Econometrica, Vol. 77, No. 2 (March, 2009), 537–560

NOTES AND COMMENTS

THE OPTIMAL INCOME TAXATION OF COUPLES

BY HENRIK JACOBSEN KLEVEN, CLAUS THUSTRUP KREINER, AND EMMANUEL SAEZ[1]

This paper analyzes the general nonlinear optimal income tax for couples, a multidimensional screening problem. Each couple consists of a primary earner who always participates in the labor market, but makes an hours-of-work choice, and a secondary earner who chooses whether or not to work. If second-earner participation is a signal of the couple being better (worse) off, we prove that optimal tax schemes display a positive tax (subsidy) on secondary earnings and that the tax (subsidy) on secondary earnings *decreases* with primary earnings and converges to zero asymptotically. We present calibrated microsimulations for the United Kingdom showing that decreasing tax rates on secondary earnings is quantitatively significant and consistent with actual income tax and transfer programs.

KEYWORDS: Optimal income tax, multidimensional screening.

1. INTRODUCTION

THIS PAPER EXPLORES the optimal income taxation 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. Primary earners differ in ability and make a continuous labor supply decision as in the Mirrlees (1971) model. Secondary earners differ in opportunity costs of work and make a binary labor supply decision (work or not work). We consider a fully general nonlinear tax system allowing us to study the central question of couple taxation: how should the tax rate on one individual vary with the earnings of the spouse. This creates a multidimensional screening problem. We show that if second-earner labor force participation is a signal of the couple being better off (as when second-earner entry reflects high labor 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 (as when second-earner entry reflects low home production ability), 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. These results imply that, in either case, the tax distortion on

[1]We thank the co-editor, Mark Armstrong, Richard Blundell, Mike Brewer, Raj Chetty, Steven Durlauf, Nada Eissa, Kenneth Judd, Botond Koszegi, Etienne Lehmann, Randall Mariger, Jean-Charles Rochet, Andrew Shephard, four anonymous referees, and numerous seminar and conference participants for very helpful comments and discussions. Financial support from NSF Grant SES-0134946 and an Economic Policy Research Network (EPRN) Grant is gratefully acknowledged.

 DOI: 10.3982/ECTA7343

the secondary earner is *declining* in primary earnings, which is therefore a general property of an optimum. We also prove that the second-earner tax distortion tends to zero asymptotically as primary earnings become large. Although this result may seem reminiscent of the classic no-distortion-at-the-top result, our result rests on a completely different reasoning and proof.

Previous work on couple taxation assumed separability in the tax function and, hence, could not address the optimal form of jointness, which we view as central to the optimal couple tax problem.[2] The separability assumption also sidesteps the complexities associated with multidimensional screening. In fact, very few studies in the optimal tax literature have attempted to deal with multidimensional screening problems.[3] The nonlinear pricing literature in industrial organization has analyzed such problems extensively. A central complication of multidimensional 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 with different types are making the same choices. Our framework with a binary labor supply outcome for the secondary earner along with continuous earnings for the primary earner avoids the bunching complexities and offers a simple understanding of the shape of optimal taxes based on graphical exposition.

Our key results are obtained under a number of strong simplifying assumptions:[4] (i) We adopt the unitary model of family decision making. (ii) We assume that the government knows a priori the identity of the primary and secondary earner in the couple. (iii) We consider only couples and do not model the marriage decision. (iv) We assume uncorrelated abilities between spouses. (v) 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. 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.

[2]Boskin and Sheshinski (1983) considered linear taxation of couples, allowing for different marginal tax rates on husband and wife. The linearity assumption effectively implies separable and hence individual-based (albeit gender-specific) tax treatment. More recently, Schroyen (2003) extended the Boskin–Sheshinski framework to the case of nonlinear taxation but kept the assumption of separability in the tax treatment.

[3]Mirrlees (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 multidimensional screening model assuming a discrete number of types. Brett (2006) and Cremer, Lozachmeur, and Pestieau (2007) considered the issue of couple taxation in discrete-type models. They showed that, in general, incentive compatibility constraints bind in complex ways, making it difficult to obtain general properties.

[4]We refer to Kleven, Kreiner, and Saez (2006) for a discussion of robustness and generalizations.

Section 2 sets out our model and Section 3 derives our theoretical results. Section 4 presents a numerically calibrated illustrative simulation based on U.K. micro data. Some proofs are presented in Appendices A and B, while some supplemental material is available on the journal's website (Kleven, Kreiner, and Saez (2009)).

2. THE MODEL

2.1. *Family Labor Supply Choice*

We consider a population of couples, the size of which is normalized to 1. 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 distributed on $(\underline{n}, \bar{n})$ in the population. The cost of earning z for a primary earner with ability n is given by $n \cdot h(z/n)$, where $h(\cdot)$ 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 redistributes based on observed earnings using a nonlinear tax $T(z, wl)$. 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.[5]

The tax system is separable iff 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

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

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. Het-

[5]Like the rest of the literature, we assume that the government observes the identity of the primary and secondary earner in each couple, and is allowed to use this information in the tax system. If identity could not be used in the calculation of taxes (a so-called anonymous tax system), a symmetry constraint $T(z, w) = T(w, z)$ would have to be added to the problem. However, this symmetry constraint can be ignored if the secondary earner is always the lower-earnings spouse in the couple. In the context of our simple model (where w is uniform), this assumption is equivalent to $w < z(\underline{n})$. When identity is perfectly aligned with earnings, an earnings-based and anonymous tax can be made dependent on identity de facto without being identity-specific de jure. This is important in countries where an identity-specific (e.g., gender-specific) tax system would be unconstitutional.

erogeneity 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 they will be better off than one-earner couples. This creates a motive for the government to tax the income of the secondary earner so as 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 they will 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 with the underlying notion in the existing optimal tax literature that higher income is a signal of higher well-being.[6] On the other hand, the existing literature did not consider two-person households where home production (including child-bearing and child-caring) is more important. We therefore analyze both models symmetrically. The online supplemental material has a discussion of the general case with both types of heterogeneity.

If T_0 and T_1 are differentiable, the first-order condition for z (conditional on $l = 0, 1$) is $h'(z_l/n) = 1 - T_l'(z)$.[7] 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.[8] Positive marginal tax rates depress actual earnings z below potential earnings n. If the tax system is nonseparable 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

[6]It is this notion that drives the result in the Mirrlees model that optimal marginal tax rates are positive. If differences in market earnings were driven by home production ability instead of market ability, the Mirrlees model would generate negative optimal tax rates as high-earnings individuals are those with low ability and utility. Ramey (2008) showed that primary earners provide significant home production but the main question is whether this effect is strong enough to make the poor better off than the rich, and thereby reverse the traditional results.

[7]If the tax system is not differentiable, we can still define the implicit marginal tax rate T_l' (with slight abuse of notation) as $1 - h'(z_l/n)$, where z_l is the utility maximizing choice of earnings conditional on l.

[8]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 $n \cdot (1 - T'(z)) = h'(z/n)$. Our results carry over to this case but n would no longer reflect potential earnings and the interpretation of optimal tax formulas would be less transparent (Saez (2001)).

elasticity of primary earnings with respect to the net-of-tax rate $1 - T_l'$ as

$$\varepsilon_l \equiv \frac{1 - T_l'}{z_l} \frac{\partial z_l}{\partial(1 - T_l')} = \frac{h'(z_l/n)}{(z_l/n)h''(z_l/n)}.$$

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 nonparticipation. Let us denote by

$$(2) \qquad V_l(n) = z_l - T_l(z_l) - nh\left(\frac{z_l}{n}\right) + w \cdot l$$

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

$$(3) \qquad \dot{V}_l(n) = -h\left(\frac{z_l}{n}\right) + \frac{z_l}{n} \cdot h'\left(\frac{z_l}{n}\right) \geq 0.$$

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 $z_l > 0$, that is, if $T_l' < 1$. The participation constraint for secondary earners is given by

$$(4) \qquad q \leq V_1(n) - V_0(n) \equiv \bar{q}(n),$$

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

The couple characteristics (n, q) are distributed according to a continuous density distribution defined over $[\underline{n}, \bar{n}] \times [0, \infty)$. We denote by $P(q|n)$ the cumulative distribution function of q conditional on n, by $p(q|n)$ the density function of q conditional on n, and by $f(n)$ 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 $\eta = \bar{q} \cdot p(\bar{q}|n)/P(\bar{q}|n)$.

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 the tax rate on secondary earnings 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 $\tau = (T_1 - T_0)/w$. If taxation is nonseparate, 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.

The central optimal couple tax question we want to tackle is whether the tax rate on one person should depend on the earnings of the spouse. We may define the possible forms of couple taxation as follows:

DEFINITION 1: At any point n, we have either (i) positive jointness, $T_1' > T_0'$ and $\dot{\tau} > 0$, (ii) separability, $T_0' = T_1'$ and $\dot{\tau} = 0$, or (iii) negative jointness, $T_1' < T_0'$ and $\dot{\tau} < 0$.[9]

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. If the secondary earner starts working, optimal primary earnings shift from $z_0(n)$ to $z_1(n)$ but the key first-order condition (3) continues to apply. As in the Mirrlees model, a given path for $(z_0(n), z_1(n))$ can be implemented via a truthful mechanism or, equivalently, by a nonlinear tax system if and only if $z_0(n)$ and $z_1(n)$ are nonnegative and nondecreasing in n (a formal proof is provided in the online supplemental material).

2.2. *Government Objective*

The government sets $T_0(z)$ and $T_1(z)$ to maximize social welfare

$$W = \int_{n=\underline{n}}^{\bar{n}} \int_{q=0}^{\infty} \Psi(V_l(n) - q^w \cdot l + q^h \cdot (1-l))\, p(q|n) f(n)\, dq\, dn, \tag{5}$$

where $\Psi(\cdot)$ is an increasing and concave transformation (representing either the government redistributive preferences or individual concave utilities) subject to the budget constraint

$$\int_{n=\underline{n}}^{\bar{n}} \int_{q=0}^{\infty} T_l(z_l)\, p(q|n) f(n)\, dq\, dn \geq 0 \tag{6}$$

and subject to $\dot{V}_0(n)$ and $\dot{V}_1(n)$ in equation (3).

Let $\lambda > 0$ be the multiplier associated with the budget constraint (6). The government's redistributive tastes may be represented by social marginal welfare weights on different couples. We denote by $g_l(n)$ the (average) social marginal welfare weight for couples with primary-earner ability n and secondary-earner participation status l. For the work cost model ($q^w > 0$, $q^h = 0$), we have $g_1(n) = \int_0^{\bar{q}} \Psi'(V_1(n) - q^w) p(q|n)\, dq / (P(\bar{q}|n) \cdot \lambda)$ 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 $g_0(n) = \int_{\bar{q}}^{\infty} \Psi'(V_0(n) + q^h) p(q|n)\, dq / ((1 - P(\bar{q}|n)) \cdot \lambda)$.

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

[9]Using equations (2)–(4), 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. Notice that T_0' and T_1' are evaluated at the same ability level n but not at the same earnings level when $T_0' \neq T_1'$ because this implies $z_0(n) \neq z_1(n)$.

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, we have $V_1(n) - q^w > V_0(n)$, which implies (as Ψ is concave) that $g_0(n) - g_1(n) > 0$. By contrast, in the home production model, we have $V_0(n) + q^h > V_1(n)$ and hence $g_0(n) - g_1(n) < 0$. 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 Mirlees* (1971)

The simple model described above makes it possible to derive explicit optimal tax formulas as in the individualistic Mirrlees (1971) model. We introduce the following assumption:

ASSUMPTION 1: *The function* $x \longrightarrow (1 - h'(x))/(x \cdot h''(x))$ *is decreasing*.

Assumption 1 ensures that the marginal deadweight loss $\varepsilon \cdot \frac{T'}{1-T'} = \frac{1-h'(z/n)}{(z/n)h''(z/n)}$ is increasing in T'. When Assumption 1 fails, ε falls so quickly with T' that the marginal deadweight loss falls with T', and such a point can never be optimum.[10] Assumption 1 is satisfied, for example, for isoelastic utilities $h(x) = x^{1+1/\varepsilon}/(1 + 1/\varepsilon)$ or any utility function such that the elasticity $\varepsilon = h'/(x \cdot h'')$ is decreasing in x. We prove the following proposition in Appendix A:

PROPOSITION 1: *Under Assumption 1, an optimal solution exists such that* (z_0, z_1, T_0', T_1') *is continuous in n and satisfies*

$$\frac{T_0'}{1-T_0'} = \frac{1}{\varepsilon_0} \cdot \frac{1}{nf(n)(1-P(\bar{q}|n))} \cdot \int_n^{\bar{n}} \big\{(1-g_0)(1-P(\bar{q}|n')) + [T_1 - T_0]p(\bar{q}|n')\big\} f(n')\,dn', \tag{7}$$

$$\frac{T_1'}{1-T_1'} = \frac{1}{\varepsilon_1} \cdot \frac{1}{nf(n)P(\bar{q}|n)} \cdot \int_n^{\bar{n}} \{(1-g_1)P(\bar{q}|n') - [T_1 - T_0]p(\bar{q}|n')\} f(n')\,dn', \tag{8}$$

[10]Mathematically, Assumption 1 is required to ensure that the first-order condition of the government problem generates a maximum (instead of a minimum); see Appendix A.

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, that is, 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.

Kleven, Kreiner, and Saez (2006) presented 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

$$(9) \qquad (1-P(\bar{q}|n))\cdot\varepsilon_0\cdot\frac{T_0'}{1-T_0'}+P(\bar{q}|n)\cdot\varepsilon_1\cdot\frac{T_1'}{1-T_1'} = \frac{1}{nf(n)}\cdot\int_n^{\bar{n}}(1-\bar{g}(n'))f(n')\,dn',$$

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 (without income effects), implying 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 nonworking 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 from the transversality conditions (see Appendix A).[11]

3.2. *Asymptotic Properties of the Optimal Schedule*

Let the ability distribution of primary earners $f(n)$ have an infinite tail ($\bar{n}=\infty$). As top tails of income distributions are well approximated by the Pareto distribution (Saez (2001)), we assume that $f(n)$ has a Pareto tail with parameter $a>1$ ($f(n)=C/n^{1+a}$). We also assume that the distribution of work costs $P(q|n)$ converges to $P^{\infty}(q)$. We can then show the next proposition:

PROPOSITION 2: *Suppose T_1-T_0, T_0', T_1', and $\bar{q}$ converge to ΔT^{∞}, ${T'}_0^{\infty}<1$, ${T'}_1^{\infty}<1$, and $\bar{q}^{\infty}$ as $n\to\infty$. Then* (i) *g_0 and g_1 converge to the same value $g^{\infty}\geq 0$,* (ii) *the second-earner tax converges to zero, $\Delta T^{\infty}=\tau^{\infty}=0$, and* (iii) *the marginal tax rates on primary earners converge to ${T'}_0^{\infty}={T'}_1^{\infty}=(1-g^{\infty})/(1-g^{\infty}+a\cdot\varepsilon^{\infty})>0$, where ε^{∞} is the asymptotic elasticity.*

[11]As is well known, these results have limited relevance because (i) the bottom result does not apply when there is an atom of nonworkers, and (ii) the top rate drops to zero only for the single topmost earner (Saez (2001)).

PROOF: $V_0(n)$ and $V_1(n)$ are increasing in n without bound (as T_0', T_1' converge to values below 1). As $\Psi' > 0$ is decreasing, it must converge to $\bar{\psi} \geq 0$. Therefore, in the work cost model, $g_0 = \Psi'(V_0)/\lambda$ and $g_1 = \int_0^{\bar{q}} \Psi'(V_0 + \bar{q} - q) p(q|n)\, dq/[\lambda \cdot P(\bar{q}|n)]$ both converge to $g^\infty = \bar{\psi}/\lambda \geq 0$.[12] Because $T_1 - T_0$ converges, it must be the case that $T_0'^\infty = T_1'^\infty = T'^\infty$. Hence, as $h'(z_l/n) = 1 - T_l'$, z_l/n converge for both $l = 0, 1$ and $\varepsilon_l = h'(z_l/n)/(h''(z_l/n)z_l/n)$ also converges to ε^∞.

Because $P(\cdot|n)$ and $\bar{q}$ converge, $P(\bar{q}|n)$ and $p(\bar{q}|n)$ converge to $P^\infty(\bar{q}^\infty)$ and $p^\infty(\bar{q}^\infty)$. The Pareto assumption implies that $(1 - F(n))/(nf(n)) = 1/a$ for large n. Taking the limit of (7) and (8) as $n \to \infty$, we obtain, respectively, $T'^\infty/(1-T'^\infty) = (1/\varepsilon^\infty)(1/a)[1-g^\infty+\Delta T^\infty p^\infty/(1-P^\infty)]$ and $T'^\infty/(1-T'^\infty) = (1/\varepsilon^\infty)(1/a)[1-g^\infty-\Delta T^\infty p^\infty/(1-P^\infty)]$. Hence, we must have $\Delta T^\infty = 0$, and the formula for T'^∞ then follows. *Q.E.D.*

It is quite striking that the spouses of 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. In fact, 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 converges 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)).[13]

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. Consider then a reform that 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.[14] 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 g_0 and g_1 have converged to g^∞, 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 $T_1 - T_0$ has converged to a constant for large n, the marginal tax rates on one- and

[12]In the home production model, we also have $\bar{\psi}/\lambda \leq g_0 < g_1 = \Psi'(V_1)/\lambda \to \bar{\psi}/\lambda$.

[13]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.

[14]Because $\bar{q}$ and hence $P(\bar{q}|n)$ 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.

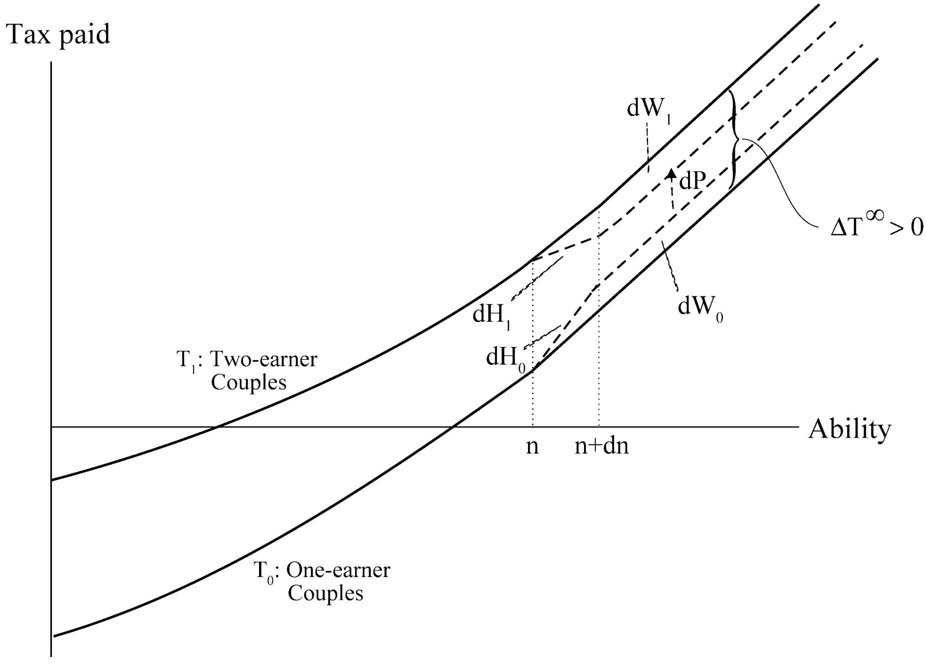

FIGURE 1.—Zero second-earner tax at the top.

two-earner couples are identical, $T'^{\infty}_0 = T'^{\infty}_1$, which implies $z_0/n = z_1/n$ and hence identical primary-earner elasticities $\varepsilon_0 = \varepsilon_1$. Thus, the negative fiscal effect dH_0 exactly offsets the positive fiscal effect dH_1. Third, there is a participation effect as some secondary earners are induced to join the labor force in response to the lower $T_1 - T_0$. Because $T_1 - T_0$ is initially positive, this response generates a positive fiscal effect, $dP > 0$. Since all other effects were zero, $dP > 0$ is the net total welfare effect of the reform, implying that the original schedule with $\Delta T^{\infty} > 0$ cannot be optimal.[15]

3.3. *Optimal Jointness*

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

ASSUMPTION 2: *The function $V \longrightarrow \Psi'(V)$ is strictly convex.*

This is satisfied for standard CRRA or CARA social welfare functions. In consumer theory, convexity of marginal utility of consumption is a common

[15]The opposite situation with $\Delta T^{\infty} < 0$ cannot be optimal either, because the reverse reform would then improve welfare.

assumption, because 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 of the spousal interaction 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, let us consider a tax reform introducing a little bit of jointness around the optimal *separable* tax system. For the work cost model, we will argue that the optimal separable schedule can be improved by introducing a little bit of *negative* jointness.[16]

A separable schedule is one where $T_0' = T_1'$, implying that $T_1 - T_0$, $\bar{q}$, 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 a 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, 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. Formally, 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

$$\text{(10)} \qquad \frac{d[g_0(n) - g_1(n)]}{dn} = \left[\frac{\Psi''(V_0)}{\lambda} - \frac{\int_0^{\bar{q}} \Psi''(V_0 + \bar{q} - q)\,p(q)\,dq}{\lambda \cdot P(\bar{q})}\right] \cdot \dot{V}_0 < 0,$$

where we have used $V_1 = V_0 + \bar{q}$ from equation (4). Since $\Psi''(\cdot)$ is increasing (by Assumption 2) and V_0 is increasing in n, it follows that the expression in (10) 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.

[16]In the home production model, reversed arguments show that some positive jointness is welfare improving.

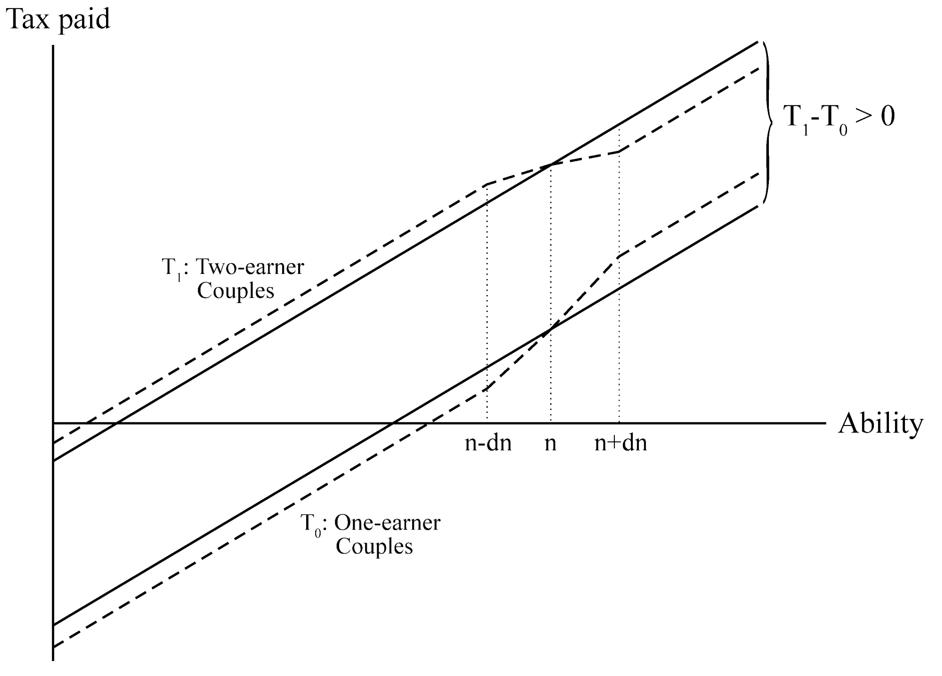

FIGURE 2.—Desirability of negative jointness.

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 $dT_0^a = dT/[(1-F(n))(1-P(\bar{q}))]$ and the tax change for two-earner couples is $dT_1^a = -dT/[(1-F(n))P(\bar{q})]$. Below n, the tax change for one-earner couples is $dT_0^b = dT/[F(n)(1-P(\bar{q}))]$ and the tax change for two-earner couples is $dT_1^b = dT/[F(n)P(\bar{q})]$. There are three effects.

First, there is a direct welfare effect created by the redistribution across couples at each n':

$$(11)\qquad dW = \frac{dT}{F(n)} \cdot \int_{\underline{n}}^{n} [g_0(n') - g_1(n')] f(n')\, dn' - \frac{dT}{1-F(n)} \cdot \int_{n}^{\bar{n}} [g_0(n') - g_1(n')] f(n')\, dn' > 0.$$

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. Equation (10) implies that the gain dominates the loss at the top, so that $dW > 0$.

Second, there are fiscal effects associated with earnings responses by primary earners induced by the changes in T_0' and T_1' around n. Since the reform increases the marginal tax rate for one-earner couples around n and reduces it for two-earner couples, the earnings responses are opposite. As we start from separable taxation, $T_0' = T_1'$, and hence identical primary-earner elasticities, $\varepsilon_0 = \varepsilon_1$, the fiscal effects of primary earner responses cancel out exactly.

Third, the reform creates participation responses by secondary earners. Above n, nonworking spouses will be induced to join the labor force. Below n, working spouses have an incentive to drop out. Because spouse characteristics q and n are independent, and since we start from a separable tax system, the participation elasticity $\eta = \bar{q} \cdot p(\bar{q})/P(\bar{q})$ and $T_1 - T_0$ are initially constant. Therefore, the fiscal implications of these responses also cancel out exactly.

Therefore, $dW > 0$ 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 suggests that, for the work cost model, the *optimal* incentive scheme will be associated with negative jointness, a point we will prove formally after introducing a final technical assumption:

ASSUMPTION 4: *The function* $x \longrightarrow x \cdot p(w-x)/[P(w-x) \cdot (1-P(w-x))]$ *is increasing and* $p(q)/P(q) \leq P(q)/\int_0^q P(q')\,dq'$ *for all* q.

This assumption is satisfied for isoelastic work cost distributions, $P(q) = (q/q_{\max})^{\eta}$, where the participation elasticity of secondary earners is constant and equal to η.[17]

PROPOSITION 3: *Under Assumptions* 1–4 *and if the optimal solution is not associated with bunching, the tax system is characterized by the following models*:

Work Cost Model: 1a. *Positive tax on secondary-earner income*, $\tau > 0$ *for all* $n \in [\underline{n}, \bar{n}]$. 1b. *Negative jointness*, $T_1' \leq T_0'$ *and* $\dot{\tau} \leq 0$ *for all* $n \in [\underline{n}, \bar{n}]$.

Home Production Model: 2a. *Negative tax on secondary-earner income*, $\tau < 0$ *for all* $n \in [\underline{n}, \bar{n}]$. 2b. *Positive jointness*, $T_1' \geq T_0'$ *and* $\dot{\tau} \geq 0$ *for all* $n \in [\underline{n}, \bar{n}]$.

PROOF: We consider the work cost model.[18] Suppose by contradiction that $T_1' > T_0'$ for some n. Then, because T_0' and T_1' are continuous in n 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 $z_1 < z_0$

[17]Assumption 4 can be seen as a counterpart to Assumption 1 for the participation margin. It ensures that the participation response does not decrease too fast with the tax rate. It was not needed for the small reform argument, because in that case the efficiency effects from participation responses cancel out to the first order.

[18]Results 2a and 2b may be established by reversing all inequalities in the proof below.

on (n_a, n_b) with equality at the end points. Assumption 1 implies

$$\begin{aligned}\varepsilon_1 T_1'/(1-T_1') &= \left(1-h'\left(\frac{z_1}{n}\right)\right)\Big/\left(\left(\frac{z_0}{n}\right)h''\left(\frac{z_1}{n}\right)\right)\\ &> \left(1-h'\left(\frac{z_0}{n}\right)\right)\Big/\left(\left(\frac{z_1}{n}\right)h''\left(\frac{z_0}{n}\right)\right)\\ &= \varepsilon_0 T_0'/(1-T_0')\end{aligned}$$

on (n_a, n_b). Then, because of our no bunching assumption, (7) and (8) imply

$$\begin{aligned}\Omega_0(n) &\equiv \frac{1}{1-P}\int_n^{\bar{n}}[(1-g_0)(1-P)+\Delta T\cdot p]f(n')\,dn'\\ &< \frac{1}{P}\int_n^{\bar{n}}[(1-g_1)P-\Delta T\cdot p]f(n')\,dn' \equiv \Omega_1(n)\end{aligned}$$

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 $\dot{\Omega}_0(n_a) \le \dot{\Omega}_1(n_a)$ and $\dot{\Omega}_0(n_b) \ge \dot{\Omega}_1(n_b)$. At the end points, we have $T_1' = T_0'$, $z_0 = z_1$, and $\dot{V}_0 = \dot{V}_1$, which implies $\dot{\bar{q}} = 0$ and $\dot{P} = 0$. Hence, the inequalities in derivatives can be written as

$$1-g_0+\Delta T\cdot p/(1-P)\begin{cases}\ge 1-g_1-\Delta T\cdot p/P & \text{at } n_a,\\ \le 1-g_1-\Delta T\cdot p/P & \text{at } n_b.\end{cases}$$

Combining these inequalities, we obtain

$$\left.\frac{\Delta T\cdot p}{P(1-P)}\right|_{n_a} \ge g_0(n_a)-g_1(n_a) > g_0(n_b)-g_1(n_b) \ge \left.\frac{\Delta T\cdot p}{P(1-P)}\right|_{n_b}.$$

From our small reform argument, the middle inequality is intuitive and we prove it formally in Appendix B. Using that $\bar{q} = w - \Delta T$ 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, given $T_1' > T_0'$ and hence $z_1 < z_0$, we have $\dot{\bar{q}} < 0$ on the interval (n_a, n_b). This implies $\bar{q}(n_a) \ge \bar{q}(n_b)$ and thus $\Delta T(n_a) \le \Delta T(n_b)$. This generates a contradiction, which proves that $T_1' \le T_0'$ for all n.

Property 1a follows easily from 1b. Since we now have $T_1' \le T_0'$ on $(\underline{n}, \bar{n})$ with equality at the end points, we obtain $\Omega_0(n) \ge \Omega_1(n)$ on $(\underline{n}, \bar{n})$ with equality at the end points. Then we have that $\dot{\Omega}_0(\bar{n}) \le \dot{\Omega}_1(\bar{n})$, which implies $1 - g_0 + \Delta T\cdot p/(1-P) \ge 1 - g_1 - \Delta T\cdot p/P$ at $\bar{n}$. Because $g_0(\bar{n}) - g_1(\bar{n}) > 0$, we have $\Delta T(\bar{n}) > 0$. Finally, $T_1' \le T_0'$ and hence $z_1 \ge z_0$ implies $\dot{\bar{q}} = \dot{V}_1 - \dot{V}_0 \ge 0$ from equation (3). Hence, $\tau(n) = (w - \bar{q}(n))/w \ge (w - \bar{q}(\bar{n}))/w = \Delta T(\bar{n})/w > 0$ for all n, where the last equality follows from $T_1' = T_0' = 0$ at $\bar{n}$. *Q.E.D.*

We may summarize our findings as follows. In the *work cost model*, second-earner participation is a signal of low work costs and hence being better off than one-earner couples. This implies $g_0(n) > g_1(n)$, which makes 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, $\tau < 0$.[19]

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, that is, $g_0(n) - g_1(n)$ 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.[20]

Instead of working with a social welfare function $\Psi(\cdot)$, if we assume exogenous Pareto weights $(\lambda_0(n), \lambda_1(n))$, then the social marginal welfare weights $g_0(n) = \lambda_0(n)/\lambda$ and $g_1 = \lambda_1(n)/\lambda$ would be fixed a priori. Optimal tax formulas (7) and (8) would carry over. Positive versus negative second-earner tax rates would depend on the sign of $\lambda_0(n) - \lambda_1(n)$, and positive versus negative jointness would depend on the profile of $\lambda_0(n) - \lambda_1(n)$ with respect to n. The asymptotic zero tax result would be true iff $\lambda_0(n) - \lambda_1(n) \to 0$ as $n \to \infty$. Hence, all results would depend on the assumptions made on the exogenous Pareto weights.

Unlike our reform argument, the negative jointness result in Proposition 3 relies on an assumption of no bunching. As we discuss in the online supplemental material, when redistributive tastes are weak, the optimal solution is close to the no-tax situation and therefore should display no bunching.[21] For strong redistributive tastes, our numerical simulations show that there is no bunching in a wide set of cases.

[19]In a more general model with both costs of work and home production, there should be a tax (subsidy) on secondary earnings if there is more (less) heterogeneity in work costs than in home production abilities (see the online supplemental material for a discussion).

[20]If Ψ is quadratic, then $g_0 - g_1$ is constant in n and the optimal tax system is separable. If Ψ' is concave, then $g_0 - g_1$ increases in n and the distortion on spouses actually increases with n. As discussed above, the case Ψ' convex (Assumption 2) fits best with the intuition that secondary earnings affect marginal social utility less when primary earnings are higher.

[21]This is also true in the one-dimensional model. We provide a simple formal proof of this in the online supplemental material.

4. NUMERICAL CALIBRATION FOR THE UNITED KINGDOM

Numerical simulations are conceptually important (i) to assess whether our no bunching assumption in Proposition 3 is reasonable, (ii) to assess how quickly the second-earner tax rate decreases to zero (scope of Proposition 2), and (iii) to analyze if and to what extent optimal schedules resemble real-world schedules.

We focus on the more realistic and traditional *work cost model* and make the following parametric assumptions: (a) $h(x) = \varepsilon/(1+\varepsilon)x^{1+1/\varepsilon}$ so that the elasticity of primary earnings ε is constant; (b) q is distributed as a power function on the interval $[0, q_{\max}]$ with distribution function $P(q) = (q/q_{\max})^{\eta}$, implying a constant second-earner participation elasticity η; (c) 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 $q_{\max}$ 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. Figure 3A depicts the ac-

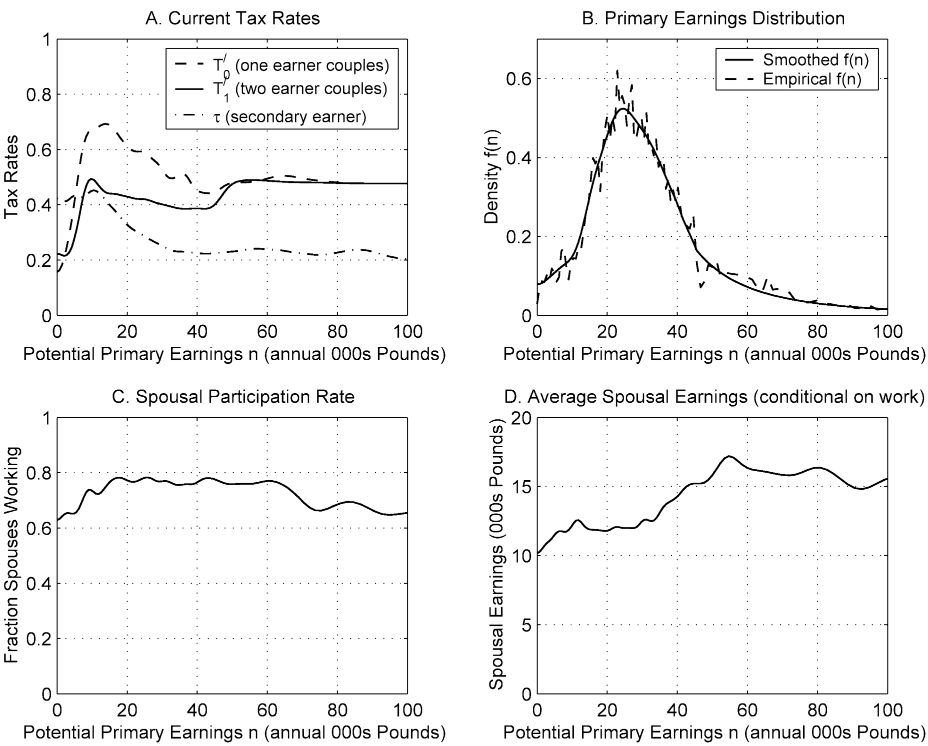

FIGURE 3.—Numerical simulations: current system. Computations are based on the British Family Resource Survey for 2004/05 and TAXBEN tax/transfer calculator.

tual tax rates T_0', T_1', and τ 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 ($n \geq$ £46,000) is approximated by a Pareto distribution with coefficient $a = 2$, a good approximation according to Brewer, Saez, and Shephard (2008). Figure 3B depicts the calibrated density distribution $f(n)$. The dashed line is the raw density distribution and the solid line is the smoothed density that we use 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. We therefore assume (except when we explore the effects of assortative matching below) that $q_{\max}$ (and hence q) is independent of n. We calibrate $q_{\max}$ 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.[22]

Based on the empirical labor supply literature for the United Kingdom (see Brewer, Saez, and Shephard (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 γ 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 so that there is no bunching. All technical details of the simulations are described in the online supplemental material.

Figure 4A plots the optimal T_0', T_1', and τ as a function of n in our benchmark case. Consistent with the theoretical results, we have $T_1' < T_0'$ and τ 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 $66.7\% = 1/(1 + a \cdot \varepsilon)$. The difference between T_1' and T_0' is about 8 percentage points on average, and τ is almost 40% at the bottom and then declines toward zero fairly quickly. This suggests that the negative jointness property as well as the zero second-earner tax at the top are quantitatively

[22]Positive correlation in abilities across spouses with income effects could also generate those empirical patterns. Analyzing a calibrated case with income effects is beyond the scope of this paper and is left for future work.

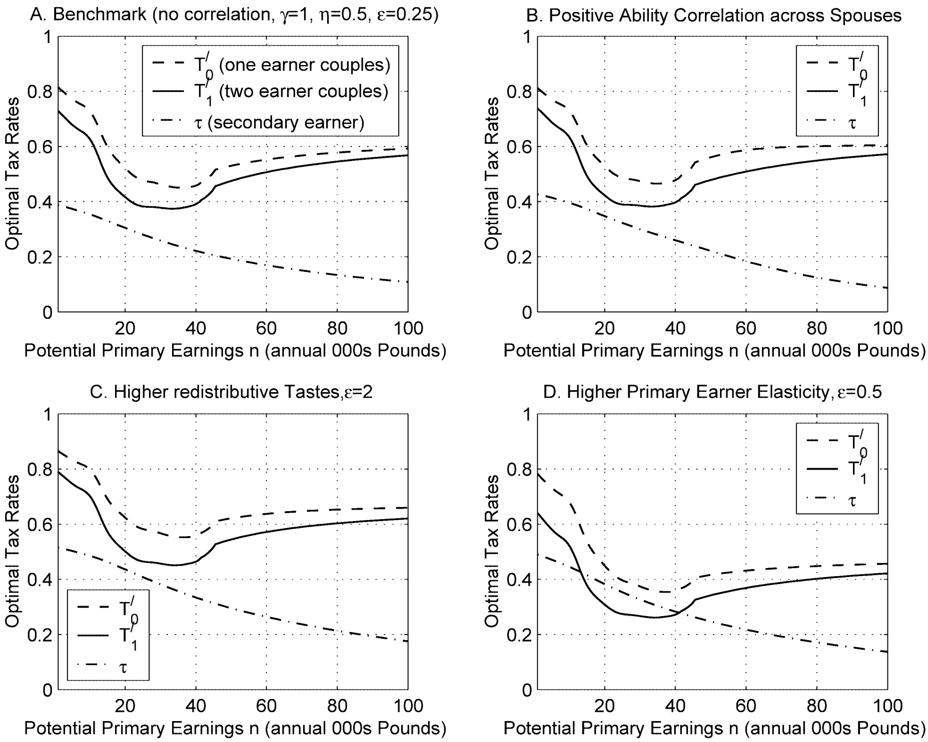

FIGURE 4.—Optimal tax simulations. Computations are based on the British Family Resource Survey for 2004/05 and TAXBEN tax/transfer calculator.

significant results and not just theoretical curiosities. Finally, notice that tax rates on primary earners are substantially higher than on secondary earners because the primary-earner elasticity is smaller than the secondary-earner elasticity.

Figure 4B introduces a positive correlation in spousal abilities by letting $q_{\max}$ depend on n, so that the fraction of working spouses (under the current tax system) increases smoothly from 55% to 80% across the distribution of n. This captures indirectly the positive correlation in earnings shown in Figure 3D. Figure 4B shows that introducing this amount of correlation has minimal effects on optimal tax rates. Compared to no correlation, the second-earner tax is slightly higher at the bottom, which reinforces the declining profile for τ. Figure 4C explores the effects of increasing redistributive tastes γ from 1 to 2. Not surprisingly, this increases tax rates across the board. Figure 4D considers a higher primary-earner elasticity ($\varepsilon = 0.5$). As expected, this reduces primary-earner tax rates (especially at the top).

Importantly, none of our simulations displays bunching, which suggests that there is no bunching in a wide set of cases and hence that Proposition 3 applies broadly.

Comparing the simulations with the empirical tax rates in Figure 3A is illuminating. The actual tax-transfer system also features negative jointness, with the second-earner tax rate falling from about 40% at the bottom to about 20% at the middle and upper parts of the primary earnings distribution. This may seem surprising at first glance given that the United Kingdom operates an individual income tax. However, income transfers in the United Kingdom (as in virtually all Organization for Economic Cooperation and Development 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 faces a high tax rate, whereas a wife married to a high-income husband is beyond benefit phase-out and hence faces a low tax rate because the income tax is individual. Thus, our theoretical and numerical findings of negative jointness may provide a justification for the current practice in many countries of combining family-based transfers with individual income taxation.[23,24]

Clearly, our calibration abstracts from several potentially important aspects such as income effects, heterogeneity in secondary 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.

APPENDIX A: PROOF OF PROPOSITION 1

The government maximizes

$$W = \int_{\underline{n}}^{\bar{n}} \left\{ \int_0^{\bar{q}} \Psi(V_1 - q^w) p(q|n)\, dq + \int_{\bar{q}}^{\infty} \Psi(V_0 + q^h) p(q|n)\, dq \right\} f(n)\, dn,$$

[23]Indeed, Immervoll, Kleven, Kreiner, and Verdelin (2008) showed that most European Union countries feature negative jointness at the bottom driven by family-based transfers.

[24]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 (Saez (2002)), not incorporated in our model. Above £6–7K, the current U.K. tax system does display a weak U-shape with the highest marginal rates at the bottom and modest increases above £40K.

where $\bar{q} = V_1 - V_0$, $q = q^w + q^h$ and either $q^w = 0$ or $q^h = 0$. The objective is maximized subject to the budget constraint

$$\int_{\underline{n}}^{\bar{n}} \left\{ \left[z_1 + w - nh\left(\frac{z_1}{n}\right) - V_1 \right] P(\bar{q}|n) + \left[z_0 - nh\left(\frac{z_0}{n}\right) - V_0 \right] (1 - P(\bar{q}|n)) \right\} f(n)\, dn \geq 0$$

and the constraints from household optimization, $\dot{V}_l = -h(z_l/n) + z_l/nh'(z_l/n)$ for $l = 0, 1$. Let λ, $\mu_0(n)$, and $\mu_1(n)$ be the associated multipliers, and let $H(z_0, z_1, V_0, V_1, \mu_0, \mu_1, \lambda, n)$ be the Hamiltonian.

We demonstrate the existence of a *measurable* solution $n \to z(n)$ in the online supplemental material. The Pontryagin maximum principle then provides *necessary* conditions that hold at the optimum:

(i) There exist absolutely continuous multipliers $(\mu_0(n), \mu_1(n))$ such that on $(\underline{n}, \bar{n})$, $\dot{\mu}_l(n) = -\partial H/\partial V_l$ almost everywhere in n with transversality conditions $\mu_l(\underline{n}) = \mu_l(\bar{n}) = 0$ for $l = 0, 1$.

(ii) We have $H(z(n), V(n), \mu(n), \lambda, n) \geq H(z, V(n), \mu(n), \lambda, n)$ for all z almost everywhere in n. The first-order conditions associated with this maximization condition are

$$\text{(A1)} \qquad \frac{\partial H}{\partial z_0} = \frac{\mu_0}{n} \cdot \frac{z_0}{n} \cdot h''\left(\frac{z_0}{n}\right) + \lambda \cdot \left(1 - h'\left(\frac{z_0}{n}\right)\right) \cdot (1 - P(\bar{q}|n)) \cdot f(n) = 0,$$

$$\text{(A2)} \qquad \frac{\partial H}{\partial z_1} = \frac{\mu_1}{n} \cdot \frac{z_1}{n} \cdot h''\left(\frac{z_1}{n}\right) + \lambda \cdot \left(1 - h'\left(\frac{z_1}{n}\right)\right) \cdot P(\bar{q}|n) \cdot f(n) = 0.$$

By Assumption 1, $\varphi(x) \equiv (1 - h'(x))/(xh''(x))$ is decreasing in x. Rewriting (A1) as $\varphi(z_0/n) = -\mu_0(n)/[\lambda(1 - P(\bar{q}|n))nf(n)]$, Assumption 1 implies that (A1) has a *unique* solution $z_0(n)$, and that $\partial H/\partial z_0 > 0$ for $z_0 < z_0(n)$ and $\partial H/\partial z_0 < 0$ for $z_0 > z_0(n)$. This ensures that $z_0(n)$ is indeed the global maximum for H as required in the Pontryagin maximum principle. Obviously, the state variable $V(n)$ is continuous in n. Thus, $\varphi(z_0(n)/n) = -\mu_0(n)/[\lambda(1 - P(V_1(n) - V_0(n)|n))nf(n)]$ implies that $z_0(n)$ is *continuous* in n.[25] Similarly, $z_1(n)$ is continuous in n.[26] By defining $T_l' \equiv 1 - h'(z_l(n)/n)$, we have that (T_0', T_1') is also *continuous* in n.[27]

[25]The assumption that $n \to f(n)$ and $x \to h''(x)$ are continuous is required here.

[26]Those continuity results also apply to the one-dimensional case and were explicitly derived by Mirrlees (1971) under a condition equivalent to our Assumption 1. The subsequent literature almost always *assumes* continuity.

[27]Notice that we adopt this definition of T_l' everywhere, including points where $z \to T(z)$ has a kink.

The conditions $\dot{\mu}_l(n) = -\partial H/\partial V_l$ for $l = 0, 1$ imply

$$-\dot{\mu}_0(n) = \int_{\bar{q}}^{\infty} \Psi'(V_0 + q^h) p(q|n) f(n)\, dq - \lambda(1 - P(\bar{q}|n)) f(n) - \lambda[T_1 - T_0] p(\bar{q}|n) f(n), \tag{A3}$$

$$-\dot{\mu}_1(n) = \int_0^{\bar{q}} \Psi'(V_1 - q^w) p(q|n) f(n)\, dq - \lambda P(\bar{q}|n) f(n) T_1 + \lambda[-T_0] p(\bar{q}|n) f(n). \tag{A4}$$

Using the definition of welfare weights, $g_0(n)$ and $g_1(n)$, we integrate (A3) and (A4) using the upper transversality conditions so as to obtain

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

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

Inserting these two equations into (A1) and (A2), noting that $T_l' = 1 - h_l'$, and using the elasticity definition $\varepsilon_l = h'(z_l/n)/[z_l/nh''(z_l/n)]$, we obtain equations (7) and (8) in Proposition 1.

The transversality conditions $\mu_0 = \mu_1 = 0$ at $\underline{n}$ and $\bar{n}$ combined with (A1) and (A2) imply that $h'(z_0/n) = h'(z_1/n) = 1$ and hence $T_1' = T_0' = 0$ at $\underline{n}$ and $\bar{n}$.

As shown in the online supplemental material, a necessary and sufficient condition for implementability is that z_0 and z_1 are weakly increasing in n (exactly as in the one-dimensional Mirrlees model). If (7) and (8) generate decreasing ranges for z_0 or z_1, there is bunching and the formulas do not apply on the bunching portions. It is straightforward to include the constraints $\dot{z}_l(n) \geq 0$ in the maximization problem (as in Mirrlees (1986)).[28] On a bunching portion, $z_l(n)$ is constant (say equal to z^*) and hence $T_l' = 1 - h'(z^*/n)$ remains continuous in n as stated in Proposition 1, but $z \to T_l'(z)$ jumps discontinuously at z^* and $z \to T_l(z)$ displays a kink at z^*. Hence the optimal solution $z \to T(z)$ is continuous and $z \to T'(z)$ is piecewise continuous.

We do not establish that the solution is unique, but uniqueness is not required for our results. Uniqueness would follow from the concavity of $(z, V) \to H(z, V, \mu(n), \lambda, n)$, but this is a very strong assumption. In the simulations, we can check numerically that, under our parametric assumptions, the stronger concavity assumptions required for uniqueness hold in the domain of interest so that we are sure the numerical solution we find is indeed the global optimum.

[28]We do not include such constraints formally so as to simplify the exposition and because our main Proposition 3 assumes no bunching and our simulations never involve bunching.

APPENDIX B: PROOF OF LEMMA IN PROPOSITION 3

LEMMA B1: *Under Assumptions 1–4, if $T_1' > T_0'$ on (n_a, n_b) with equality at the end points, then $g_0(n_a) - g_1(n_a) > g_0(n_b) - g_1(n_b)$.*

PROOF: We have $\bar{q} = V_1 - V_0$ and $g_0 - g_1 = \Psi'(V_0)/\lambda - \int_0^{\bar{q}} \Psi'(V_1 - q)p(q)\,dq/(\lambda \cdot P(\bar{q})) > 0$ (inequality follows from Ψ' decreasing). Differentiating with respect to n, we obtain

$$\dot{g}_0 - \dot{g}_1 = \dot{V}_0 \cdot \frac{\Psi''(V_0)}{\lambda} - \dot{V}_1 \cdot \frac{\int_0^{\bar{q}} \Psi''(V_1 - q)p(q)\,dq}{\lambda \cdot P(\bar{q})} + \frac{p(\bar{q})\dot{\bar{q}}}{P(\bar{q})} \cdot \left[\frac{\int_0^{\bar{q}} \Psi'(V_1 - q)p(q)\,dq}{\lambda \cdot P(\bar{q})} - \frac{\Psi'(V_0)}{\lambda}\right],$$

which can be rewritten as

$$\dot{g}_0 - \dot{g}_1 = \dot{V}_1 \cdot \left[\frac{\Psi''(V_0)}{\lambda} - \frac{\int_0^{\bar{q}} \Psi''(V_0 + \bar{q} - q)p(q)\,dq}{\lambda \cdot P(\bar{q})}\right] + \dot{\bar{q}} \cdot \left[-(g_0 - g_1) \cdot \frac{p(\bar{q})}{P(\bar{q})} - \frac{\Psi''(V_0)}{\lambda}\right]. \tag{B1}$$

The first term in (B1) is negative, because $\dot{V}_1 > 0$ and Ψ'' is increasing (by Assumption 2) so that the term inside the first square brackets is negative. On (n_a, n_b), $z_1 < z_0$ and hence $\dot{\bar{q}} < 0$. Moreover, convexity of Ψ' implies $\Psi'(V_0) - \Psi'(V_0 + \bar{q} - q) \leq -\Psi''(V_0) \cdot (\bar{q} - q)$ and hence

$$g_0 - g_1 = \frac{\int_0^{\bar{q}} [\Psi'(V_0) - \Psi'(V_0 + \bar{q} - q)]p(q)\,dq}{\lambda \cdot P(\bar{q})} \leq -\Psi''(V_0) \cdot \frac{\int_0^{\bar{q}} P(q)\,dq}{\lambda \cdot P(\bar{q})}, \tag{B2}$$

where we have used that $\int_0^{\bar{q}} (\bar{q} - q)p(q)\,dq = \int_0^{\bar{q}} P(q)\,dq$ by integration by parts and $P(0) = 0$. Combining (B2) and the second part of Assumption 4, we have $(g_0 - g_1) \cdot p(\bar{q})/P(\bar{q}) \leq -\Psi''(V_0)/\lambda$. Thus, the second term in square brackets in (B1) is nonnegative, making the entire second term in (B1) nonpositive. As a result, $\dot{g}_0(n) - \dot{g}_1(n) < 0$ on (n_a, n_b) and the lemma is proven. *Q.E.D.*

REFERENCES

ARMSTRONG, M. (1996): "Multiproduct Nonlinear Pricing," *Econometrica*, 64, 51–75. [538]

BOSKIN, M., AND E. SHESHINSKI (1983): "Optimal Tax Treatment of the Family: Married Couples," *Journal of Public Economics*, 20, 281–297. [538]

BRETT, C. (2006): "Optimal Nonlinear Taxes for Families," *International Tax and Public Finance*, 14, 225–261. [538]

BREWER, M., E. SAEZ, AND A. SHEPHARD (2008): "Means Testing and Tax Rates on Earnings," IFS Working Paper. Forthcoming in *Reforming the Tax System for the 21st Century*, Oxford University Press, 2009. [553]

CHETTY, R. (2006): "A New Method of Estimating Risk Aversion," *American Economic Review*, 96, 1821–1834. [553]

CREMER, H., J. LOZACHMEUR, AND P. PESTIEAU (2007): "Income Taxation of Couples and the Tax Unit Choice," CORE Discussion Paper No. 2007/13. [538]

CREMER, H., P. PESTIEAU, AND J. ROCHET (2001): "Direct versus Indirect Taxation: The Design of the Tax Structure Revisited," *International Economic Review*, 42, 781–799. [538]

IMMERVOLL, H., H. J. KLEVEN, C. T. KREINER, AND N. VERDELIN (2008): "An Evaluation of the Tax-Transfer Treatment of Married Couples in European Countries," Working Paper 2008-03, EPRU. [555]

KLEVEN, H. J., C. T. KREINER, AND E. SAEZ (2006): "The Optimal Income Taxation of Couples," Working Paper 12685, NBER. [538,544]

——— (2009): "Supplement to 'The Optimal Income Taxation of Couples'," *Econometrica Supplementary Material*, 77, http://www.econometricsociety.org/ecta/Supmat/7343_Proofs.pdf and http://www.econometricsociety.org/ecta/Supmat/7343_Data and programs.zip. [539]

MIRRLEES, J. A. (1971): "An Exploration in the Theory of Optimal Income Taxation," *Review of Economic Studies*, 38, 175–208. [537,543,556]

——— (1976): "Optimal Tax Theory: A Synthesis," *Journal of Public Economics*, 6, 327–358. [538]

——— (1986): "The Theory of Optimal Taxation," in *Handbook of Mathematical Economics*, Vol. 3, ed. by K. J. Arrow and M. D. Intrilligator. Amsterdam: Elsevier. [538,557]

RAMEY, V. A. (2008): "Time Spent in Home Production in the 20th Century: New Estimates From Old Data," Working Paper 13985, NBER. [540]

ROCHET, J., AND C. PHILIPPE (1998): "Ironing, Sweeping, and Multi-Dimensional Screening," *Econometrica*, 66, 783–826. [538]

SAEZ, E. (2001): "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 68, 205–229. [540,544,545,553]

——— (2002): "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses," *Quarterly Journal of Economics*, 117, 1039–1073. [555]

SCHROYEN, F. (2003): "Redistributive Taxation and the Household: The Case of Individual Filings," *Journal of Public Economics*, 87, 2527–2547. [538]

Dept. of Economics, London School of Economics, Houghton Street, London WC2A 2AE, U.K. and Economic Policy Research Unit, Dept. of Economics, University of Copenhagen, Copenhagen, Denmark and Centre for Economic Policy Research, London, U.K.; h.j.kleven@lse.ac.uk,

Dept. of Economics, University of Copenhagen, Studiestraede 6, # 1455 Copenhagen, Denmark and Economic Policy Research Unit, Dept. of Economics, University of Copenhagen, Copenhagen, Denmark and CESifo, Munich, Germany,

and

Dept. of Economics, University of California–Berkeley, 549 Evans Hall 3880, Berkeley, CA 94720, U.S.A. and NBER; saez@econ.berkeley.edu.

Manuscript received August, 2007; final revision received August, 2008.