Optimal Family Taxation and Income Inequality

Patricia Apps
University of Sydney Law School and IZA

Ray Rees
University of Munich, University of Oslo and Sydney Law School

November 17, 2016

Abstract

This paper presents the properties of optimal piecewise linear income tax systems for families based on joint and individual incomes respectively. It models the interaction between the wage rates of mothers as "second earners" and variation in child care prices and productivities as determinants of across-household heterogeneity in second earner labour supply. We find that individual taxation welfare dominates joint taxation not only on the well-known grounds of efficiency but also of equity. An important driver of this result is the sharp rise in wage rates in the top percentiles of the primary wage distribution. In addition to reducing the intra-household net-of-tax wage gap, individual taxation removes the opportunity for tax avoidance income splitting makes available to high wage primary earners, leading to a much fairer distribution of the tax burden.

Corresponding author:
Patricia Apps, Faculty of Law, University of Sydney
NSW 2006, Australia
Email: patricia.apps@sydney.edu.au

1

1 Introduction

There is a long history of criticism of the system of joint taxation, commonly known as joint filing or income-splitting, by tax economists[1]. The early critique was based on a straightforward, and still valid, efficiency argument. Under a progressive rate scale, joint taxation equalises the marginal tax rates on the incomes of a two-earner couple whereas under individual taxation the second earner on a lower wage, typically the female partner, is very likely to face a lower marginal tax rate than the primary earner. The empirical evidence[2] 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.[3] Now, at least five decades after increased levels of female labour force participation made the issue of the taxation of couples a central one in tax policy, income-splitting still seems to be firmly enshrined in the personal income tax systems of two of the world's largest economies, Germany and the USA, with no obvious indications of the likelihood of change.[4]

A strong intuition in support of joint taxation is that 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.[5] This would seem to violate the principles of both vertical and horizontal equity. However, it is not just money income that determines a household's real living standard. We also have to take into account the value of household goods and services produced and consumed within the household. Where these are produced primarily by the (potential) second earner using time that could alternatively be spent earning market income, this *untaxed*

---

[1] See for example Rosen (1977), Boskin and Sheshinski (1983), and Apps and Rees (1988).

[2] See Heckman and Killingsworth (1986) for a survey of the literature of that period. More recent empirical studies confirm these earlier results, see for example LaLumia (2008) and Steiner and Wrohlich (2004, 2008).

[3] Individual taxation, by imposing a lower tax rate on the second income, lowers the net-of-tax gender pay gap, and therefore the gender gap in "outside opportunities". Since Alesina et al. (2011) a tax system with a lower rate on the second income has been referred to as "gender-based taxation".

[4] Forms of income splitting vary across countries. The US and Germany are two 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.

[5] We 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.

contribution to the household's living standard may actually vary inversely with its total money income.

In general terms, the central issue is the extent of the gains in equity and efficiency in moving from joint to individual taxation in an economy where domestic production is a significant form of time use[6] 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, and show that in fact there is a further, so far unappreciated strand in the equity argument. This goes beyond the recognition of the real patterns of time use within households and second earner labour supply heterogeneity, and has become of increasing significance in the last few decades. This has been a period characterised by increasing wage and income inequality[7] and falling top tax rates[8]. 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 show in this paper that income splitting imposes a severe constraint on the extent to which income can be redistributed from households with very high standards of living to those that are in the low to middle ranges of the distribution of wellbeing. Income-splitting could in fact be viewed as a means of tax avoidance: having the second earner substitute household for market work is a perfectly legal means by which a high income primary earner can significantly reduce his tax bill. We show here that 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,[9] which up until now has focused either on linear taxation,[10] tax reform, or on nonlinear taxation[11] 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 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

---

[6] For earlier analyses of income taxation with household production see Alesina et al. (2011), Apps and Rees (1988), (1999), Boskin (1975) and Sandmo (1990).

[7] For empirical work on this see Atkinson (2015), Atkinson et al. (2011) and Piketty and Saez (2003).

[8] This very important development in the tax structures of high income countries is thoroughly documented in Peter et al. (2010).

[9] The 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).

[10] As in the seminal paper by Boskin and Sheshinski (1983).

[11] See Apps and Rees (2009) for a literature survey.

the number of dimensions of private information that can be handled severely limits the type specification and general structure of the household model.

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 functions[12] 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

We present a 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. Recent studies have shown that the quality of investment in child development must be perceived as an additional dimension of across-household inequality.[13] This is very relevant for the appraisal of alternative tax systems in terms of their implications for distributional equity. The numerical simulations presented in section 4 constitute, we believe, the first attempt to integrate this dimension into an analysis of optimal family taxation.

Households are assumed to consist of couples with the same number of children, normalised at one.[14] 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.[15]

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

---

[12] Detailed derivations of these are given in Appendix A.

[13] This 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, Almond and Currie (2011), Lundborg et al. (2014), and, in particular, Lundberg et al. (2016).

[14] In 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.

[15] Nothing 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.

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,[16] 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.[17]

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.[18] 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

$$z = z(kc, qb) \qquad (1)$$

where k and q are strictly positive measures of the productivity/quality in child

---

[16] Ranging from grandparents and other family members, neighbours or an *au pair*, through day-care centres and private child-minders, to highly trained tutors.

[17] Variation in the price of bought in child care may also 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.

[18] We 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.

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

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

The $\hat{u}_1(.)$ 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 $\hat{u}_2(.)$ 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,[19] normalised at 1, and so we have

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

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

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

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

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

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 supply[20] and, at a given choice of quality, variations in w_{2h} and p_h have equal but opposite effects. 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}.

---

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

[20] Rather 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.

2.1 Household equilibrium

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

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

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 condition[21] determining l^*_{2h} is rather more interesting than the conventional expressions determining second earner labour supply. We can write this as

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

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.

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 model[22] 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. Furthermore, given that child care 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.

---

[21] See Appendix A for the full set of first order conditions.

[22] Given in full in Appendix A.

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 labour supplies with respect to tax rates and the lump sum.[23]

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 $v(.)$ and individual earnings functions by $y_i(.)$, $i = 1, 2$. 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 sum[24] funded by revenue from taxes on the labour incomes of the two earners aggregated across 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 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.[25]

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

---

[23] However, the economic interpretation and empirical estimation of these sufficient statistics will closely depend on the underlying structural model.

[24] This could be thought of as a standard child benefit.

[25] Apps 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.

overwhelmingly of the individual rather than selective kind.[26] 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.

3.2 Joint taxation

There is a two-bracket piecewise linear tax[27] 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:

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

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

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,[28] which give a partition $\{\mathcal{H}_0, \mathcal{H}_1, \mathcal{H}_2\}$ of the index set $\mathcal{H}$ defined as follows:

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

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

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

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 $h \in \mathcal{H}_1 \cup \mathcal{H}_2$

---

[26] At 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.

[27] For an extension of optimum piecewise linear taxation to the case of an arbitrary number $m \geq 2$ of tax brackets with only single-earner households, see Andrienko et al. (2016).

[28] See Appendix A for details.

- 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 income[29]

3.2.1   Optimal tax analysis

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

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

subject to the public sector budget constraint[30]

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

where $S(.)$ 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 variables[31] we can derive:

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

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

$$\tau_1^* = \frac{\int_{\mathcal{H}_0} (\sigma_h - 1)y_h^* dF + \eta^* \int_{\mathcal{H}_1 \cup H_2} (\sigma_h - 1)dF}{\int_{\mathcal{H}_0} (\partial y_h / \partial \tau_1)dF} \tag{16}$$

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

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

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 theory[32]: Denoting the shadow price of the government

[29] Again the details are in Appendix A.

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

[31] Of 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.

[32] See Sheshinski (1972).

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 $S(.)$ 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 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.[33] 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

---

[33] It 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).

increases in top incomes.[34]

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.[35] 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:

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

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

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 $\{H_0, H_1, ..., H_5\}$ of the index set $\mathcal{H}$, defined by

$$H_0 = \{ \ h \ | \ 0 \leq y_{ih}^* < y, \ i = 1, 2\} \tag{21}$$

$$H_1 = \{ \ h \ | \ y_{2h}^* < y = y_{1h}^*\} \tag{22}$$

$$H_2 = \{ \ h \ | \ y_{ih}^* = y, \ i = 1, 2\} \tag{23}$$

$$H_3 = \{ \ h \ | \ y_{2h}^* < y < y_{1h}^*\} \tag{24}$$

---

[34] See Andrienko et al. (2016) for more on this.
[35] For the details again see Appendix A.

$$H_4 = \{ \; h \;\; | \;\; y_{2h}^* = y < y_{1h}^* \} \tag{25}$$

$$H_5 = \{ \; h \;\; | \; y_{ih}^* > y, \;\; i = 1, 2 \} \tag{26}$$

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.

There are two important differences to the joint taxation case, resulting from the obviously finer partition of households based on individual reactions to the tax system. First, one or both individuals in households with a total income large enough to place them in the upper tax bracket under joint taxation may be in the lower bracket under individual taxation; and secondly, high wage primary earners in households with a low enough second income may be placed in the lower bracket of the joint tax system, while in the individual system the primary earner will be in the upper bracket and the second earner in the lower.

That is, the individual tax system corrects two types of errors that the joint tax system makes. Taking joint income as the tax base assigns to lower wage two-earner households a rank in the joint income distribution that is too high relative to their position in the utility ranking; and at the same time the income splitting advantage to a high-wage primary earner allows the household to place itself in a position in the income ranking that is too low relative to its position in the utility ranking. This second type of error is particularly significant when there is a high degree of inequality in primary earner wages.

Of course, neither of these errors could arise if household wellbeing increased monotonically with joint income, but the purpose of the model in the previous section was to show that this cannot in general be assumed. The numerical analysis in the following section, which calculates optimal taxes on the basis of two versions of that model, shows how the equity effects arising from the correction of these errors can reinforce the efficiency effects and ensure a welfare dominance of individual taxation that is robust to a wide range of assumptions consistent with the empirical evidence.

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}, $i, j, k = 0, ..., 5$, $i \neq j$, $i, j \neq k$. 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

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

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.[36]

---

[36] See Appendix A.

13

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 $\sigma_h = 1$.

From the first order conditions for an optimal solution[37] we derive:

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

$$\int_{\mathcal{H}} \delta_h dF = 0 \tag{28}$$

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

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

$$\int_{H_{12}} (\sigma_h \mu_{1h} + t_1)dF + \int_{H_{24}} (\sigma_h \mu_{2h} + t_1)dF =$$

$$-(t_2 - t_1)[\int_{H_{34}} \delta_h dF + 2\int_{H_5} \delta_h dF] \tag{31}$$

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 integrals 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 the tax structure. First note that the denominator in the expression for t_1^* will tend to contain more lower income second earners, since the subset H_{013} contains second earners who, because they are in households with high wage primary earners, would under joint taxation be in the upper tax bracket. The subset H_{345} in the denominator for t_2^* will tend to include more high wage primary earners, who have lost the income-splitting advantage they obtain under joint taxation. Other things equal therefore, 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 fact that labour supply elasticities are lower for primary than for second earners.[38]

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 (17) 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

---

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

[38] We should note that the empirical estimates of elasticities are gender-based - female labour supply elasticities are 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. Also, as pointed out earlier, it is still the case that the large majority of second earners are women.

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 subset H_{345} contains only the highest income primary earners, and H_5 only the highest income primary and second earners, with (as shown by the empirical wage distributions in Figure 1 below) very large differences $(y_{ih}^* - y^*)$ between 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 the differences in marginal rate progressivity between the two tax systems are striking. We highlight these differences as a major reason for the welfare superiority of individual over joint taxation systems, given the very high degree of inequality exhibited by the empirical primary earner wage distribution illustrated in Figure 1. Joint taxation allows high wage primary earners to reduce their tax burden by having the second earner in the household reduce her labour supply or leave the labour force altogether. Individual taxation removes this opportunity and so allows the tax system the possibility of significantly more redistribution across the highly unequal primary 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, we draw on household survey data for a sample of two-parent families to calibrate models nested within the following general empirical specification of the theoretical model set out in Section 2. We then use the models to solve for the optimal tax parameters of joint and individual taxation and to compute aggregate measures of social welfare under each system.

4.1   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

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

with $\alpha_1 = (1 + e_1)/e_1 > 1$, where e_1 is the elasticity of labour supply with

15

respect to the net wage of the primary earner, and $\alpha_2 \in (0, 1)$.[39] All households are assumed to have the same preferences and therefore the same parameters in the household utility function.

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

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

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 $\beta_h \in (0, 1)$.

Following the convention of the labour supply literature, we set market productivities to respective wage rates and we define market consumption, x, as a Hicksian composite good and choose it as numeraire with price set to 1.

We solve for the optimal tax parameters under joint and individual taxation for two models, labeled Model 1 and Model 2, that are nested within the general empirical specification in (32) - (33). The models incorporate different assumptions on the productivity variable, k_h, for which information is missing in available household surveys data sets, and on the exogenous factors driving heterogeneity. The assumptions of each model and the parameters values selected to generate labour supplies that broadly match the data are set out below.

4.2 Data

The sample of two-parent families is drawn from the Australian Bureau of Statistics (ABS) 2009-10 Survey of Income and Housing (SIH).[40] We first construct a percentile primary wage distribution from the data for primary earnings and hours of work. 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 average of around 8 hours per day for a five day working week. We find that $e_1 = 0.1$ generates a primary hours profile across the primary wage distribution that broadly matches

---

[39] The 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 avalable from the authors.

[40] See Australian Bureau of Statistics (2009-10). The 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.

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

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

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.

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. Very little of the variation in second hours can be explained by wage rates or demographics at a given primary wage.

As a reference for the data on second earner labour supplies, we construct profiles of second hours based on splitting the sample according to median second hours across the primary wage distribution. Records with second hours below the median as labeled "H1" and those with hours at or above the median, "H2". Parameter values are selected for the simulations which can generate labour supplies that approximate the average hours of H1 and H2 households by falling broadly within the range of 1.0 and 2.2 hours for the former and 7.0 and 8.0 hours for the latter. The time constraint is set to 10 hours per day.

4.3 Model 1

All households have the same production parameters γ, β and ρ, in addition to the same preference parameters. Consistent with the treatment of market time, the productivity of home child care time is set to the wage, that is, we set $k_h = w_{2h}$ for $\tilde{\kappa} = 0$. The quality of market child care chosen by the household, q_h, is assumed to match that of second earner's own child care time.[43] Heterogeneity in second earner labour supply is introduced by random variation in either k_h or p_h.[44] We therefore have $k_h = w_{2h} + \kappa_h$ and $p_h = w_{2h} + \varepsilon_h$, with κ_h and ε_h the household's realisations of $\tilde{\kappa}$ and $\tilde{\varepsilon}$ respectively.

We define p_z as the implicit price of a 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 $\beta = 0.5$. We also set $\gamma = 2$ to give $p_z = 1$ at the benchmark equilibrium. Given that introducing labour supply heterogeneity by varying k_h will tend to bias the results towards individual

---

[43] These asumptions are consistent with empirical studies showing that child outcomes improve with maternal human capital and that parental investment in quality child care and education increases with family resources, as noted in Section 2.

[44] In contrat to k_h for which data are entirely missing, there is information on p_h and q_h. However, the data for these variables indicate a very high degree of variation in the price of an hour of child care, p_h, of a given quality, as noted previously.

taxation, by making the H1 household better off at any given wage pair, we report results for variation in p_h only.

We perturb the benchmark labour supplies by varying p_h above and below its benchmark value, $p_h = w_{2h}$, by the same percentage at each wage pair,[45] 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 benchmark median, corresponding to the households facing lower/higher prices. This provides us with a population of heterogenous households for which we then compute optimal piecewise linear tax systems based on joint and individual incomes respectively. Figure 2 illustrates this procedure for variation in p_h.

Figure 2 about here

Given the above specifications, it is easy to show that, at any wage pair, an equilibrium of the household 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 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.

Table 1 report results for a price variations of $\pm 25\%$. 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 $\rho = 0.7$, a value for ρ which generates second earner labour supplies that approximate the data means of the H1 and H2 subsamples. The results show that individual taxation is consistently superior to joint taxation for $\pi = 0.2$ to 0.6, and becomes increasingly superior as the degree of inequality aversion, π, rises.

Table 1 about here

---

[45]We restrict the variations in p_h above and below w_{2h} to the same percentage to facilitate comparisons across the results. This restriction, of course, rules out the possibility of a perfect match with data means.

Extending the analysis by combining two degrees of price variation gives four rather than two relative prices for child care in each percentile, that is, $p_h \in \{p^1, ..., p^4\}$, which increases the total number of household records from 200 to 400. Table 2 reports the results for price variations of $\pm 25\%$ and $\pm 50\%$ above and below the benchmark price. To match the data when the degree of price variation rises $\pm 50\%$ it is necessary to reduce ρ from 0.7 to 0.5. The explanation is straightforward. To obtain labour supplies within ranges consistent with the data as the degree of price variation rises, we need to assume that home and bought in care are less close substitutes, as implied by the lower value of ρ. Again we find individual taxation remains consistently superior to joint taxation, and becomes increasingly so as π rises from 0.2 to 0.6.

Table 2 about here

The most striking feature of the results, apart from the complete dominance of individual taxation, 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, with the bracket limit at which tax rates change invariably above the 85'th percentile. This reflects the shape of the wage profiles in Figure 1. The profiles are relatively flat up to around the 85th percentile and rise steeply thereafter.[46]

The variation in the price of bought in care, p_h, in each table generates second earner labour supplies consistent with the data and creates a correspondingly large gap between the incomes of H1 and H2 households that also matches the data. However, given the degree of substitution of the two time inputs associated with each value of ρ the impact of the price gap on the implicit prices of child care, p_z, for H1 and H2, is minimal. The consequent minimal impact on household utilities is then completely swamped by high degree of inequality in the primary wage distribution. In other words, because inequality *across* the distribution is large relative to that between households *at a given wage pair*, individual taxation is consistently welfare superior to joint taxation.[47]

Thus, given the shape of the wage distribution, a progressive individual 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. Furthermore, as discussed in Section 3, including second earners with relatively higher compensated elasticities in the higher tax bracket under joint taxation reduces the marginal rate progressivity of the tax system, since it

---

[46] The effect of the high degree of wage inequality at the top end of the distribution on the stucture 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 50'th percentile, the second around the 90'th and the third at the 99'th, with sharply increasing marginal tax rates.

[47] It can be shown that to obtain a result in support of joint taxation when heterogeneity is driven by variation in the price of child care, it is necessary to specify a virtually flat primary wage distribution. The results for this case are available from the authors.

raises the absolute value of the denominator in the expression determining the optimal tax rate. It is counter-productive in terms of both equity and efficiency 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 the marginal tax rate on high primary wage earners is set at relatively low levels.

4.4   Model 2

Model 2 incorporates the assumptions on domestic time productivity and quality and price of bought in child care time adopted in the estimation of two-person family labour supply models underlying much of the literature[48]. A key assumption is that the productivity of home child care time, k_h (typically labeled "leisure"), is constant across all households, with one hour priced at the wage. In addition, with market consumption, x, defined as a Hicksian composite good and selected as numeraire with price of one, the standard empirical model sets $k_h = 1$ across all households,[49] in contrast to $k_h = w_{2h}$ in Model 1. Since this literature also defines bought-in child care as expenditure on the market good, we can denote expenditure on child care by x_{bh} and rewrite the household's budget constraint as

$$x_h + x_{bh} \leq \sum_{i=1}^{2} w_{ih} l_{ih} \quad h \in \mathcal{H} \tag{34}$$

Heterogeneity in second earner labour supply is driven by variation in the production parameter, ρ_h, rather than by preferences, which allows us to make welfare comparisons. The child care production function can therefore be rewritten as

$$z_h = \gamma[\beta c_h^{\rho_h} + (1 - \beta) x_{bh}^{\rho_h}]^{1/\rho_h} \quad h \in \mathcal{H} \tag{35}$$

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, which is however an endogenous variable.[50]

Again we 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 look very similar to that shown in

---

[48] For an outline of the theoretical framework of the standard labour labour supply model and survey of empirical applications, see Blundell and MaCurdy (1999), Section 7.

[49] The arbitariness of this assumption was recognised by Stern (1976) who noted the potential sensitivity of estimated parameters to the assumed productivity or quality of "leisure".

[50] Heckman (1974) notes that data 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.These assumptions also appear in extensions of the unitary model to collective empirical applications with child care. See, for example, Cherchye et al. (2012) and the literature cited there.

Figure 2 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 match their data means. This outcome is achieved by setting $\rho_{1h} = -0.7$ and $\rho_{2h} = 0.3$. Table 3 presents results for the optimal marginal tax rates and lump sums, bracket points and aggregate social welfare values, SW, under piecewise linear joint and individual taxation for $\pi = 0.2$ to 0.6.

Table 3 about here

The values selected for ρ_{1h} and ρ_{2h} 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, p_z. Even though the gap between the price, p_z, faced by H1 and H2 households is significantly wider than that for Model 1 due to setting due to setting $k_h = 1$, 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 consistently superior to joint taxation, and increasingly so as we increase the degree of inequality aversion, π, from 0.2 to 0.6. Again 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.

As in the preceding Model 1, 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 H1 and H2 household within a given primary wage percentile. In other words, together with 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 the population live in households that 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.

This paper has presented a new model of the household that focuses on child care as the form of household production.[51] 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.

We have also shown that it is essential to take account of the shape of the primary earner wage distribution, since the degree of inequality in this, which has greatly increased over the past few decades, determines the fairness with which each tax system distributes the tax burden across households. We have argued that in the light of this 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, we consistently find that the optimal piecewise linear tax system based on individual income yields a higher social welfare than the optimal system based on joint income.

A central result of both the theoretical and numerical analyses is the deeper understanding of the deficiencies of income-splitting from the equity point of view. These have become accentuated, not only by the growth in gross wage and income inequality, but also by the reductions in top tax rates, further increasing net income inequality, that have occurred in the past few decades. The opportunity to reduce one's tax liability that the joint taxation system offers to very high wage-earners severely constrains the extent to which the tax system can achieve an equitable allocation of the tax burden between households in the low-to-middle region of the distribution of well-being and households at the top. The sharp differences in the marginal rate structures of the two tax systems reflect the compositions of the subsets of individuals in the respective brackets, with second earners with high elasticities and primary earners with low elasticities being pooled under joint taxation and separated under individual taxation. Removal of the income-splitting advantage leads to high wage primary earners being placed in the upper tax bracket regardless of the incomes of their spouses, with a consequential sharp increase in the redistributional possibilities made available by the tax system. In this way the tax system would counteract, rather than exacerbate, the effects of the trend in inequality of pre-tax wages and incomes.

The household model we have specified is based on a "unitary" approach to the household, in which we effectively assume that the welfare weights the

---

[51] In carrying out the numerical calculations in Section 4 we have used labour supply elasticities. There is a recent literature 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.

household attaches to the wellbeing of the two individuals correspond to those applied by the tax authority or "planner".[52] 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,[53] 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.

References

[1] Alesina, Alberto, Andrea Ichino, and Loukas Karabarbounis. 2011. "Gender Based Taxation and the Division of Family Chores". *American Economic Journal: Economic Policy* 3 (2): 1-40.

[2] Almond, Douglas, and Janet Currie. 2011. "Human Capital Development Before Age Five." In *Handbook of Labor Economics*, Vol 4b, edited by Orley C. Ashenfelter and David Card, Ch. 5, 1315-1686. Amsterdam: Elsevier.

[3] Andrienko, Yuri, Patricia Apps, and Ray Rees. 2016. "Optimal Taxation and Top Incomes." *International Tax and Public Finance* 23 (6): 981-1003.

[4] Apps, Patrica. 1982. Institutional Inequality and Tax Incidence. *Journal of Public Economics*, 18 (2), 217-42.

[5] Apps, Patricia, Ngo Van Long, and Ray Rees. 2014. "Optimal Piecewise Linear Income Taxation." *Journal of Public Economic Theory* 16 (4): 523-545.

[6] Apps, Patricia, and Ray Rees. 2009. *Public Economics and the Household*. Cambridge: Cambridge University Press.

[7] Apps, Patricia, and Ray Rees. 1999. "On the Taxation of Trade within and between Households." *Journal of Public Economics* 73 (2): 241-63.

[8] Apps, Patricia, and Ray Rees. 1988. "Taxation and the Household." *Journal of Public Economics* 35 (3): 355–69.

[9] Atkinson, Anthony B. 2015. *Inequality: What Can Be Done*. Cambridge, Mass: Harvard University Press.

---

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

[53] This 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).

[10] Atkinson, Anthony B, Thomas Piketty, and Emmanuel Saez. 2011. "Top incomes in the Long Run of History." *Journal of Economic Literature* 49 (1): 3-71.

[11] Attanasio, Orazio, Hamish Low, and Virginia Sanchez-Marcos. 2008. "Explaining Changes in Female Labor Supply in a Life-Cycle Model." *American Economic Review* 98 (4): 1517-52.

[12] Australian Bureau of Statistics. 2009-10. *Household Expenditure Survey Australia 2009-10.* Australian Bureau of Statistics – MiCRO. http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Microdata.

[13] Basu, Kaushik. 2006. Gender and Say: A Model of Household Behaviour with Endogenously Determined Balance of Power. *The Economic Journal*, 116, 558-580.

[14] Gugl, E., 2009. Income splitting, specialization and Intrafamily distribution, Canadian Journal of Economics, 42(3), 1050-1071.

[15] Blundell, Richard, and Thomas MaCurdy. 1999. "Labor Supply: A Review of Alternative Approaches." In *Handbook of Labor Economics*, Vol 3., edited by Orley C. Ashenfelter and David Card, 1559-1695. Amsterdam: Elsevier.

[16] Boskin, Michael. 1975. "Efficiency Aspects of the Differentiated Tax Treatment of Market and Household Activity." *Journal of Public Economics* 4 (1): 1-25.

[17] Boskin, Michael, and Eytan Sheshinski. 1983. "Optimal Tax Treatment of the Family: Married Couples." *Journal of Public Economics* 20 (3): 281-297.

[18] Cherchye, Laurens, Bram De Roch, and Frederic Vermeulen. 2012. "Married with Children: A Collective Labor Supply Model with Detailed Time Use and Intrahousehold Expenditure Information." *American Economic Review* 102 (7): 3377-3405.

[19] Chetty, Raj. 2009. "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods." *Annual Review of Economics* 1: 451-488.

[20] Dahlby, Bev. 1998. "Progressive Taxation and the Marginal Social Cost of Public Funds." *Journal of Public Economics* 67 (1): 105-122.

[21] Dahlby, Bev. 2008. *The Marginal Cost of Public Funds.* Cambridge: MIT Press.

[22] Gugl, Elizabeth. 2009. "Income Splitting, Specialization and Intrafamily Distribution, *Canadian Journal of Economics* 42 (3), 1050-1071.

[23] Heckman James J. 1974. "Effects of Child Care Programs on Women's Work Effort." *Journal of Political Economy* 82 (2): S236-S163.

[24] Heckman James J., and Mark Killingsworth. 1986. "Female Labor Supply: A Survey." In *Handbook of Labor Economics*, Vol. 1., edited by Orley Ashenfelter and Richard Layard, 103-204, Amsterdam and New York: North Holland.

[25] LaLumia, Sara. 2008. "The Effects of Joint Taxation of Married Couples on Labor Supply and Non-Wage Income, *Journal of Public Economics* 92 (7): 1698-1791.

[26] Lundberg, Shelly, Robert Pollak, and Jenna Stearns. (2016). "Family Inequality: Diverging Patterns in Marriage, Cohabitation and Childbearing." *Journal of Economic Perspectives* 30(2): 79-102.

[27] Lundborg, Petter, Anton Nilsson, and Dan-Olof Rooth. 2014. "Parental Education and Offspring Outcomes: Evidence from the Swedish Compulsory Schooling Reform." *American Economic Journal: Applied Economics* 6 (1): 253-278.

[28] Mirrlees, James. 1971. An Exploration in the Theory of Optimum Income Taxation, *Review of Economic Studies* 38: 175-208.

[29] Moffitt, Robert A., and Mark O. Wilhelm. 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.

[30] Peter, Klara S., Steve Buttrick, and Denvil Duncan. 2010. "Global Reform of Personal Income Taxation, 1981-2005: Evidence from 189 Countries." *National Tax Journal* 63(3): 447-478.

[31] Piketty, Thomas, and Emmanuel Saez. 2003. "Income Inequality in the United States, 1913-1998." *Quarterly Journal of Economics* 118 (1): 1-39.

[32] Piketty, Thomas, Emmanuel Saez, and Stefanie Stantcheva. 2014. Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities." *American Economic Journal: Economic Policy* 6 (1): 230-271.

[33] Saez, Emmanuel, Joel Slemrod, and 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.

[34] Sandmo, A., 1990. "Tax Distortions and Household Production." *Oxford Economic Papers* 42: 78-90.

[35] Sheshinski, Eytan. 1972. "The Optimal Linear Income Tax." *Review of Economic Studies* 39: 297-302.

[36] Sheshinski, Eytan. 1989. "Note on the Shape of the Optimum Income Tax Schedule." *Journal of Public Economics* 40 (2): 201-215.

[37] Slemrod, Joel, Shlomo Yitzhaki, Joran Mayshar, and Michael Lundholm. 1994. "The Optimal Two-Bracket Linear Income Tax." *Journal of Public Economics* 53 (2): 269-290.

[38] Steiner, Viktor, and Katharina Wrohlich. 2008. Introducing Family Tax Splitting in Germany: How Would It Affect the Income Distribution, Work Incentives, and Household Welfare?" *FinanzArchiv* 64(1): 115-142.

[39] Steiner, Viktor, and Katharina Wrohlich. 2004. :Household Taxation, Income Splitting and Labor Supply Incentive: A Microsimulation Study for Germany." *CESifo Economic Studies* 50: 541-568.

[40] Stern, Nicholas. 1976. "On the Specification of Optimum Income Taxation." *Journal of Public Economics* 6 (1-2): 123-162.

[41] Tuomala, Matti. 1984. "On the Optimal Income Taxation: Some Further Numerical Results." *Journal of Public Economics* 23 (3): 351-366.

FIGURES

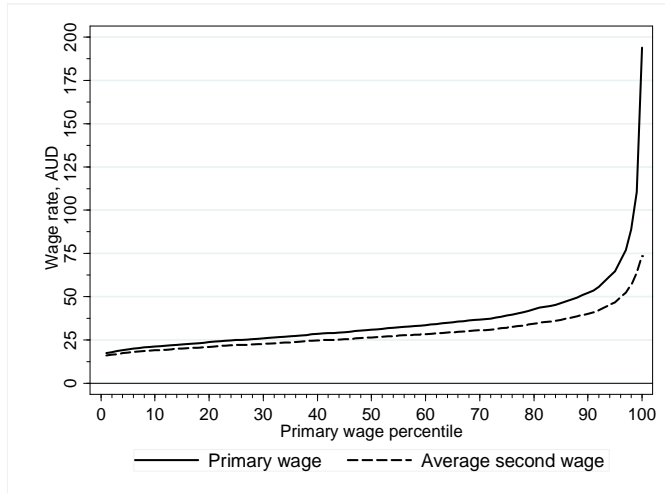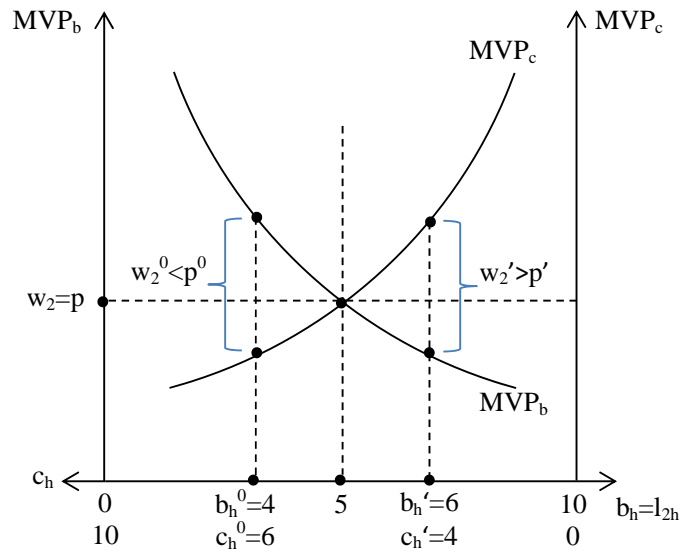## Figure 1   Percentile wage distributions

Figure 1   Percentile wage distributions

Figure 2   Model 1

Figure 2   Model 1

TABLES

MODEL 1: Child care price variations

Table 1:   Model 1 with ρ = 0.7 and p = p±0.25w$_2$

π	Tax system	τ_1, t_1	τ_2, t_2	α, a	Bkt[+]	SW	H1 hr[a]	H2 hr[b]
0.2	Joint	0.05	0.14	5778	97	87302	1.80	7.41
	Individual	0.04	0.43	6908	95	87491	1.87	7.50
0.3	Joint	0.07	0.19	7876	97	223061	1.67	7.20
	Individual	0.05	0.48	8885	95	224236	1.80	7.41
0.4	Joint	0.09	0.21	9875	97	781724	1.55	7.00
	Individual	0.06	0.52	10263	92	786120	1.73	7.31
0.5	Joint	0.11	0.22	11796	97	4539888	1.44	6,79
	Individual	0.06	0.53	11639	87	4574157	1.73	7.31
0.6	Joint	0.13	0.34	13704	99	63786656	1.35	6.55
	Individual	0.07	0.56	12875	87	64393808	1.67	7.22

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

Table 2:   Model 1 with ρ = 0.5 and p = p±0.25 and p±0.50w$_2$

π	Tax system	τ_1, t_1	τ_2, t_2	α, a	Bkt[+]	SW	H1 hr[a]	H2 hr[b]
0.2	Joint	0.07	0.22	8252	97	209981	1.64	7.90
	Individual	0.06	0.44	8671	96	210313	1.61	7.86
0.3	Joint	0.10	0.24	11218	97	607343	1.51	7.73
	Individual	0.08	0.50	11286	95	609026	1.58	7.82
0.4	Joint	0.14	0.34	15281	97	2511256	1.40	7.52
	Individual	0.10	0.56	13558	95	2521040	1.15	7.73
0.5	Joint	0.16	0.36	17216	97	18383842	1.35	7.42
	Individual	0.11	0.56	15204	93	18475026	1.49	7.67
0.6	Joint	0.18	0.37	19097	97	365424644	1.31	7.30
	Individual	0.13	0.59	17301	92	367662752	1.44	7.57

+: Income percentile of bracket point; **a**: median H1 second hours for p = p±0.50w$_2$;
b: median H2 second hours for p=p±0.50w$_2$

MODEL 2: Variation in ρ_h

Table 3:   Model 2 with ρ_1 = -0.7, ρ_2 = 0.3

π	Tax system	τ_1, t_1	τ_2, t_2	α, a	Bkt[+]	SW	H1 hr[a]	H2 hr[b]
0.2	Joint	0.04	0.09	4569	96	57706	1.66	7.54
	Individual	0.03	0.44	5949	95	57947	1.75	7.68.
0.3	Joint	0.06	0.11	6574	97	147168	1.52	7.24
	Individual	0.03	0.48	7064	92	148267	1.75	7.68
0.4	Joint	0.07	0.12	7566	97	514909	1.45	7.08
	Individual	0.04	0.50	9120	89	520577	1.66	7.54
0.5	Joint	0.08	0.14	8538	97	2986014	1.39	6.92
	Individual	0.04	0.52	9821	87	3030118	1.66	7.54
0.6	Joint	0.09	0.16	9477	98	41897056	1.34	6.72
	Individual	0.05	0.55	11077	87	42671180	1.59	7.39

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

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

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

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

where:

$$k_h = k(w_{2h}) + \kappa_h > 0, \quad p_h = p(q_h) + \varepsilon_h > 0 \tag{3}$$

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

Writing the Lagrange function as

$$\mathcal{L}_h = u_h + \lambda_h(w_{1h} l_{1h} + (w_{2h} - p_h) l_{2h} - x_h) \tag{4}$$

the first order conditions are

$$\frac{\partial \mathcal{L}_h}{\partial x_h} = 1 - \lambda_h = 0 \tag{5}$$

$$\frac{\partial \mathcal{L}_h}{\partial l_{1h}} = \hat{u}_1'(l_{1h}^*) - \lambda_h w_{1h} = 0 \tag{6}$$

$$\frac{\partial \mathcal{L}_h}{\partial l_{2h}} = (-\frac{\partial z}{\partial c_h} + \frac{\partial z}{\partial b_h})\hat{u}_2' + \lambda_h(w_{2h} - p_h) = 0 \tag{7}$$

$$\frac{\partial \mathcal{L}_h}{\partial q_h} = [\hat{u}_2' z_2 - \lambda_h p_h'(q_h^*)] l_{2h}^* = 0 \tag{8}$$

together with the budget constraint. The household's marginal utilty 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 conditions[1]

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

---

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

1

$$\varphi_2(l_{2h}^*, q_h^*, k_h) - p'(q_h^*)l_{2h}^* = 0 \tag{9}$$

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:

$$\frac{\partial l_{2h}^*}{\partial w_{2h}} = \frac{-(\varphi_{22} - p''l_{2h}^*)}{D} > 0; \tag{10}$$

$$\frac{\partial q_h^*}{\partial w_{2h}} = \frac{(\varphi_{21} - p')}{D} \gtreqless 0 \tag{11}$$

$$\frac{\partial l_{2h}^*}{\partial \varepsilon_h} = -\frac{\partial l_{2h}^*}{\partial w_{2h}} < 0; \quad \frac{\partial q_h^*}{\partial \varepsilon_h} = -\frac{\partial q_h^*}{\partial w_{2h}} \gtreqless 0 \tag{12}$$

where

$$D \equiv \varphi_{11}\varphi_{22} - (\varphi_{12})^2 \tag{13}$$

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

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

$$\varphi_{22} = \hat{u}_2''(z_2 l_{2h}^*)^2 + \hat{u}_2' z_{22} l_{2h}^{*2} < 0 \tag{14}$$

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

$$\varphi_{12} = \varphi_{21} = [(z_{22}q_h^* - z_{12}k_h)l_{2h}^* + z_2]\hat{u}_2' + z_2 l_{2h}^*(z_2 q_h^* - z_1 k_h)\hat{u}_2'' \gtreqless 0 \tag{15}$$

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

$$\frac{\partial l_{2h}^*}{\partial w_{2h}} = \frac{\partial l_{2h}^*}{\partial w_{2h}}\Big|_{k_h=const} + \frac{\partial l_{2h}^*}{\partial k_h}k'(w_{2h}) \tag{16}$$

$$\frac{\partial q_h^*}{\partial w_{2h}} = \frac{\partial q_h^*}{\partial w_{2h}}\Big|_{k_h=const} + \frac{\partial q_h^*}{\partial k_h}k'(w_{2h}) \tag{17}$$

2

implying:

Proposition A2:

$$\frac{\partial l_{2h}^*}{\partial w_{2h}} = \frac{-(\varphi_{22} - p'' l_{2h}^*) - [\varphi_{13}(\varphi_{22} - p'' l_{2h}^*) - \varphi_{23}(\varphi_{21} - p')]k'(w_{2h})}{D} \gtreqqless 0$$

$$\frac{\partial q_h^*}{\partial w_{2h}} = \frac{(\varphi_{21} - p') - [\varphi_{11}\varphi_{23} - \varphi_{13}(\varphi_{21} - p')]k'(w_{2h})}{D} \gtreqqless 0 \qquad (18)$$

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:

$$\varphi_{11} = (z_{11}k_h^2 - 2z_{12}k_h q_h^* + z_{22}q_h^{*2})\hat{u}_2' + (z_2 q_h^* - z_1 k_h)^2 \hat{u}_2'' < 0 \qquad (19)$$

$$\varphi_{13} = \varphi_{31} = [z_{12}(1-l_{2h}^*)q_h^* - z_{11}(1-l_{2h}^*)k_h - z_1]\hat{u}_2' + z_1(1-l_{2h}^*)(z_2 q_h^* - z_1 k_h)\hat{u}_2'' \gtreqqless 0 \qquad (20)$$

$$\varphi_{23} = \varphi_{32} = z_1 z_2 (1 - l_{2h}^*)l_{2h}^* \hat{u}_2'' + z_{12}l_{2h}^*(1 - l_{2h}^*)\hat{u}_2' \gtreqqless 0 \qquad (21)$$

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

To present the results of the analysis[2] 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

$$\hat{u}_2[z(k_h c_h, q_h b_h)] \equiv \hat{u}_2[z(k_h(1 - y_{2h}/w_{2h}), q_h y_{2h}/w_{2h})] \equiv -\psi_{2h}(y_{2h}; w_{2h}) \qquad (22)$$

where the $\psi_{ih}(.)$ are strictly increasing and convex in y_{ih}. We retain the assumption of identical preferences $\hat{u}_1(.)$, $\hat{u}_{2h}(.)$ across households, as is usual in optimal tax analysis.[3]

---

[2] Throughout 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.

[3] In 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

$$\max_{x_h, y_{ih}} u_h = x_h - \sum_{i=1}^{2} \psi_i(y_{ih}; w_{ih}) \tag{23}$$

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:

$$x_h = \alpha + (1 - \tau_1) \sum_i y_{ih} \tag{24}$$

and the first order conditions imply:

$$\frac{\partial \psi_i}{\partial y_{ih}} = 1 - \tau_1 \quad i = 1, 2, \tag{25}$$

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

$$\frac{\partial y_{ih}(\tau_1)}{\partial \tau_1} < 0, \quad i = 1, 2, \tag{26}$$

where, note, this is a compensated derivative.

We write the household indirect utility function[4] as $v_h(\alpha, \tau_1)$, with, by the Envelope Theorem,

$$\frac{\partial v_h}{\partial \alpha} = 1; \quad \frac{\partial v_h}{\partial \tau_1} = -y_h^* = -\sum_i y_{ih}(\tau_1) \quad i = 1, 2, \tag{27}$$

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.[5] We can write the first order conditions as

$$(1 - \tau_1) - \frac{\partial \psi_{ih}}{\partial y_{ih}} - \mu_h = 0 \quad i = 1, 2, \tag{28}$$

$$y_h \leq \eta \quad \mu_h \geq 0 \quad \mu_h[\, y_h - \eta] = 0 \tag{29}$$

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

---

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

[5] Case 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,

$$\frac{\partial v_h}{\partial \alpha} = 1; \quad \frac{\partial v_h}{\partial \tau_1} = -\eta; \quad \frac{\partial v_h}{\partial \eta} = (1 - \tau_1) - \frac{\partial \psi_i}{\partial y_{ih}} \geq 0 \tag{30}$$

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

$$x_h \leq \alpha + (1 - \tau_2)y_h + (\tau_2 - \tau_1)\eta \tag{31}$$

and the first order conditions imply

$$\frac{\partial \psi_{ih}}{\partial y_{ih}} = 1 - \tau_2 \quad i = 1, 2, \tag{32}$$

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

$$\frac{\partial y_{ih}(\tau_2, w_{ih})}{\partial \tau_2} < 0, \quad \frac{\partial y_{ih}(\tau_2, w_{ih})}{\partial w_{ih}} > 0 \quad i = 1, 2, \tag{33}$$

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

$$\frac{\partial v_h}{\partial \alpha} = 1; \quad \frac{\partial v_h}{\partial \tau_1} = -\eta; \quad \frac{\partial v_h}{\partial \tau_2} = -(y_h^* - \eta); \quad \frac{\partial v_h}{\partial \eta} = \tau_2 - \tau_1 > 0 \tag{34}$$

In all three cases, it follows from the properties of the function $\psi_{ih}(.)$ 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

$$\frac{\partial v_h}{\partial a} = 1; \quad \frac{\partial v_h}{\partial t_1} = -(y + y_{2h}^*); \quad \frac{\partial v_h}{\partial y} = (1 - t_1) - \frac{\partial \psi_1}{\partial y_{1h}} \qquad (35)$$

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

$$\frac{\partial v_h}{\partial a} = 1; \quad \frac{\partial v_h}{\partial t_1} = -2y; \quad \frac{\partial v_h}{\partial y} = 2(1 - t_1) - \sum_i \frac{\partial \psi_i}{\partial y_{ih}} \qquad (36)$$

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

$$x_h \leq a + (t_2 - t_1)y + (1 - t_2)y_{1h} + (1 - t_1)y_{2h} \qquad (37)$$

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

$$\frac{\partial v_h}{\partial a} = 1; \quad \frac{\partial v_h}{\partial t_1} = -(y + y_{2h}^*); \quad \frac{\partial v_h}{\partial t_2} - (y_{1h}^* - y); \quad \frac{\partial v_h}{\partial y} = t_2 - t_1 \qquad (38)$$

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

$$\frac{\partial v_h}{\partial a} = 1; \quad \frac{\partial v_h}{\partial t_1} = -2y; \quad \frac{\partial v_h}{\partial t_2} - (y_{1h}^* - y); \quad \frac{\partial v_h}{\partial y} = t_2 - t_1 + (1 - t_1) - \frac{\partial \psi}{\partial y_{2h}} \qquad (39)$$

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

$$\frac{\partial v_h}{\partial a} = 1; \quad \frac{\partial v_h}{\partial t_1} = -2y; \quad \frac{\partial v_h}{\partial t_2} - \sum_i (y_{ih}^* - y); \quad \frac{\partial v_h}{\partial y} = 2(t_2 - t_1) \qquad (40)$$

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