Working Papers

www.cesifo.org/wp

# Optimal Taxation, Child Care and Models of the Household

Patricia Apps
Ray Rees

CESIFO WORKING PAPER NO. 4578
CATEGORY 1: PUBLIC FINANCE
JANUARY 2014

CESifo *Working Paper No. 4578*

# Optimal Taxation, Child Care and Models of the Household


## Abstract

This paper presents the properties of optimal piecewise linear tax systems for two-earner households, based on joint and individual incomes respectively. A key contribution is the analysis of the interaction between second earner wage differences, variation in the price of child care and domestic productivity differences as determinants of across-household heterogeneity in second earner labour supply, and of the resulting relationship between household income and the wellbeing of household members. A central result is that taking account of a richer and more realistic specification of household time use widens the set of cases in which individual taxation is welfare-superior.


JEL-Code: H210, J220, H310, H240.

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

*Patricia Apps*
*University of Sydney Law School*
*Sydney / Australia*
*p.apps@usys.edu.au*

*Ray Rees**
*Professor Emeritus, University of Munich*
*& Associate Fellow, University of Warwick*
*ray.rees@lrz.uni-muenchen.de*

*corresponding author

December 13, 2013
We would like to thank Craig Brett and Vidar Christiansen for helpful comments on an earlier draft of this paper, and Yuri Andrienko for excellent research assistance. The work was supported under the Australian Research Council's Discovery Project funding scheme (DP1094021).

*The current [income taxation] system is considered unfair because it imposes the same tax burden on a married couple with one earner as it does on a two-earner couple with the same income. The two-earner couple will in general have more total hours of work and less of the untaxed home services of the second earner.*[1]

# 1 Introduction

The early critique[2] of joint income as the base for taxing two-earner households in the US focused on a straightforward application of the Ramsey principle of optimal taxation: If women have higher labour supply elasticities than men then they should be subject to lower tax rates. However, since optimal income tax rates are meant to reflect considerations of equity as well as efficiency this is not a conclusive argument. The proposition advanced by Feldstein and Feenberg in the above quotation takes the argument a step further, as well as placing the discussion in a broader context. Joint taxation can also be criticised on equity grounds, if we take into account the existence of untaxed household production. Moreover, it suggests the idea that a household's income is not an accurate indicator of its achieved utility or standard of living. The statement implies that two households with the same income may have widely different living standards, and, by extension, that a household with a higher labor income could actually be worse off than one with a lower.

This paper explores these ideas formally in the context of an analysis of optimal piecewise linear income tax systems. Real tax systems are almost universally of the piecewise linear kind, in which marginal tax rates are constant within but vary between a small number of specified income brackets. Yet there has been relatively little analysis of their optimal structure,[3] and none at all of the two-earner household case.[4] This paper analyses the optimal two-bracket piecewise linear tax system for two-earner households with the aim of bringing out the importance of the structural form of the underlying household model, and in particular the extent to which it captures the idea underlying Feldstein and Feenberg's proposition, in determining the main features of the system. Given its empirical importance, we take child care as the canonical form of household production.[5]

The paper proceeds in two steps. First we present two alternative structural household models and characterise the optimal piecewise linear tax systems

[1] From Feldstein and Feenberg (1996).

[2] See for example Rosen (1977), Munnell (1980) and Boskin and Sheshinski (1983).

[3] The main references are Sadka (1976), Sheshinski (1989), Dahlby (1998), (2008) and Apps, Long and Rees (2011), all of which deal only with single-person households.

[4] See the comprehensive survey of modern tax theory in Boadway (2012) for confirmation of this.

[5] We focus on child care as the specific form of household production because the data show that the phase of the life cycle in which young children are present in the household is the one in which crucial decisions on second earner labour supply are taken, that then have persistent effects over the remainder of the life cycle. See Apps and Rees (2009) for further discussion.

for the cases of joint and individual taxation respectively, in the context of a "reduced form" household model in which both structural models are nested. These results are new to the literature. We then show how the comparison of the welfare properties of the two tax systems depends on the specific assumptions of the underlying structural model on the productivity of non-market time and the price of bought in substitutes.

Model 1 is the standard household labour supply model.[6] In this each individual's time is divided between market work and leisure, with the latter measured as the time not spent in market work, and treated as the same consumption good across all households. The implicit assumption is that the productivity of non-market time is constant across households.[7] Across-household heterogeneity in labour supply decisions is therefore driven solely by wage rates, and maximised household utility is strongly positively associated with household income. When this model is used as the basis for an optimal tax analysis, and with standard stylised facts on the compensated labour supply elasticities of primary and second earners, there will be gains in efficiency in moving from optimal joint to optimal individual taxation. However, since this move tends to redistribute the tax burden from households with a higher to those with a lower second earner labour supply, the equity effects may be adverse and outweigh the efficiency gains.

In Model 2 we seek to reflect the data on the time use and expenditure decisions of two-earner couples with at least one young child present in the household. In such households parental child care is a major form of time use and bought in child care can be a large component of household expenditure. In contrast to Model 1, but consistent with the findings of empirical studies on the relationship between child outcomes and parental human capital,[8] we allow the productivity of parental child care to rise with the wage. We show that the inputs to household production can vary widely across households with the same wage rates and demographic characteristics in response to varying productivities and child care prices. The result is that maximised household utilities may no longer track household incomes. The adverse equity effects of a move from optimal joint to optimal individual taxation under Model 1 may therefore be replaced by distributional improvements. In general, we find that the analysis of marginal rate progressive piecewise linear tax systems in the presence of a realistic system of household production supports the case for individual taxation. It is in this sense that we confirm the Feldstein-Feenberg proposition.

The paper is set out as follows. In the next section we present the two structural household models that provide the analytical basis for the indirect utility and labour supply functions used in the tax analysis. In the following two sections we define the tax systems, characterise households' optimal allocations under each of them, and carry out the optimal tax analysis for joint and

[6] As used for example by Boskin and Sheshinski (1983).

[7] As Stern (1976) notes, the absence of across-household variation in the quality (or productivity) of "leisure" is a crucial assumption of the standard labor supply model.

[8] See, for example, the survey by Almond and Currie (2011).

individual taxation respectively. Section 5 interprets the results of the optimal tax analysis for each of the two household models. In Section 6 we present the results of an illustrative numerical analysis of the optimal tax systems. Section 7 concludes.

# 2 Two Household Models

In Model 1 the two adults divide their time between market work and leisure, where the latter has a constant quality or, equivalently, productivity across all households, but a variable price given by the net wage. In contrast, the productivity of market time is given by the gross wage, and the price of every market good of a given quality is the same across all households. "Household type" is therefore defined by gross wage pairs alone.

In Model 2 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.[9] We define child care broadly, to denote not just physically looking after the child, but also to include all the activities that contribute to the child's welfare and development of human capital. In contrast to Model 1, the productivity of the second earner's time input to child care varies exogenously across households with the same second earner wage, with a distribution that shifts upward with her wage. There is in addition a bought in child care time input, the quality of which increases with the second earner wage, to capture the idea that the second earner will prefer to substitute for her own child care input a market input of similar quality. The price of this input at each quality also varies exogenously across households, while increases in quality shift the distribution of prices upward. Thus second earner productivity and price of the market input are further dimensions of household type. We now set out the models more formally.

## 2.1 Model 1

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 $y$ from their labour supply. $P$ types of primary and $S$ types of second earners are defined by their wage rates, with

$$w_1 \in \{w_1^1, w_1^2, ..., w_1^P\}, w_2 \in \{w_2^1, w_2^2 ..., w_2^S\}, w_2^1 < w_1^1, w_2^S < w_1^P \tag{1}$$

and in every household $w_2 < w_1$. Subject to this restriction, household type is then defined by the pair $(w_1, w_2)$. Let $h$ index these pairs $(w_{1h}, w_{2h})$ lexico-

[9] The pronouns reflect the data, in OECD countries typically more than three quarters of second earners are women. 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.

graphically so that, for any pair of indices $h, h'$,

$$h > h' \Leftrightarrow w_{1h} > w_{1h'} \text{ or } w_{1h} = w_{1h'} \text{ and } w_{2h} > w_{2h'} \; i = 1, 2, \quad h = 1, ..., H$$

with $h = 1$ for $(w_1^1, w_2^1)$. This convention determines how household welfare,[10] labour supply and income will vary with $h$. Note that it does not imply that household income increases monotonically with $h$, since one household may have a higher primary wage than another but a sufficiently lower second wage that household income is lower.

The household's utility function[11] is

$$u_h = x_h - \sum_{i=1}^{2} u_i(l_{ih}) \quad h = 1, ..., H \tag{2}$$

where the $u_i(.)$ are identical across households for given $i$, strictly increasing and strictly convex in labour supplies $l_{ih}$. Since $y_{ih} = w_{ih} l_{ih}$ we rewrite the utility function as

$$u_h = x_h - \sum_{i=1}^{2} u_i(y_{ih}/w_{ih}) = \; x_h - \sum_{i=1}^{2} \psi_i(y_{ih}, w_{ih}) \; h = 1, ..., H \tag{3}$$

where the $\psi_i(.)$ are strictly increasing and convex and possess the single-crossing property

$$\frac{\partial}{\partial w_{ih}} [\frac{\partial \psi_i}{\partial y_{ih}}] < 0 \quad i = 1, 2, \quad h = 1, ..., H \tag{4}$$

This says that the higher the wage type, the lower the marginal effort cost to $i$ of achieving a given increase in labour earnings.

The household budget constraint is given by

$$x_h \leq \sum_{i=1}^{2} y_{ih} - T(y_{1h}, y_{2h}) \quad h = 1, ..., H \tag{5}$$

where the tax function $T(y_{1h}, y_{2h})$ is further specified below.

We retain the assumption of identical preferences across households, as is usual in optimal tax analysis, by assuming all primary earners have the same preferences and similarly for second earners. However we allow the preferences of primary and second earners within a household to differ.[12] Given identical preferences for second earners, heterogeneity across households in second earner labour supply and income at a given primary earner wage is driven entirely by variation in the second earner wage. A household with lower second earnings, and therefore a lower household income, than another with the same primary income must have a lower second wage and must therefore be worse off.

[10] Of course, only individuals can have "welfare", but we use this term as a shorthand to refer to the utility pair that a household can achieve.

[11] The quasilinear and additively separable form assumed here, though special, is very convenient, since it eliminates income effects and greatly simplifies the presentation of the optimal tax formulas.

[12] Emprically, male and female labour supplies differ to an extent that cannot plausibly be explained by wage variation with identical preferences.

## 2.2 Model 2

In addition to the market consumption good $x$, household utility depends on child care, $z$, which is produced using the second earner's time input, $c$, and a bought-in child care time input, $b$, according to a standard strictly quasiconcave and increasing production function

$$z_h = z(k_h c_h, q_h b_h) \tag{6}$$

where $k_h$ and $q_h$ are measures of the productivity/quality in child care of $c_h$ and $b_h$ respectively. As just mentioned, for each $w_2 \in \{w_2^1, w_2^2..., w_2^S\}$ there is a quality $q_j$ and a distribution of qualities $\{k_j^1, ..., k_j^{n_j}\}$, while for each quality $q_j$ there is a distribution of prices $\{p_j^1, ..., p_j^m\}$ $j = 1, .., S$. Thus this model adds two further dimensions to household type, which now depends on the vector of variables $(w_1, w_2, p, k)$. We extend the previous method for defining the type index $h$ by again taking a lexicographic ordering such that, for any pair $h$, $h'$

$$\begin{aligned} h &> h' \Leftrightarrow w_{1h} > w_{1h'} && (7)\\ \text{or } w_{1h} &= w_{1h'} \text{ and } w_{2h} > w_{1h'} && (8)\\ \text{or } w_{1h} &= w_{1h'} \text{ and } w_{2h} = w_{1h'} \text{ and } p_h > p_{h'} && (9)\\ \text{or } w_{1h} &= w_{1h'} \text{ and } w_{2h} = w_{1h'} \text{ and } p_h = p_{h'} \text{ and } k_h > k_{h'} && (10) \end{aligned}$$

with $h = 1$ for $(w_1^1, w_2^1, p^1, k_1)$. Thus, in this model, at any given primary earner wage rate, across-household heterogeneity is driven by price and productivity variation as well as by second earner's wage variation.

The household's maximised utility increases *ceteris paribus* monotonically with increasing wage rates and productivity and decreasing child care price, as shown in Section 4 below. However, the relationship between household income and maximised utility is no longer necessarily positive or monotonic. It depends on exactly how changes in a wage rate, productivity or price of child care of a given quality cause changes in labour supply, income and maximised household utility. We explore this further in Section 4.

The household utility function is now given by

$$u_h = x_h - \psi_1(y_{1h}, w_{1h}) + \hat{u}(z_h) \quad h = 1, ..., H \tag{11}$$

The $\hat{u}(.)$ function is strictly increasing and strictly concave. For the second earner, the time spent in market work and child care must sum to the total time endowment, normalised at 1, and so we have

$$c_h + l_{2h} = 1 \qquad h = 1, ..., H \tag{12}$$

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

There is however a further 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 $b_h = l_{2h}$. Recalling that

$y_{2h} = w_{2h}l_{2h}$, we can use these time constraints to eliminate $c_h$ and $b_h$ and rewrite $\hat{u}(.)$ as

$$\hat{u}[z(k_h c_h, q_h b_h)] \equiv \hat{u}[z(y_{2h}/w_{2h}, 1 - y_{2h}/w_{2h}; q_h, k_h)] \equiv -\psi_2(y_{2h}; w_{2h}, q_h, k_h) \tag{13}$$

Writing the household budget constraint as

$$x_h + p_h b_h = x_h + p_h y_{2h}/w_{2h} \leq \sum_{i=1}^{2} y_{ih} - T(y_{1h}, y_{2h}) \quad h = 1, ..., H \tag{14}$$

we again have a model that can be used to derive the household's indirect utility function with the tax parameters as arguments.

## 2.3 Application to tax analysis

In applying these two household models to the optimal tax analysis, the key relationships are households' indirect utility functions and their derivatives with respect to the tax parameters. The specifics of these will depend on whether we have individual or joint taxation. However, we can show that for Model 2 it is possible to write the expressions for the derivatives of indirect utility with respect to the tax parameters in each case in exactly the same form as for Model 1, despite the radical differences in the underlying structural forms of the two models.[13] This leads to a considerable economy of effort in deriving the optimal tax conditions, but, as we emphasise, this should not be at the cost of drawing the false conclusion that the results of the two models are "essentially" the same.

To see this, consider the Lagrange functions corresponding to the household optimisation problems in the cases of Model 1 and Model 2 respectively:

**Model 1:**

$$L_h = u_h + \lambda_h[\sum_{i=1}^{2} y_{ih} - T(y_{1h}, y_{2h}) - x_h] \tag{15}$$

**Model 2:**

$$\mathcal{L}_h = u_h + \lambda_h[\sum_{i=1}^{2} y_{ih} - T(y_{1h}, y_{2h}) - x_h - p_h y_{2h}/w_{2h}] \tag{16}$$

Since the tax parameters do not enter the utility functions $u_h$ in either problem, and $p_h$ is taken as exogenously given throughout,[14] by the Envelope Theorem

[13] See also Sandmo (1990). This is because the "sufficient statistics" for the optimal taxes, in the sense of Chetty (2009), are just the derivatives of earnings/labour supplies with respect to the tax parameters, the marginal social utilities of incomes of the various household types and their proportions in the population. The "reduced forms" of the tax conditions are the same for the different structural models, but these structural differences do matter profoundly, not least in determining the basis for the empirical measurement of the reduced form parameters.

[14] This would change of course if $p_h$ were to be an instrument of tax policy, which is an interesting possibility but not one we pursue in this paper.

the derivatives of the indirect utility functions $v_h(.)$ will take the same form in each model whenever the tax function $T(y_{1h}, y_{2h})$ is also of the same form. Since these derivatives are all we use in the optimal tax analysis we obtain precisely the same general form of conditions on the tax parameters whether we take Model 1 or Model 2 as the household model. What is important however is that because of the underlying model structure, both the interpretation of the optimal tax conditions and their policy implications change fundamentally.

# 3 Taxation systems

The tax system pays households a uniform lump sum funded[15] by revenue from taxes on the labour incomes of the two earners. As well as the issue of the choice of tax base, 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 space of gross income-net income/consumption to which they give rise. For the purposes of this paper we focus on the convex case of a two-bracket piecewise linear system.[16]

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 "selective taxation", under which separate optimal tax schedules are found for primary and second earners respectively.[17] The main reason for constraining the rate schedules to be identical under individual taxation is that in practice, piecewise linear tax systems that are not joint are in fact overwhelmingly of the individual rather than selective kind.[18] Moreover, if individual taxation yields higher social welfare than joint taxation under realistic assumptions, this result applies *a fortiori* to selective taxation, since removing the constraint that tax schedules must be identical cannot reduce the maximised value of social welfare and would be expected to increase it. It is not difficult to extend the results of this paper to the selective taxation case, at the cost however of a large step up in notational complexity.

[15] This could be thought of as a standard child benefit. Effectively, in this model households are assumed to have the same number of children, normalised at 1, and $z$ represents Beckerian child quality.

[16] Apps, Long and Rees (2011) show that for wage distributions such as those currently prevailing in many OECD countries convex systems are highly likely to be welfare optimal.

[17] Boskin and Sheshinski (1983) and Apps and Rees (2009) analyse this problem in the context of linear taxation. Apps and Rees (1999), (2009) also analyse the tax reform problem with two-earner households and household production. See also Alesina et al (2011), where the more expressive term "gender-based taxation" is introduced.

[18] 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, and Feldstein and Feenberg (1996). In this paper we focus on the formal tax system, leaving the issue of implicit modifications to it created by social benefit payments and withdrawal rates for future work.

## 3.1 Tax functions

The tax functions $T(y_{1h}, y_{2h})$ are specified as follows.

**Joint Taxation:**

There is a two-bracket piecewise linear tax on total household labour earnings, the parameters of which are $(\alpha, \tau_1, \tau_2, \eta)$, where $\alpha$ is a uniform lump sum paid to every household, $\tau_1$,$\tau_2$ are the marginal tax rates in the lower and upper brackets of the tax schedules, and $\eta$ is the value of *joint earnings* defining the bracket limit. Thus the household tax function $T(y_{1h}, y_{2h}) \equiv T(y_h)$, with $y_h = \sum_{i=1}^{2} y_{ih}$, is defined by:

$$T(y_h) = -\alpha + \tau_1 y_h \quad y_h \leq \eta \tag{17}$$

$$T(y_h) = -\alpha + \tau_2 y_h + (\tau_1 - \tau_2)\eta \quad y_h > \eta \quad h = 1, ..., H \tag{18}$$

**Individual Taxation:**

There is a two-bracket piecewise linear tax system now applied to individual labour earnings, the parameters of which are $(a, t_1, t_2, y)$, 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:

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

$$\hat{T}(y_{ih}) = t_2 y_{ih} + (t_1 - t_2) y \quad y_{ih} > y \quad h = 1, ..., H \tag{20}$$

and the household tax function is $T(y_{1h}, y_{2h}) \equiv -a + \sum_{i=1}^{2} \hat{T}(y_{ih})$. Note that this specification of the tax function implies that $\partial^2 T(y_{1h}, y_{2h})/\partial y_{1h} \partial y_{2h} = 0$, and so does not allow the marginal tax rate paid by one earner in the household to depend on the income of the other.[19] In what follows, as mentioned earlier, we assume that we have the convex case, in which at the tax optima $\tau_1 < \tau_2$ and $t_1 < t_2$. Every household faces the same convex budget set.

## 3.2 Household Allocations

We present the analysis of the household's choice of consumption and wage earnings under each of the two alternative tax systems, first joint and then individual taxation.

### 3.2.1 Joint Taxation

A household $h$ solves the problem

$$\max_{x_h, y_{ih}} u_h = x_h - \sum_{i=1}^{2} \psi_i(y_{ih}, w_{ih}) \tag{21}$$

[19] The analysis of optimal nonlinear taxation of couples shows that in general the marginal tax rate of one earner in the household will depend on the wage type of the other (see for example Apps and Rees (2009), and the literature cited there. Thus restriction to a piecewise linear tax system implies sacrificing some social welfare in exchange for a more practicable and implementable tax system.

subject to a budget constraint determined by the tax system, as just described. 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:

$$x_h = \alpha + (1 - \tau_1) \sum_i y_{ih} \tag{22}$$

and the first order conditions imply:

$$\frac{\partial \psi_i}{\partial y_{ih}} = 1 - \tau_1 \quad i = 1, 2, \tag{23}$$

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

$$\frac{\partial y_{ih}(\tau_1, w_{ih})}{\partial \tau_1} < 0, \quad i = 1, 2, \tag{24}$$

where, note, this is a compensated derivative.

We write the household indirect utility function[20] as $v_h(\alpha, \tau_1)$, with, by the Envelope Theorem,

$$\frac{\partial v_h}{\partial \alpha} = 1; \quad \frac{\partial v_h}{\partial \tau_1} = -y_h^* = -\sum_i y_{ih}(\tau_1, w_{ih}) \quad i = 1, 2, \tag{25}$$

**Case 2.** The household is effectively constrained at the bracket limit $\eta$, 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 $\tau_1$, but not if it would be taxed at the rate $\tau_2$. We formulate its allocation problem by adding the constraint $y_h \leq \eta$, noting that this will be binding at the optimum.[21] We can write the first order conditions as

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

$$y_h \leq \eta \quad \mu_h \geq 0 \quad \mu_h[\, y_h - \eta] = 0 \tag{27}$$

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

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

$$\frac{\partial v_h}{\partial \alpha} = 1; \quad \frac{\partial v_h}{\partial \tau_1} = -\eta; \quad \frac{\partial v_h}{\partial \eta} = (1 - \tau_1) - \frac{\partial \psi_i}{\partial y_{ih}} \geq 0 \tag{28}$$

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

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

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

exceeds for almost every individual the marginal cost of effort $\partial\psi_i/\partial y_{ih}$. In diagrammatic terms, the household is at the kink in its budget constraint at the bracket limit $\eta$. The term is zero only if $i$'s marginal rate of substitution happens to equal $(1-\tau_1)$ at the kink. Note that condition (26) implies that the individuals' marginal effort costs are equalised also in this type of equilibrium.

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

$$x_h \leq \alpha + (1-\tau_2)y_h + (\tau_2 - \tau_1)\eta \tag{29}$$

and the first order conditions imply

$$\frac{\partial\psi_i}{\partial y_{ih}} = 1 - \tau_2 \quad i = 1, 2, \tag{30}$$

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

$$\frac{\partial y_{ih}(\tau_2, w_{ih})}{\partial\tau_2} < 0, \; \frac{\partial y_{ih}(\tau_2, w_{ih})}{\partial w_{ih}} > 0 \;\; i = 1, 2, \tag{31}$$

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

$$\frac{\partial v_h}{\partial\alpha} = 1; \;\; \frac{\partial v_h}{\partial\tau_1} = -\eta; \;\; \frac{\partial v_h}{\partial\tau_2} = -(y_h^* - \eta); \;\; \frac{\partial v_h}{\partial\eta} = \tau_2 - \tau_1 > 0 \tag{32}$$

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

Given these three cases, we define a partition$\{\mathcal{H}_0, \mathcal{H}_1, \mathcal{H}_2\}$ of the index set $\{1, 2, ..., H\}$ as follows:

$$\mathcal{H}_0 = \{\ h \ \mid 0 \leq y_h^* < \eta\} \tag{33}$$

$$\mathcal{H}_1 = \{\ h \mid y_h^* = \eta)\} \tag{34}$$

$$\mathcal{H}_2 = \{\ h \ \mid y_h^* > \eta\} \tag{35}$$

where $y_h^*$ is the household's optimal income under the given tax structure. 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 and therefore, in this model, household utility are increasing as we move from $\mathcal{H}_0$ to $\mathcal{H}_1$ to $\mathcal{H}_2$, though these may not increase monotonically with $h$ as pointed out earlier. Important points to note are that:

- $\tau_1$ is a marginal tax rate for $h \in \mathcal{H}_0$ but defines an intra-marginal, non-distortionary tax for $h \in \mathcal{H}_1 \cup \mathcal{H}_2$
- A small increase in $\eta$ 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 earnings, because the budget constraint is defined only on total household income.[22]

#### 3.2.2 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

$$\frac{\partial v_h}{\partial a} = 1; \quad \frac{\partial v_h}{\partial t_1} = -(y + y^*_{2h}); \quad \frac{\partial v_h}{\partial y} = (1 - t_1) - \frac{\partial \psi_1}{\partial y_{1h}} \tag{36}$$

**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

$$\frac{\partial v_h}{\partial a} = 1; \quad \frac{\partial v_h}{\partial t_1} = -2y; \quad \frac{\partial v_h}{\partial y} = 2(1 - t_1) - \sum_i \frac{\partial \psi_i}{\partial y_{ih}} \tag{37}$$

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

$$x_h \leq a + (t_2 - t_1)y + (1 - t_2)y_{1h} + (1 - t_1)y_{2h} \tag{38}$$

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

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

**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

$$\frac{\partial v_h}{\partial a} = 1; \quad \frac{\partial v_h}{\partial t_1} = -2y; \quad \frac{\partial v_h}{\partial t_2} - (y^*_{1h} - y); \quad \frac{\partial v_h}{\partial y} = t_2 - t_1 + (1 - t_1) - \frac{\partial \psi}{\partial y_{2h}} \tag{40}$$

[22] 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.

**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

$$\frac{\partial v_h}{\partial a} = 1; \quad \frac{\partial v_h}{\partial t_1} = -2y; \quad \frac{\partial v_h}{\partial t_2} - \sum_i (y^*_{ih} - y); \quad \frac{\partial v_h}{\partial y} = 2(t_2 - t_1) \tag{41}$$

We define the partition of the index set corresponding to these six cases, $\{H_0, H_1, ..., H_5\}$, as follows:

$$H_0 = \{\ h\ \mid 0 \leq y^*_{ih} < y,\ i = 1, 2\} \tag{42}$$
$$H_1 = \{\ h\ \mid y^*_{2h} < y = y^*_{1h}\} \tag{43}$$
$$H_2 = \{\ h\ \mid y^*_{ih} = y,\ i = 1, 2\} \tag{44}$$
$$H_3 = \{\ h\ \mid y^*_{2h} < y < y^*_{1h}\} \tag{45}$$
$$H_4 = \{\ h\ \mid\ y^*_{2h} = y < y^*_{1h}\} \tag{46}$$
$$H_5 = \{\ h\ \mid y^*_{ih} > y,\ i = 1, 2\} \tag{47}$$

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

# 4 Optimal Tax Analysis

## 4.1 Joint Taxation

The planner solves

$$\max_{\alpha, \tau_1, \tau_2, \eta} W = \sum_{h=1}^{H} \phi_h S(v_h) \tag{48}$$

subject to the public sector budget constraint[23]

$$\sum_{h \in \mathcal{H}_0} \phi_h \tau_1 y_h + \sum_{h \in \mathcal{H}_1} \phi_h \tau_1 \eta + \sum_{h \in \mathcal{H}_2} \phi_h [\tau_2 y_h + (\tau_1 - \tau_2)\eta] \geq \alpha \tag{49}$$

where $\phi_h$ is the proportion of households of type $h = 1, 2, ..., H$, and $S(.)$ is a strictly concave and increasing function expressing the planner's preferences over household utilities. From the first order conditions characterising the optimal tax parameters[24] we can derive:

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

[24] 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.

**Proposition 1**: The optimal tax parameters satisfy the conditions:

$$\sum_{h=1}^{H} \phi_h(\sigma_h - 1) = 0 \tag{50}$$

$$\tau_1 = \frac{\sum_{\mathcal{H}_0} \phi_h(\sigma_h - 1)y_h^* + \eta \sum_{\mathcal{H}_1 \cup \mathcal{H}_2} \phi_h(\sigma_h - 1)}{\sum_{\mathcal{H}_0} \phi_h \partial y_h / \partial \tau_1} \tag{51}$$

$$\tau_2 = \frac{\sum_{\mathcal{H}_2} \phi_h(\sigma_h - 1)(y_h^* - \eta)}{\sum_{\mathcal{H}_2} \phi_h \partial y_h / \partial \tau_2} \tag{52}$$

$$\sum_{\mathcal{H}_1} \phi_h \{\sigma_h[(1-\tau_1) - \frac{\partial \psi}{\partial y_h}] + \tau_1\} = -(\tau_2 - \tau_1) \sum_{\mathcal{H}_2} \phi_h(\sigma_h - 1) \tag{53}$$

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

We first interpret and discuss those properties of these conditions which are common to both the models 1 and 2. In the following subsection we examine how the choice of model affects the interpretation of the conditions.

Condition (50) is familiar from linear tax theory[25]: the optimal lump sum $\alpha$ equalises the average of the marginal social utilities of household income, $\sigma_h$, in terms of the numeraire, to the marginal cost of one unit of the lump sum, which of course is 1. Denoting the shadow price of the government budget constraint by $\lambda$, $\sigma_h \equiv S'(v_h)/\lambda$, and so the concavity of $S(.)$ implies that $\sigma_h$ falls with the utility level of the household. From now on we denote $\sigma_h - 1$ by $\delta_h$. Then $\delta_h > (<)0$ according as household $h$ is relatively worse (better) off than the average *in utility terms*.[26]

The two conditions corresponding to the tax rates $\tau_1, \tau_2$, are analogous, though not of course identical, to those obtained in optimal linear tax theory. The denominators are the 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. The numerators give the equity effects. The two terms in the numerator of (51) correspond to the two ways in which the lower bracket tax rate affects the contributions households make to funding the lump sum payment $\alpha$. Given their optimal earnings $y_h^*$, the first term aggregates over subset $\mathcal{H}_0$, which is the subset with relatively lower incomes, the effect of a marginal tax rate change on welfare net of its marginal contribution to tax revenue, all in terms of the numeraire. 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 $\eta$, while leaving their (compensated) labour supply unchanged.

[25] See Sheshinski (1972).

[26] Whether this corresponds to the household having a relatively lower or higher income depends on which model, 1 or 2, underlies the analysis.

Only the first of these two effects is present in the condition (52) corresponding to the second tax rate. The portion of the income of the households in the higher tax bracket that is taxed at the rate $\tau_2$ 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 $\sigma_h$ over each of the subsets is not 1. However, intuitively they can still be thought of as measures of the strength of the relationship between the marginal social utility of income and household incomes, which determines the effectiveness of the tax rate on *income* in redistributing *utility* across households. In other words, the goal of taxation is to redistribute utility, but the available instruments are the lump sum payment and marginal tax rates on income, and so the strength of the relationship between the marginal social utility of income and income determines the effectiveness of the income tax system in redistributing utility.

It is interesting to rewrite this numerator term as

$$\sum_{\mathcal{H}_2} \phi_h \delta_h y_h^* - \eta \sum_{\mathcal{H}_2} \phi_h \delta_h \tag{54}$$

where the second term is seen to be the negative of the second term in the numerator of (51), net of the lump sum tax contribution of the subset $\mathcal{H}_1$. This suggests that the greater the contribution of the lump sum tax on upper income bracket households arising from the tax rate $\tau_1$, the smaller is the tax rate $\tau_2$, and so the smaller is the distortionary effect on labour supplies in this bracket, other things being equal.[27]

Condition (53), the condition on the bracket limit $\eta$, 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 who are effectively constrained at $\eta$. 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. If however the right hand term was not positive, then this condition could not be satisfied and this would make untenable the assumption that $(\tau_2 - \tau_1) > 0$, in other words, that the optimal

[27] 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, Long and Rees (2011).

piecewise linear tax system is indeed convex. We have ruled this possibility out by assumption, though strictly speaking it is an empirical question as to whether this is really the case.

## 4.2 Individual Taxation

The planner solves

$$\max_{a,t_1,t_2,y} \sum_{h=1}^{H} \phi_h S(v_h) \tag{55}$$

subject now to the public sector budget constraint

$$\sum_{\cup_{i=0}^{2} H_i} \phi_h t_1 y_h + \sum_{\cup_{i=3}^{4} H_i} \phi_h [t_2 y_{1h} + t_1 y_{2h} + (t_1 - t_2) y] + \sum_{H_5} \phi_h [t_2 y_h + 2(t_1 - t_2) y] \geq a \tag{56}$$

where again $y_h = \sum_{i=1}^{2} y_{ih}$. In what follows it will be useful to denote by $\mu_{ih}$ the term $(1 - t_1) - \partial\psi/\partial y_{ih}$, the value of a relaxation of the bracket limit to an individual at the kink in the budget constraint. Then from the first order conditions for an optimal solution[28] we derive:

**Proposition 2**: The optimal tax parameters in the case of individual taxation are characterised by the following conditions.

$$\sum_{h=1}^{H} \phi_h \delta_h = 0 \tag{57}$$

$$t_1^* = \frac{\sum_{H_0} \phi_h \delta_h y_h^* + \sum_{H_1 \cup H_3} \phi_h \delta_h y_{2h}^* + y^* [\sum_{H_1 \cup H_3} \phi_h \delta_h + 2 \sum_{H_2 \cup H_4 \cup H_5} \phi_h \delta_h]}{\sum_{H_0} \phi_h \partial y_{1h}/\partial t_1 + \sum_{H_0 \cup H_1 \cup H_3} \phi_h \partial y_{2h}/\partial t_1} \tag{58}$$

$$t_2^* = \frac{\sum_{H_3 \cup H_4 \cup H_5} \phi_h \delta_h (y_{1h}^* - y^*) + \sum_{H_5} \phi_h \delta_h (y_{2h}^* - y^*)}{\sum_{H_3 \cup H_4 \cup H_5} \phi_h \partial y_{1h}/\partial t_2 + \sum_{H_5} \phi_h \partial y_{2h}/\partial t_2} \tag{59}$$

$$\sum_{H_1 \cup H_2} \phi_h (\sigma_h \mu_{1h} + t_1) + \sum_{H_2 \cup H_4} \phi_h (\sigma_h \mu_{2h} + t_1) = -(t_2 - t_1) [\sum_{H_3 \cup H_4} \phi_h \delta_h + 2 \sum_{H_5} \phi_h \delta_h] \tag{60}$$

The first condition, since it involves the entire population, is exactly as for joint taxation. The remaining three conditions have basically the same interpretation as before, but of course the relevant sums are now over subsets of individuals reflecting the partition defined in the previous section. In particular, both numerator and denominator of the expression for $t_1^*$ contain terms corresponding to lower wage second earners in households with higher wage primary earners who are in the higher tax bracket. Such households may well have lower total incomes than households with both earners in the lower tax bracket, but be paying more tax. The welfare interpretation of this will depend however

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

on which of the two models is the basis for the analysis, as we further discuss in the next section.

By comparing the denominators of the expressions in (51), (52), (58), and (59), and given the stylised fact that second earners' labour supplies are significantly more sensitive to net wage rate changes than those of primary earners, we see that as between the cases of joint and individual taxation the denominators of the lower tax rate will tend to increase and those of the higher tax rate to fall as a result of the switch of second earners to the lower tax bracket. This implies, other things being equal, a fall in the lower bracket tax rate relative to that in the higher bracket, and so an increase in the progressivity of the tax system. On the other hand, the interpretation of the equity effects arising from the change will depend closely on the underlying structural model, and so we now turn to an explicit analysis of these.

# 5 Tax system comparisons

The equity effects of the finer matching of individuals with tax brackets that results from the switch between tax systems are complex, not least because when we maintain revenue neutrality the changes in tax parameters and bracket limits are discrete.[29] For this reason we present a numerical analysis in the next section. The precise distribution of gains and losses will depend on how the tax rates and bracket limit change. The central theoretical question however is: How good is income as an indicator of household welfare when joint incomes are equal, or at least fairly close together, but second earner incomes are very unequal? The answer depends on the model we use.

## 5.1 Equity effects in Model 1

As we showed earlier, in this model we have $\partial y_{ih}/\partial w_{ih} > 0$ and we can also show using (3), (15) and the Envelope Theorem that

$$\frac{\partial v_h}{\partial w_{ih}} = l_{ih}^* > 0 \tag{61}$$

and so maximised household utility increases whenever one or both individual wage rates and household incomes increase. Thus, in this model, under joint taxation the criteria of horizontal and vertical equity, which could be interpreted as requiring, respectively, equal tax burdens for equal incomes and tax burdens that increase with household income, appear to be met. Individual taxation, on the other hand, may appear to result in a violation of this notion of equity, since a two-earner household may have a larger income than another but be paying less tax. This then suggests the intuition that moving from joint to individual taxation necessarily involves an equity-efficiency trade-off.

[29] Apps and Rees (1999) takes a tax reform approach to this problem, with a change from joint to selective linear taxation. In that case a marginal analysis is possible.

It is interesting however to note that even in this model, this intuition is not entirely correct. In contrast to the standard individual labour supply model, relative household income here is not an exact indicator of relative household utility, since households may have the same income but different achieved utility levels. There is therefore a violation of horizontal and vertical equity in this case, if this is interpreted, as it should be, in terms of household utility rather than income - households with different utility levels may pay the same amount of tax, and those with the same utility level may pay different amounts of tax. In fact we can prove:

**Proposition 3**: *For Model 1, any subset of households with the same gross incomes will necessarily have the same utility levels if and only if labour supply elasticities of primary and second earners are identical.*

Proof: See the Appendix.

Intuitively, as we move through the subset of households by increasing the primary wage, the second earner's wage has to be reduced to hold income constant, but, given a higher labour supply elasticity of second earners, by proportionately less than the increase in the primary wage, and the net effect is to increase household utility.

If all labour supply elasticities were identical, there would be no case for individual taxation on either efficiency or equity grounds. However, as we start raising second earner elasticities relative to those of primary earners, we not only create potential efficiency gains from a switch to individual taxation, but also create inequities in the joint taxation system, which become greater as the ratio of primary to second earner elasticities increases. The "equity penalty" of a switch from joint to individual taxation falls and the efficiency gain increases as this ratio increases, even in the case of Model 1. In the next section we give numerical examples of such elasticity ratio "switching points".[30]

## 5.2 Equity effects in Model 2

The quotation from Feldstein and Feenberg given at the beginning of this paper suggests that when household production is taken into account, inequity in joint taxation arises because two households with the same labour income but widely different utility levels pay the same amount of tax. We have just shown that this can be the case even when household production is not explicitly taken into account, though undoubtedly, as we now show, realistic modelling of household production strengthens the argument considerably. Individual taxation may actually improve distributional outcomes over joint taxation by imposing a lower tax burden on a household with a higher second earner labour supply at a given level of total household income. This can be viewed as implicitly compensating a household for its lower level of household production or, equivalently, implicitly taxing the higher level of household output in the other household. In the next

[30] This suggests that the numerical example used by Boskin and Sheshinski (1983) to argue for selective taxation with lower rates for females effectively assumed an elasticities ratio that lay above the switching point.

section we use parametrised versions of Models 1 and 2 to illustrate this. Here we use Model 2 to provide an analytical basis for this proposition.

The argument can be put in general terms as follows. In Model 1 the fact that in one household the second earner supplies more time to child care and less to the market than in another household with the same income must be due to her lower wage. In Model 2 on the other hand she may well have the same or even a higher wage, and the difference in earnings is due to differences in productivities and/or prices of bought-in care. If domestic and bought-in child care are close substitutes, a household with a high wage primary earner and possibly, given positive associative matching, also a high wage second earner who supplies little or no labour to the market, may have the same market income as a low- to middle-wage household where both earners work full time, but enjoy a much higher standard of living. This is because the lower wage household produces less child care and buys in more of its market substitute, which, if productivity of the inputs increases with wage type, may also be of a lower quality than that in the higher wage household. A switch from joint to individual taxation then shifts some of the tax burden from the lower- to the higher-wage household and this is a progressive rather than regressive change. The resulting distribution of the tax burden is more equitable, but this is not perceived to be the case if one views the world through the lens of Model 1, where household income is a good indicator of household utility.

To develop this argument, consider first what Model 2 says about the determinants of second earner labour supply. The first order conditions for the household's optimal allocation in tax bracket $j = 1, 2$ under joint taxation are (recalling the definition of the function $\psi_2(.)$ in this model given in Section 2):

$$\frac{\partial \psi_1}{\partial y_{1h}} = 1 - \tau_j \quad j = 1, 2 \tag{62}$$

$$\hat{u}'.[\frac{\partial z}{\partial c_h} - \frac{\partial z}{\partial b_h}] = (1 - \tau_j)w_{2h} - p_h \quad j = 1, 2 \tag{63}$$

where $j = 1, 2$ indicates the household's tax bracket. The condition (63) on the second earner's time use is of main interest here. The left hand side of (63) gives the marginal value product of the second earner's time spent in child care net of that of bought in child care, thus giving the opportunity cost to the household of a diversion of a unit of the second earner's time to the labour market. The right hand side gives the return to an hour of market work net of tax and child care cost. This suggests that even where the price of bought-in child care is higher than the second earner's net of tax wage,[31] she may still work in the market and buy in child care if it is sufficiently more productive than her own at the margin.

Overall, since this condition determines the second earner's labour supply, it emphasises not only preferences (the marginal utility of child care $\hat{u}'$ in terms of

[31] A case which is empirically observed and impossible to rationalise in the standard model. Of course there may be other explanations not captured in the present model, for example investment in maintaining and extending work-related human capital. See Apps and Rees (2009) for further discussion.

consumption) and the net of tax wage rate, but also the relative productivities of parental and market child care and the price of the latter as the underlying determinants of labour supply elasticities that are relevant for the tax analysis. This gives a much richer theory of second earner labour supply than the standard labour supply model.

We write the indirect utility function derived from this model as $v_h(w_{1h}, w_{2h}, k_h, p_h)$. Now consider any subset of households with the same household income $y_h$. We move between households within this subset by changing two of the arguments in $v_h(w_{1h}, w_{2h}, k_h, p_h)$ and note what happens to $v_h$ in each case. All proofs are given in the Appendix.[32]

**Proposition 4**: *Achieved utility $v_h$ rises as we move across the given subset of households by raising the primary wage and the price of bought-in child care, if and only if*

$$\frac{e^h_{l_2 p}}{1 + e^h_{l_1 w_1}} > \frac{w_{2h}}{p_h} \tag{64}$$

The intuition here is that when the primary wage increases across households the second earner labour supply must fall to hold joint income constant, requiring therefore the price of bought-in care to rise. If the elasticity of second earner labour supply with respect to this price relative to the primary earner's wage elasticity of labour supply is sufficiently high, the net effect of the wage and price increases on maximised utility, holding joint income constant, is positive. If not, it is negative, except in a knife-edge case in which the effects could cancel and utility also remains constant.

**Proposition 5**: *Achieved utility rises unconditionally as we raise the primary wage and the productivity of domestic child care $k_h$.*

In this case to induce the fall in second earner labour supply we need to raise her child care productivity, causing her to substitute her own for bought-in child care. Both the wage and productivity increases then increase achieved household utility.

Note that in this model, changes in wage rates have no income effects on the demand for child care, because of the quasilinearity of the utility function. More generally, increases in the primary earner wage could be expected to increase demand for child care and therefore, other things being equal, reduce second earner market labour supply, therefore strengthening the effects shown in the above propositions.

**Proposition 6**: *Achieved utility rises unconditionally as we raise the productivity of domestic child care and reduce the price of bought in child care.*

With wage rates constant joint income remains constant only if labour supplies do not change. Reducing the price of bought-in child care has an offsetting

[32] The expression $e^h_{ab}$ denotes the elasticity in household $h$ of the variable $a$ with respect to the variable $b$, where all elasticities are defined in such a way as to be positive. For example $e^h_{l_1 w_1} = \partial \log l_{1h} / \partial \log w_{1h}$ while $e^h_{l_2 p} = -\partial \log l_{2h} / \partial \log p_h$.

effect on second earner labour supply to the increase in productivity, and both these changes increase maximised utility. This effect will be larger, the larger the elasticity of second earner labour supply with respect to the price of bought-in child care, i.e. the greater the elasticity of substitution between the two forms of child care.

The overall effect of these propositions is to support the argument that household income is a poor indicator of household well-being, and so using it as a tax base tends to create inequity, which may be significantly reduced by a change to individual taxation.

We now relate this discussion to the equity terms in the numerators of the expressions for optimal tax rates in (51), (52), (58) and (59) by arguing that the change is likely to lead to an increase in progressivity of the tax system, thus reinforcing the effects of the changes in denominators, the efficiency terms, as already pointed out.

First note that, using (50), the numerator in (51) can be written as

$$\sum_{\mathcal{H}_0} \phi_h(\sigma_h - 1)(y_h^* - \eta) \tag{65}$$

The values of the marginal social utilities of income, $\sigma_h$, depend on achieved utility $v_h$ rather than income, so that, if indeed this tax bracket contains a significant proportion of households with relatively lower incomes but high utility, the negative components $(\sigma_h - 1)$ of this sum will have a greater weight than would be the case if joint income were a perfect indicator of utility. This in turn makes the lower bracket tax rate higher than it would be in that case. A similar argument for the higher bracket suggests that it would be lower than it should be, essentially because ranking households on the basis of joint income is poorly correlated with the ranking on the basis of utility. The switch to individual taxation effectively moves primary earners in households with high achieved utilities into the upper tax bracket and second (and possibly primary) earners in households with lower achieved utilities into the lower bracket, thus raising the former and lowering the latter tax rates, and so increasing the progressivity of the tax system as well as its redistributive equity. We now explore these issues in the context of a numerical example.

# 6 Numerical Analysis

We use specific functional forms and numerical values for the key parameters to explore how the relationship between the maximised values of social welfare under joint and individual taxation changes when we replace Model 1 by Model 2. In each case we solve the model for the optimal parameters of the tax system by maximising a social welfare function (SWF) of the form $[\sum_{i=1}^{n} v_i^{1-\pi}]^{1/(1-\pi)}$, with $\pi$ a measure of inequality aversion.

## 6.1 Model 1

We choose as the household utility function the simple quasilinear form

$$u_h = x_h - \gamma_1 (y_{1h}/w_{1h})^{\alpha_1} - \gamma_2 (y_{2h}/w_{2h})^{\alpha_2} \quad h = 1, .., H \tag{66}$$

with $\gamma_i > 0$ and $\alpha_i = (1 + e_i)/e_i$ where $e_i$ is the elasticity of labour supply with respect to the net wage. With preference variation between primary and second earners, but not across households, we calibrate

$$\gamma_i = (1/\alpha_i)(\overline{y}_i / \bar{l}_i^{\alpha_i}) \quad i = 1, 2 \tag{67}$$

where $\overline{y}_i$, $\overline{w}_i$, and $\bar{l}_i$ are representative values of earnings, the gross wage rate and labour supply respectively.

As shown in Apps, Long and Rees (2011), the assumed wage distribution is of central importance to the results for the structure of optimal tax rates. The distribution indicated by household survey data for a number of the major OECD countries is one in which the wage rates of full time employed primary earners grow slowly and virtually linearly up to around the 80th percentile, and then increase sharply beyond the 90th percentile.[33] Here we derive a percentile "primary wage" distribution based on data for a sample of young families selected from the Australian Bureau of Statistics (ABS) 2010 Survey of Income and Housing (SIH).[34] The wage in each percentile is calculated as average hourly earnings, with hours smoothed across the distribution, to give the distribution depicted in Figure 1. The distribution of the "average second wage" shown in Figure 1 is also constructed to reflect the data. In each primary wage percentile the average second wage is set as a proportion of the primary wage, beginning at 80% in percentile 1 and declining to 50% in the top percentile.

**Figure 1 about here**

We present two sets of simulation results. In the first, Set 1, there are 200 household types constructed by associating a lower and a higher second wage with respect to the average second wage in each primary wage percentile. We therefore have $w_1 \in \{w_1^1, w_1^2, ..., w_1^{100}\}$ and $w_2 \in \{w_2^1, w_2^2\}$ where $w_2^1$ is taken from a distribution that is 25 per cent below, and $w_2^2$ from a distribution 25 per cent above, the average second wage in each primary wage percentile. Figure 1 plots the wage profiles of these two types, labelled "S1" and "S2", respectively.

In Set 2 of the simulations the number of household types is increased to 400 by attaching two lower and two higher second wage rates at each primary wage percentile. Thus we have $w_2 \in \{w_2^1, ..., w_2^4\}$. The two lower (higher) second wage rates are set at 25% and 10% below (above) the average second wage in each percentile.

[33] This is evident, for example, in UK, US and Australian household survey data on the earnings and hours of prime aged adults in full time employment (see Apps and Rees, 2009).

[34] The sample is selected on the criteria that a young child (0 - 9 years) is present, partners are aged from 25 to 59 years and the primary earner works at least 30 hours per week for a wage of at least $15.0 (the minimum wage in 2010). The sample contains 1872 couple income unit records.

In all simulations for Model 1 we set $e_1 = 0.1$, $\bar{l}_1 = 2000$ and $\overline{w}_1$ to the data mean of the primary earner wage distribution. These parameter values generate primary earner annual labour supplies that are broadly consistent with the SIH data. We also set $\overline{w}_2$ to the data mean of the second earner wage distribution. Given that the data indicate that average second hours of work are around 50 per cent of primary hours while the gap between the primary and average second wage is much smaller, we introduce within household preference heterogeneity by setting $\bar{l}_2 = 0.5\bar{l}_1 = 1000$, implying that the second earner has a greater preference for leisure.

Simulation results are presented for three tax regimes: a linear tax with a lump sum and single marginal tax rate; a two-bracket piecewise linear tax on joint income; and a two-bracket piecewise linear tax on individual incomes. In each of the latter cases, we find the optimal tax parameters by solving for the optimal marginal rates for an initial given bracket limit. We then vary the bracket limit by increments of $1000 to find the value that maximises the SWF overall.[35] At any set of tax parameters, the values of consumption and earnings are of course those that maximise each household's utility subject to its budget constraint at that set of parameters. We present results for $e_2$ taking the values 0.1, 0.12, 0.13, 0.2 and 0.3. The values of 0.12 and 0.13 are included to indicate the point at which switching from joint to individual taxation becomes optimal.

Tables 1 and 2 present the results for Set 1 and Set 2, respectively, for three values of the inequality aversion parameter, $\pi = 0.1$, 0.2, and 0.3. The SWF for each of the second earner wage elasticities gives the social welfare ranking of the three tax regimes, given a primary earner wage elasticity of 0.1 in each case.

**Tables 1 and 2 about here**

In both tables we see that:

- Linear taxation is always dominated by a piecewise linear system, arguing therefore against a "flat tax". This is to be expected given the expressions that characterise the optimal tax rates in the theoretical analysis presented in the earlier sections and the steeply rising wage rates in the upper percentiles.
- When the labour supply elasticity of the second earner is sufficiently close to that of the primary earner, for example at $e_2 = 0.12$ and 0.13 in Table 1,[36] joint taxation yields higher social welfare values than individual taxation, which is due to the superiority of the former on equity grounds.
- However, as the elasticity of the second earner rises relative to that of the primary earner, the efficiency gains, even given inequality aversion, start to outweigh whatever equity losses arise, and individual taxation increasingly dominates joint taxation.

[35] Two methods were used: general grid search and global optimisation software. They gave virtually identical results.

[36] The switching point occurs at a second earner elasticity slightly above 0.13 for $\pi = 0.1$ and 0.2, and slightly above 0.14 for $\pi = 0.3$.

The consistency of the results for the two sets of simulations shows that increasing the number of second earner wage rates at each primary earner wage rate from two to four, thus increasing the total number of household types from 200 to 400, leaves the SWF ranking of the three tax regimes essentially unchanged.

As we expect from the theoretical analysis, marginal tax rates rise with the degree of inequality aversion and they fall as the second earner's wage elasticity rises. These results are also reflected in the extent of income redistribution, as indicated by the terms $\alpha$ and $a$.

In comparison with the optimal tax parameters for Model 2 presented in Tables 3 and 4 below, marginal tax rates in the first bracket are relatively high and rise less steeply in the second, and the bracket point for each system is almost constant for all values of $e_2$ and $\pi$, at percentiles 80 and 81 for joint taxation and 86 and 88 for individual taxation. These results reflect the strong correlation between second earner labour supplies and household income, and therefore the equity losses that would be associated with a more progressive rate scale under individual taxation, as discussed in the preceding section. Indeed, on the basis of the comparison of the two tax systems in Tables 1 and 2 one would be inclined to conclude that there is little difference between the tax rates under two systems, certainly not enough to justify incurring the costs of a change from a joint to an individual tax system. This reinforces our emphasis on the importance of the underlying structural model.

## 6.2 Model 2

For the main purpose of this paper it is sufficient to focus on non-wage sources of heterogeneity across households, and so we simplify by assuming perfect assortative matching - the second earner wage increases with, though as a decreasing proportion of, the primary earner wage. The household's utility function is

$$u_h = x_h - \gamma_1 (y_{1h}/w_{1h})^{\alpha_1} + z_h^{\kappa} \quad h = 1, ..., H \tag{68}$$

with $\kappa \in (0, 1)$. Child care is produced with the CES production function[37]

$$z_h = [\beta (k_h c_h)^{\rho} + (1-\beta)(q_h b_h)^{\rho}]^{1/\rho} \quad h = 1, ..., H \tag{69}$$

where the parameter $\rho$ determines the elasticity of substitution between the second earner's child care time input and bought in child care. The analysis is based on the primary and average second wage distributions illustrated earlier in Figure 1.

We focus on the effects of exogenous variation in the price of child care. For this purpose we assume that the quality of bought in care matches that of the second earner, and that both are perfectly correlated with the second earner's wage rate.[38] On the other hand, at each wage rate the price of bought

[37] The notation is as in Section 2 earlier.

[38] As noted previously, these assumptions are broadly consistent with studies that find parental investment in child care and education rises with family resources and that child outcomes improve with maternal human capital.

in child care of a given quality varies exogenously across households. Thus the determinant of across household heterogeneity at a given wage pair $(w_{1h}, w_{2h})$ is the price of bought in child care of quality $q_h$, in conjunction with the elasticity of substitution, as determined by $\rho$.

We again present two sets of simulation results. In Set 1 we introduce two prices for bought in child care at a given second wage, that is, we have $p_h \in \{p^1, p^2\}$ and therefore 200 household types. We present the results with $p^1, p^2$ varying first by $\pm 10\%$ and then by $\pm 20\%$ above and below the second wage. Each price is expressed as a proportion of the second earner's wage. In Set 2 we combine the $\pm 10\%$ and $\pm 20\%$ price variations to obtain four prices at each second wage, $p_h \in \{p^1, .., p^4\}$, and we have therefore 400 household types. Results for both sets of simulations are reported for close substitutes, $\rho = 0.9$, weak complements, $\rho = -0.1$, and strong complements, $\rho = -10.0$. Under these degrees of price variation we find that, after controlling for wage differences, Model 2, with parental and bought in care close substitutes, generates the kind of second earner labour supply heterogeneity observed in the data.[39]

The degree of substitutability between parental and bought in care is important in understanding the welfare comparison of joint and individual taxation in this model. When they are strong complements, the labour supply of the second earner is virtually unaffected by the price of child care relative to the wage, and her labour income varies essentially with her wage rate. Thus household income reflects primarily wage type and, given that productivities of child care inputs also depend on her wage, is a good measure of household utility possibilities. However, in the more realistic case in which they are substitutes,[40] labour incomes can vary widely in response to a small variation in the child care price at a given wage rate, as noted in the analysis of the previous section. As a result household income is not a close or reliable measure of welfare because it omits the value of parental child care. Put differently, two households with the same labour incomes may be of widely differing wage types if the prices of market child care that they face differ, the more so, the greater the elasticity of substitution between domestic and market child care.

This is illustrated in Figure 2, which plots household income across the primary wage distribution for $\rho = 0.9$ and $p^1, p^2$ varying by $\pm 10\%$. The incomes of households facing the higher price, labelled H1, lie well below those of households facing the lower price, labelled H2. However, when the additional value of parental child care in the H1 household is added to its labour income to obtain "full income", there is virtually no difference between the resulting full income profiles of the two household types.

[39] As pointed out earlier, this is impossible to achieve in Model 1 without introducing second earner preference differences across households, which however are normally ruled out in optimal tax analysis because of the problems it creates for interpersonal welfare comparisions.

[40] This is the more realistic assumption for countries in which the focus of bought in child care is on child minding rather than on development and learning. In contrast to the ABS SIH data used here, the data for countries that invest heavily in child care as the first stage of their education system typically show higher secondary earner hours and a lower degree of heterogeneity, indicating that the two time inputs tend to become complements with this type of investment.

**Figure 2 about here**

Table 3 presents the simulation results for the 200 households in Set 1 for $\pi = 0.1, 0.2$ and $0.3$ and $\kappa = 0.9$. The upper panel contains the results for the $\pm10\%$ variation in price and the lower panel, for the $\pm20\%$ variation in price. Table 4 reports the results for the 400 households in Set 2 for the same values of $\pi$ and $\kappa$.

**Tables 3 and 4 about here**

From the SWF rankings we see that individual taxation dominates joint taxation for all cases of close substitutes and weak complements.[41] The ranking is reversed when we move to the case of strong complements because, under this assumption, time use choices are very insensitive to relative prices, and so there is virtually no labour supply heterogeneity across households with the same wage rates. The higher price of bought in child care for the H1 household simply makes it relatively worse off, thus supporting redistribution from H2 to H1 households. However, strong complementarity is inconsistent with the observed heterogeneity in second earner labour supplies.

The results for the optimal tax parameters under the two tax systems differ sharply from those for Model 1. When parental and bought in child care are close substitutes there is very little scope for redistribution under joint taxation. This reflects the efficiency losses that would follow from a high marginal rate on second earnings, driven by the primary earner's income, under a more progressive joint income tax, together with the limited distributional gains when household income as the tax base is poorly correlated with household welfare. This is consistent with the discussion of the determinants of the optimal tax rates in the previous section.

Very different results are obtained when we switch to individual taxation. The optimal marginal rate in the first bracket is extremely low, ranging from $3-10\%$, while in the second bracket, which begins at a percentile that excludes most second earners, the rate rises to something well above 50%. Because heterogeneity in second earner labour supplies tends to disappear as we move to the strong complements case, we obtain a closer match between the optimal parameters for joint and individual taxation. However, as noted above, strong complementarity is not supported by the data.

# 7 Conclusions

This paper analyses the problem of optimal income taxation for two-earner households when the tax system is constrained to take the piecewise linear form that is typical of virtually all real-world tax systems. One aim is to characterise and compare the structure of the optimal tax system for the alternative tax bases

[41] The small absolute differences in the SWF values are a result of the simplifying assumption of quasilinearity in the individual utility functions. Introducing more concavity into these functions would increase the measure of utility differences, but would introduce income effects and thus greatly complicate the optimal tax analysis. Thus we should only attach significance to the directions of change in the SWF measures.

of joint and individual incomes. A second aim is to put forward the argument that the welfare superiority of individual over joint taxation is substantially increased when we take a model of the household which, structurally speaking, is much closer to reality than those used up until now to explore these issues. The central point is that the strong positive relationship between household income and achieved utility which characterises both the single individual household model and the standard two-person model of household labour supply, here called Model 1, does not hold in an empirically more relevant setting, and this has important implications for the equity effects of the alternative tax systems.

In Model 1, with standard stylised facts on the compensated labour supply elasticities of primary and second earners respectively, there will be gains in efficiency in moving from optimal joint to optimal individual taxation. However, since this move tends to redistribute the tax burden from two-earner to single-earner households, and, in this model, household utility is strongly positively correlated with household income, the equity effects tend to be adverse and may outweigh the efficiency gains. Model 2 on the other hand seeks to reflect the data on the time use and expenditure decisions of two-earner households with at least one young child present. In such households parental child care is a major form of time use and bought in care can be a large component of household expenditure. When these inputs to household production vary widely across households with the same wage rates and demographic characteristics, this can have significant implications for the nature of the across-household relationships among second earner labor supply, household income and achieved utility. In this model, the adverse distributional effects of a move from joint to individual taxation are much weaker, or may actually be replaced by distributional improvements. In other words, the analysis of marginal rate progressive piecewise linear tax systems in the presence of a realistic system of household production strengthens the case for individual taxation, even when not selective, still further. It is in this sense that we confirm the Feldstein/Feenberg proposition quoted at the beginning of the paper.

The numerical analysis of specific versions of the two models also brings out the importance of the elasticity of substitution between parental and bought in child care, together with the price of bought in care, in determining the across household relationship between household income and utility. This suggests new directions for the empirical work required to provide the basis for the design of real-world piecewise linear tax systems.

# Appendix

Propositions 3-6 consider how maximised utility varies when we vary type parameters for a subset of households with equal joint incomes. Proposition 3 shows that even in Model 1, household income is not a perfect indicator of achieved utility, while the remaining propositions show that in Model 2, the relationship between them may be highly imperfect. Since the issue is the relationship between gross income and utility, we ignore taxation. It is also useful to assume a continuum of household types. The approach is the same for each proposition. We characterise the relationship between the changes in a given pair of type parameters implied by holding joint income constant, and then use this to examine how maximised utility varies within this subset of households.

**Proof of Proposition 3:**

We have a continuum of wage pairs $(w_{1h}, w_{2h})$ yielding at the household optimum equal household incomes $y_h = \sum_{i=1}^{2} w_{ih} l_{ih}(w_{ih})$ and so within this subset

$$dw_{2h} = -\frac{l_{1h} + w_{1h} l'_{1h}(w_{1h})}{l_{2h} + w_{2h} l'_{2h}(w_{2h})} dw_{1h} \tag{70}$$

Given the indirect utility functions $v_h(w_{1h}, w_{2h})$ with derivatives $\partial v_h / \partial w_{ih} = l_{ih}$ we have within this subset

$$dv_h = [l_{1h} - l_{2h} \frac{l_{1h} + w_{1h} l'_{1h}(w_{1h})}{l_{2h} + w_{2h} l'_{2h}(w_{2h})}] dw_{1h} \tag{71}$$

$$= l_{1h}(1 - \frac{1 + e_1}{1 + e_2}) dw_{1h} \tag{72}$$

where $e_i = w_{ih} l'_{ih}(w_{ih}) / l_{ih} \quad i = 1, 2$ is $i$'s labour supply elasticity.

Then we have that $dv_h = 0$, so that all households within the subset are equally well off, if and only if at every wage pair primary and second earner labour supply elasticities are identical. Since empirically we have $e_1 < e_2$, this implies that household utility is rising as we move through the subset of households with equal incomes by increasing the primary wage, and the greater the elasticity difference, the greater the rate of increase.

**Proof of Proposition 4:**

From the condition for constant $y_h$ we obtain (recalling that elasticities are defined to be positive)

$$dw_{1h} = \frac{w_{2h} l_{2h} e^h_{l_2 p}}{p_h l_{1h} (1 + e^h_{l_1 w_1})} dp_h \tag{73}$$

where $w_{2h}$ is constant and $e^h_{l_1 w_1} > 0$, $e^h_{l_2 p} < 0$. Substituting into the expression for $dv_h$ gives

$$dv_h = l_{2h} (\frac{w_{2h} e^h_{l_2 p}}{p_h (1 + e^h_{l_1 w_1})} - 1) dp_h \tag{74}$$

from which the result follows.

**Proof of Proposition 5:**

From the condtion for constant $y_h$ we obtain

$$dw_{1h} = \frac{w_{2h} l_{2h} e^h_{l_2 k}}{l_{1h} (1 + e^h_{l_1 w_1})} dk_h \tag{75}$$

and substituting into the expression for $dv_h$ with $w_{2h}$ constant gives

$$dv_h = (\frac{\partial v_h}{\partial k_h} + \frac{y_{2h} e^h_{l_2 k}}{1 + e^h_{l_1 w_1}}) dk_h \tag{76}$$

and the term in brackets is positive since $e^h_{l_2 k} > 0$ and $\partial v_h / \partial k_h > 0$.

**Proof of Proposition 6:**
From the condition for constant $y_h$ we obtain

$$dk_h = -\frac{\partial l_{2h}/p_h}{\partial l_{2h}/k_h} dp_h \tag{77}$$

where both derivatives are negative, so $k_h$ and $p_h$ must vary inversely. Substituting then gives

$$dv_h = -(\frac{\partial v_h}{\partial k_h}\frac{\partial l_{2h}/p_h}{\partial l_{2h}/k_h} + l_{2h})dp_h \tag{78}$$

which gives the result.

# References


[1] Alesina, A., Ichino, A. and Karabarbounis, L. (2011). "Gender based taxation and the division of family chores", *American Economic Journal: Economic Policy*, vol. 3(2), pp. 1-40.

[2] Almond, D. and Currie, J. (2011). "Human capital development before age five", in O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, vol. 4, Part B, ch. 15, pp. 1315-1480.

[3] Apps, P., Long, N. and Rees, R. (2011). "Optimal piecewise linear income taxation", *Journal of Public Economic Theory*, forthcoming, IZA Discussion Paper No. 6007.

[4] Apps, P. and Rees, R. (1988). "Taxation and the household", *Journal of Public Economics*, vol. 35(3), pp. 355-69.

[5] Apps, P. and Rees, R. (1999). "On the taxation of trade within and between households", *Journal of Public Economics*, vol. 73(2), pp. 241-63.

[6] Apps, P. and Rees, R. (2009). "*Public Economics and the Household*", Cambridge: Cambridge University Press.

[7] Boadway, R. (2012). "*From Optimal Tax Theory to Tax Policy*", Cambridge: MIT Press.

[8] Boskin M. and Sheshinski, E. (1983). "Optimal tax treatment of the family: Married couples", *Journal of Public Economics*, vol. 20(3), pp. 281-297.

[9] Chetty, R. (2009). "Sufficient statistics for welfare analysis: A bridge between structural and reduced form methods", *Annual Review of Economics*, vol.1(1), pp. 451-488.

[10] Dahlby, B. (1998). "Progressive taxation and the marginal social cost of public funds", *Journal of Public Economics*, vol. 67(1), pp. 105-122.

[11] Dahlby, B. (2008). "*The Marginal Cost of Public Funds'*, Cambridge: MIT Press.

[12] Feldstein, M. and Feenberg, D. (1996). "The taxation of two-earner families", in M. Feldstein and J. Poterba (eds), *Empirical Foundations of Household Taxation*, University of Chicago Press, Chicago, ch. 2, pp. 39-75.

[13] Lundborg, P., Nilsson, A. and Rooth, D. (2012). "Parental education and offspring outcomes: Evidence from the Swedish compulsory schooling reform", IZA Discussion Paper No. 6570.

[14] Munnell, A. (1980). "The couple versus the individual under the federal personal income tax", in H. Aaron and M. Boskin (eds),*The Economics of Taxation,* The Brookings Institution, pp. 247-280.

[15] Rosen, H. (1977). "Is it time to abandon joint filing?", *National Tax Journal*, vol. 30(4), pp. 423-428.

[16] Sadka, E. (1976). "On income distribution, incentive effects and optimal income taxation", *Review of Economic Studies,* vol. 43(2), pp. 261-267.

[17] Sandmo, A. (1990). "Tax distortions and household production", *Oxford Economic Papers*, vol. 42(1), pp. 78-90.

[18] Sheshinski, E. (1972). "The optimal linear income tax', *Review of Economic Studies*, vol. 39, pp. 297-302.

[19] Sheshinski, E. (1989). "Note on the shape of the optimum income tax schedule", *Journal of Public Economics*, vol. 40, pp. 201-215.

[20] Stern, N. (1976). "On the specification of optimum income taxation", *Journal of Public Economics*, vol. 6(1-2), pp. 123-162.

**Figure 1 Wage distributions**

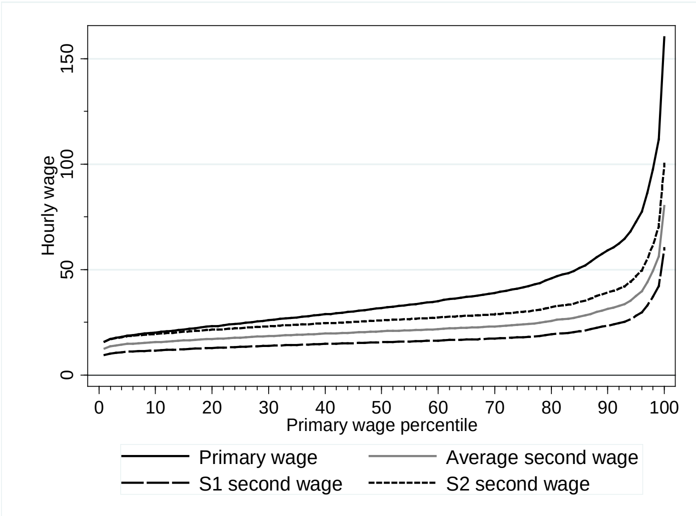


**Figure 2 Model 2, household income and parental child care**

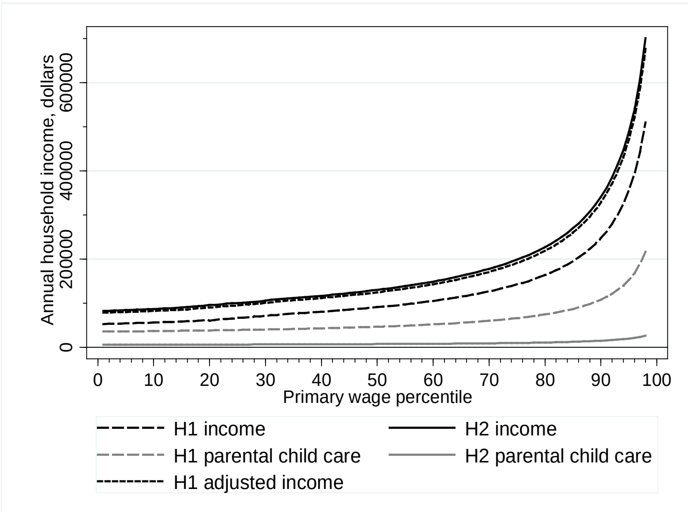

**Table 1 Model 1, Set 1 simulations with $w_2 \in \{w_2^1, w_2^2\}$**

| $\pi$ | $\varepsilon_2$ | Tax system | $\tau_1, t_1$ | $\tau_2, t_2$ | $\alpha, a$ | Bracket** | SWF/$10^3$ |
|---|---|---|---|---|---|---|---|
| 0.1 | 0.1 | linear | 0.17 | 0.17 | 16520 | - | 32083 |
| | | pw* joint | 0.12 | 0.24 | 13220 | 80 | 32092 |
| | | pw individual | 0.13 | 0.24 | 13989 | 88 | 32090 |
| | 0.12 | linear | 0.16 | 0.16 | 15562 | - | 31961 |
| | | pw joint | 0.11 | 0.23 | 12269 | 80 | 31971 |
| | | pw individual | 0.12 | 0.24 | 13150 | 88 | 31970 |
| | 0.2 | linear | 0.15 | 0.15 | 14598 | - | 31522 |
| | | pw joint | 0.10 | 0.21 | 11171 | 81 | 31532 |
| | | pw individual | 0.10 | 0.24 | 11413 | 88 | 31535 |
| | 0.3 | linear | 0.13 | 0.13 | 12691 | - | 31063 |
| | | pw joint | 0.09 | 0.19 | 10128 | 81 | 31072 |
| | | pw individual | 0.09 | 0.23 | 10653 | 86 | 31079 |
| 0.2 | 0.1 | Linear | 0.26 | 0.26 | 24977 | - | 66450 |
| | | pw joint | 0.19 | 0.35 | 20254 | 81 | 66495 |
| | | pw individual | 0.20 | 0.35 | 21034 | 88 | 66488 |
| | 0.12 | linear | 0.25 | 0.25 | 24029 | - | 66187 |
| | | pw joint | 0.18 | 0.34 | 19315 | 81 | 66233 |
| | | pw individual | 0.19 | 0.35 | 20202 | 88 | 66230 |
| | 0.2 | Linear | 0.23 | 0.23 | 22113 | - | 65238 |
| | | pw joint | 0.17 | 0.32 | 18256 | 81 | 65285 |
| | | pw individual | 0.17 | 0.35 | 18495 | 88 | 65301 |
| | 0.3 | Linear | 0.21 | 0.21 | 20214 | - | 64244 |
| | | pw joint | 0.15 | 0.30 | 16395 | 81 | 64290 |
| | | pw individual | 0.14 | 0.34 | 16161 | 86 | 64327 |
| 0.3 | 0.1 | Linear | 0.32 | 0.32 | 30482 | - | 170112 |
| | | pw joint | 0.24 | 0.42 | 25082 | 81 | 170285 |
| | | pw individual | 0.25 | 0.42 | 25866 | 88 | 170256 |
| | 0.13 | linear | 0.31 | 0.31 | 29521 | - | 169076 |
| | | pw joint | 0.23 | 0.41 | 24140 | 81 | 169253 |
| | | pw individual | 0.24 | 0.42 | 25028 | 88 | 169252 |
| | 0.2 | Linear | 0.29 | 0.29 | 27605 | - | 166891 |
| | | pw joint | 0.22 | 0.39 | 23068 | 81 | 167073 |
| | | pw individual | 0.19 | 0.38 | 23318 | 88 | 167135 |
| | 0.3 | Linear | 0.27 | 0.27 | 25691 | - | 164248 |
| | | pw joint | 0.20 | 0.36 | 21126 | 81 | 164430 |
| | | pw individual | 0.19 | 0.41 | 21019 | 86 | 164572 |

*piecewise linear; ** Income percentile of bracket point

**Table 2 Model 1, Set 2 simulations $w_2 \in \{w_2^1, \ldots, w_2^4\}$**

| $\pi$ | $\varepsilon_2$ | Tax system | $\tau_1, t_1$ | $\tau_2, t_2$ | α, $a$ | Bracket** | SWF/$10^3$ |
|---|---|---|---|---|---|---|---|
| 0.1 | 0.1 | Linear | 0.17 | 0.17 | 16514 | - | 69283 |
| | | pw* joint | 0.12 | 0.23 | 13135 | 79 | 69303 |
| | | pw individual | 0.12 | 0.24 | 13140 | 87 | 69300 |
| | 0.2 | Linear | 0.15 | 0.15 | 14588 | - | 68048 |
| | | pw joint | 0.10 | 0.21 | 11241 | 79 | 68069 |
| | | pw individual | 0.10 | 0.24 | 11368 | 87 | 68075 |
| | 0.3 | Linear | 0.13 | 0.13 | 11578 | - | 67032 |
| | | pw joint | 0.09 | 0.19 | 10166 | 79 | 67052 |
| | | pw individual | 0.09 | 0.23 | 10636 | 86 | 67068 |
| 0.2 | 0.1 | Linear | 0.26 | 0.26 | 24968 | - | 158006 |
| | | pw joint | 0.18 | 0.34 | 19633 | 76 | 158114 |
| | | pw individual | 0.20 | 0.35 | 21022 | 87 | 158096 |
| | 0.2 | Linear | 0.23 | 0.23 | 22097 | - | 155070 |
| | | pw joint | 0.16 | 0.32 | 17532 | 79 | 155181 |
| | | pw individual | 0.17 | 0.35 | 18397 | 87 | 155217 |
| | 0.3 | Linear | 0.21 | 0.21 | 20190 | - | 152652 |
| | | pw joint | 0.15 | 0.29 | 16334 | 79 | 152763 |
| | | pw individual | 0.14 | 0.34 | 16136 | 86 | 152849 |
| 0.3 | 0.1 | Linear | 0.32 | 0.32 | 30472 | - | 457801 |
| | | pw joint | 0.23 | 0.41 | 24431 | 77 | 458270 |
| | | pw individual | 0.25 | 0.42 | 25852 | 87 | 458193 |
| | 0.2 | Linear | 0.29 | 0.29 | 27585 | - | 448978 |
| | | pw joint | 0.21 | 0.38 | 22427 | 77 | 449471 |
| | | pw individual | 0.22 | 0.42 | 23209 | 87 | 449628 |
| | 0.3 | Linear | 0.27 | 0.27 | 25662 | - | 441707 |
| | | pw joint | 0.19 | 0.35 | 20433 | 77 | 442201 |
| | | pw individual | 0.18 | 0.41 | 20249 | 85 | 442585 |

*piecewise linear; ** Income percentile of bracket point

**Table 3 Model 2, Set 1 simulations with $p_h \in \{p^1,p^2\}$**

| $\pi$, $p_h$ | $\rho$ | Tax system | $\tau_1$, $t_1$ | $\tau_2$, $t_2$ | α, $a$ | Bracket* | SWF/$10^3$ |
|---|---|---|---|---|---|---|---|
| **Panel A** | | | | | | | |
| $\pi$ = 0.1<br>$p_h$ = ±10%$w_2$ | 0.9 | Joint | 0.03 | 0.04 | 4276 | 43 | 47582 |
| | | Individual | 0.03 | 0.58 | 8206 | 93 | 47659 |
| | -0.1 | Joint | 0.01 | 0.20 | 6780 | 53 | 47085 |
| | | Individual | 0.01 | 0.27 | 6331 | 81 | 47089 |
| | -10.0 | Joint | 0.05 | 0.27 | 12537 | 59 | 47142 |
| | | Individual | 0.08 | 0.27 | 14965 | 76 | 47122 |
| $\pi$ = 0.2<br>$p_h$ = ±10%$w_2$ | 0.9 | Joint | 0.03 | 0.04 | 4276 | 43 | 98370 |
| | | Individual | 0.04 | 0.47 | 10361 | 89 | 98839 |
| | -0.1 | Joint | 0.09 | 0.20 | 14790 | 53 | 97555 |
| | | Individual | 0.17 | 0.58 | 24823 | 94 | 97716 |
| | -10.0 | Joint | 0.12 | 0.47 | 21216 | 80 | 97886 |
| | | Individual | 0.17 | 0.47 | 26160 | 89 | 97794 |
| $\pi$ = 0.3<br>$p_h$ = ±10%$w_2$ | 0.9 | Joint | 0.04 | 0.06 | 5243 | 95 | 251038 |
| | | Individual | 0.04 | 0.47 | 10361 | 89 | 253001 |
| | -0.1 | Joint | 0.11 | 0.20 | 16690 | 55 | 249537 |
| | | Individual | 0.20 | 0.58 | 28222 | 94 | 250417 |
| | -10.0 | Joint | 0.23 | 0.54 | 33519 | 87 | 250944 |
| | | Individual | 0.23 | 0.58 | 32813 | 93 | 250687 |
| **Panel B** | | | | | | | |
| $\pi$ = 0.1<br>$p_h$ = ±20%$w_2$ | 0.9 | Joint | 0.04 | 0.10 | 6628 | 65 | 48771 |
| | | Individual | 0.10 | 0.67 | 15550 | 96 | 48792 |
| | -0.1 | Joint | 0.07 | 0.27 | 10421 | 94 | 47267 |
| | | Individual | 0.07 | 0.27 | 13747 | 77 | 47345 |
| | -10.0 | Joint | 0.05 | 0.27 | 12537 | 59 | 47134 |
| | | Individual | 0.08 | 0.27 | 14965 | 76 | 47115 |
| $\pi$ = 0.2<br>$p_h$ = ±20%$w_2$ | 0.9 | Joint | 0.10 | 0.11 | 12766 | 97 | 100937 |
| | | Individual | 0.10 | 0.58 | 16255 | 93 | 101224 |
| | -0.1 | Joint | 0.07 | 0.27 | 12154 | 80 | 97899 |
| | | Individual | 0.07 | 0.54 | 13610 | 93 | 98187 |
| | -10.0 | Joint | 0.12 | 0.45 | 21181 | 79 | 97845 |
| | | Individual | 0.17 | 0.50 | 25993 | 90 | 97764 |
| $\pi$ = 0.3<br>$p_h$ = ±20%$w_2$ | 0.9 | Joint | 0.10 | 0.11 | 12766 | 97 | 257879 |
| | | Individual | 0.10 | 0.58 | 16308 | 93 | 259189 |
| | -0.1 | Joint | 0.07 | 0.27 | 13403 | 66 | 250427 |
| | | Individual | 0.07 | 0.54 | 13610 | 93 | 251340 |
| | -10.0 | Joint | 0.17 | 0.46 | 26801 | 80 | 250782 |
| | | Individual | 0.23 | 0.58 | 32813 | 93 | 250574 |

*Income percentile of bracket point

**Table 4 Model 2, Set 2 simulations with $p_h \in \{p^1,..,p^4\}$**

| $\pi$, $p_h$ | $\rho$ | Tax system | $\tau_1$, $t_1$ | $\tau_2$, $t_2$ | $\alpha$, $a$ | Bracket* | SWF/$10^3$ |
|---|---|---|---|---|---|---|---|
| $\pi$ = 0.1 $p_h$ = ±10%$w_2$ & ±20%$w_2$ | 0.9 | Joint | 0.03 | 0.04 | 4269 | 44 | 104023 |
| | | Individual | 0.03 | 0.58 | 8184 | 93 | 104151 |
| | -0.1 | Joint | 0.07 | 0.20 | 10182 | 92 | 101819 |
| | | Individual | 0.01 | 0.27 | 6964 | 78 | 101978 |
| | -10.0 | Joint | 0.05 | 0.27 | 12537 | 59 | 101824 |
| | | Individual | 0.08 | 0.27 | 14965 | 77 | 101782 |
| $\pi$ = 0.2 $p_h$ = ±10%$w_2$ & ±20%$w_2$ | 0.9 | Joint | 0.06 | 0.09 | 7710 | 96 | 236826 |
| | | Individual | 0.04 | 0.58 | 9292 | 93 | 237768 |
| | -0.1 | Joint | 0.07 | 0.20 | 11836 | 67 | 232187 |
| | | Individual | 0.07 | 0.47 | 14057 | 89 | 232870 |
| | -10.0 | Joint | 0.12 | 0.45 | 21181 | 79 | 232750 |
| | | Individual | 0.17 | 0.50 | 25993 | 91 | 232558 |
| $\pi$ = 0.3 $p_h$ = ±10%$w_2$ & ±20%$w_2$ | 0.9 | Joint | 0.06 | 0.09 | 7742 | 95 | 684316 |
| | | Individual | 0.06 | 0.47 | 12516 | 89 | 688755 |
| | -0.1 | Joint | 0.07 | 0.20 | 12639 | 53 | 671995 |
| | | Individual | 0.08 | 0.47 | 15214 | 89 | 674829 |
| | -10.0 | Joint | 0.17 | 0.46 | 26801 | 80 | 675209 |
| | | Individual | 0.23 | 0.58 | 32813 | 94 | 674648 |

*Income percentile of bracket point