earner, we can define six possible cases for the household equilibrium. In each case we present the earnings and indirect utility functions and partial derivatives of the latter with respect to the tax instruments.

Case 1: $y_{ih}^* < y$, $i = 1, 2$. In this case the household's budget constraint, earnings and indirect utility functions are identical to those in Case 1 of joint taxation.

Case 2: $y_{2h}^* < y = y_{1h}^*$. The results here are derived by imposing the constraint $y_{1h} \leq y$ on the problem and noting that it is binding at the optimum. Thus we have $y_{2h}^* = y_{2h}(t_1, w_{2h})$, and $v_h(a, t_1, y)$, with

$$\frac{\partial v_h}{\partial a} = 1; \quad \frac{\partial v_h}{\partial t_1} = -(y + y_{2h}^*); \quad \frac{\partial v_h}{\partial y} = (1 - t_1) - \frac{\partial \psi_1}{\partial y_{1h}} \quad (71)$$

Case 3: $y_{ih}^* = y$, $i = 1, 2$. Here we impose the two constraints $y_{ih} \leq y$ and take them as both binding at the optimum, giving $v_h(a, t_1, y)$ and

$$\frac{\partial v_h}{\partial a} = 1; \quad \frac{\partial v_h}{\partial t_1} = -2y; \quad \frac{\partial v_h}{\partial y} = 2(1 - t_1) - \sum_i \frac{\partial \psi_i}{\partial y_{ih}} \quad (72)$$

Case 4: $y_{2h}^* < y < y_{1h}^*$. In this case the budget constraint becomes

$$x_h \leq a + (t_2 - t_1)y + (1 - t_2)y_{1h} + (1 - t_1)y_{2h} \quad (73)$$

and we have $y_{1h}^* = y_{1h}(t_2, w_{1h})$, $y_{2h}^* = y_{2h}(t_1, w_{2h})$ and the indirect utility function $v_h(a, t_1, t_2, y)$ with

$$\frac{\partial v_h}{\partial a} = 1; \quad \frac{\partial v_h}{\partial t_1} = -(y + y_{2h}^*); \quad \frac{\partial v_h}{\partial t_2} = (y_{1h}^* - y); \quad \frac{\partial v_h}{\partial y} = t_2 - t_1 \quad (74)$$

Case 5: $y_{2h}^* = y < y_{1h}^*$. We now have $y_{1h}^* = y_{1h}(t_2, w_{1h})$ and the indirect utility function $v_h(a, t_1, t_2, y)$ with

$$\frac{\partial v_h}{\partial a} = 1; \quad \frac{\partial v_h}{\partial t_1} = -2y; \quad \frac{\partial v_h}{\partial t_2} = (y_{1h}^* - y); \quad \frac{\partial v_h}{\partial y} = t_2 - t_1 + (1 - t_1) - \frac{\partial \psi}{\partial y_{2h}} \quad (75)$$

Case 6: $y_{ih}^* > y$, $i = 1, 2$. This gives $y_{ih}^* = y_{ih}(t_2, w_{ih})$, $i = 1, 2$, and $v_h(a, t_1, t_2, y)$ with

$$\frac{\partial v_h}{\partial a} = 1; \quad \frac{\partial v_h}{\partial t_1} = -2y; \quad \frac{\partial v_h}{\partial t_2} = \sum_i (y_{ih}^* - y); \quad \frac{\partial v_h}{\partial y} = 2(t_2 - t_1) \quad (76)$$

These results then feed directly into the specification of the optimal tax conditions presented in the paper.

TABLES**MODEL 1****Table 1 Model 1: home productivity variation, $k=k\pm 0.15w_2$**

ρ	π	Tax system	τ_1, t_1	τ_2, t_2	α, a	Bkt ⁺	SW	H1 hr ^a	H2 hr ^b	H1 p _z ^c	H2 p _z ^d
0.85	0.1	Joint	0.01	0.04	1229	97	41361	2.90	6.91	0.930	1.092
		Individual	0.01	0.30	3076	95	41405	2.90	6.91	0.930	1.092
	0.3	Joint	0.03	0.13	3419	99	218403	2.51	6.40	0.932	1.098
		Individual	0.01	0.43	6361	86	219746	2.90	6.91	0.932	1.098
0.5	0.1	Joint	0.02	0.10	2577	96	41160	4.50	5.32	0.937	1.091
		Individual	0.02	0.27	4497	92	41198	4.50	5.23	0.937	1.091
	0.3	Joint	0.07	0.26	8085	97	217673	4.13	4.79	0.941	1.100
		Individual	0.03	0.43	8158	86	218700	4.43	5.14	0.939	1.095

+: Income percentile of bracket point; **a**: median H1 second hours; **b**: median H2 second hours;
c: median H1 p_z; **d**: median H2 p_z

Table 2 Model 1: child care price variation, $p=p\pm 0.15w_2$

ρ	π	Tax system	τ_1, t_1	τ_2, t_2	α, a	Bkt ⁺	SW	H1 hr ^a	H2 hr ^b	H1 p _z ^c	H2 p _z ^d
0.85	0.1	Joint	0.02	0.06	2316	98	41886	1.47	7.97	1.084	0.912
		Individual	0.01	0.32	2640	97	41922	1.60	8.12	1.084	0.912
	0.3	Joint	0.04	0.12	4499	97	221282	1.26	7.62	1.088	0.918
		Individual	0.03	0.47	6976	92	222364	1.37	7.80	1.084	0.912
0.5	0.1	Joint	0.03	0.12	3611	98	41404	3.59	5.96	1.088	0.921
		Individual	0.02	0.30	4103	95	41439	3.67	6.04	1.088	0.921
	0.3	Joint	0.07	0.27	8090	98	218993	3.32	5.64	1.102	0.928
		Individual	0.04	0.43	9081	86	219914	3.52	5.88	1.093	0.923

+: Income percentile of bracket point; **a**: median H1 second hours; **b**: median H2 second hours;
c: median H1 p_z; **d**: median H2 p_z

Table 3 Model 1: child care price variation, $p = p\pm 0.25w_2$

ρ	π	Tax system	τ_1, t_1	τ_2, t_2	α, a	Bkt ⁺	SW	H1 hr ^a	H2 hr ^b	H1 p _z ^c	H2 p _z ^d
0.75	0.1	Joint	0.03	0.09	2542	97	42365	1.56	7.96	1.130	0.842
		Individual	0.02	0.32	3688	97	42395	1.63	8.06	1.130	0.842
	0.3	Joint	0.07	0.19	7800	98	223953	1.31	7.56	1.143	0.580
		Individual	0.05	0.51	8368	95	224801	1.42	7.77	1.139	0.847
0.5	0.1	Joint	0.03	0.12	3656	97	41815	2.93	6.69	1.144	0.859
		Individual	0.02	0.30	3987	96	41846	2.99	6.75	1.140	0.858
	0.3	Joint	0.08	0.23	9088	97	221149	2.64	6.33	1.160	0.865
		Individual	0.05	0.47	8983	92	221949	2.80	6.54	1.150	0.861

+: Income percentile of bracket point; **a**: median H1 second hours; **b**: median H2 second hours;
c: median H1 p_z; **d**: median H2 p_z

Table 4 Model 1: child care price variation: $p = p \pm 0.50w_2$

ρ	π	Tax system	τ_1, t_1	τ_2, t_2	α, a	Bkt ⁺	SW	H1 hr ^a	H2 hr ^b	H1 p _z ^c	H2 p _z ^d
0.5	0.1	Joint	0.05	0.14	6035	94	43523	1.68	7.94	1.265	0.670
		Individual	0.05	0.33	6712	98	43544	1.68	7.94	1.260	0.669
	0.3	Joint	0.13	0.29	14489	97	230206	1.44	7.57	1.287	0.678
		Individual	0.12	0.53	14753	97	230674	1.46	7.63	1.281	0.676
0.2	0.1	Joint	0.05	0.15	6045	95	42742	2.58	7.08	1.280	0.699
		Individual	0.05	0.32	6826	97	42763	2.54	7.08	1.274	0.698
	0.3	Joint	0.13	0.36	14703	97	226071	2.33	6.76	1.305	0.704
		Individual	0.11	0.51	14225	96	226537	2.39	6.84	1.299	0.702

+ : Income percentile of bracket point; **a**: median H1 second hours; **b**: median H2 second hours;
c: median H1 p_z; **d**: median H2 p_z

Table 5 Model 1: child care price variations: $p = p \pm 0.25w_2$ and $p = p \pm 0.50w_2$

ρ	π	Tax system	τ_1, t_1	τ_2, t_2	α, a	Bkt ⁺	SW	H1 hr ^a	H2 hr ^b	H1 p _z ^c	H2 p _z ^d
0.5	0.1	Joint	0.04	0.14	4802	97	92163	1.71	7.99	1.238	0.667
		Individual	0.03	0.32	4738	97	92218	1.75	8.03	1.236	0.667
	0.3	Joint	0.10	0.24	11218	97	607343	1.51	7.73	1.253	0.668
		Individual	0.08	0.50	11286	95	609026	1.58	7.82	1.247	0.667
0.2	0.1	Joint	0.04	0.12	4931	93	91073	2.62	7.10	1.252	0.697
		Individual	0.04	0.32	5785	97	91124	2.68	7.10	1.252	0.697
	0.3	Joint	0.12	0.30	13428	97	600381	2.35	6.80	1.273	0.697
		Individual	0.09	0.50	12359	95	601927	2.45	6.91	1.265	0.697

+ : Income percentile of bracket point; **a**: median H1 lower second hours; **b**: median H2 higher second hours
c: median H1 higher p_z; **d**: median H2 lower p_z

MODEL 2

Table 6 Model 2: Price and productivity settings of standard model: $p_x=1$ and $k = 1.0$

ρ_1, ρ_2	π	Tax system	τ_1, t_1	τ_2, t_2	α, a	Bkt ⁺	SW	H1 hr ^a	H2 hr ^b	H1 p _z ^c	H2 p _z ^d
0.8 &	0.1	Joint	0.02	0.06	2398	96	27914	1.68	7.80	9.359	3.176
		Individual	0.02	0.34	3759	97	27953	1.68	7.80	9.356	3.176
0.3	0.3	Joint	0.06	0.11	6549	97	146992	1.40	7.24	9.607	3.369
		Individual	0.04	0.48	7997	92	148084	1.52	7.54	9.472	3.259
0.4 &	0.1	Joint	0.01	0.06	1226	97	27339	2.66	6.23	7.961	4.407
		Individual	0.01	0.31	3128	95	27390	2.66	6.23	7.891	4.344
0.15	0.3	Joint	0.04	0.40	4475	99	143905	2.27	5.68	8.125	4.549
		Individual	0.02	0.43	7117	86	145281	2.52	6.05	7.961	4.407

+ : Income percentile of bracket point; **a**: median H1 second hours; **b**: median H2 second hours;
c: median H1 p_z; **d**: median H2 p_z

FIGURES

Figure 1 Percentile wage distributions

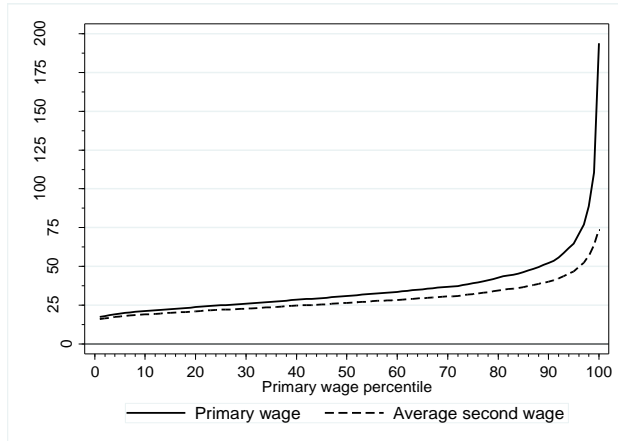


Figure 2 Household income and imputed expenditure on home child care

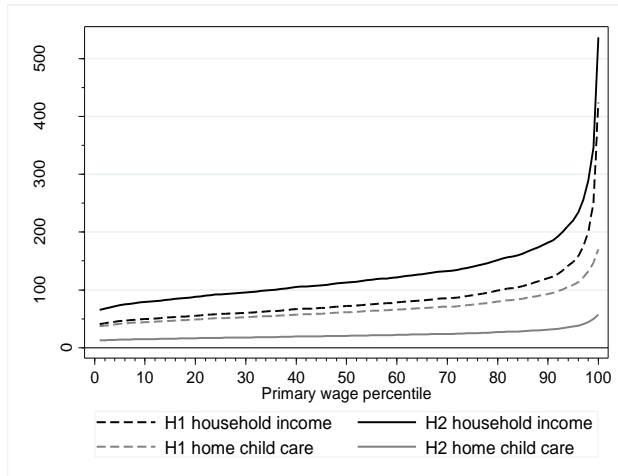


Figure 3 Model 1

