The Optimal Income Taxation of Couples\textsuperscript{1}

Henrik Jacobsen Kleven, London School of Economics, EPRU, and CEPR  
Claus Thustrup Kreiner, University of Copenhagen, EPRU, and CESifo  
Emmanuel Saez, UC Berkeley and NBER

April 2008

1 Introduction

This paper explores the optimal income tax treatment of couples. Each couple is modelled as a unitary agent supplying labor along two dimensions: the labor supply of a primary earner and the labor supply of a secondary earner. The primary earner makes a continuous labor supply decision as in the Mirrlees (1971) optimal income tax model, whereas the secondary earner makes a binary labor supply decision (work or not work). We impose no a priori restrictions on the income tax system, allowing it to depend on the earnings of each spouse in any nonlinear fashion. This creates a multi-dimensional screening problem. Under our modelling assumptions, if, conditional on primary earnings, two earner couples are better off than one earner couples (for example, if second earners face different market opportunities), optimal tax schemes display positive tax rates on secondary earnings along with negative jointness whereby the tax rate on one person decreases with the earnings of the spouse. Conversely, if second-earner participation is a signal of the couple being worse off (for example, if second earners differ in their home production abilities), we obtain a negative tax rate on the secondary earner along with positive jointness: the second-earner subsidy is being phased out with primary earnings. The above results imply that, in either case, the distortion on the secondary earner is declining with primary earnings, which is therefore the general property of optimal tax schedules. We also prove that the second-earner distortion tends to zero asymptotically as the earnings of the primary earner becomes large. Although this result may seem reminiscent to the classic no-distortion-at-the-top

\footnotesize{\textsuperscript{1}We thank Mark Armstrong, Richard Blundell, Mike Brewer, Raj Chetty, Steven Durlauf, Nada Eissa, Kenneth Judd, Botond Koszegi, David Lannes, Etienne Lehmann, Randall Mariger, Jean-Charles Rochet, Andrew Shephard, and numerous seminar and conference participants for very helpful comments and discussions. We are also grateful for improvements suggested by three anonymous reviewers and co-editor Daron Acemoglu. Corresponding author: Emmanuel Saez, University of California, Department of Economics, 549 Evans Hall #3880, Berkeley, CA 94720, email: saez@econ.berkeley.edu. Financial support from NSF Grant SES-0134946 and a grant from the Economic Policy Research Network (EPRN) is gratefully acknowledged.}
result, our result rests on a completely different reasoning and proof.

Previous work on couple taxation usually assumed separability in the tax function, and hence could not fully address the desirability of joint versus individual taxation, nor investigate the optimal form of jointness. The separability assumption also sidesteps the complexities associated with multi-dimensional screening. In fact, very few studies in the optimal tax literature have attempted to deal with multi-dimensional screening problems. The nonlinear pricing literature in Industrial Organization has analyzed such problems extensively. A central complication of multi-dimensional screening problems is that first-order conditions are often not sufficient to characterize the optimal solution. The reason is that solutions usually display 'bunching' at the bottom (Armstrong, 1996; Rochet and Choné, 1998), whereby agents of different types are forced to make the same choices. By considering a framework with a binary labor supply outcome for the secondary earner along with continuous earnings for the primary earner, we are able to avoid the complexities associated with bunching and to obtain an intuitive understanding of the shape of optimal schedules based on graphical exposition.

Our key results are obtained under a number of strong simplifying assumptions: (i) we use a unitary model of family decision making. (ii) We consider only couples and do not model the marriage decision. (iii) We assume uncorrelated abilities between spouses. (iv) We assume no income effects on labor supply and separability in the disutility of working for the two members of the household, implying that there is no jointness in the family utility function per se. Instead, jointness in our model arises solely because the social welfare function depends on family utilities rather than individual utilities. Our assumptions allow us to zoom in on the role of equity concerns for the jointness of the tax system.

Sections 2 and 3 present our model and theoretical results respectively, while Section 4 presents a numerically calibrated illustrative simulation based on UK micro data.

---


3Mirrlees (1976, 1986) set out a general framework to study such problems and derived first-order optimality conditions. More recently, Cremer, Pestieau and Rochet (2001) revisited the issue of commodity versus income taxation in a multi-dimensional screening model assuming a discrete number of types. Brett (2006) and Cremer et al. (2006) consider the issue of couple taxation in discrete-type models. They show that, in general, incentive compatibility constraints bind in complex ways making it difficult to obtain general properties.

4We refer to Kleven et al. (2006, 2007) for a discussion of robustness and generalizations.
2 The Model

2.1 Family Labor Supply Choice

We consider a population of couples, the size of which is normalized to one. In each couple, there is a primary earner who always participates in the labor market and makes a choice about the size of labor earnings \( z \). The primary earner is characterized by a scalar ability parameter \( n \) which is heterogeneous in the population. The cost of earning \( z \) for a primary earner with ability \( n \) is given by \( n \cdot h(z/n) \), where \( h(.) \) is an increasing and convex function of class \( C^2 \) and normalized so that \( h(0) = 0 \) and \( h'(1) = 1 \). Secondary earners choose whether or not to participate in the labor market, \( l = 0, 1 \), but hours worked conditional on working are fixed. Their labor income is given by \( w \cdot l \), where \( w \) is a uniform wage rate, and they face a fixed cost of participation \( q \) which is heterogeneous across the secondary earners.

The government cannot observe \( n \) and \( q \) and has to base redistribution solely on observed earnings using a non-linear tax system \( T(z, wl) \). Because \( l \) is binary and \( w \) is uniform, this tax system simplifies to a pair of schedules, \( T_0(z) \) and \( T_1(z) \), depending on whether the spouse works or not. Assuming that \( T_0 \) and \( T_1 \) are differentiable, the tax system is separable iff \( T_0' = T_1' \) everywhere or, equivalently, if \( T_0 \) and \( T_1 \) differ by a constant. Net-of-tax income for a couple with earnings \( (z, wl) \) is given by \( c = z + w \cdot l - T_l(z) \).

We consider two sources of heterogeneity across secondary earners, differences in market opportunities and differences in home production abilities, as reflected in the utility function

\[
u(c, z, l) = c - n \cdot h\left(\frac{z}{n}\right) - q^w \cdot l + q^h \cdot (1 - l),
\]

where \( q^w + q^h \equiv q \) is the total cost of second-earner participation, the sum of a direct work cost \( q^w \) and an opportunity cost of lost home production \( q^h \). Heterogeneity in \( q^w \) creates differences in household utility across couples with \( l = 1 \) (heterogeneity in market opportunities), whereas heterogeneity in \( q^h \) generates differences in household utility across couples with \( l = 0 \) (heterogeneity in home production abilities).

As we shall see, the two types of heterogeneity pull optimal redistribution policy in opposite directions. To isolate the impact of each type of heterogeneity, we consider them in turn. In the work cost model \( (q = q^w > 0, q^h = 0) \), at a given primary earner ability \( n \), two-earner couples will be those with low work costs and hence be better off than one-earner couples. This creates
a motive for the government to tax the income of the secondary earner in order to redistribute from two-earner to one-earner couples. By contrast, in the home production model \((q^w = 0, q = q^h > 0)\), two-earner couples will be those with low home production abilities and therefore be worse off than one-earner couples, creating the reverse redistributive motive.

The work cost model is more consonant with the tradition in applied welfare and poverty measurement which assumes that secondary earnings contribute positively to family well-being and rarely adjusts for lost home production. It is also consonant with the underlying notion in the existing optimal tax literature that higher income is a signal of being better off.\(^5\) On the other hand, the existing literature did not consider a secondary earner in the household, which is exactly where home production (including child-bearing and child-caring) becomes important. We therefore analyze both models symmetrically.

The first-order condition for primary earnings \(z\) (conditional on \(l = 0, 1\)) is given by

\[
h'(\frac{z}{n}) = 1 - T'_l(z) = n \cdot h'(\frac{z}{n}) + w \cdot l.
\]

In the case of no tax distortion, \(T'_l(z) = 0\), our normalization \(h'(1) = 1\) implies \(z = n\). Hence, it is natural to interpret \(n\) as potential earnings.\(^6\) Positive marginal tax rates depress actual earnings \(z\) below potential earnings \(n\). If the tax system is non-separable such that \(T'_0 \neq T'_1\), primary earnings \(z\) depend on the labor force participation decision \(l\) of the spouse. We denote by \(z_l\) the optimal choice of \(z\) at a given \(l\).

We define the elasticity of primary earnings with respect to the net-of-tax rate \(1 - T'_l\) as

\[
\varepsilon_l \equiv \frac{1 - T'_l}{z_l} \frac{\partial z_l}{\partial (1 - T'_l)} = \frac{nh'(z_l/n)}{z_l h''(z_l/n)}.\]

Under separable taxation where \(T'_0 = T'_1\), we have \(z_0 = z_1\) and \(\varepsilon_0 = \varepsilon_1\).

Secondary earners work if the utility from participation is greater than or equal to the utility from non-participation. Let us denote by

\[
V_l(n) = z_l - T_l(z_l) - nh'(\frac{z_l}{n}) + w \cdot l.
\]

\(^5\)It is this notion which underlies the result, in the Mirrlees model, that optimal marginal tax rates are positive. A modified Mirrlees model where differences in market earnings are driven by differences in home production ability instead of market ability would generate negative optimal tax rates as high-ability individuals have lower market earnings but are nonetheless better off.

\(^6\)Typically, economists consider models where \(n\) is a wage rate and utility is specified as \(u = c - h(z/n)\), leading to a first-order condition \(1 - T'(z) = n \cdot h'(z/n)\). Our results carry over to this specification but \(n\) would no longer reflect potential earnings and the interpretation of optimal tax formulas would be less transparent (Saez, 2001).
the indirect utility of the couple (exclusive of the fixed cost \(q\)) at a given \(l\). Differentiating with respect to \(n\) (denoted by an upper dot from now on), and using the envelope theorem, we obtain

\[
\dot{V}_l(n) = -h \left( \frac{z_l}{n} \right) + \frac{z_l}{n} \cdot h' \left( \frac{z_l}{n} \right) \geq 0. 
\] (5)

The inequality follows from the fact that \(x \to -h(x) + x \cdot h'(x)\) is increasing (as \(h'' > 0\) and null at \(x = 0\). The inequality is strict if \(z_l > 0\) i.e., if \(T'_l < 1\). The participation constraint for secondary earners is given by

\[
q \leq V_1(n) - V_0(n) \equiv \bar{q}, 
\] (6)

where \(\bar{q}\) is the net gain from working exclusive of the fixed cost \(q\). For families with a fixed cost below (above) the threshold-value \(\bar{q}\), the secondary earner works (does not work).

The couple characteristics \((n, q)\) are distributed according to a continuous density distribution defined over \([n, \bar{n}] \times [0, \infty)\). We denote by \(P(q|n)\) the cumulative distribution function of \(q\) conditional on \(n\), \(p(q|n)\) the density function of \(q\) conditional on \(n\), and \(f(n)\) the unconditional density of \(n\). The probability of labor force participation for the secondary earner at a given ability level \(n\) of the primary earner is \(P(\bar{q}|n)\). We define the participation elasticity with respect to the net gain from working \(\bar{q}\) as \(\eta = \bar{q}p(\bar{q}|n)/P(\bar{q}|n)\).

To complete the description of the household, we need to define a tax rate on second-earner participation. Since \(w\) is the gross gain from working, and \(\bar{q}\) has been defined as the (money metric) net utility gain from working, we can define this tax rate as \(\tau = (w - \bar{q})/w\). Notice that, if taxation is separate so that \(T'_0 = T'_1\) and \(z_0 = z_1\), we have \(\tau = (T_1 - T_0)/w\). On the other hand, if taxation is non-separate, then \(T_1 - T_0\) reflects the total tax change for the family when the secondary earner starts working and the primary earner makes an associated earnings adjustment, whereas \(w - \bar{q}\) reflects the tax burden on second-earner participation per se. Using the definition of \(\tau\), we may define the possible forms of couple taxation as

**Definition 1** At any point \(n\), we have either (i) positive jointness, \(T'_1 > T'_0\) and \(\dot{\tau} > 0\), or (ii) separability, \(T'_0 = T'_1\) and \(\dot{\tau} = 0\), or (iii) negative jointness, \(T'_1 < T'_0\) and \(\dot{\tau} < 0\).\(^7\)

Finally, notice that double-deviation issues are taken care of in our model, because we consider earnings at a given \(n\) and allow \(z\) to adapt optimally when \(l\) changes. That is, if the secondary

\(^7\)Using (4)-(6), it is easy to prove that \(\text{sign}(T'_1 - T'_0) = \text{sign}(\dot{\tau})\). This is simply another way of stating the theorem of equality of cross-partial derivatives.
earners starts working, optimal primary earnings shift from \(z_0(n)\) to \(z_1(n)\) but the key first-order condition (5) continues to apply. More precisely, we show in electronic appendix B.1 that, similarly to the Mirrlees model, a given path for \((z_0(n), z_1(n))\) can be implemented via a truthful mechanism or equivalently by a non-linear tax system if and only if \(z_0(n)\) and \(z_1(n)\) are non-negative and non-decreasing in \(n\).

2.2 Government Objective

The government sets taxes as a function of earnings, \(T_0(z)\) and \(T_1(z)\), to maximize a social welfare function defined as the sum of concave and increasing transformations \(\Psi(.)\) of the couples’ utilities subject to a government budget constraint and the constraints imposed by household utility maximization. Formally, the government maximizes

\[
W = \int_{\bar{q}}^{\bar{q}_0} \Psi \left( V_1(n) - q^w \cdot l + q^h \cdot (1 - l) \right) p(q|n)f(n)dqdn, \tag{7}
\]

subject to the budget constraint

\[
\int_{\bar{q}}^{\bar{q}_0} T_l(z_1)p(q|n)f(n)dqdn \geq 0, \tag{8}
\]

and subject to \(\dot{V}_0(n)\) and \(\dot{V}_1(n)\) in eq. (5).

We denote by \(\lambda\) the multiplier associated with the budget constraint (8). The redistributive tastes of the government may be represented by social marginal welfare weights on different couples. We denote by \(g_0(n)\) the (average) social marginal welfare weight for couples with primary-earner ability \(n\) and secondary-earner participation status \(l\). Formally, for the work cost model \((q^w > 0, q^h = 0)\), we have \(g_1(n) = \int_0^q \Psi'(V_1(n) - q^w)p(q|n)dp/(P(\bar{q}|n) \cdot \lambda)\) and \(g_0(n) = \Psi'(V_0(n))/\lambda\). For the home production model \((q^w = 0, q^h > 0)\), we have \(g_1(n) = \Psi'(V_1(n))/\lambda\) and \(g_0(n) = \int_{\bar{q}}^{\infty} \Psi'(V_0(n) + q^h)p(q|n)dp/((1 - P(\bar{q}|n)) \cdot \lambda)\).

Optimal redistribution depends crucially on the evolution of weights \(g_0(n)\) and \(g_1(n)\) through the ability distribution. In particular, we will show that the optimal tax scheme depends on properties of \(g_0(n) - g_1(n)\), which reflects the preferences for redistribution between one- and two-earner couples. At this stage, notice that the sign of \(g_0(n) - g_1(n)\) depends on whether second-earner heterogeneity is driven by work costs or by home production ability. In the work cost model, \(V_1(n) - q^w > V_0(n)\), implying, as \(\Psi\) is concave, that \(g_0(n) - g_1(n) > 0\). By contrast, in the home production model, \(V_0(n) + q^h > V_1(n)\), implying that \(g_0(n) - g_1(n) < 0\). As we
shall see, whether \(g_0(n) - g_1(n)\) is positive or negative determines whether the optimal tax on secondary earners is positive or negative.

3 Characterization of the Optimal Income Tax Schedule

3.1 Optimal Tax Formulas and their Relation to Mirrless (1971)

The simple model described above makes it possible to derive explicit optimal tax formulas as in the individualistic Mirrles (1971) framework. As shown in appendix A.1, under

**Assumption 1** The function \(x \rightarrow \frac{(1 - h'(x))}{(x \cdot h''(x))}\) is decreasing,

the Hamiltonian of the government maximization program in concave in \(z = (z_0, z_1)\). Assumption 1 is satisfied, for example, for iso-elastic utilities \(h(x) = x^{1+k}/(1 + k)\) where the labor supply elasticity \(\varepsilon\) is constant and equal to \(1/k\) or for any utility function such that the elasticity \(\varepsilon = h'/x \cdot h''\) is decreasing in \(x\).

**Proposition 1** Under assumption 1, the optimal solution \((z_0, z_1, T_0', T_1')\) is continuous in \(n\). The first-order conditions for the optimal marginal tax rates \(T_0'\) and \(T_1'\) at ability level \(n\) can be written as

\[
\frac{T_0'}{1-T_0'} = \frac{1}{\varepsilon_0} \cdot \frac{1}{nf(n)(1-P(\bar{q}|n))} \int_n^{\bar{n}} \left\{ (1-g_0) \left( 1 - P(\bar{q}|n') \right) + [T_1 - T_0]p(\bar{q}|n') \right\} f(n') dn' \quad (9)
\]

\[
\frac{T_1'}{1-T_1'} = \frac{1}{\varepsilon_1} \cdot \frac{1}{nf(n)p(\bar{q}|n)} \int_n^{\bar{n}} \left\{ (1-g_1)P(\bar{q}|n') - [T_1 - T_0]p(\bar{q}|n') \right\} f(n') dn', \quad (10)
\]

where all the terms outside the integrals are evaluated at ability level \(n\) and all the terms inside the integrals are evaluated at \(n'\). These conditions apply at any point \(n\) where there is no bunching, i.e., where \(z_l(n)\) is strictly increasing in \(n\). If the conditions generate segments over which \(z_0(n)\) or \(z_1(n)\) are decreasing, then there is bunching and \(z_0(n)\) or \(z_1(n)\) are constant over a segment.

**Proof:** See Appendix A1.

Kleven et al. (2006) presents a detailed discussion of these formulas. Let us here remark on just two aspects. First, the (weighted) average marginal tax rate faced by primary earners in one- and two-earner couples equals

\[
(1 - P(\bar{q}|n)) \varepsilon_0 \cdot \frac{T_0'}{1-T_0'} + P(\bar{q}|n)\varepsilon_1 \cdot \frac{T_1'}{1-T_1'} = \frac{1}{nf(n)} \int_n^{\bar{n}} (1 - \bar{g}(n')) f(n') dn', \quad (11)
\]
where \( \bar{g}(n') = (1 - P(\bar{q}|n')) g_0(n') + P(\bar{q}|n') g_1(n') \) is the average social marginal welfare weight for couples with ability \( n' \). This result is identical to the Mirrlees formula in the case with no income effects (Diamond, 1998). This implies that redistribution across couples with different primary earners follows the standard logic in the literature. The introduction of a secondary earner in the household creates a potential difference in the marginal tax rates faced by primary earners with working and non-working spouses, which we explore in detail below.

Second, the famous results that optimal marginal tax rates are zero at the bottom and at the top carry over to the couple model, and follow directly from the transversality conditions (see Appendix A.1). As is well-known, these results have limited practical relevance, because the bottom result does not apply when there is an atom of non-workers, and because the top rate drops to zero only for the single topmost earner (Saez, 2001).

### 3.2 Asymptotic Properties of the Optimal Schedule

Suppose that the ability distribution of primary earners \( f(n) \) has an infinite tail so that \( \bar{n} = \infty \). Since top tails of income distributions are well approximated by Pareto distributions (Saez, 2001), we assume that \( f(n) \) has a Pareto tail with parameter \( a > 1 \) \( f(n) = C/n^{1+a} \). As \( n \) tends to infinity, the contribution of the secondary earner to household welfare becomes infinitesimal relative to the contribution of the primary-earner. For any reasonable welfare function, we would then have that \( g_0(n) \) and \( g_1(n) \) converge to the same value \( g^{\infty} < 1 \). It is also natural to assume that primary-earner elasticities \( \varepsilon_0 \) and \( \varepsilon_1 \) converge to \( \varepsilon^{\infty} \), and that the distribution of fixed work costs \( P(q|n) \) converges to a distribution \( P^{\infty}(q) \). We can then prove the following result:

**Proposition 2** Suppose \( T_1 - T_0, T_0', T_1', \bar{q}, \tau \) converge to \( \Delta T^{\infty}, T_0^{\infty}, T_1^{\infty}, \bar{q}^{\infty}, \tau^{\infty} \) as \( n \to \infty \). Then we have (i) the second-earner tax converges to zero, \( \Delta T^{\infty} = \tau^{\infty} = 0 \), and (ii) the marginal tax rates on primary earners converge to \( T_0^{\infty} = T_1^{\infty} = (1 - g^{\infty}) / (1 - g^{\infty} + a \cdot \varepsilon^{\infty}) > 0 \).

**Proof:** Because \( T_1 - T_0 \) converges as \( n \) goes to infinity, it must be the case that \( T_0^{\infty} = T_1^{\infty} = T^{\infty} \). Because \( \bar{q} \) converges, we have that \( P(\bar{q}) \) and \( p(\bar{q}) \) also converge, and we denote their limits by \( P^{\infty} \) and \( p^{\infty} \). The Pareto assumption implies that \( (1 - F(n))/(nf(n)) = 1/a \) for large \( n \). Taking the limit of (9) and (10) as \( n \to \infty \), we obtain

\[
\frac{T^{\infty}}{1 - T^{\infty}} = \frac{1}{\varepsilon^{\infty}} \cdot \frac{1}{a} \cdot \left[ 1 - g^{\infty} + \Delta T^{\infty} \frac{p^{\infty}}{1 - P^{\infty}} \right],
\]
\[
\frac{T'_{\infty}}{1 - T'_{\infty}} = \frac{1}{\varepsilon_{\infty}} \cdot \frac{1}{a} \left[ 1 - g^\infty - \Delta T^\infty \frac{p^\infty}{P^\infty} \right].
\]

For this to be satisfied, we must have \(\Delta T^\infty = 0\), and the formula for \(T'_{\infty}\) then follows. □

It is quite striking that the spouses of the very high earners should be exempted from taxation as \(n\) tends to infinity, even in the case where the government tries to extract as much tax revenue as possible from high-income couples \((g^\infty = 0)\). Although this result may seem similar to the classic no-distortion-at-the-top result reviewed above, the logic behind our result is completely different. Indeed, in the present case with an infinite tail for \(n\), Proposition 2 shows that the marginal tax rate on primary earners does not converge to zero. Instead, the marginal tax rates converge to the positive constant \((1 - g^\infty) / (1 - g^\infty + a\varepsilon^\infty)\), exactly as in the individualistic Mirrlees model when \(n \to \infty\) (Saez, 2001).^8

To grasp the intuition behind the zero second-earner tax at the top, consider a situation where \(T_1 - T_0\) does not converge to zero but instead converges to \(\Delta T^\infty > 0\) as illustrated on Figure 1. We want to establish a contradiction by arguing that, in this situation, it is possible to increase welfare by reducing \(T_1 - T_0\) a little bit at the top. Consider specifically a reform which increases the tax on one-earner couples and decreases the tax on two-earner couples above some high \(n\), and in such a way that the net mechanical effect on government revenue is zero.\(^9\) These tax burden changes are achieved by increasing the marginal tax rate for one-earner couples in a small band \((n, n + dn)\), and lowering the marginal tax rate for two-earner couples in this band.

What are the welfare effects of the reform? First, there are direct welfare effects as the reform redistributes income from one-earner couples (who lose \(dW_0\)) to two-earner couples (who gain \(dW_1\)). However, because social marginal welfare weights for one- and two-earner couples have converged to \(g^\infty\), these direct welfare effects cancel out. Second, there are fiscal effects due to earnings responses of primary earners in the small band where marginal tax rates have been changed \((dH_0\) and \(dH_1\)). Because \(T_1 - T_0\) have converged to a constant for large \(n\), the marginal tax rates on one- and two-earner couples are identical, \(T'_{0\infty} = T'_{1\infty}\), which implies \(z_0 = z_1\) and hence identical primary-earner elasticities \(\varepsilon_0 = \varepsilon_1\). As a consequence, the negative fiscal effect \(dH_0\) exactly offsets the positive fiscal effect \(dH_1\). Third, there is a participation effect as some

---

\(^8\)Conversely, in the case of a bounded ability distribution, the top marginal tax rate on primary earnings would be zero, but then the tax on the secondary earner would be positive.

\(^9\)Because \(\bar{q}\) and hence \(P(\bar{q})\) have converged, revenue-neutrality requires that the tax changes on one- and two-earner couples are \(dT_0 = dT/(1 - P(\bar{q}))\) and \(dT_1 = -dT/P(\bar{q})\), respectively.
secondary earners are induced to join the labor force, since the extra tax on two-earner families has been reduced. Because $T_1 - T_0$ is initially positive, this response will generate a positive fiscal effect, $dP > 0$. Since all other effects were zero, $dP > 0$ is the net total welfare effect of the reform. The reform then increases welfare, implying that the original schedule with $\Delta T^\infty > 0$ cannot be optimal. Of course, the opposite situation with $\Delta T^\infty < 0$ cannot be optimal either, because the reverse reform would then improve welfare.

3.3 Optimal Jointness

To analyze the optimal form of jointness, we introduce two additional assumptions.

**Assumption 2** The function $V \rightarrow \Psi'(V)$ is strictly convex.

This is a very natural assumption on social preferences which is satisfied for all standard social welfare functions such as the CRRA and CARA forms. In the context of consumer theory, convexity of marginal utility of consumption is a common assumption, since it captures the notion of prudence and generates precautionary savings. As shown below, this assumption captures the central idea that secondary earnings matter less and less for social marginal welfare as primary earnings increase.

**Assumption 3** $q$ and $n$ are independently distributed.

Abstracting from correlation in spouse characteristics (assortative matching) allows us to isolate the implications for the optimal tax system of the interaction between spouses occurring through the social welfare function. In Section 4, we examine numerically how assortative matching affects our results.

To establish an intuition on the optimal form of jointness, it is helpful to start by considering a tax reform introducing a little bit of jointness around the optimal separable tax system. In particular, for the work cost model, we will argue that the optimal separable schedule can be improved by introducing a little bit of negative jointness. For the home production model, the arguments can be reversed to show that a little bit of positive jointness is welfare improving.

A separable schedule is one where $T_0' \equiv T_1'$ and hence $T_1 - T_0$, $\bar{q}$, and $P(\bar{q})$ are constant in $n$. In the work cost model, we would have $T_1 - T_0 > 0$ due to the property $g_0 - g_1 > 0$. As discussed above, this property follows from the fact that, at given $n$, being a two-earner couple is a signal of low work costs and being better off than one-earner couples. Moreover, under assumptions 2 and
3, and starting from a separable tax system, $g_0 - g_1$ is decreasing in $n$. Intuitively, this property derives from the fact that, as primary-earner ability increases, the contribution of secondary earnings to couple utility is declining in relative terms, and therefore the value of redistribution from two- to one-earner couples is declining. To see this formally, notice that, under separable taxation and Assumption 3, we have that $\bar{q} = w - (T_1 - T_0)$, $P(\bar{q}|n) = P(\bar{q})$, and $p(q|n) = p(q)$ are constant in $n$. Then, from the definitions of $g_0(n)$ and $g_1(n)$, we obtain

$$\frac{d [g_0(n) - g_1(n)]}{dn} = \left[ \frac{\Psi''(V_0)}{\lambda} - \int_0^\bar{q} \frac{\Psi''(V_0 + \bar{q} - q)p(q) dq}{\lambda \cdot P(\bar{q})} \right] \cdot \dot{V}_0 < 0, \quad (12)$$

where we have used $V_1 = V_0 + \bar{q}$ from eq. (6). Since $\Psi''(.)$ is increasing (by Assumption 2) and $V_0$ is increasing in $n$, it follows that the expression in (12) is negative.

Now, consider a tax reform introducing a little bit of negative jointness as shown in Figure 2. The tax reform has two components. Above ability level $n$, we increase the tax on one-earner couples and decrease the tax on two-earner couples. Below ability level $n$, we decrease the tax on one-earner couples and increase the tax on two-earner couples. These tax burden changes are associated with changes in the marginal tax rates on primary earners around $n$.

To ensure that the reform is revenue-neutral (absent any behavioral responses), let the size of the tax change on each segment be inversely proportional to the number of couples on the segment. That is, above $n$, the tax change for one-earner couples is $dT_a^0 = \frac{dT}{(1 - F(n))(1 - P(\bar{q}))}$ and the tax change for two-earner couples is $dT_a^1 = -\frac{dT}{(1 - F(n))P(\bar{q})}$. Below $n$, the tax change for one-earner couples is $dT_b^0 = \frac{dT}{F(n)(1 - P(\bar{q}))}$ and the tax change for two-earner couples is $dT_b^1 = \frac{dT}{F(n)P(\bar{q})}$.

What are the effects? First, there is a direct welfare effect created by the redistribution across the different types of couples. This effect can be written as

$$dW = \frac{dT}{F(n)} \cdot \int_n^n [g_0(n') - g_1(n')] f(n') dn' - \frac{dT}{1 - F(n)} \cdot \int_n^n [g_0(n') - g_1(n')] f(n') dn' > 0. \quad (13)$$

The first term reflects the gain created at the bottom by redistributing from two-earner to one-earner couples, and the second term reflects the loss created at the top from the opposite redistribution. Our result in (12) implies that the gain at the bottom dominates the loss at the top, and therefore $dW > 0$.

Second, there are fiscal effects associated with earnings responses by primary earners induced by the changes in $T_0^a$ and $T_1^a$ around $n$. Since the reform increases the marginal tax rate one-
earner couples around \( n \), and reduces it in two-earner couples, the earnings responses are going in opposite directions. In fact, since we start from a situation with separate taxation, \( T'_0 = T'_1 \), and hence identical primary-earner elasticities, \( \varepsilon_0 = \varepsilon_1 \), the fiscal effects of primary earner responses offset one another exactly.

Third, the reform creates participation responses by secondary earners. Above \( n \), non-working spouses will be induced to join the labor force, whereas below \( n \), working spouses have an incentive to drop out. The fiscal implications of these responses also offset exactly. To see this, notice that, because spouse characteristics \( q \) and \( n \) are independent, and since we start from a separable tax system, the participation elasticity \( \eta = \bar{q}p(\bar{q})/P(\bar{q}) \) is constant in \( n \). Then, as the second-earner tax \( T_1 - T_0 \) is initially constant, the fiscal implications of the positive responses at the top are going to offset exactly the negative responses at the bottom.

We can then conclude that \( dW > 0 \) is the net total welfare effect of the reform. Hence, under Assumptions 1-3, introducing a little bit of negative jointness increases welfare. This perturbation argument indicates that, for the work cost model, the optimal incentive scheme will be associated with negative jointness. This stronger statement turns out to be true, although the proof is more complicated than the perturbation argument because, as we move away from separability, efficiency effects from behavioral responses no longer cancel out to the first order. In order to establish our main result, we introduce a final technical assumption:

**Assumption 4** The function \( x \rightarrow x \cdot p(w - x)/[P(w - x) \cdot (1 - P(w - x))] \) is increasing and \( p(q)/P(q) \leq P(q)/\int_q^1 P(q')dq' \) for all \( q \).

This assumption is satisfied for iso-elastic cost of work distributions of the type \( P(q) = (q/q_{max})^\eta \) where the participation elasticity of secondary earners is constant and equal to \( \eta \).

**Proposition 3** Under assumptions 1-4, and assuming that the optimal solution has no bunching, then the tax system is characterized by

**Work cost model:** (1a) Negative jointness, \( T'_1 < T'_0 \) and \( \dot{\tau} < 0 \) for all \( n \in [\bar{n}, \bar{n}] \). (1b) Positive tax on secondary-earner income, \( \tau > 0 \) for all \( n \in [\bar{n}, \bar{n}] \).  

**Home production model:** (2a) Positive jointness, \( T'_1 > T'_0 \) and \( \dot{\tau} > 0 \) for all \( n \in [\bar{n}, \bar{n}] \). (2b) Negative tax on secondary-earner income, \( \tau < 0 \) for all \( n \in [\bar{n}, \bar{n}] \).

**Proof:** Below we prove result (1a) and (1b) for the work cost model. Result (2a) and (2b) may be established by reversing all inequalities in the proof below.

12
Suppose by contradiction that \( T'_1 > T'_0 \) for some \( n \). Then, because \( T'_0 \) and \( T'_1 \) are continuous in \( n \) (see Appendix A.1) and because \( T'_1 = T'_0 \) at the top and bottom skills, there exists an interval \( (n_a, n_b) \) where \( T'_1 > T'_0 \) and where \( T'_1 = T'_0 \) at the end points, \( n_a \) and \( n_b \). This implies that \( z_1 < z_0 \) on \( (n_a, n_b) \) with equality at the end points. Assumption 1 implies that \( (1 - h')/(x \cdot h'') \) is decreasing in \( x \) and hence \( \varepsilon_1 T'_1/(1 - T'_1) = (1 - h'_1)/(h''_1, z_1/n) > (1 - h'_0)/(h''_0, z_0/n) = \varepsilon_0 T'_0/(1 - T'_0) \) on \( (n_a, n_b) \). Then, using the first-order conditions (9) and (10) which apply everywhere because of our no bunching assumption, we obtain

\[
\Omega_0(n) \equiv \frac{1}{1 - P} \int_n^n [(1 - g_0)(1 - P) + \Delta T \cdot p(n')]dn' < \frac{1}{P} \int_n^n [(1 - g_1)P - \Delta T \cdot p(n')]dn' \equiv \Omega_1(n)
\]
on \( (n_a, n_b) \) with equality at the end points. This implies that the derivatives of the above expressions with respect to \( n \), at the end points, obey the inequalities \( \dot{\Omega}_0(n_a) \leq \dot{\Omega}_1(n_a) \) and \( \dot{\Omega}_0(n_b) \geq \dot{\Omega}_1(n_b) \). At the end points, we have \( T'_1 = T'_0, z_0 = z_1, \) and \( \dot{V}_0 = \dot{V}_1 \), which implies \( \dot{q} = 0 \) and \( \dot{P} = 0 \). Hence, the inequalities in derivatives can be written as

\[
1 - g_0 + \Delta T \cdot p/(1 - P) \geq 1 - g_1 - \Delta T \cdot p/P \text{ at } n_a,
\]

\[
1 - g_0 + \Delta T \cdot p/(1 - P) \leq 1 - g_1 - \Delta T \cdot p/P \text{ at } n_b.
\]

Combining these inequalities, we obtain

\[
\frac{\Delta T \cdot p}{P(1 - P)}|_{n_a} \geq g_0(n_a) - g_1(n_a) > g_0(n_b) - g_1(n_b) \geq \frac{\Delta T \cdot p}{P(1 - P)}|_{n_b}.
\]

The middle inequality is intuitive and follows formally from Assumptions 1-4 as shown in Appendix A.2. Using that \( \dot{q} = w - \Delta T \) at \( n_a \) and \( n_b \), along with the first part of Assumption 4, we obtain \( \Delta T(n_a) > \Delta T(n_b) \). However, \( T'_1 > T'_0 \) and hence \( z_1 < z_0 \) implies that \( \dot{q} < 0 \) on the interval \( (n_a, n_b) \). Then we have \( \dot{q}(n_a) \geq \dot{q}(n_b) \) and hence \( \Delta T(n_a) \leq \Delta T(n_b) \). This generates a contradiction, which proves that \( T'_1 \leq T'_0 \) for all \( n \).

The second part of the proposition follows easily from the first part. Since we now have \( T'_1 \leq T'_0 \) on \( (n, \bar{n}) \) with equality at the end points, we obtain \( \Omega_0(n) \geq \Omega_1(n) \) on \( (n, \bar{n}) \) with equality at the end points. Then we have that \( \dot{\Omega}_0(\bar{n}) \leq \dot{\Omega}_1(\bar{n}) \), which implies \( 1 - g_0 + \Delta T \cdot p/(1 - P) \geq 1 - g_1 - \Delta T \cdot p/P \) at \( \bar{n} \). Because \( g_0(\bar{n}) - g_1(\bar{n}) > 0 \), we have \( \Delta T(\bar{n}) > 0 \).

Finally, \( T'_1 \leq T'_0 \) and hence \( z_1 \geq z_0 \) implies \( \dot{q} = \dot{V}_1 - \dot{V}_0 \geq 0 \) from equation (5). Hence, \( \tau(n) = (w - \bar{q}(n))/w \geq (w - \bar{q}(\bar{n}))/w = \Delta T(\bar{n})/w > 0 \) for all \( n \), where the last equality follows from \( T'_1 = T'_0 = 0 \) at \( \bar{n} \). □
We may summarize our findings as follows. In the work cost model, second-earner participation is a signal of low fixed work costs and hence being better off than one-earner couples. This implies $g_0(n) > g_1(n)$ making it optimal to tax secondary earnings: $\tau > 0$. In the home production model, second-earner participation is a signal of low ability in home production and hence being worse off than one-earner couples. In this model, it is therefore optimal to subsidize secondary earnings, $\tau < 0$.

In either model, the redistribution between one- and two-earner couples gives rise to a distortion in the entry-exit decision of secondary earners, creating an equity-efficiency trade-off. The size of the efficiency cost does not depend on the ability of the primary earner because spousal characteristics $q$ and $n$ are independently distributed. An increase in $n$ therefore influences the optimal second-earner distortion only through its impact on the equity gain as reflected by $g_0(n) - g_1(n)$. Because the contribution of the secondary earner to couple utility is declining in relative terms, the value of redistribution between one- and two-earner couples is declining in $n$, i.e. $g_0(n) - g_1(n)$ is decreasing in $n$. Therefore, the second-earner distortion is declining with primary earnings. As shown in Proposition 2, if the ability distribution of primary earners is unbounded, the secondary-earner distortion tends to zero at the top. It is important to note that if $\Psi$ is quadratic, then $g_0 - g_1$ is constant in $n$ and the optimal tax system is separable. If $\Psi'$ is concave, then $g_0 - g_1$ increases in $n$ and the distortion on spouses actually increases with $n$. As discussed above, the case $\Psi'$ convex (Assumption 1) fits best with the intuition that secondary earnings affect marginal social utility less when primary earnings are higher.

4 Numerical Calibration for the United Kingdom

We focus on the work cost model and make the following parametric assumptions. First, we assume $h(x) = \frac{x^{\epsilon}}{1 + \epsilon} x^{\frac{1}{1 + \epsilon}}$, implying a constant elasticity of primary earnings $\epsilon$. Second, we assume that $q$ is distributed as a power function on the interval $[0, q_{\text{max}}]$ with distribution function $P(q) = (q/q_{\text{max}})^\eta$. This implies a constant elasticity of second-earner participation $\eta$. Third, we assume that the social welfare function is CRRA, $\Psi(V) = V^{1-\gamma}/(1-\gamma)$, where $\gamma > 0$ measures preferences for equity.

We calibrate the ability distribution $F(n)$ and $q_{\text{max}}$ using the British Family Resource Survey for 2004/5 linked to the tax-benefit microsimulation model TAXBEN at the Institute for Fiscal
Studies. We define the primary earner as the husband and the secondary earner as the wife in each family. Figure 3A depicts the actual tax rates $T_0^\prime$, $T_1^\prime$, $\tau$ faced by couples in the United Kingdom. As in Saez (2001), $f(n)$ is calibrated such that, at the actual marginal tax rates, the resulting distribution of primary earnings matches the empirical earnings distribution for married men. The top quintile of the distribution (above $n = £46,000$) is approximated by a Pareto distribution with coefficient $a = 2$, which is a good approximation according to Brewer et al. (2008). Figure 3B depicts the calibrated density distribution $f(n)$. The dashed line depicts the raw un-smoothed density distribution. The solid line is the smoothed density we use in order to obtain smooth optimal schedules.

Figure 3C shows that the participation rate of wives conditional on husbands’ earnings is fairly constant across the earnings distribution and equal to 75% on average. Figure 3D shows that average female earnings, conditional on participation, are slightly increasing in husbands’ earnings. Our model with homogenous secondary earnings does not capture this feature. Therefore, we assume (except when we explore the effects of positive correlation in spousal abilities) that $q_{\text{max}}$ and hence $q$ is independent of $n$. $q_{\text{max}}$ is calibrated so that the average participation rate (under the current tax system) matches the empirical rate. The $w$-parameter is set equal to average female earnings conditional on participation. It is important to note that the combination of positive correlation in abilities across spouses and income effects could also generate the empirical patterns. Analyzing a calibrated case with income effects is beyond the scope of this note and left for future work.

Based on the empirical labor supply literature for the U.K. (see Brewer et al. 2008), we assume $\varepsilon = 0.25$ and $\eta = 0.5$ in our benchmark case. Based on estimates of the curvature of utility functions consistent with labor supply responses, we set $\gamma$ equal to 1 (see e.g., Chetty 2006). Finally, we assume that the simulated optimal tax system (net of transfers) must collect as much tax revenue (net of transfers) as the actual U.K. tax system, which we compute using TAXBEN and the empirical data. In all simulations, we check that the implementation conditions ($z_l(n)$ increasing in $n$) are satisfied. All the technical details of the simulations are described in an electronic appendix to the paper.

Figure 4A plots the optimal $T_0^\prime$, $T_1^\prime$, and $\tau$ as a function of $n$ in our benchmark case. Consistent with our theoretical results, we have $T_1^\prime < T_0^\prime$ and $\tau$ declining in $n$. Consistent with earlier work on the single-earner model (e.g., Saez, 2001), optimal marginal tax rates on primary earners...
follow a U-shape, with very high marginal rates at the bottom corresponding to the phasing out of welfare benefits, lower rates at the middle, and increasing rates at the top converging to 66.7% = 1/(1 + a · ε). The difference between $T'_1$ and $T'_0$ is about 8 percentage points on average, and $\tau$ is almost 40 percent at the bottom and then declines gradually towards zero. This suggests that the negative jointness property is not a negligible phenomenon. The simulations generate substantially higher tax rates on primary earners than on secondary earners because the primary-earner elasticity is smaller than the secondary-earner elasticity.

Figure 4B introduces positive correlation between spouses ability. In order to capture indirectly the positive spousal correlation in earnings displayed on Figure 3D, we assume that $q_{\max}$ depends on $n$, and that the fraction of spouses working (under the current tax system) increases smoothly from 55% to 80% across the distribution of primary earnings $n$. Figure 4B shows that introducing this (crudely) calibrated amount of correlation has minimal effect on optimal tax rates. Relative to no correlation, the secondary tax rate is slightly higher at the bottom so that the declining pattern for $\tau$ is actually strengthened. Figures 4C explores the effects of increasing redistributive tastes $\gamma$ from 1 to 2. Unsurprisingly, this increases tax rates across the board. Figure 4D considers the case with a higher primary earner elasticity $\varepsilon$ equal to 0.5 (instead of 0.25). As expected, this reduces the primary earner tax rates (especially at the top of the distribution).

Comparing the simulations with the empirical tax rates from Figure 3A is illuminating. The actual tax-transfer system also generates negative jointness, with the second-earner participation tax rate falling from about 40% at the bottom to about 20% at the middle and upper part the primary earnings distribution. This may seem surprising at first glance given that the U.K. operates an individual income tax system. However, income transfers in the U.K. (as in virtually all OECD countries) are means-tested based on family income. The combination of an individual income tax and a family-based, means-tested welfare system generates negative jointness: a wife married to a low-income husband will be in the phase-out range of welfare programs and hence face a high tax rate, whereas a wife married to a high-income husband is beyond the welfare benefit phase-out and hence faces a low tax rate because the income tax is individual. Therefore, our theoretical and numerical findings of negative jointness are roughly consistent with the current U.K. tax-transfer schedule. As for the size and profile of primary-earner tax rates, the current U.K. schedule displays lower rates at the very bottom (below
£6-7K) than the simulations. This might be justified by participation responses for low-income primary earners (as in Saez, 2002), which are not incorporated in our model. Above £6-7K, the current UK tax system does display a weak U-shape with the highest marginal rates at the bottom and modest increases above £40K.

Clearly, our calibration abstracts from several potentially important aspects such as income effects, heterogeneity in second-earner earnings, and endogenous marriage. Hence, our simulations should be seen as an illustration of our theory rather than actual policy recommendation. More complex and comprehensive numerical calibrations are left for future work.

A Appendix

A.1 Proof of Proposition 1

The government maximizes

\[ W = \int_{\frac{z}{n}}^{\bar{z}} \left\{ \int_{0}^{q} \Psi(V_{1}(n) - q^n)p(q|n) dq + \int_{q}^{\infty} \Psi(V_{0}(n) + q^n)p(q|n) dq \right\} f(n) dn, \]

where \( \bar{q} = V_{1}(n) - V_{0}(n) \), \( q = q^n + q^h \) and where either \( q^n = 0 \) or \( q^h = 0 \). The objective is maximized subject to the budget constraint

\[ \int_{\frac{z}{n}}^{\bar{z}} \left\{ \left[z_{1}(n) + w - nh \left( \frac{z_{1}}{n} \right) - V_{1}(n) \right] P(\bar{q}|n) + \left[z_{0}(n) - nh \left( \frac{z_{0}}{n} \right) - V_{0}(n) \right] (1 - P(\bar{q}|n)) \right\} f(n) dn \geq 0, \]

and the constraints arising from the couples utility maximization: \( \dot{V}_{l}(n) = -h(z_{1}(n)/n) + (z_{l}(n)/n)h'(z_{l}(n)/n) \) for \( l = 0, 1 \). Let us denote by \( \lambda, \mu_{0}(n), \mu_{1}(n) \), the associated multipliers and let \( H(z_{0}, z_{1}, V_{0}, V_{1}, \mu_{0}, \mu_{1}, \lambda, n) \) denote the Hamiltonian function. The transversality conditions are \( \mu_{0}(\bar{z}) = \mu_{1}(\bar{z}) = \mu_{0}(\bar{\bar{z}}) = \mu_{1}(\bar{\bar{z}}) = 0 \). We abbreviate \( h(z_{1}(n)/n) \) into \( h_{1} \), etc. The first order conditions with respect to \( z_{0}(n) \) and \( z_{1}(n) \) equal

\[ H'_{z_{0}} = \mu_{0} \cdot \frac{z_{0}}{n^2} h_{0}^{n} + \lambda \cdot (1 - h_{0}') \cdot (1 - P(\bar{q}|n)) \cdot f(n) = 0, \]  

\[ H'_{z_{1}} = \mu_{1} \cdot \frac{z_{1}}{n^2} h_{1}^{n} + \lambda \cdot (1 - h_{1}') \cdot P(\bar{q}|n) \cdot f(n) = 0. \]  

Routine differentiation implies that:

\[ H''_{z_{0}z_{0}} = \frac{\mu_{0}}{n^2} \cdot \left(h_{0}'' + \frac{z_{0}}{n} h_{0}''' \right) - \frac{\lambda h_{0}'}{n} (1 - P) f = -\frac{\lambda (1 - P) f}{z_{0}} \left[ (1 - h_{0}') \left(1 + \frac{z_{0}}{n} h_{0}' \right) + \frac{z_{0}}{n} h_{0}'' \right], \]

where the last equality is obtained substituting \( \mu_{0} \) from equation (14). It is straightforward to show that the expression in square brackets being positive is equivalent to Assumption 1 (namely
$x \rightarrow (1 - h')/(x \cdot h'')$ decreasing. Therefore $H''_{z_{0}z_{0}} < 0$. Similarly $H''_{z_{1}z_{1}} < 0$. Furthermore, $H''_{z_{0}z_{1}} = 0$ implying that the Hamiltonian $z \rightarrow H(z, V(n), \mu(n), \lambda, n)$ is strictly concave in $z$. This implies that the optimal solution $z(n)$ (and hence $T'_0, T'_1$) is continuous in $n$ (see e.g., Seierstad and Sydsæter 1987, pp. 85-86, Note 2(b)).

The first order conditions in $V_0, V_1$ are $-\dot{\mu}_0 = H'_{V_0}$ and $-\dot{\mu}_1 = H'_{V_1}$ where

\[
H'_{V_0} = \int_{q}^{\infty} \Psi'(V_0(n) + q^{h})p(q|n)f(n)\,dq - \lambda(1 - P(\bar{q}|n))f(n) - \lambda[T_1 - T_0]p(\bar{q}|n)f(n), \quad (16)
\]

\[
H'_{V_1} = \int_{0}^{q} \Psi'(V_1(n) - q^{h''})p(q|n)f(n)\,dq - \lambda P(\bar{q}|n)f(n) + \lambda[T_1 - T_0]p(\bar{q}|n)f(n). \quad (17)
\]

Using the social marginal welfare weights $g_0(n)$ and $g_1(n)$, we can integrate those two equations using the upper transversality conditions and obtain:

\[
-\frac{\mu_0(n)}{\lambda} = \int_{n}^{\bar{n}} \{ [1 - g_0(n')] \lambda - (1 - P(\bar{q}|n'))f(n') + [T_1 - T_0]p(\bar{q}|n')f(n') \} \,dn',
\]

\[
-\frac{\mu_1(n)}{\lambda} = \int_{n}^{\bar{n}} \{ [1 - g_1(n')] \lambda - (1 - P(\bar{q}|n'))f(n') - [T_1 - T_0]p(\bar{q}|n')f(n') \} \,dn'.
\]

Plugging these two equations into the first order conditions for $z_0$ and $z_1$, noting that $T'_l = 1 - h'_l$, and using the definition of the labor supply intensive elasticity ($3$), $\varepsilon_l = h'_l/(h''_l \cdot z_l/n)$, we obtain the expressions (9) and (10) in Proposition 1.

The transversality conditions imply that $T'_l = T'_0 = 0$ at the end points $n_2$ and $\bar{n}$.

As discussed in electronic appendix B.1, $z_0$ and $z_1$ weakly increasing in $n$ is a necessary and sufficient condition for implementability (exactly as in the one dimensional Mirrlees model). If (9) and (10) generate decreasing ranges for $z_0$ or $z_1$ then there is bunching and the formulas do not apply on the bunching portions.

### A.2 Proof of Lemma in Proposition 3

**Lemma 1** Under Assumptions 1-4, if $T'_1 > T'_0$ on $(n_a, n_b)$ with equality at the end points, then $g_0(n_a) - g_1(n_a) > g_0(n_b) - g_1(n_b)$.

We have $\bar{q} = V_1 - V_0$ and $g_0(n) - g_1(n) = \Psi'(V_0)/\lambda - \int_{0}^{\bar{q}} \Psi'(V_1 - q)p(q)\,dq/\lambda \cdot P(\bar{q}) > 0$ as $\Psi'$ decreasing. Differentiating with respect to $n$, we have:

\[
\dot{g}_0(n) - \dot{g}_1(n) = \dot{V}_0 \cdot \frac{\Psi''(V_0)}{\lambda} - \dot{V}_1 \cdot \frac{\int_{0}^{\bar{q}} \Psi''(V_1 - q)p(q)\,dq}{\lambda \cdot P(\bar{q})} + \frac{p(\bar{q})\dot{\bar{q}}}{P(\bar{q})} \cdot \left[ \frac{\int_{0}^{\bar{q}} \Psi'(V_1 - q)p(q)\,dq}{\lambda \cdot P(\bar{q})} - \frac{\Psi'(V_0)}{\lambda} \right].
\]

18
Rearranging, we obtain:

\[
\dot{g}_0(n) - \dot{g}_1(n) = \dot{V}_1 \cdot \left[ \frac{\Psi''(V_0)}{\lambda} - \frac{\int_0^\eta \Psi''(V_0 + \bar{q} - q)p(q) dq}{\lambda \cdot P(\bar{q})} \right] + \dot{q} \left[ - (g_0(n) - g_1(n)) \cdot \frac{p(\bar{q})}{P(\bar{q})} - \frac{\Psi''(V_0)}{\lambda} \right].
\]

(18)

The first term in expression is negative because \(\dot{V}_1 > 0\) and by Assumption 2, \(\Psi''\) is increasing and hence the term inside the first square brackets is negative. In the segment \((n_a, n_b)\), \(z_1 < z_0\) and hence \(\dot{q} < 0\). Furthermore, \(\Psi'\) convex implies that \(\Psi'(V_0) - \Psi'(V_0 + \bar{q} - q) \leq -\Psi''(V_0) \cdot (\bar{q} - q)\) and hence:

\[
g_0(n) - g_1(n) = \int_0^\eta \frac{[\Psi'(V_0) - \Psi'(V_0 + \bar{q} - q)]p(q) dq}{\lambda \cdot P(\bar{q})} \leq -\Psi''(V_0) \cdot \frac{\int_0^\eta P(q) dq}{\lambda \cdot P(\bar{q})}.
\]

(19)

using that \(\int_0^\eta (\bar{q} - q)p(q) dq = \int_0^\eta P(q) dq\) by integration by parts and \(P(0) = 0\). Therefore, combining (19) and the second part of Assumption 4 implies that \((g_0(n) - g_1(n)) \cdot p(\bar{q})/P(\bar{q}) \leq -\Psi''(V_0)/\lambda\). Therefore, the second term in square brackets in expression (18) above is non-negative. Thus, the second term in (18) is non-positive. As a result, \(\dot{g}_0(n) - \dot{g}_1(n) < 0\) on \((n_a, n_b)\) and the Lemma is proven.

References


Figure 1. Zero second-earner tax at the top

Tax paid

$T_1$: Two-earner Couples

$T_0$: One-earner Couples

$\Delta T^\infty > 0$
Figure 2. Desirability of Negative Jointness

Tax paid

T₁: Two-earner Couples

T₀: One-earner Couples

\[ T_1 - T_0 > 0 \]
Figure 3: Numerical Simulations: Current System

A. Current Tax Rates

B. Primary Earnings Distribution

C. Spousal Participation Rate

D. Average Spousal Earnings (conditional on work)

Notes: Computations are based on the British Family Resource Survey for 2004/05 and TAXBEN tax/transfer calculator.
Figure 4: Optimal Tax Simulations

A. Benchmark (no correlation, $\gamma = 1$, $\eta = 0.5$, $\varepsilon = 0.25$)

B. Positive Ability Correlation across Spouses

C. Higher redistributive Tastes, $\gamma = 2$

D. Higher Primary Earner Elasticity, $\varepsilon = 0.5$

Potential Primary Earnings n (annual 000s Pounds) Optimal Tax Rates

Notes: Computations are based on the British Family Resource Survey for 2004/05 and TAXBEN tax/transfer calculator.
B Electronic Appendix (not for publication)

B.1 Implementation

As in the one-dimensional mechanism design theory, we define implementability as follows: An action profile \((z_0(n), z_1(n))_{n \in \{0, 1\}}\) is implementable if and only if there exists transfer functions \((c_0(n), c_1(n))_{n \in \{0, 1\}}\) such that \((z_l(n), c_l(n))_{l \in \{0, 1\}, n \in \{0, 1\}}\) is a simple truthful mechanism.\(^{10}\) The central implementability theorem of the one-dimensional case carries over to our model.

**Lemma 2** An action profile \((z_0(n), z_1(n))_{n \in \{0, 1\}}\) is implementable if and only if \(z_0(n)\) and \(z_1(n)\) are both non-decreasing in \(n\).

**Proof:**

The utility function \(c - nh(z/n)\) satisfies the classic single crossing (Spence-Mirrlees) condition. Hence, from the one-dimensional case, we know that \(z(n)\) is implementable, i.e., there is some \(c(n)\) such that \(c(n) - nh(z(n)/n)) \geq c(n') - nh(z(n')/n))\) for all \(n, n'\), if and only if \(z(n)\) is non-decreasing.

Suppose \((z_0(n), z_1(n))\) is implementable, implying that there exists \((c_0(n), c_1(n))\) such that \((z_l(n), c_l(n))_{l \in \{0, 1\}, n \in \{0, 1\}}\) is a simple truthful mechanism. That implies in particular that \(c_l(n) - nh(z_l(n)/n)) \geq c_l(n') - nh(z_l(n')/n))\) for all \(n, n'\) and for \(l = 0, 1\). Hence, the one dimensional result implies that \(z_0(n)\) and \(z_1(n)\) are non-decreasing.

Conversely, suppose that \(z_0(n)\) and \(z_1(n)\) are non-decreasing. Because \(z_0(n)\) is non-decreasing, the one dimensional result implies there is \(c_0(n)\) such that \(c_0(n) - nh(z_0(n)/n)) \geq c_0(n') - nh(z_0(n')/n))\). Similarly, there is \(c_1(n)\) such that \(c_1(n) - nh(z_1(n)/n)) \geq c_1(n') - nh(z_1(n')/n))\).

It is easy to show that the mechanism \((z_l(n), c_l(n))_{l \in \{0, 1\}, n \in \{0, 1\}}\) is actually truthful. Define \(V_l(n) = c_l(n) - nh(z_l(n)/n))\) for \(l = 0, 1\) and \(\bar{q}(n) = V_1(n) - V_0(n)\). We only need to prove the cross-inequalities. For all \(n, n', q \geq \bar{q}(n)\),

\[
\begin{align*}
  u(z_0(n), 0, c_0(n), (n, q)) &= V_0(n) \geq V_1(n) - q \geq u(z_1(n'), 1, c_1(n'), (n, q)).
\end{align*}
\]

For all \(n, n', q < \bar{q}(n)\),

\[
\begin{align*}
  u(z_1(n), 1, c_1(n), (n, q)) &= V_1(n) - q \geq V_0(n) \geq u(z_0(n'), 0, c_0(n'), (n, q)).
\end{align*}
\]

---

\(^{10}\) A mechanism is defined a truthful if there is a \(\bar{q}(n)\) so that: (1) When \(q < \bar{q}(n)\), \((l' = 1, n' = n)\) maximizes \(u(z_{l'}(n'), l', c_{l'}(n'), (n, q))\) over all \((l', n')\). (2) When \(q \geq \bar{q}(n)\), \((l' = 0, n' = n)\) maximizes \(u(z_{l'}(n'), l', c_{l'}(n'), (n, q))\) over all \((l', n')\).
The key assumption that allows us to obtain those simple results is the fact that \( q \) is separable in our utility specification. □

**B.2 Numerical Simulations**

Simulations are performed with Matlab software and our programs are available upon request. We select a grid for \( n \), from \( n \) to \( \bar{n} \) with 1000 elements: \((n_k)_k\). Integration along the \( n \) variable is carried out using the trapezoidal approximation. All integration along the \( q \) variable is carried out using explicit closed form solutions using the incomplete \( \beta \) function:

\[
\int_{V_0}^{V_1} \Psi'(V_1 - q)p(q) dq = \int_{V_0}^{V_1} \frac{1}{(V_1 - q)\gamma} \frac{\eta \cdot q^{\eta-1}}{\eta_{\text{max}}} dq
\]

where the incomplete beta function \( \beta \) is defined as (for \( 0 \leq x \leq 1 \)):

\[
\beta(x, a, b) = \int_0^x t^{a-1}(1-t)^{b-1} dt.
\]

Matlab does not compute it directly for \( \gamma \geq 1 \) (\( b \leq 0 \)) but we have used the development in series to compute it very accurately and quickly with a subroutine:

\[
\beta(x, a, b) = 1 + \sum_{n=1}^{\infty} \frac{(1-b)(2-b)...(n-b)}{n!} \cdot \frac{x^{n+a}}{n+a}.
\]

Simulations proceed by iteration:

We start with given \( T_0', T_1' \) vectors, derive all the vector variables \( z_0, \ z_1, V_0, V_1, \bar{q}, T_0, T_1, \lambda, \) etc. which satisfy the government budget constraint and the transversality conditions.\(^{11}\) This is done with a sub-iterative routine that adapts \( T_0 \) and \( T_1 \) as the bottom \( n \) until those conditions are satisfied. We then use the first order conditions (9), (10) from Proposition 1 to compute new vectors \( T_0', T_1' \). In order to converge, we use adaptive iterations where we take as the new vectors \( T_0', T_1' \), a weighted average of the old vectors and newly computed vectors. The weights are adaptively adjusted down when the iteration explodes. We then repeat the algorithm.

This procedure converges to a fixed point in most circumstances. The fixed point satisfies all the constraints and the first order conditions. We check that the resulting \( z_0 \) and \( z_1 \) are

\(^{13}\)The adjust the constants for \( T_l(n) \) until all those constraints are satisfied. This is done using a secondary iterative procedure.
non-decreasing so that the fixed point solution is implementable. So the fixed point is expected to be the optimum.\footnote{We also compute total social welfare and verify on examples that it is higher than social welfare generated by other tax rates $T'_1, T'_0$ satisfying the government budget constraint.}

The central advantage of our method is that the optimal solution can be approximated very closely and quickly. In contrast, direct maximization where we search the optimum over a large set of parametric tax systems by computing directly social welfare would be much slower and less precise.