

Kleven, Henrik Jacobsen; Kreiner, Claus Thustrup; Saez, Emmanuel

**Working Paper**

## The optimal income taxation of couples as a multi-dimensional screening problem

CESifo Working Paper, No. 2092

**Provided in Cooperation with:**

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

*Suggested Citation:* Kleven, Henrik Jacobsen; Kreiner, Claus Thustrup; Saez, Emmanuel (2007) : The optimal income taxation of couples as a multi-dimensional screening problem, CESifo Working Paper, No. 2092, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/26137>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

THE OPTIMAL INCOME TAXATION OF COUPLES  
AS A MULTI-DIMENSIONAL  
SCREENING PROBLEM

HENRIK JACOBSEN KLEVEN  
CLAUS THUSTRUP KREINER  
EMMANUEL SAEZ

CESIFO WORKING PAPER NO. 2092  
CATEGORY 1: PUBLIC FINANCE  
SEPTEMBER 2007

PRESENTED AT CESIFO AREA CONFERENCE ON PUBLIC SECTOR ECONOMICS, APRIL 2007

*An electronic version of the paper may be downloaded*

- *from the SSRN website:* [www.SSRN.com](http://www.SSRN.com)
- *from the RePEc website:* [www.RePEc.org](http://www.RePEc.org)
- *from the CESifo website:* [www.CESifo-group.org/wp](http://www.CESifo-group.org/wp)

# THE OPTIMAL INCOME TAXATION OF COUPLES AS A MULTI-DIMENSIONAL SCREENING PROBLEM

## Abstract

This paper explores the optimal income tax treatment of couples. Each couple is modelled as a single agent supplying labor along two dimensions: primary-earner and secondary-earner labor supply. We consider fully general nonlinear income tax schedules which creates a multi-dimensional screening problem. We prove that, under regularity and separability assumptions for utility functions and for a wide class of social welfare functions, optimal tax schemes display negative jointness such that the tax rate on one person decreases in the earnings of the spouse. We also show that the tax on the secondary earner tends to zero asymptotically as the earnings of the primary earner becomes large. These results are valid both in models where secondary earners make only a binary labor supply choice (work or not work), and in models where both spouses make continuous labor supply decisions. In the latter case and in contrast to the multi-dimensional screening monopoly model, the optimal tax system is regular everywhere with no bunching for a wide set of parameters.

JEL Code: H21, H31.

*Henrik Jacobsen Kleven*  
*London School of Economics*  
*Houghton Street*  
*London WC2A 2AE*  
*United Kingdom*  
*h.j.kleven@lse.ac.uk*

*Claus Thustrup Kreiner*  
*Department of Economics*  
*University of Copenhagen*  
*Studivestraede 6*  
*Denmark – 1455 Copenhagen K*  
*ctk@econ.ku.dk*

*Emmanuel Saez*  
*University of California, Berkeley*  
*Department of Economics*  
*549 Evans Hall 3880*  
*USA - Berkeley, CA 94720-3880*  
*saez@econ.berkeley.edu*

August 2007

We thank Mark Armstrong, Richard Blundell, Raj Chetty, Steven Durlauf, Nada Eissa, Kenneth Judd, Botond Koszegi, David Lannes, Jean-Marie Lozachmeur, Randall Mariger, Jean-Charles Rochet, Andrew Shephard, and numerous seminar and conference participants for very helpful comments and discussions. Financial support from NSF Grant SES-0134946 is gratefully acknowledged. The project has been supported by a grant from the Economic Policy Research Network (EPRN).

# 1 Introduction

The purpose of this paper is to explore the optimal income taxation of couples. Following the seminal contribution of Mirrlees (1971), optimal income tax theory has focused almost exclusively on redistribution across individuals. The small set of papers which do consider couples usually assume separability in the couple tax function, and hence cannot fully address the desirability of joint versus individual taxation, nor investigate the optimal form of jointness.<sup>1</sup> In this paper, 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 is a multi-dimensional screening problem where agents (couples) are characterized by a multi-dimensional parameter (ability and taste-for-work parameters of each spouse) that are unobserved by the principal (the government which maximizes social welfare).

Due to the technical difficulties involved, there are very few studies in the optimal tax literature attempting to deal with multi-dimensional screening problems. Mirrlees (1976, 1986) set out a general framework to study such problems and derived first-order conditions for an incentive scheme to be optimal, but he did not attempt to characterize the shape of optimal tax schedules and he did not consider specifically the important case of family taxation.<sup>2</sup> The nonlinear pricing literature in the field of Industrial Organization has investigated a number of aspects of multi-dimensional screening problems. Wilson (1993), Armstrong and Rochet (1999), Rochet and Stole (2003), Basov (2005) survey this literature. A central complication of multi-dimensional screening problems is that, in contrast to one-dimensional problems, first-order conditions are not always 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.

Our paper tackles these complexities in the following ways. First, we consider a framework with a binary labor supply outcome (work or not work) for the secondary earner along with

---

<sup>1</sup>Rosen (1977) and Pechman (1987) provide informal arguments about the issue. Boskin and Sheshinski (1983) considered a formal linear taxation problem of couples allowing for the possibility of selective marginal tax rates on husband and wife. The linearity assumption effectively implies separable and hence individual-based (albeit gender specific) tax treatment. Their problem is formally identical to a many-person Ramsey optimal tax problem. More recently, Schroyen (2003) and Alesina and Ichino (2007) have extended the Boskin-Sheshinski framework to the case of nonlinear taxation but keeping the assumption of separability in the tax treatment.

<sup>2</sup>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, Lozachmeur and Pestieau (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. Cremer et al. (2006) show that fully joint taxation is optimal only under very restrictive assumptions.

continuous earnings for the primary earner, allowing us to obtain an intuitive understanding of the shape of optimal schedules based on graphical exposition. Second, in a model featuring continuous earnings for both spouses, we show analytically that there is no bunching when redistributive tastes are moderate. Third, in both the continuous and the binary settings, we are able to obtain qualitative properties of optimal schedules which are relevant to tax-transfer policy and which, because of the bunching complications mentioned above, have not been obtained in nonlinear pricing theory.

As in the nonlinear pricing literature, we make a number of simplifying assumptions to be able to make progress in our understanding of optimal schedules. In particular, our framework is based on the unitary approach whereby each couple is modelled as a single agent supplying labor along two dimensions: the labor supply of a primary earner and the labor supply of a secondary earner.<sup>3</sup> We consider only couples and do not model the marriage decision.<sup>4</sup> 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 as such. Instead, jointness effects in our model arise because the social welfare function depends on family utilities rather than individual utilities, and because of a potential correlation in spouse abilities (assortative mating). As we shall see, our assumptions allow us to zoom in on the role of equity concerns for the jointness of the tax system. We obtain the following two main results.

First, assuming uncorrelated abilities across spouses, we show that optimal incentive schemes feature *negative jointness* defined as a situation where the tax rate on one person depends negatively on the earnings of the spouse. In the binary model, this implies that the participation tax rate on the secondary earner is decreasing in primary earnings. The intuition can be understood as follows. At a given level of primary earnings, the government values redistribution from two-earner couples to one-earner couples, because two-earner couples have a higher total income and tend to be better off. This requires a positive participation tax on the secondary earner. However, because the second-earner contribution to couple utility is declining in importance as the primary earner ability becomes larger, the redistributive virtue of taxing secondary earnings is also declining. As a result, the optimal second-earner tax is declining in primary earnings. This negative jointness result carries over to the continuous model, where we present a proof that, in any no-bunching solution, the couple tax liability as a function of spousal earnings displays a

---

<sup>3</sup>We adopt the unitary approach because it is the simplest tool of analysis, acknowledging that this model conflicts with empirical evidence (e.g. Lundberg et al., 1997). In Section 4, we discuss the implications of adopting a more realistic model of family labor supply.

<sup>4</sup>We discuss the implications of endogenous marriage briefly in Section 4.

negative cross-derivative everywhere. We are able to demonstrate that bunching does not occur as long as redistributive tastes are moderate.

Second, we analyze the asymptotics of optimal tax schemes as primary earnings become large, and show that, for a wide class of social welfare objectives, the tax distortion on the secondary earner tends to zero in the limit. In other words, the earnings of wives married to the highest-income husbands should be exempted from taxation. Although this statement may seem reminiscent of the classic result that optimal tax schemes display no distortion at the top, our result rests on a different logic and may be seen as an extension of the negative jointness result. A positive tax on secondary earners amounts to redistributing from two-earner couples to one-earner couples. But for couples with very large primary earnings, second-earner participation has a negligible effect on family utility, implying that redistribution from two-earner to one-earner couples has no value to the government in the limit.

The desirability of negative jointness may seem surprising at first glance. Indeed, a progressive family based income tax system, as used in for example the United States, is associated with positive jointness and progressive individually based income tax system is associated with zero jointness. However, it is important to note that most OECD countries, including those which have moved to individual income tax filing, also operate family-based means-tested welfare programs with transfers being phased out with joint family income. The combination of an individual income tax and a joint welfare system creates negative jointness. To see this, consider a secondary earner, say the wife, deciding about labor market entry. If she is married to a low-income husband, the family is in the phase-out range of transfer programs, and she will face a high effective tax rate. On the other hand, if she is married to a high-income husband, the family is beyond the phase-out range of transfer programs, and she will face a low effective tax rate because the income tax is individual. Hence, the wife's tax is declining in the husband's earnings.

The rest of the paper is organized as follows. Section 2 analyzes the binary model where secondary earners respond only along the extensive margin, while Section 3 extends our results to the continuous model where both spouses are modelled symmetrically and respond along the intensive margin. Section 4 discusses model extensions and Section 5 concludes.

## 2 A Binary Labor Supply Choice for the Secondary Earner

### 2.1 Labor Supply Model

We start by setting out a simplified labor supply model for couples allowing us to derive explicit optimal tax formulas which can be compared directly to Mirrlees (1971). 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  $z$ . As in the Mirrlees (1971) model, the primary earner is characterized by a scalar ability parameter  $n$  which is heterogeneous in the population and cannot be observed by the government. 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  $w \cdot l$ , and they face a fixed cost of work  $q$  if  $l = 1$ . In this simplified model, we assume that secondary earners are identical with respect to the wage rate  $w$ , but allow for heterogeneity with respect to the fixed cost  $q$  which is unobserved by the government. Our model implies that primary earners respond to taxes only along the intensive margin, whereas secondary earners respond only along the extensive margin. The main reason for introducing this asymmetric model is for simplicity of exposition and to allow us to understand the intuition behind our key negative jointness result.<sup>5</sup>

We assume that 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  $n$ ,  $p(q|n)$  the density function of  $q$  conditional on  $n$ , and  $f(n)$  the unconditional density of  $n$ . We normalize the size of the total population to one.

Because the government cannot observe  $n$  and  $q$ , it 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. The tax system is separable if and only if  $T'_0 = T'_1$  everywhere or, equivalently, if  $T_0$  and  $T_1$  differ by a constant. Net-of-tax income for a couple with earnings  $(z, w \cdot l)$  is given by  $c = z + w \cdot l - T_l(z)$ . The utility function for a couple whose primary earner has ability

---

<sup>5</sup>We show in Section 3 that our results extend to a symmetric model where both spouses respond along the extensive margin. It should be noted, however, that because of fixed costs of work (due to child care for example), secondary earners' labor supply responds primarily along the extensive margin (see Blundell and MaCurdy, 1999 for a recent survey).

$n$  and whose secondary earner has a fixed cost of work  $q$  takes the quasi-linear form

$$u(c, z, l) = c - n \cdot h\left(\frac{z}{n}\right) - q \cdot l. \quad (1)$$

The couple chooses  $(z, l)$  so as to maximize utility (1) subject to its budget constraint  $c = z + w \cdot l - T_l(z)$ .

A number of important assumptions are embodied in this specification of the couple's problem. First, the quasi-linear utility specification implies no income effects on the labor supply of either spouse. As is well known from the nonlinear multi-product pricing literature (e.g., Wilson, 1993), and shown more recently by Diamond (1998) in the context of optimal nonlinear income taxation, ruling out income effects simplifies greatly the theoretical analysis.<sup>6</sup> Second, we assume that the disutility of work is separable for the two spouses. This assumption would be violated if spouses like to spend leisure time together, and it may be violated if the husband's and wife's time are combined in household production processes to generate commodities within the home.<sup>7,8</sup> Third, our model is equivalent to a single decision maker optimizing along two dimensions,  $z$  and  $l$ , implying that there is no conflict in the family regarding consumption or labor supply choices.<sup>9</sup> Fourth, since we consider a model with only couples, we do not account for the potential effect of taxes on marriage decisions.<sup>10</sup> To be sure, this is a set of very strong assumptions. However, the simplicity of our model allows us to zoom in on reasons for jointness driven by social preferences for equity. In Section 4, we discuss in some detail how relaxing a number of these assumptions would affect our results.

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

$$h'\left(\frac{z}{n}\right) = 1 - T'_l(z). \quad (2)$$

In the case of no tax distortion,  $T'_l(z) = 0$ , our normalization assumption  $h'(1) = 1$  implies

---

<sup>6</sup>The empirical labor market literature tends to find small income effects (e.g., Blundell and MaCurdy, 1999), but the empirical identification of income effects is not as compelling as the identification of substitution effects. In particular, it is perceivable that primary earnings have important income effects on secondary earners' work decisions.

<sup>7</sup>Notice also that, since assumptions one and two together imply independence between spouses in the utility function, we are stacking the cards in favor of separable taxation.

<sup>8</sup>Piggott and Whalley (1996) extended the Boskin-Sheshinski linear tax model to incorporate home production, making the point that selective marginal tax rates on spouses leads to a distortion in the household production input mix.

<sup>9</sup>This stands in contrast to the recent literature on collective labor supply decisions (following the seminal contributions by Chiappori 1988, 1992) modelling couples as two individual utility maximizers interacting with one another. The single decision maker hypothesis provides a useful and simpler benchmark for our analysis.

<sup>10</sup>However, the empirical magnitude of such effects seems to be quite modest (Alm and Whittington, 1999; Eissa and Hoynes, 2000).



$z = n$ . Hence, it is natural to interpret  $n$  as potential earnings.<sup>11</sup> Positive marginal tax rates depress actual earnings  $z$  below potential earnings  $n$ . If the tax system is non-separable such that  $T'_0 \neq T'_1$ , there will be an interdependence between the earnings choice  $z$  of the primary earner and the labor force participation decision  $l$  of the spouse. We denote by  $z_l$  the optimal choice of  $z$  at a given  $l$ . If the tax system is separable,  $T'_0 = T'_1$ , we have  $z_0 = z_1$ .

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

$$\varepsilon_l \equiv \frac{1 - T'_l}{z_l} \frac{\partial z_l}{\partial (1 - T'_l)} = \frac{nh'(z_l/n)}{z_l h''(z_l/n)}. \quad (3)$$

Because we have assumed away income effects, the compensated and uncompensated elasticities of labor supply are of course identical. With separable taxation so that  $z_0 = z_1$ , we have  $\varepsilon_0 = \varepsilon_1$ .

For the secondary earner to enter the labor market and work, the utility from participation must be greater than or equal to the utility from non-participation. Let us denote by

$$V_l(n) = z_l - T_l(z_l) - nh\left(\frac{z_l}{n}\right) + w \cdot l, \quad (4)$$

the indirect utility of the couple (exclusive of the fixed work cost  $q$ ) at a given  $l$ . Differentiating with respect to  $n$  (which we denote by an upper dot from now on), and using the envelope theorem, we obtain

$$\dot{V}_l(n) = -h\left(\frac{z_l}{n}\right) + \frac{z_l}{n} \cdot h'\left(\frac{z_l}{n}\right). \quad (5)$$

The participation constraint for secondary earners is given by

$$q \leq V_1(n) - V_0(n) \equiv \bar{q}, \quad (6)$$

where  $\bar{q}$  is the net gain from working exclusive of the fixed work cost  $q$ . For families with a fixed cost below (above) the threshold-value  $\bar{q}$ , the secondary earner works (does not work).<sup>12</sup> The probability of labor force participation for the secondary earner at a given ability level  $n$  of the primary earner is given by  $P(\bar{q}|n)$ . We define the participation elasticity with respect to the net gain from working  $\bar{q}$  as

$$\eta \equiv \frac{\bar{q}}{P(\bar{q}|n)} \frac{\partial P(\bar{q}|n)}{\partial \bar{q}}. \quad (7)$$

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

---

<sup>11</sup>In general, economists consider models where  $n$  is a wage rate and where  $u = c - h(z/n)$ , which leads to a first order condition  $1 - T'(z) = n \cdot h'(z/n)$ . Our results would carry over to this standard model but  $n$  could no longer be interpreted as potential earnings and the interpretation of optimal tax formulas would be less transparent (see Saez (2001)).

<sup>12</sup>If the tax function is non-separable (so that  $z_0 \neq z_1$ ), the value of  $\bar{q}$  and hence the participation decision of the secondary earner will depend on the earnings choice of the primary earner.

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 does an associated earnings adjustment, whereas the  $w - \bar{q}$  reflects the tax burden on second-earner participation as such. It is easy to prove the following (using eqs 4-6):

**Lemma 1** *At any point  $n$ , we have:*

- $T'_0 > T'_1 \iff z_0 < z_1 \iff \dot{\tau} < 0$
- $T'_0 = T'_1 \iff z_0 = z_1 \iff \dot{\tau} = 0$
- $T'_0 < T'_1 \iff z_0 > z_1 \iff \dot{\tau} > 0$

This lemma is simply another way of stating the theorem of equality of cross-partial derivatives. We naturally say that a tax system has *positive jointness* if  $\tau$  is increasing in  $n$  and *negative jointness* if  $\tau$  is decreasing in  $n$ . If  $\tau$  is constant, the tax system is separable. These definitions can be either local (at a given  $n$ ) or global (for every  $n$ ).

It is important to note 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 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, it is easy to show, exactly as in the Mirrlees (1971) model, that a given path for  $(z_0(n), z_1(n))$  can be implemented via a truthful mechanism or equivalently with a non-linear tax system if and only if  $z_0(n)$  and  $z_1(n)$  are non-negative and non-decreasing in  $n$  (see Kleven et al., 2006, for details).

## 2.2 Government Objective

As usual in optimal income tax models, the government maximizes a social welfare function defined as the sum of concave and increasing transformations  $\Psi(\cdot)$  of the couples' utilities subject to a government budget constraint and the constraints imposed by household utility maximization. Formally, the government maximizes

$$W = \int_{n=\underline{n}}^{\bar{n}} \int_{q=0}^{\infty} \Psi(V_l(n) - q \cdot l) p(q|n) f(n) dq dn, \quad (8)$$

subject to the budget constraint

$$\int_{n=\underline{n}}^{\bar{n}} \int_{q=0}^{\infty} T_l(z_l) p(q|n) f(n) dq dn \geq 0, \quad (9)$$

and subject to  $\dot{V}_0(n)$  and  $\dot{V}_1(n)$  in eq. (5). We denote by  $\lambda$  the multiplier of the budget constraint (9). Nothing would change in the analysis if we assumed a positive exogenous revenue requirement for the government.

We may capture the redistributive tastes of the government by social marginal welfare weights across 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  $l$ . Formally, we have  $g_0(n) = \Psi'(V_0(n))/\lambda$  and  $g_1(n) = \int_0^{\bar{q}} \Psi'(V_1(n) - q)p(q|n)dq/(P(\bar{q}|n) \cdot \lambda)$ . The profile for these  $g$ -weights in the population is crucial for the properties of optimal tax schedules.

Figure 1 illustrates curves for  $g_0(n)$  and  $g_1(n)$  satisfying four ‘natural’ properties. First, because of our assumption of no income effects, the average of  $g_0$  and  $g_1$  across the full population is one.<sup>13</sup> Second, the concavity of  $\Psi$  tend to make  $g_0$  and  $g_1$  decreasing in  $n$ .<sup>14</sup> Third, we have  $g_0(n) - g_1(n) > 0$  because, at a given  $n$ , one-earner couples are worse off than two-earner couples and  $\Psi$  is concave. To see why one-earner couples are worse off (at a given  $n$ ), notice that the reason for second-earner non-participation is a high work cost  $q$ . More precisely, the utility of any one-earner couple is  $V_0(n)$ , and this must be lower than the utility of a two-earner couple,  $V_1(n) - q$ , given that this couple has decided to let the spouse work (from eq. 6).<sup>15</sup> Fourth, the difference in weights  $g_0 - g_1$  is naturally decreasing in  $n$  as the contribution of secondary earnings becomes relatively smaller as  $n$  becomes larger. As we shall see below, this property is closely related to  $\Psi'$  being convex. In the limit when  $n$  goes to infinity, we would expect  $g_0 - g_1$  to converge to zero.

## 2.3 The Optimal Income Tax Schedule and its Properties

### 2.3.1 Explicit Tax Formulas and their relation to Mirrlees (1971)

The simple model described above makes it possible to derive explicit optimal tax formulas as in the individualistic Mirrlees (1971) framework. In appendix A.1, we show that the optimal tax scheme satisfies the following.

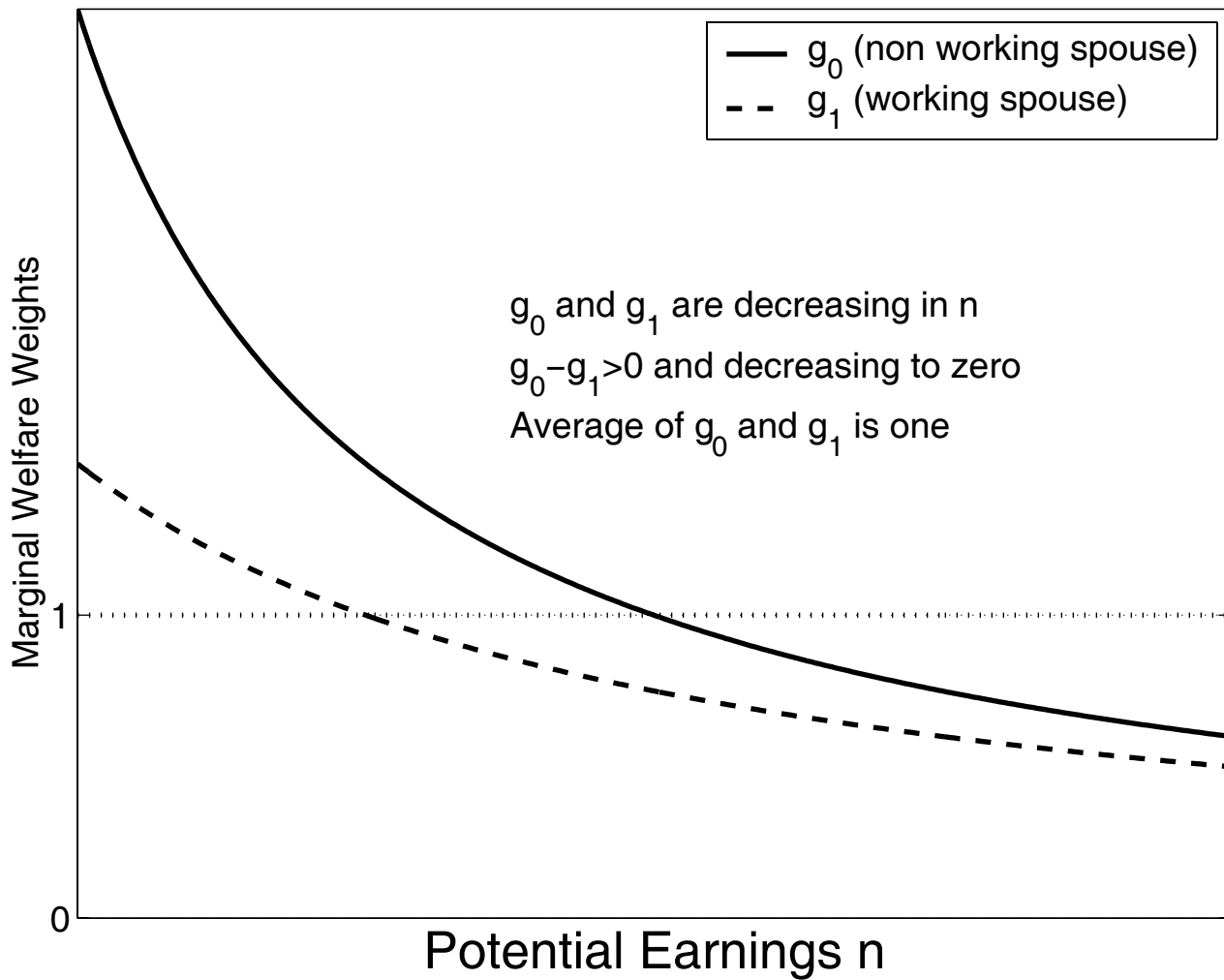
---

<sup>13</sup>Because of no income effects, it costs exactly \$1 to redistribute \$1 uniformly across all couples. The social marginal value (expressed in terms of government funds) of redistributing \$1 to every couple is exactly the sum of the  $g$ 's across the full population.

<sup>14</sup>As  $V_0(n)$  is increasing in  $n$ ,  $g_0(n)$  is obviously decreasing in  $n$ . As we will see,  $g_1(n)$  will in general be decreasing in  $n$  as well.

<sup>15</sup>Conceivably, we may alternatively have defined  $q$  as the value of home production—say, the spouse’s ability in cooking or child care—by adding  $q$  in the utility function such that  $u = z + w \cdot l - T - n \cdot h(z/n) + q \cdot (1 - l)$ . The work decision would be identical in this model, but one-earner couples would be better off than two-earner couples and hence  $g_0 - g_1 < 0$ . It is easy to show that our negative jointness result would become a positive jointness result in this context. However, we believe that inequality in work opportunities is much more important than inequality in home production abilities.

Figure 1: Marginal Welfare Weights:  $g_0(n)$  and  $g_1(n)$



**Proposition 1** *The first-order conditions for the optimal marginal tax rates  $T'_0$  and  $T'_1$  at ability level  $n$  can be written as*

$$\frac{T'_0}{1 - T'_0} = \frac{1}{\varepsilon_0} \cdot \frac{1}{nf(n)(1 - P(\bar{q}|n))} \cdot \int_n^{\bar{n}} \{(1 - g_0)(1 - P(\bar{q}|n')) + [T_1 - T_0]p(\bar{q}|n')\} f(n')dn' \quad (10)$$

$$\frac{T'_1}{1 - T'_1} = \frac{1}{\varepsilon_1} \cdot \frac{1}{nf(n)P(\bar{q}|n)} \cdot \int_n^{\bar{n}} \{(1 - g_1)P(\bar{q}|n') - [T_1 - T_0]p(\bar{q}|n')\} f(n')dn', \quad (11)$$

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.

Kleven et al. (2006) presents a detailed discussion of the relation between these tax formulas and classic Mirrlees-type formulas. At the present moment, let us remark on just two aspects.

First, the average marginal tax rate faced by primary earners in one- and two-earner couples is identical to the optimal marginal tax rate in the Mirrlees framework. By taking the sum of (10) and (11), we obtain a weighted average of  $T'_0$  and  $T'_1$  which is exactly identical to the Mirrlees formula in the case with no income effects (as in Diamond, 1998). This implies that redistribution across couples with different primary earners follows the standard logic in the literature. The role of introducing a secondary earner in the household is to create 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 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 top earner in empirical earnings distributions (Saez, 2001).

### 2.3.2 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, we assume that  $f(n)$  has a Pareto tail with parameter  $a > 1$ .

As  $n$  tends to infinity, the additional income generated by the secondary earner becomes infinitesimal relative to primary-earner income in the limit. For any reasonable welfare function,

we would then have that  $g_0(n)$  and  $g_1(n)$  converge to the same value  $g^\infty$ .<sup>16</sup> 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  $\Delta T^\infty, T'^\infty_0, T'^\infty_1, \bar{q}^\infty, \tau^\infty$  as  $n \rightarrow \infty$ .*

*Then we have*

- $\Delta T^\infty = \tau^\infty = 0$ , *i.e., the second-earner tax rate goes to zero as  $n$  tends to infinity.*
- $T'^\infty_0 = T'^\infty_1 = (1 - g^\infty) / (1 - g^\infty + a \cdot \varepsilon^\infty) > 0$ , *exactly as in the Mirrlees model.*

**Proof:**

Because  $T_1 - T_0$  converges as  $n$  goes to infinity, it must be the case that  $T'^\infty_0 = T'^\infty_1 = T'^\infty$ . 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 (10) and (11) as  $n \rightarrow \infty$ , we obtain

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

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

For this to be satisfied, we must have  $\Delta T^\infty = 0$ , and the formula for  $T'^\infty$  then follows.  $\square$

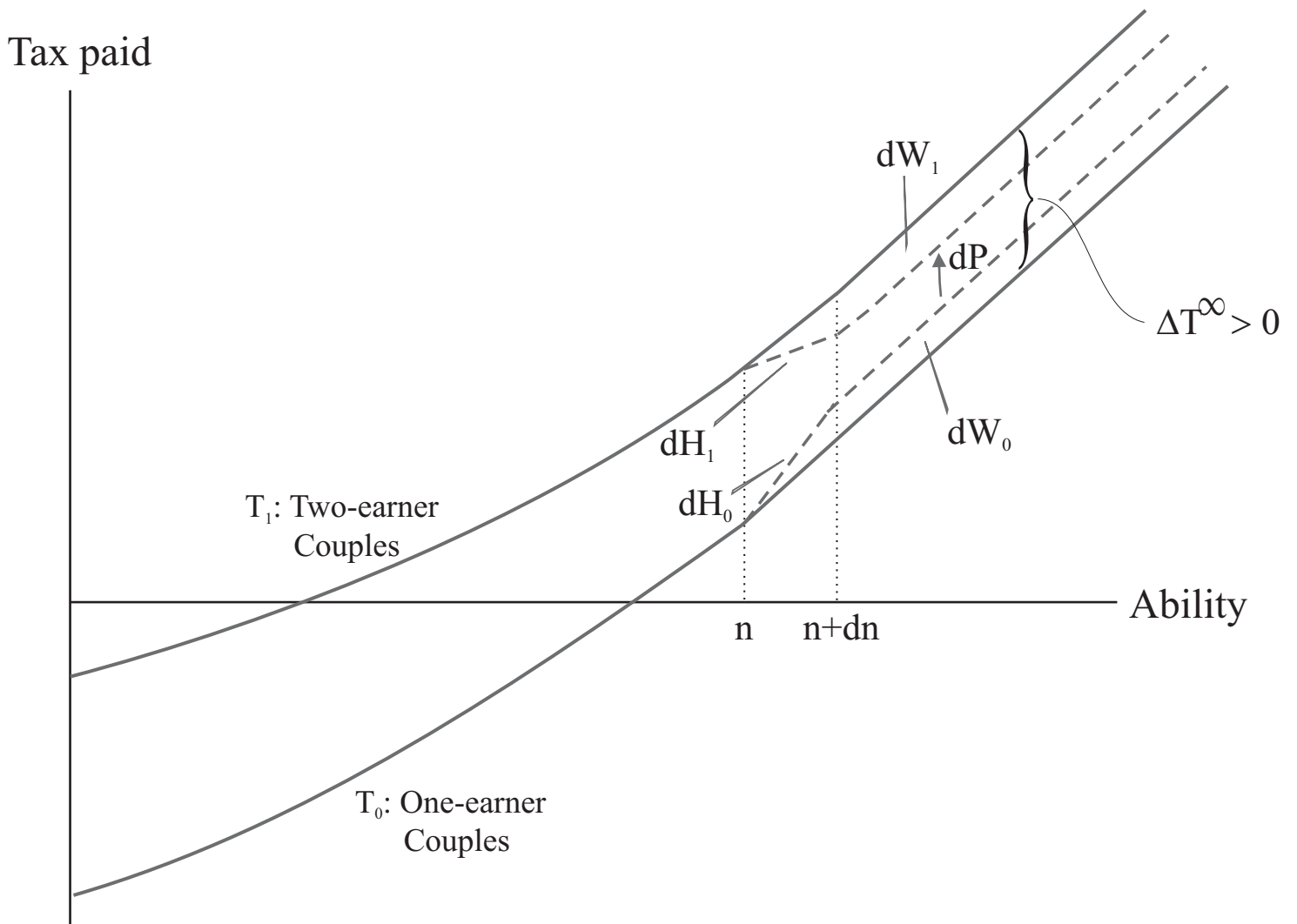
The result in Proposition 2 is quite striking. The earnings of spouses to the highest-income earners should be exempted from taxation, even in the case where the government tries to extract as much tax revenue as possible from high-income couples ( $g^\infty = 0$ ). Although the result may seem similar to the classic result of no distortion at the top, the logic behind our result is completely different. Indeed, in the present case with an infinite tail for  $n$ , the traditional result does not apply and we have  $T'^\infty_0 = T'^\infty_1 > 0$ .<sup>17</sup>

To understand the economic intuition for this result, consider a situation where  $T_1 - T_0$  does not converge to zero. This is illustrated in Figure 2 which shows the two schedules  $T_0, T_1$  as a function of ability  $n$ , assuming that  $T_1 - T_0$  converges to  $\Delta T^\infty > 0$ . We want to establish a contradiction by arguing that, in this situation, it is always possible to increase welfare by reducing  $T_1 - T_0$  a little bit at the top. Consider specifically a reform which increases the tax

<sup>16</sup>In the case where  $g^\infty = 0$ , the optimal tax system extracts as much tax revenue as possible from the very rich ('soaking the rich').

<sup>17</sup>Conversely, in the case of a bounded ability distribution, the top marginal tax rate on the primary earner would be zero, but then the tax on the secondary earner would be positive.

Figure 2. Zero second-earner tax at the top



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.<sup>18</sup> 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 redistribute 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,  $T'_0 = T'_1$ , which implies  $z_0 = z_1$  and hence identical primary-earner elasticities  $\varepsilon_0 = \varepsilon_1$ . As a consequence, the negative fiscal effect  $dH_0$  offsets the positive fiscal effect  $dH_1$ . Third, there is a participation effect as some 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. Since the reform increases welfare, the original schedule with  $\Delta T^\infty > 0$  cannot be optimal.<sup>19</sup>

### 2.3.3 Desirability of Negative Jointness

A key point of this paper is to demonstrate that optimal schedules are characterized by negative jointness. To show this, we introduce two additional assumptions.

**Assumption 1** *The function  $V \rightarrow \Psi'(V)$  is convex.*

This is a very natural assumption on social preferences which is satisfied for all standard social welfare functions such as the CRRA form,  $\Psi(V) = V^{1-\gamma}/(1-\gamma)$  with  $\gamma > 0$ , and the CARA form. As we show formally below, the assumption is directly related to the property that  $g_0 - g_1$  is decreasing in  $n$  which, as we discussed above, is intuitively appealing. Notice also that, 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 (e.g. Deaton, 1992).

---

<sup>18</sup>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.

<sup>19</sup>Of course, the opposite situation with  $\Delta T^\infty < 0$  cannot be optimal either, because then the opposite reform would improve welfare.



**Assumption 2**  $q$  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. Obviously, we do not expect this assumption to hold in practice and in Section 2.4 we examine numerically how assortative matching affects our results.

The most transparent way to demonstrate the desirability of negative jointness is by a tax reform argument starting from the optimal *separable* tax system. Under separable tax treatment, the primary-earner marginal tax rate is identical in one- and two-earner couples,  $T'_0 = T'_1 \equiv T'$ , and it is straightforward to show that  $T'$  is given by the standard Mirrlees formula with no income effects (as in Diamond, 1998). Moreover, separable tax treatment implies that  $T_1 - T_0$  is constant in  $n$ , and its value can be obtained by shifting the  $T_1$ - and  $T_0$ -schedules uniformly by  $dT$ . For the  $T_1$ -schedule, this generates the formula

$$(T_1 - T_0) \cdot \frac{p(\bar{q})}{P(\bar{q})} = 1 - \int_{\underline{n}}^{\bar{n}} g_1(n) f(n) dn,$$

and for the  $T_0$ -schedule, we obtain

$$(T_1 - T_0) \cdot \frac{p(\bar{q})}{1 - P(\bar{q})} = \int_{\underline{n}}^{\bar{n}} g_0(n) f(n) dn - 1.$$

Summing these two equations implies

$$(T_1 - T_0) \cdot \frac{p(\bar{q})}{P(\bar{q}) \cdot (1 - P(\bar{q}))} = \int_{\underline{n}}^{\bar{n}} [g_0(n) - g_1(n)] f(n) dn > 0, \quad (12)$$

where the positive sign follows from the property  $g_0(n) - g_1(n) > 0 \forall n$ . As pointed out above, this property derives from the fact that one-earner couples are worse off than two-earner couples at any  $n$  along with  $\Psi$  being concave. Hence, the optimal separable tax schedule involves  $T_1 - T_0 > 0$ . Any separable tax system also satisfies the following important property.

**Lemma 2** *Under Assumptions 1 and 2 and with a separable tax system,  $g_0(n) - g_1(n)$  is (weakly) decreasing in  $n$ .*

**Proof:**

Because the tax system is separable, we have that  $\bar{q} = w - (T_1 - T_0)$  is constant in  $n$ . Moreover, by Assumption 2, we also have that  $p(q|n) = p(q)$  and  $P(\bar{q}|n) = P(\bar{q})$  are constant in  $n$ . Then, by using the definitions of  $g_0(n)$  and  $g_1(n)$ , we obtain

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

where we have used  $V_1 = V_0 + \bar{q}$  from eq. (6). By Assumption 1,  $\Psi''$  is increasing and hence the expression in square brackets is negative. Moreover,  $V_0$  is increasing in  $n$ , which demonstrates the Lemma.  $\square$

Starting from the optimal separable tax schedule, consider a tax reform introducing a little bit of negative jointness as shown in Figure 3. 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. As shown the figure, 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. This implies that, above  $n$ , the tax change for one-earner couples is  $dT_0^a = dT/[(1 - F(n))(1 - P(\bar{q}))]$ , while 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  $dT_1^b = dT/[F(n)P(\bar{q})]$ . These changes imply that the direct welfare effect of redistributing income across the different types of couples can be written as

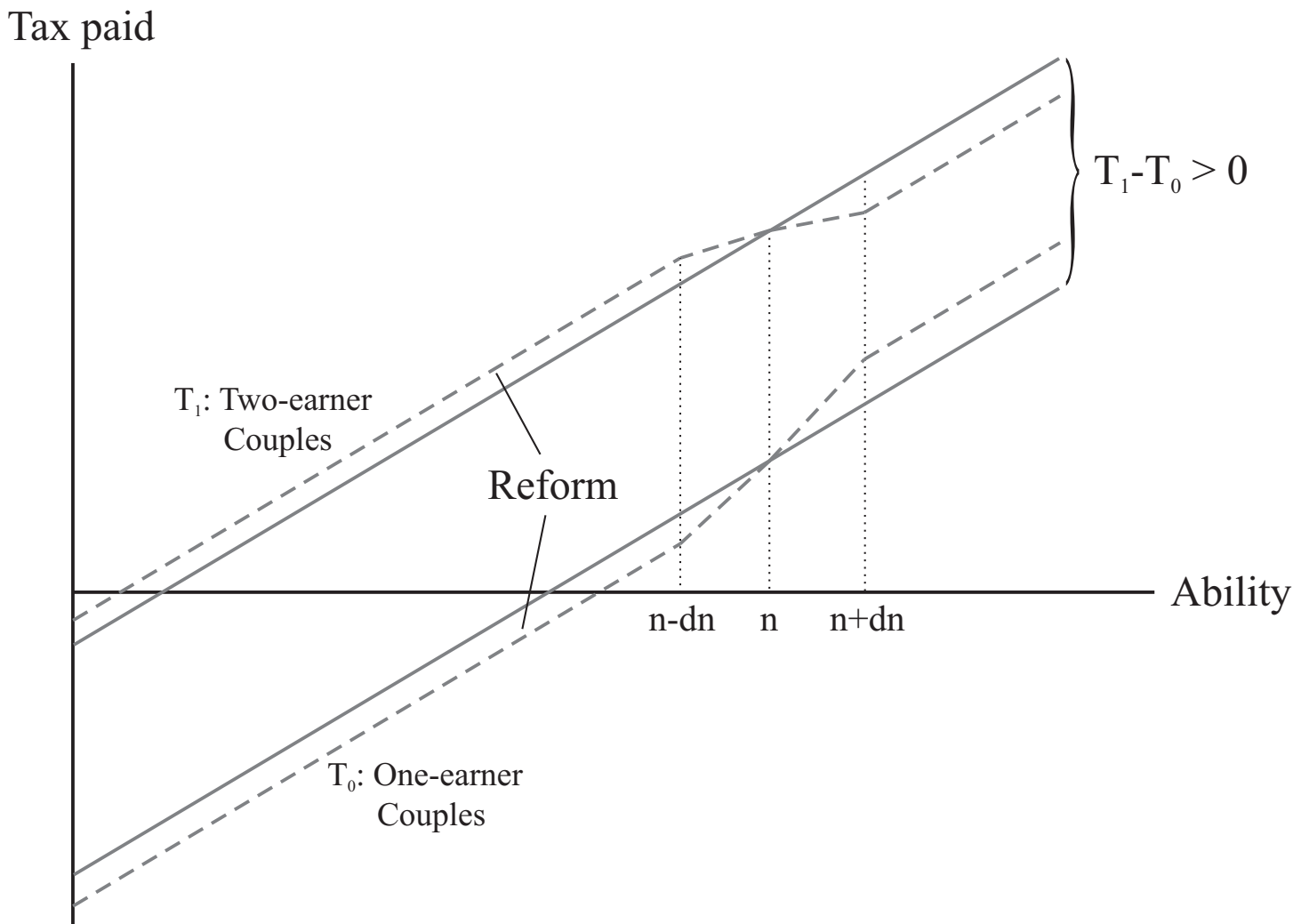
$$dW = \frac{dT}{F(n)} \cdot \int_n^n [g_0(n') - g_1(n')]f(n')dn' - \frac{dT}{1 - F(n)} \cdot \int_n^{\bar{n}} [g_0(n') - g_1(n')]f(n')dn'. \quad (13)$$

Lemma 2 implies that  $dW > 0$ . That is, the gain created at the bottom by redistributing from two-earner to one-earner couples (the first term in 13) dominates the loss created at the top from the opposite redistribution (the second term in 13), because  $g_0 - g_1$  is higher at the bottom as second-earner participation is relatively more important in low-income families.

Besides the direct welfare effect, the tax reform gives rise to behavioral responses along the intensive and extensive margins. First, since the reform increases (reduces) the marginal tax rate on the primary earner in one-earner (two-earner) couples around  $n$ , there are earnings responses going in opposite directions in the two types of couples. Since we start from a situation with separable taxation,  $T_0' = T_1'$ , we have identical primary-earner elasticities  $\varepsilon_0 = \varepsilon_1$ . This implies that the fiscal effect of these intensive responses offset one another exactly.

Second, the tax reform induces some secondary earners to change labor force participation status. Above  $n$ , non-working spouses will be induced to join the labor force, whereas below  $n$ , working spouses have an incentive to drop out. Because spouse characteristics  $q$  and  $n$  are independent, and because we start from a separable tax system, the participation elasticity is constant in  $n$  (from eq. 7) and the positive and negative participation effects then cancel out.

Figure 3. Desirability of Negative Jointness



To see this more formally, note that the number of switchers above  $n$  is  $(1 - F(n))p(\bar{q})d\bar{q}^a$  where  $d\bar{q}^a = dT_0^a - dT_1^a = dT/[(1 - F(n))P(\bar{q})(1 - P(\bar{q}))]$ . Symmetrically, the number of switchers below  $n$  is  $F(n)p(\bar{q})d\bar{q}^b$  where  $d\bar{q}^b = dT_0^b - dT_1^b = -dT/[F(n)P(\bar{q})(1 - P(\bar{q}))]$ . Since the positive and negative participation effects have the same magnitude, and because  $T_1 - T_0$  is initially constant in  $n$ , the net fiscal effect of participation responses is zero.

We can then conclude that  $dW > 0$  is the net total welfare effect of the reform, allowing us to state the following proposition.

**Proposition 3** *Under Assumptions 1 and 2, and starting from the optimal separable schedule, introducing some negative jointness always increases welfare.*

This tax-reform result represents a first step in establishing that negative jointness is a feature of fully optimized schedules. In Kleven et al. (2006), we show formally that, under additional regularity assumptions on the functions  $h(\cdot)$  and  $P(\cdot)$ , the optimum schedule does indeed display negative jointness everywhere, i.e.,  $T'_0 > T'_1 \forall n$  and  $\tau$  is decreasing in  $n$ . We omit the formal proof here because it does not provide any additional economic insight, and because the proof in the double continuous model in Section 3 is mathematically more elegant.

Although our results may seem surprising at first glance, they obey a simple redistributive logic. The government wants to support one-earner families because they are less well-off than two-earner families. If the tax schedule for two-earner couples is seen as the base schedule, the schedule for one-earner couples is obtained from this base by giving a tax break for having a dependent spouse. Because the importance of second-earner participation declines with primary earnings, the dependent spouse tax allowance should be declining in primary earnings. In the limit where primary earnings go to infinity, the allowance converges to zero.

## 2.4 Numerical Simulations

We make the following simple parametric assumptions. First, we assume that  $h(x) = x^{1+k}/(1+k)$ , so that we have a constant primary earner elasticity  $\varepsilon = 1/k$ . Second, we assume that  $F(n)$  is distributed over  $[\underline{n}, \bar{n}]$  as a truncated Pareto distribution with parameter  $a > 1$ . Third, we assume that  $q$  is distributed as a power function on the interval  $[0, q_{max}]$  with distribution function  $P(q) = (q/q_{max})^\eta$  and density function  $p(q) = \eta \cdot (q^{\eta-1})/q_{max}^\eta$  so that the elasticity of participation with respect to net gain of working is constant and equal to  $\eta$ . Fourth, we assume that the social welfare function  $\Psi$  is CRRA with coefficient of risk aversion  $\gamma > 0$ , i.e.,  $\Psi(V) = V^{1-\gamma}/(1-\gamma)$ . We set  $\underline{n} = 1$ ,  $\bar{n} = 4$ ,  $w = 1$ ,  $q_{max} = 2 \cdot w$ , and  $a = 2$ . For our benchmark

case, we assume  $\gamma = 2$ ,  $\varepsilon = 0.5$ ,  $\eta = 0.5$ . In all cases, we check that the implementation conditions ( $z_l(n)$  increasing in  $n$ ) are satisfied. Details about our simulations are presented in Appendix A.7.

Figure 4 plots the optimal  $T'_0$ ,  $T'_1$ , and  $\tau$  as a function of  $n$ . Consistent with our theoretical results, we have  $T'_0 = T'_1 = 0$  at the end points and  $T'_1 < T'_0$  everywhere else. The difference between  $T'_1$  and  $T'_0$  is about 7 percentage points which makes  $T'_0$  about 30% percent larger than  $T'_1$ . The graph also shows that the tax on secondary earners  $\tau$  is decreasing in  $n$  from about 37 percent at  $\underline{n}$  to 22 percent at  $\bar{n}$ . This suggests that the negative jointness property is not a negligible phenomenon and that it generates a significant difference in marginal tax rates between one- and two-earner couples.

Figure 5 examines the sensitivity of optimal tax rates with respect to alternative parameter values. It shows optimal tax rates  $T'_0$ ,  $T'_1$ , and  $\tau$  in four situations. In Panel A, we increase the participation elasticity  $\eta$  to one. We find that this decreases the level of the tax on secondary earners by about 10 percentage points but the decreasing slope for  $\tau$  (or, equivalently, the gap between  $T'_0$  and  $T'_1$ ) remains significant and fairly close to the benchmark case. In Panel B, we increase the intensive elasticity  $\varepsilon$  to one. We find that this decreases the level of marginal tax rates on primary earners by about 10 percentage points but again the decreasing slope for  $\tau$  (and the gap between  $T'_0$  and  $T'_1$ ) remains significant as a proportion of tax rate levels. In panel C, we increase both  $\eta$  and  $\varepsilon$  to one. This reduces  $T'_0$ ,  $T'_1$ , and  $\tau$  but the negative jointness pattern remains. Taken together, results from Panels A, B, C show that levels of tax rates obey the traditional Ramsey principle: when the elasticity increases, the corresponding tax rate decreases. In Panel D, we increase redistributive tastes of the government to  $\gamma = 4$ . We find that all tax rates increase significantly but, again, the negative jointness pattern remains about the same in proportion to tax rates.

Figure 6 explores two other departures from our benchmark case. Panel A focuses on the Rawlsian case ( $\gamma = \infty$ ). In this case, we have that  $g_1(n) = 0$  and that  $g_0(n)$  is a Dirac distribution with all mass concentrated at  $\underline{n}$ . The optimal tax formulas from Proposition 1 continue to apply but the transversality condition  $T'_0 = 0$  is no longer true at the bottom. Indeed, the simulation shows that  $T'_0(\underline{n}) = 59\%$  in this case. Interestingly, the negative jointness result carries over to this case. The Rawlsian case is theoretically interesting because it is formally equivalent to a multi-product nonlinear pricing problem as analyzed in the Industrial Organization literature. This shows that the negative jointness result would carry over in that case as well.

Figure 6, Panel B, explores the case with a long tail. In the simulation, we set  $\bar{n} = 200$

Figure 4: Benchmark Simulation:  $\gamma=2$ ,  $\eta=0.5$ ,  $\varepsilon=0.5$

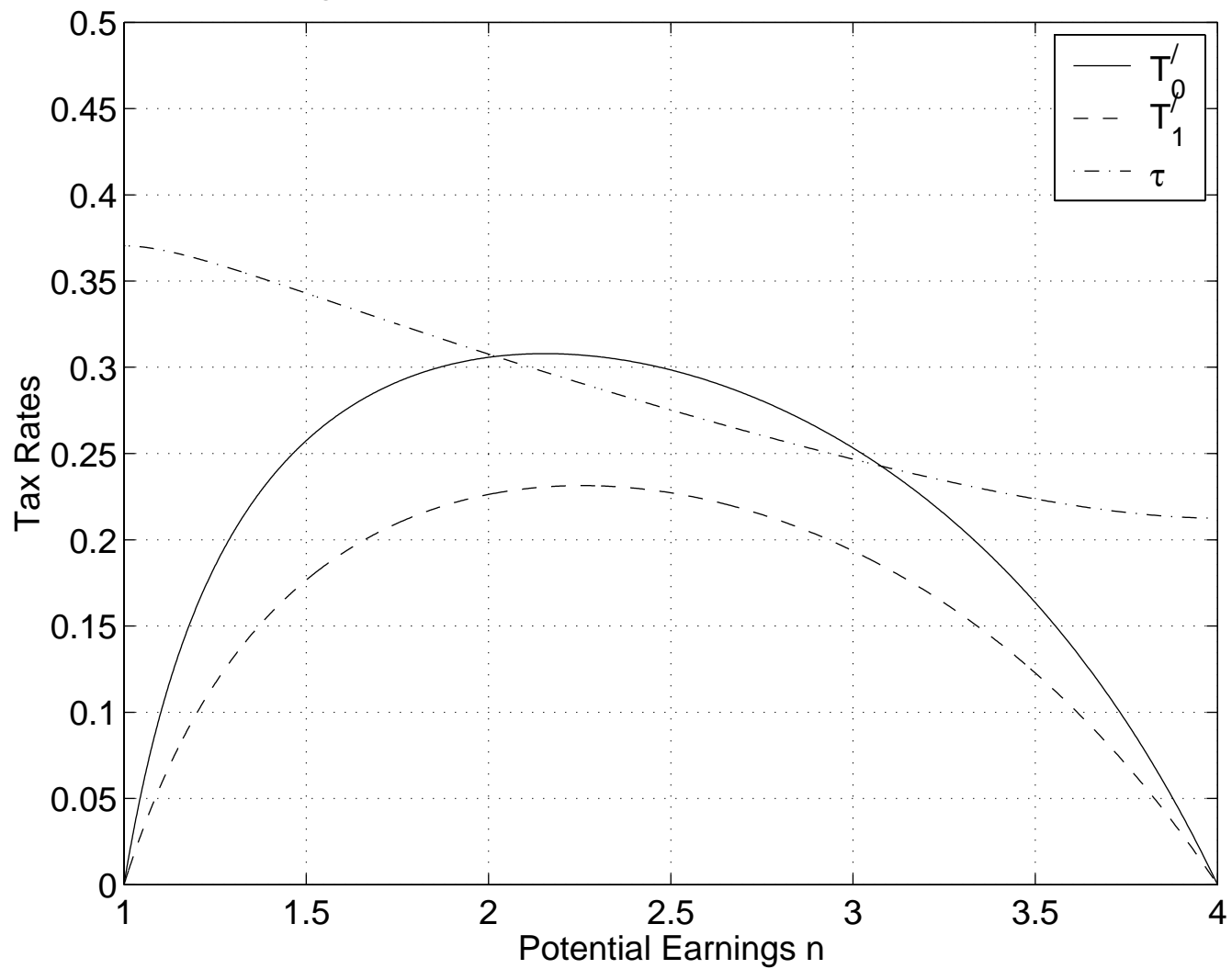


Figure 5: Sensitivity Analysis around Benchmark ( $\gamma=2, \eta=0.5, \varepsilon=0.5$ )

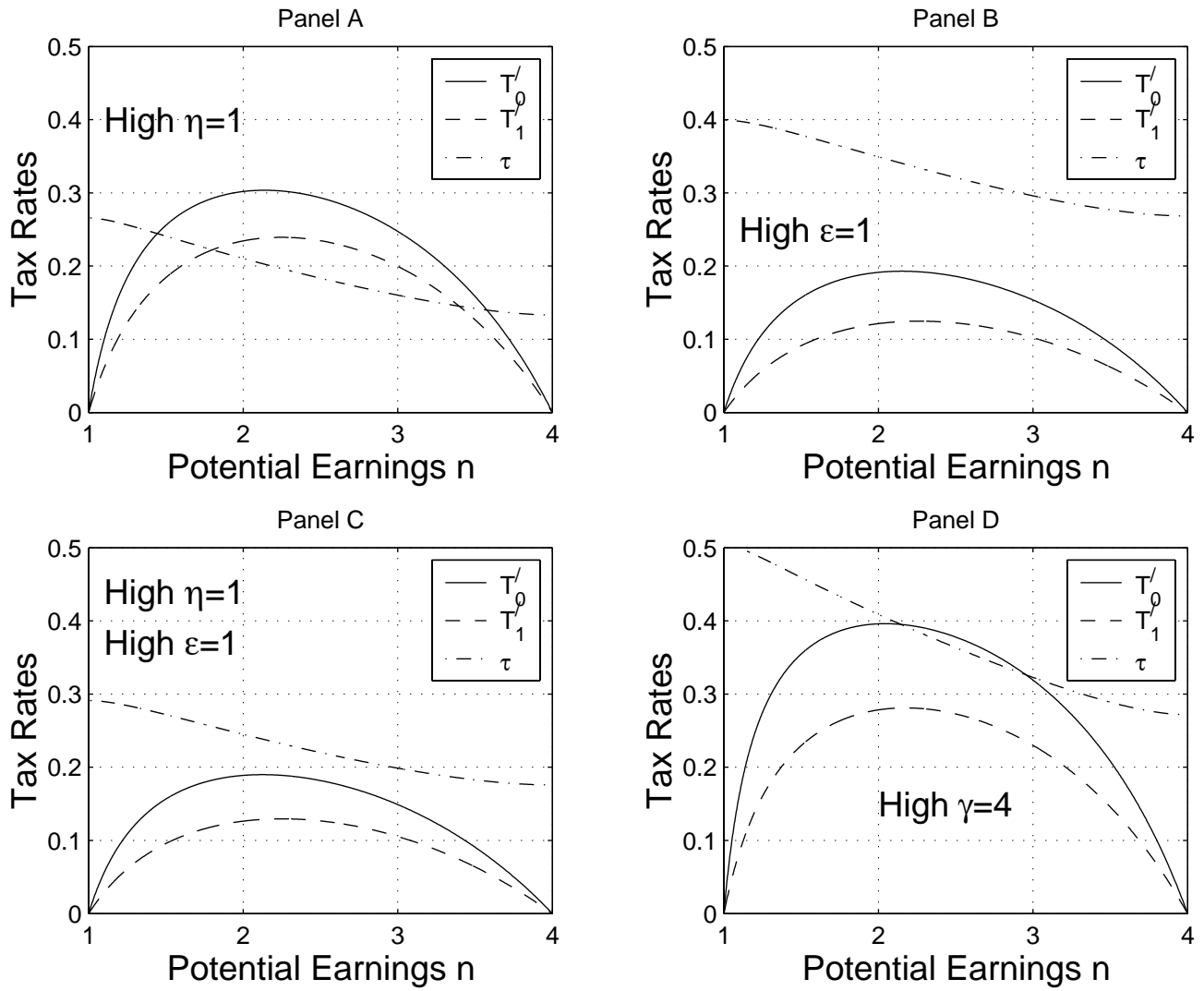
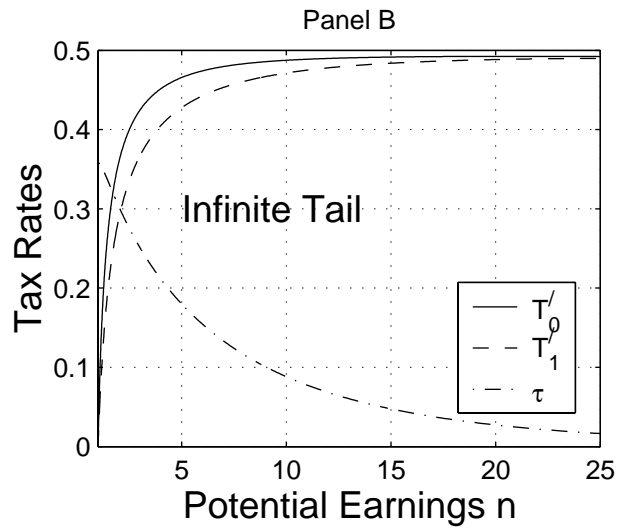
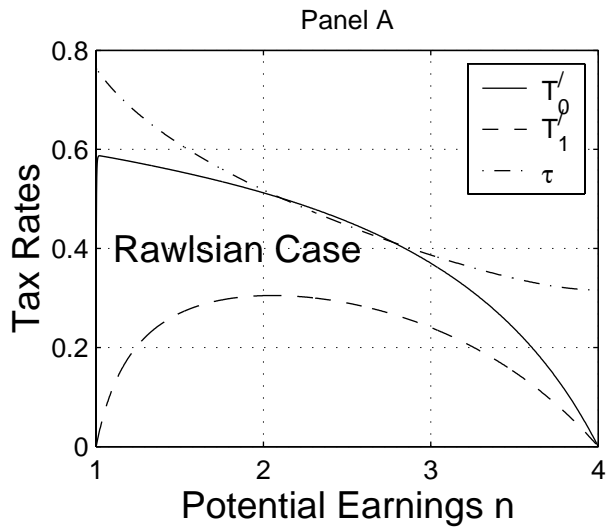


Figure 6: Two Cases of Interest





(which is a close approximation to an infinite tail). The figure shows that in this case,  $T'_0$  and  $T'_1$  converge to the theoretical asymptotic value of  $1/(1 + a \cdot \varepsilon) = 1/2$ . We also see that, as expected,  $\tau$  converges to zero.

Figure 7 examines the implications of introducing positive or negative correlation in spouse characteristics,  $n$  and  $q$ . If we think of a low  $q$  as reflecting a high ability of the secondary earner, a negative correlation in  $n$  and  $q$  would correspond to a positive correlation in ability, and vice versa. We introduce correlation by making  $q_{max}$  a function of  $n$ ; it will be a decreasing function in the case of positive ability correlation and an increasing function in the case of negative ability correlation. The correlations are calibrated so that the average participation rates of spouses remains approximately the same. Panel C displays the participation rates of spouses by potential earnings in the cases of independent abilities (benchmark), positive correlation in ability, and negative correlation in ability. Panel C shows that we have introduced significant correlation with participation rates doubling from  $\underline{n}$  to  $\bar{n}$  in the positive correlation case and decreasing by 50% from  $\underline{n}$  to  $\bar{n}$  in the negative correlation case. Panels A and B display the optimal tax rates in the positive and negative correlation case, respectively. The levels of tax rates are higher in the positive correlation case because inequality is more important in that case and hence redistribution more desirable. However, the negative jointness pattern is very similar to the cases with no correlation. This suggests that the empirical observation of positive correlation in ability across spouses (positive assortative mating) would not overturn the negative jointness result we have obtained.

### 3 A Continuous Labor Supply Choice for the Secondary Earner

#### 3.1 Model and Optimal Tax Formulas

In this Section, we model primary and secondary earners symmetrically. There is a distribution of earnings abilities  $(n_p, n_s)$  over the population of couples with density  $f(n_p, n_s)$  on the domain  $D = (\underline{n}_p, \bar{n}_p) \times (\underline{n}_s, \bar{n}_s)$ .<sup>20</sup> We define  $f_p(n_p) = \int_{\underline{n}_s}^{\bar{n}_s} f(n_p, n_s) dn_s$  as the unconditional density distribution of  $n_p$ , and  $f_s(n_s)$  symmetrically as the unconditional density distribution of  $n_s$ . We then define  $f_{p|s}(n_p|n_s) = f(n_p, n_s)/f_s(n_s)$  as the density distribution of  $n_p$  conditional on  $n_s$ , and  $f_{s|p}(n_s|n_p) = f(n_p, n_s)/f_p(n_p)$  as the density distribution of  $n_s$  conditional on  $n_p$ .

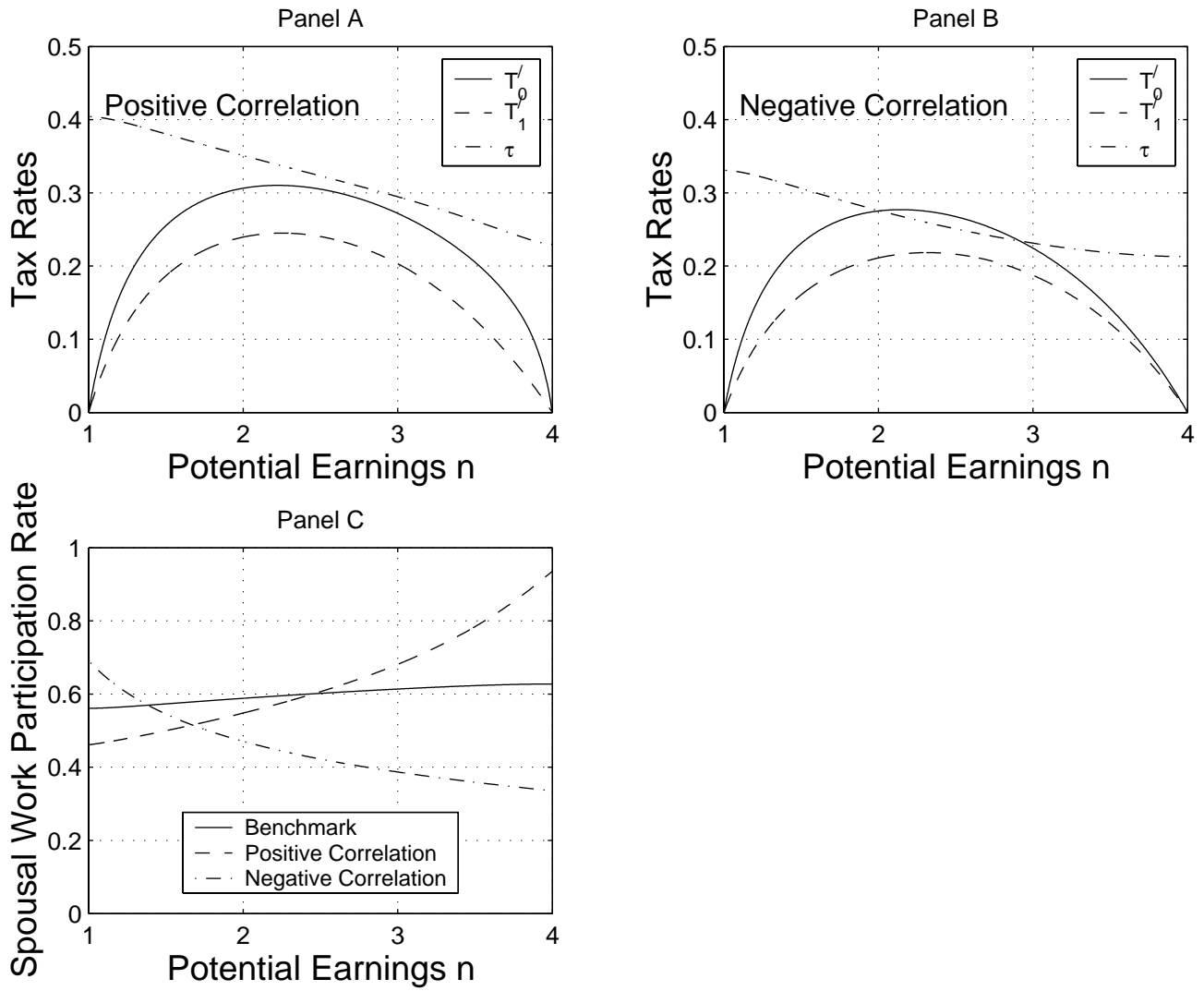
The utility function is given by

$$u(c, z_p, z_s) = c - n_p h_p(z_p/n_p) - n_s h_s(z_s/n_s),$$

---

<sup>20</sup>We assume that  $D$  is open and we denote by  $\bar{D}$  the closure of  $D$ .

Figure 7: The Effects of Spousal Correlation of Ability



with  $c = z_p + z_s - T(z_p, z_s)$ . The first-order conditions with respect to earnings  $z_p$  and  $z_s$  are given by

$$h'_p(z_p/n_p) = 1 - T'_p \quad \text{and} \quad h'_s(z_s/n_s) = 1 - T'_s. \quad (14)$$

The indirect utility function is denoted by  $V(n_p, n_s)$ , and its first-order derivatives with respect to  $n_p$  and  $n_s$  are given by (using the envelope theorem):

$$\partial_p V = -h_p + (z_p/n_p)h'_p \quad \text{and} \quad \partial_s V = -h_s + (z_s/n_s)h'_s. \quad (15)$$

The government's problem is to maximize social welfare

$$W = \int \int_D \Psi(V(n_p, n_s)) f(n_p, n_s) dn_p dn_s,$$

subject to a government budget constraint

$$\int \int_D T(z_p, z_s) f(n_p, n_s) dn_p dn_s \geq E,$$

and subject to the conditions for couple utility maximization in (15).

This is a continuous two-dimensional screening problem. There is a small literature in optimal tax theory considering this type of multi-dimensional screening models originating with Mirrlees (1976, 1986). There is a larger literature on multi-dimensional screening problems in nonlinear pricing theory (see McAfee and McMillan, 1988; Wilson, 1993; Armstrong, 1996; Armstrong 1999; Rochet and Choné, 1998; and Rochet and Stole, 2002). We explain the link to this literature in Section 3.4.

We can state the following proposition:

**Proposition 4** *The first-order conditions for the optimal marginal tax rates  $T'_p$  and  $T'_s$  at ability level  $(n_p, n_s)$  can be written as*

$$\frac{T'_p}{1 - T'_p} = \frac{1}{\varepsilon_p} \cdot \frac{1}{n_p f_{p|s}(n_p|n_s)} \cdot t_p, \quad (16)$$

$$\frac{T'_s}{1 - T'_s} = \frac{1}{\varepsilon_s} \cdot \frac{1}{n_s f_{s|p}(n_s|n_p)} \cdot t_s, \quad (17)$$

where  $t_p$  and  $t_s$  are multipliers satisfying the transversality conditions  $t_p(\underline{n}_p, n_s) = t_p(\bar{n}_p, n_s) = 0$  for all  $n_s$  and  $t_s(n_p, \underline{n}_s) = t_s(n_p, \bar{n}_s) = 0$  for all  $n_p$ , along with the divergence equation

$$\frac{\partial t_p}{\partial n_p} \cdot f_s(n_s) + \frac{\partial t_s}{\partial n_s} \cdot f_p(n_p) = [g(n_p, n_s) - 1] \cdot f(n_p, n_s), \quad (18)$$

where  $g(n_p, n_s) = \Psi'(V(n_p, n_s))/\lambda$  is the marginal welfare weight for couples with ability  $(n_p, n_s)$ .

The proof is presented in Appendix A.2.

The formulas are obtained from the first-order conditions to the Hamiltonian. The divergence equation (18) has many solutions satisfying the boundary transversality conditions. The fact that the second-order derivative of the indirect utility function  $V(n_p, n_s)$  has to be symmetric, gives an additional condition making the optimum solution unique generically. In addition, the global individual maximization conditions need to be satisfied. If those conditions fail, then there is bunching and the first-order conditions from the proposition break down. We come back to this important issue in detail in Section 3.4. In Sections 3.1-3, we always assume that those conditions are met and hence that there is no bunching.

It is easy to show that the average  $T'_p$  across  $n_s$  is the same as in the individualistic Mirrlees (1971) model. Let us define  $F_p(n_p) = \int_{\underline{n}_p}^{n_p} f_p(n'_p) dn'_p$  as the cumulated unconditional distribution of  $n_p$ . We then define  $G_p(n_p)$  as the average of marginal welfare weights  $g(n'_p, n'_s)$  above  $n_p$ :

$$G_p(n_p) \cdot [1 - F_p(n_p)] = \int_{n_p}^{\bar{n}_p} \int_{\underline{n}_s}^{\bar{n}_s} g(n'_p, n'_s) f(n'_p, n'_s) dn'_s dn'_p.$$

We can then show,

**Proposition 5**

$$\frac{T'_p}{1 - T'_p} = \frac{1}{\varepsilon_p} \cdot \frac{(1 - F_p(n_p)) \cdot (1 - G_p(n_p)) + \delta^p(n_p, n_s)}{n_p f_p(n_p)}, \quad (19)$$

where  $\delta^p(n_p, n_s)$  averages to zero when summed over  $n_s$ , i.e., for all  $n_p$

$$\int_{\underline{n}_s}^{\bar{n}_s} \delta^p(n_p, n_s) f(n_p, n_s) dn_s = 0.$$

The symmetric equations hold when substituting  $p$  for  $s$ .

**Proof:**

$\delta^p(n_p, n_s)$  is defined as:

$$\delta^p(n_p, n_s) = n_p f_p \cdot \varepsilon_p \cdot \frac{T'_p}{1 - T'_p} - (1 - F_p) \cdot (1 - G_p).$$

Hence, equation (16) implies:

$$\delta^p(n_p, n_s) \cdot f(n_p, n_s) = t_p \cdot f_p \cdot f_s - (1 - F_p) \cdot (1 - G_p) \cdot f(n_p, n_s).$$

Integrating this expression over  $(\underline{n}_s, \bar{n}_s)$ , we have:

$$\int_{\underline{n}_s}^{\bar{n}_s} \delta^p(n_p, n_s) f(n_p, n_s) dn_s = f_p(n_p) \int_{\underline{n}_s}^{\bar{n}_s} t_p(n_p, n_s) f_s(n_s) dn_s - f_p(n_p) \cdot (1 - F_p) \cdot (1 - G_p). \quad (20)$$

Integrating the divergence equation (18) over  $n_s$  and using the transversality conditions, we have:

$$\int_{\underline{n}_s}^{\bar{n}_s} \frac{\partial t_p}{\partial n_p} f_s(n_s) dn_s = \int_{\underline{n}_s}^{\bar{n}_s} [g(n_p, n_s) - 1] \cdot f(n_p, n_s) dn_s,$$

Integrating again from  $n_p$  to  $\bar{n}_p$ , we have:

$$\int_{\underline{n}_s}^{\bar{n}_s} t_p(n_p, n_s) f_s(n_s) dn_s = \int_{n_p}^{\bar{n}_p} \int_{\underline{n}_s}^{\bar{n}_s} [1 - g(n_p, n_s)] \cdot f(n_p, n_s) dn_s = (1 - G_p(n_p)) \cdot (1 - F_p(n_p)).$$

This implies that the expression (20) is zero which completes the proof.  $\square$

### 3.2 Asymptotic Properties of the Optimal Schedule

Suppose that the ability distribution of primary earners has an infinite tail,  $\bar{n}_p = \infty$ . Let us assume that  $f(n_p, n_s) = f_p(n_p) f_s(n_s)$  and that, for  $n_p$  large,  $f_p(n_p)$  is the density of a Pareto distribution with parameter  $a > 1$ .

As  $n_p$  tends to infinity, the additional income generated by the secondary earner becomes infinitesimal relative to primary-earner income in the limit. For any reasonable welfare function, we would then have that  $g(n_p, n_s)$  converge to the same value  $g^\infty$  for all  $n_s$ . It is also natural to assume that elasticities  $\varepsilon_i$  converge to  $\varepsilon_i^\infty$  when  $n_p \rightarrow \infty$  for  $i = p, s$ . We can then prove the following result:

**Proposition 6** *If  $T'_p(z(n_p, n_s))$  converges to  $\tau_p^\infty(n_s)$  and  $T'_s(z(n_p, n_s))$  converges to  $\tau_s^\infty(n_s)$  as  $n_p \rightarrow \infty$  (and assuming that the limits are bounded from below uniformly in  $n_s$ ), then we have:*

- $\tau_s^\infty(n_s) = 0$  for all  $n_s$ .
- $\tau_p^\infty(n_s) = (1 - g^\infty) / (1 - g^\infty + a \cdot \varepsilon_p^\infty)$  for all  $n_s$ .

**Proof:**

We first establish that  $\tau_p^\infty(n_s)$  is constant in  $n_s$ . By contradiction, suppose that there are  $n_s^1, n_s^2$  so that  $\tau_p^\infty(n_s^2) < \tau_p^\infty(n_s^1)$ . Using  $\partial_p V = -h_p(z_p/n_p) + (z_p/n_p)h'_p(z_p/n_p)$  which is increasing in  $z_p/n_p$  and hence in  $1 - T'_p$  from (14), we have  $\partial_p V(n_p, n_s^2) - \partial_p V(n_p, n_s^1) \rightarrow \delta > 0$  when  $n_p \rightarrow \infty$ . That implies that  $V(n_p, n_s^2) - V(n_p, n_s^1) \rightarrow +\infty$  when  $n_p \rightarrow \infty$ . However,  $\partial_s V(n_p, n_s) = -h_s(z_s/n_s) + (z_s/n_s)h'_s(z_s/n_s)$  converges to a finite limit for any  $n_s$  uniformly bounded in  $n_s$  (because  $T'_s$  converges and is uniformly bounded from below). Therefore,  $V(n_p, n_s^2) - V(n_p, n_s^1) = \int_{n_s^1}^{n_s^2} \partial_s V(n_p, n'_s) dn'_s$  converges to a finite limit when  $n_p \rightarrow \infty$  which is a contradiction. Let us now denote  $\tau_p^\infty$  the uniform limit of  $T'_p$ .

Integrating the divergence equation (18) over the NE quadrant  $(n_p, \infty) \times (n_s, \bar{n}_s)$ , and using the transversality conditions implies:

$$\int_{n_s}^{\bar{n}_s} t_p(n_p, n'_s) f_s(n'_s) dn'_s + \int_{n_p}^{\infty} t_s(n'_p, n_s) f_p(n'_p) dn'_p = \int_{n_s}^{\bar{n}_s} \int_{n_p}^{\infty} [1 - g(n'_p, n'_s)] f_p(n'_p) f_s(n'_s) dn'_p dn'_s. \quad (21)$$

The first-order condition (16) and the Pareto assumption imply that  $t_p(n_p, n'_s)/(1 - F_p(n_p))$  converges to  $a \cdot \varepsilon_p^\infty \cdot \tau_p^\infty / (1 - \tau_p^\infty)$  for any  $n'_s$ . The first-order condition (17) imply that  $t_s(n_p, n_s)$  converges to  $t_s^\infty(n_s) = n_s f_s(n_s) \cdot \varepsilon_s^\infty \cdot \tau_s^\infty(n_s) / (1 - \tau_s^\infty(n_s))$ . Dividing (21) by  $1 - F_p(n_p)$  and taking the limit when  $n_p \rightarrow \infty$ , we have:

$$a \cdot \varepsilon_p^\infty \cdot \frac{\tau_p^\infty}{1 - \tau_p^\infty} \cdot \int_{n_s}^{\bar{n}_s} f_s(n'_s) dn'_s + t_s^\infty(n_s) = (1 - g^\infty) \int_{n_s}^{\bar{n}_s} f_s(n'_s) dn'_s. \quad (22)$$

The transversality condition implies that  $t_s^\infty(\underline{n}_s) = 0$  so that the second term on the left-hand-side of (22) vanishes when  $n_s = \underline{n}_s$ . This establishes the second bullet of the lemma. Equation (22) then implies that  $t_s^\infty(n_s) = 0$  for any  $n_s$  and hence  $\tau_s^\infty(n_s) = 0$  for any  $n_s$  which proves the first bullet of the lemma.  $\square$

This proposition shows that the no tax on spouses at the top result generalizes to the double-continuous model. The intuition is the same as in the binary case.<sup>21</sup>

### 3.3 Desirability of Negative Jointness

Suppose the government implements the optimal *separable* tax schedule. It is then straightforward to show, using the standard one-dimensional approach, that the optimal schedules take the form

$$\frac{T'_p}{1 - T'_p} = \frac{1}{\varepsilon_p} \frac{(1 - F_p(n_p)) \cdot (1 - G_p(n_p))}{n_p f_p}, \quad (23)$$

$$\frac{T'_s}{1 - T'_s} = \frac{1}{\varepsilon_s} \frac{(1 - F_s(n_s)) \cdot (1 - G_s(n_s))}{n_s f_s}. \quad (24)$$

Let us introduce the equivalent of Assumption 2 in the current model:

**Assumption 2'**:  $n_p$  and  $n_s$  are independently distributed.

Now, as in the binary case, it is possible to show that under Assumptions 1 ( $\Psi'$  convex) and Assumption 2', and starting from the optimal separable schedule characterized above, a tax

---

<sup>21</sup>If both distributions of  $n_p$  and  $n_s$  have an infinite tail with the same Pareto parameter and the same asymptotic elasticity of labor supply  $\varepsilon^\infty$ , then, along the diagonal  $n_p = n_s$ , both marginal tax rates should be equal and converge to  $(1 - g^\infty)/(1 - g^\infty + 2 \cdot a \cdot \varepsilon^\infty)$  when  $n_p = n_s$  tend to infinity.

reform introducing a little bit of negative jointness increases welfare. The proof is given in Kleven et al. (2006), but we omit it here and focus instead on the properties of optimal schemes. Of course, the result that it is welfare improving to introduce a little bit of negative jointness establishes a strong intuition that negative jointness is a feature of optimal incentive schemes.<sup>22</sup> Indeed, we can show:

**Theorem 1** *Under Assumptions 2' and assuming that the optimal tax system is smooth and displays no bunching, we have:*

- If  $\Psi'$  is convex, the optimal tax scheme features negative jointness, i.e., for all  $n_p, n_s$ ,

$$\frac{\partial T'_p}{\partial n_s} \leq 0, \quad \text{and} \quad \frac{\partial T'_s}{\partial n_p} \leq 0, \quad \text{and} \quad T''_{ps} \leq 0. \quad (25)$$

- If  $\Psi'$  is concave, the optimal tax scheme features positive jointness everywhere.
- If  $\Psi'$  is linear, the optimal tax scheme is separable and the optimal tax rates are given by equations (23) and (24).

**Proof:**

The proof proceeds by contradiction. Let us consider the case where  $\Psi'$  is convex. We define as  $C$  the subset of  $D$  where property (25) is not met.

First, we note that eq. (14) and the fact that the functions  $h_p$  and  $h_s$  are convex imply that the signs of  $\partial T'_p/\partial n_s$  and  $\partial z_p/\partial n_s$  are opposite (such that one is positive/zero/negative iff the other is negative/zero/positive). Similarly, the signs of  $\partial T'_s/\partial n_p$  and  $\partial z_s/\partial n_p$  are opposite. Second, from eq. (15) we obtain  $\partial_{ps}^2 V = (z_p/n_p^2) \cdot h''_p \cdot \partial z_p/\partial n_s$  and  $\partial_{sp}^2 V = (z_s/n_s^2) \cdot h''_s \cdot \partial z_s/\partial n_p$ , and the symmetry condition  $\partial_{ps}^2 V = \partial_{sp}^2 V$  then implies that the signs of  $\partial z_s/\partial n_p$  and  $\partial z_p/\partial n_s$  are identical. Hence, the first two inequalities in (25) are equivalent.

Assumption 2' implies that  $f_{p|s}(n_p|n_s) = f_p(n_p)$  and  $f_{s|p}(n_s|n_p) = f_s(n_s)$ . Hence, the optimal tax formulas from Proposition 4 imply:

$$t_p(n_p, n_s) = \frac{T'_p}{1 - T'_p} \cdot \varepsilon_p \cdot n_p f_p(n_p) = \frac{1 - h'_p(z_p/n_p)}{h''_p(z_p/n_p) \cdot z_p/n_p} \cdot n_p f_p(n_p), \quad (26)$$

$$t_s(n_p, n_s) = \frac{T'_s}{1 - T'_s} \cdot \varepsilon_s \cdot n_s f_s(n_s) = \frac{1 - h'_s(z_s/n_s)}{h''_s(z_s/n_s) \cdot z_s/n_s} \cdot n_s f_s(n_s). \quad (27)$$

---

<sup>22</sup>The proof of the desirability of introducing negative jointness requires to assume no bunching. The analysis of bunching in the optimal *separable* tax system is the same as in the one dimensional Mirrlees (1971) model. Therefore, there will be no bunching in a wide set of cases as in the Mirrlees (1971).

We show in Appendix A.3 that the second order condition of the government maximization program imply that  $x \rightarrow (1 - h'_p(x))/(xh''_p(x))$  is decreasing at any point  $x = z_p(n)/n_p$  for  $n \in D$ . Therefore, differentiating (26) with respect to  $n_s$ , we have that  $\partial t_p/\partial n_s$  has the opposite sign of  $\partial z_p/\partial n_s$  and hence the same sign as  $\partial T'_p/\partial n_s$ . Similarly,  $\partial t_s/\partial n_p$  has the same sign as  $\partial T'_s/\partial n_p$ . We can then define our contradiction set  $C$  as follows:

$$C = \{(n_p, n_s) \in D | \partial t_s/\partial n_p > 0\} = \{(n_p, n_s) \in D | \partial t_p/\partial n_s > 0\}.$$

$C$  is an open set (because  $\partial t_s/\partial n_p$  and  $\partial t_p/\partial n_s$  are continuous functions by assumption). We denote by  $\partial C$  the boundary of  $C$ . Again, by continuity of  $\partial t_s/\partial n_p$ , we have  $\partial t_s/\partial n_p = \partial t_p/\partial n_s = 0$  for any  $(n_p, n_s) \in \partial C$ . We denote by  $C^c$  the complement of  $C$  in  $D$ .

**Lemma 3** *If  $\Psi'$  is convex then, for any  $(n_p, n_s) \in C$ , we have  $\partial^2 g/(\partial n_p \partial n_s) > 0$ .*

By definition,  $g(n_p, n_s) = \Psi'(V(n_p, n_s))/\lambda$  where  $\lambda > 0$  is the multiplier of the government budget constraint. Hence,  $\lambda \cdot \partial g/\partial n_p = \Psi''(V) \cdot \partial_p V$  and  $\lambda \cdot \partial^2 g/(\partial n_p \partial n_s) = \Psi'''(V) \cdot \partial_p V \cdot \partial_s V + \Psi''(V) \cdot \partial_{ps}^2 V$ . The first term in this expression is positive because  $\Psi'$  is convex and  $\partial_p V, \partial_s V > 0$ . In the second term, we have  $\partial_{ps}^2 V = (z_p/n_p^2) \cdot h''_p \cdot \partial z_p/\partial n_s$ . By definition of  $C$ , we have  $\partial z_p/\partial n_s < 0$  and hence  $\partial_{ps}^2 V < 0$ . Hence, since  $\Psi$  is concave such that  $\Psi'' < 0$ , the second term is also positive, and the lemma is then established.

The property  $\partial^2 g/(\partial n_p \partial n_s) > 0$  captures the notion that the difference in social marginal welfare weights between families with low- and high-ability primary earners decreases when secondary-earner ability increases. This property is directly equivalent to  $\Psi''' > 0$  when the tax system is separable (in which case  $\partial_{ps}^2 V = 0$ ). The lemma shows that it holds a-fortiori when the tax system displays positive jointness.

Using  $f(n_p, n_s) = f_p(n_p) \cdot f_s(n_s)$ , the divergence equation (18) can be rewritten to

$$\frac{1}{f_p} \cdot \frac{\partial t_p}{\partial n_p} + \frac{1}{f_s} \cdot \frac{\partial t_s}{\partial n_s} = g(n_p, n_s) - 1. \quad (28)$$

Therefore, we have

$$\frac{\partial^2 g}{\partial n_p \partial n_s} = -\frac{f'_p}{f_p^2} \cdot \frac{\partial^2 t_p}{\partial n_p \partial n_s} + \frac{1}{f_p} \cdot \frac{\partial^3 t_p}{(\partial n_p)^2 \partial n_s} - \frac{f'_s}{f_s^2} \cdot \frac{\partial^2 t_s}{\partial n_p \partial n_s} + \frac{1}{f_s} \cdot \frac{\partial^3 t_s}{\partial n_p (\partial n_s)^2}. \quad (29)$$

Let us now introduce the field vector  $\mathbf{K} = (K_p(n_p, n_s), K_s(n_p, n_s))$  defined as:

$$K_p = \frac{1}{f_p} \cdot \frac{\partial^2 t_p}{\partial n_p \partial n_s} \quad \text{and} \quad K_s = \frac{1}{f_s} \cdot \frac{\partial^2 t_s}{\partial n_p \partial n_s}.$$



Routine differentiation shows that:

$$\frac{\partial^2 g}{\partial n_p \partial n_s} = \frac{\partial K_p}{\partial n_p} + \frac{\partial K_s}{\partial n_s}. \quad (30)$$

This allows us to apply the *Divergence Theorem* relating the area integral of the divergence of a function to the boundary integral of the function:

$$\int \int_C \frac{\partial^2 g}{\partial n_p \partial n_s} dn_p dn_s = \int \int_C \left( \frac{\partial K_p}{\partial n_p} + \frac{\partial K_s}{\partial n_s} \right) dn_p dn_s = \oint_{\partial C} \mathbf{K} \cdot d\mathbf{m}, \quad (31)$$

where  $\mathbf{m}$  is the unit vector outward normal to  $C$  on  $\partial C$ . Figure 8 displays an illustration in the case of a region  $C$  with no holes with a simple curve  $\partial C$ .<sup>23</sup> Lemma 3 establishes that the left-hand side of (31) is positive. We are now going to show that  $\mathbf{K} \cdot \mathbf{m} \leq 0$  on  $\partial C$  to establish a contradiction.

**Lemma 4** *For any  $(n_p, n_s) \in \partial C$ , we have  $\mathbf{K} \cdot \mathbf{m} \leq 0$ .*

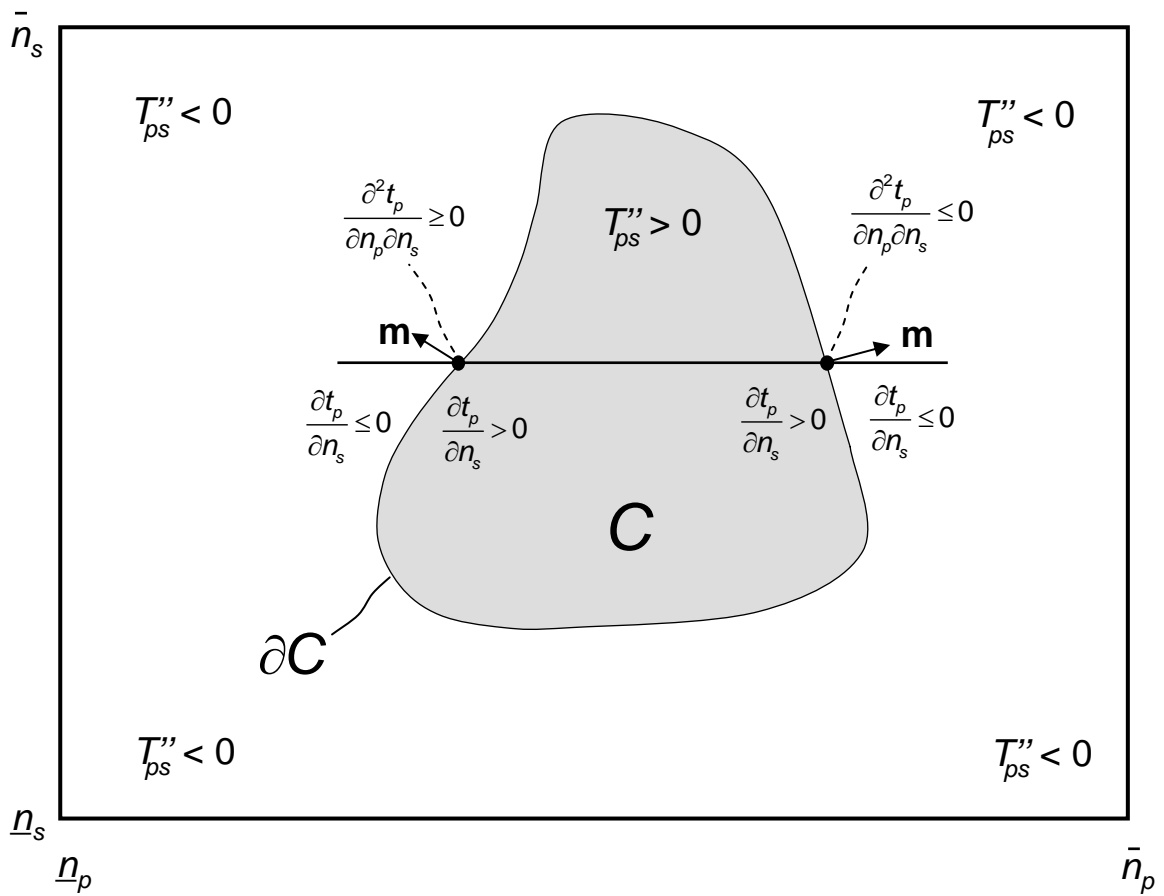
The proof of this Lemma is illustrated in Figure 8. We have  $\partial t_p / \partial n_s > 0$  inside  $C$  and  $\partial t_p / \partial n_s \leq 0$  outside  $C$ . Hence,  $\partial t_p / \partial n_s$  increases as one goes from outside  $C$  to inside  $C$  along a horizontal line (constant  $n_s$  and changing  $n_p$ ) as shown in the figure. Consider the two points where this horizontal line intersects the boundary  $\partial C$ . If area  $C$  is on the right side of the intersection point, then we have that  $\partial t_p / \partial n_s$  is increasing in  $n_p$  and hence  $\partial^2 t_p / (\partial n_p \partial n_s) \geq 0$  at the intersection point. Area  $C$  is on the right side of the intersection point if and only if  $m_p$ , defined as the  $p$  component of vector  $\mathbf{m}$ , is negative.<sup>24</sup> Conversely, if area  $C$  is on the left side of the intersection point, then we have  $\partial^2 t_p / (\partial n_p \partial n_s) \leq 0$  and  $m_p > 0$ . By definition of  $K_p$ , this means that  $K_p \cdot m_p \leq 0$  in all cases. In a similar way, we can show that  $K_s \cdot m_s \leq 0$  by using that  $\partial t_s / \partial n_p$  increases as one goes from outside  $C$  to inside  $C$  along a vertical line (constant  $n_p$  and changing  $n_s$ ). If area  $C$  is above the intersection point of this vertical line with  $\partial C$ , it means that  $\partial^2 t_p / (\partial n_p \partial n_s) \geq 0$  at  $\partial C$ . Area  $C$  being above the intersection point means that  $m_s < 0$ , and we then have  $K_s \cdot m_s \leq 0$ . When area  $C$  is instead below the intersection point, we have  $\partial^2 t_p / (\partial n_p \partial n_s) \leq 0$ ,  $m_s > 0$ , and then again  $K_s \cdot m_s \leq 0$ . Hence, we have  $\mathbf{K} \cdot \mathbf{m} = K_p \cdot m_p + K_s \cdot m_s \leq 0$  and the lemma is established.

Finally, we need to show that  $\partial t_s / \partial n_p \leq 0$  and  $\partial t_p / \partial n_s \leq 0$  implies  $T''_{ps} \leq 0$  such that the direct tax function  $T(z_p, z_s)$  has a negative cross-derivative on the image domain of the solution.

<sup>23</sup>The divergence theorem is valid for a region with holes. In that case the boundary integral is the sum of all the simple boundary integrals along all the simple closed curves defining the boundary  $\partial C$ . Note that the curves in  $\partial C$  always close because the region  $C$  cannot intersect with the boundary  $\partial D$  of the domain  $D$ . This is because the transversality conditions imply that  $\partial t_s / \partial n_p = \partial t_p / \partial n_s = 0$  on  $\partial D$ .

<sup>24</sup>Remember  $\mathbf{m}$  is pointing outward from  $C$ .

Figure 8. Proof of Lemma 4



The proof follows from manipulation of the first-order and second-order conditions from the individual maximization problem and is presented in Appendix A.4.

The proof in the case  $\Psi'$  concave follows exactly the same path by defining a symmetric contradictory set. In that case, Lemmas 3 and 4 also apply in the contradictory set but with opposite signs. Finally, the case where  $\Psi'$  is linear can be demonstrated by showing that equations (23) and (24) define an optimum that satisfies all the equations from Proposition 4.  $\square$

### 3.4 Bunching and Link with Multi-Dimensional Screening

Our main result Theorem 1 has been demonstrated assuming that the optimal tax system has no bunching. Yet, the important studies by Armstrong (1996) and Rochet and Choné (1998) have demonstrated that bunching is generic in multi-dimensional screening problems. Armstrong (1996) made the important point that bunching happens generically at the bottom. Rochet and Choné (1998) then characterized with great detail the complex nature of bunching. As they explain, bunching arises from a conflict between participation constraints and second order incentive compatibility conditions. However, in the case of social welfare maximization, there are no participation constraints.

Theoretically, we can show that for moderate redistributive tastes, the solution displays no bunching. The argument is the following: when  $\Psi(V) = V$ , the government has no concerns for redistribution and hence the optimal system is the laissez-faire situation with no taxation  $T_0 = T_1 \equiv 0$ . Obviously, the laissez-faire optimum displays no bunching. If we introduce a little bit of taste for redistribution, the optimal solution should remain close to the laissez-faire solution, implying that there should be no bunching for low levels of redistribution. This argument amounts to proving that the optimal solution varies smoothly with the redistributive tastes of the government.

In order to make a formal argument, we will make a number of parametric and regularity assumptions in order to keep the mathematical proofs reasonably simple.<sup>25</sup> First, we consider the CRRA case where  $\Psi'(V) = V^{-\gamma}$ . The case  $\gamma = 0$  is the case with no redistribution concerns. Second, we assume that  $h_p(x) = h_s(x) = x^{1+1/\varepsilon}/(1 + 1/\varepsilon)$  so that the elasticities are constant  $\varepsilon_p = \varepsilon_s = \varepsilon$ . In that case, equations (15) imply that  $\partial_j V = (z_j/n_j)^{1+1/\varepsilon}/(1 + \varepsilon)$  for  $j = p, s$ . Third, we assume that  $f$  is  $C^\infty$  and bounded away from zero on  $\bar{D}$ . Finally, we assume that  $D$

---

<sup>25</sup>We conjecture that the no bunching result can be generalized to a wider set of situations. We are particularly indebted to Jean-Charles Rochet and David Lannes for helping us with the proof in the simple case we consider.

is convex, bounded, and has a smooth ( $C^\infty$ ) boundary.<sup>26</sup>

We start by ignoring the second order condition for the couples maximization. This is called the relaxed problem in Rochet (1987) and Rochet and Choné (1998). As they do, we can then express the government maximization problem solely in terms of  $V$ :<sup>27</sup>

$$\max_V \int_D \frac{V(n)^{1-\gamma}}{1-\gamma} f(n) dn,$$

subject to the government budget constraint (with multiplier  $\lambda$ ):

$$B(V) = \int_D \left[ \sum_{i=p,s} n_i ((1+\varepsilon)\partial_i V)^{\frac{\varepsilon}{1+\varepsilon}} - \varepsilon \sum_{i=p,s} n_i \partial_i V - V \right] f(n) dn \geq 0.$$

It is straightforward to show that the objective function is concave (and strictly so when  $\gamma > 0$ ). Furthermore, the budget constraint  $B(V) \geq 0$  defines a closed and convex set.<sup>28</sup> Therefore, the maximization problem can be written as the following convex minimization problem:

$$\min_{V \in K} \Phi(V, \gamma), \tag{32}$$

where  $\Phi(V, \gamma) = - \int_D V(n)^{1-\gamma} f(n) dn$  is convex and continuous on  $K = \{V \in H^1(D) \mid 0 \leq V \leq N, 0 \leq \partial_i V \leq N, B(V) \geq 0\}$  convex, closed, and bounded<sup>29</sup> in the Sobolev (Hilbert) space  $H^1(D)$ . A standard theorem in functional analysis ensures that a solution to problem (32) exists.<sup>30</sup> The strict convexity of  $K$  ensures that the solution is unique.<sup>31</sup> Let us denote the unique solution of (32) by  $V(\gamma)$ . It is easy to show that  $V(\gamma = 0) = (n_p + n_s)/(1 + \varepsilon)$  which corresponds to the case with no redistributive tastes and no taxes where  $z_i = n_i$  for  $i = p, s$ .

**Lemma 5** *The solution  $V(\gamma)$  of the relaxed problem (32) is smooth in  $\gamma$  around  $\gamma = 0$  so that  $V(\gamma) = V(0) + \gamma \cdot U + o(\gamma)$  where  $U \in C^2(\bar{D})$  and  $o(\gamma)/\gamma \rightarrow 0$  when  $\gamma \rightarrow 0$  (in the norm  $C^2(\bar{D})$ ).*

<sup>26</sup>In our previous subsection, we assumed that  $D$  was a rectangle which is not smooth at the corners. We need to make a smoothness assumption in order to avoid the difficulties arising in elliptic problems in non-smooth domains (see e.g., Grisvard (1985)).

<sup>27</sup>This can be done by replacing  $z_j$  in the budget constraint by  $\partial_j V = (z_j/n_j)^{1+1/\varepsilon}/(1+\varepsilon)$ .

<sup>28</sup>Strict convexity is obtained under general  $h_p, h_s$  disutility of labor functions if and only if  $x \rightarrow (1 - h'_i(x))/(xh''_i(x))$  is decreasing for  $i = p, s$ . This condition is obviously met in the case of iso-elastic  $h_i$ .

<sup>29</sup>The bounded property is obtained by imposing the additional constraints that  $\partial_i V \leq N$  where  $N$  is a large and fixed constant.

<sup>30</sup>The theorem states that a convex lower semi-continuous function attains a minimum on a closed convex and bounded set of a reflexive Banach vector space (see e.g., Brezis (1983), Corollary III.19, p. 46). The Hilbert space  $H^1(D)$  is obviously a reflexive Banach vector space.

<sup>31</sup>If  $V^1$  and  $V^2$  are two solutions, then  $V^* = (V^1 + V^2)/2$  will be such that  $B(V^*) > 0$  (unless  $V^1 = V^2$  a.e.) and hence  $V^* + \delta$  will be in  $K$  for  $\delta$  small enough and will generate strictly higher social welfare than  $V^1$  and  $V^2$ .

We establish this lemma in Appendix A.5 where we show that  $U$  is the solution of a linear Elliptic PDE.  $U$  characterizes the *direction* of the optimal tax distortion when small redistributive tastes are introduced.

We now need to establish that the solution  $V(\gamma)$  displays no bunching for  $\gamma$  small. If  $V \in K$  is solution of the relaxed problem (32), we can define  $z_j(n)$  for  $j = p, s$  using  $\partial_j V = (z_j/n_j)^{1+1/\varepsilon}/(1+\varepsilon)$ . We can then define  $c(n) = V(n) + n_p h_p(z_p(n)/n_p) + n_s h_s(z_s(n)/n_s)$ . The direct utility function is defined as  $u(c, z, n) = c - n_p h_p(z_p/n_p) - n_s h_s(z_s/n_s)$ . The solution  $V$  satisfies the global individual utility maximization if and only if  $V(n) \geq u(c(n'), z(n'), n)$  for all  $n, n' \in D$ . In that case, the solution of the relaxed problem is actually the solution of the full problem and be decentralized with a tax system.

Following Mirrlees (1976, 1986), we can establish the following lemma insuring that  $V$  (and the implied  $z(n)$ ) satisfy global utility maximization:

**Lemma 6** *If  $\partial z_p/\partial n_p \geq 0$ ,  $\partial z_s/\partial n_s \geq 0$ , and  $(\partial z_p/\partial n_p) \cdot (\partial z_s/\partial n_s) \geq K \cdot (\partial z_p/\partial n_s) \cdot (\partial z_s/\partial n_p)$  for all  $n = (n_p, n_s) \in D$  where  $K = (1/4) \max_{n, m \in \bar{D}} \left[ 1 + (n_p m_s)/(n_s m_p) \right]^{1+1/\varepsilon} \cdot \left[ 1 + ((n_s m_p)/(n_p m_s))^{1+1/\varepsilon} \right]$ , then the solution  $V$  satisfies global individual utility maximization:  $V(n) \geq u(c(n'), z(n'), n)$  for all  $n, n' \in D$ .*

The Lemma is proved in appendix A.6. The lemma condition is obviously satisfied when  $\gamma = 0$  as  $z_p = n_p$  and  $z_s = n_s$  in that case. The second derivative of  $V(\gamma = 0) = (n_p + n_s)/(1+\varepsilon)$  is zero  $D^2V(\gamma = 0) = 0$  (as a two-by-two matrix), therefore, Lemma 5 implies that  $D^2V(\gamma) = \gamma D^2U + o(\gamma) \rightarrow 0$  when  $\gamma \rightarrow 0$ . As  $z_i = n_i((1+\varepsilon)\partial_i V)^{\varepsilon/(1+\varepsilon)}$  for  $i = p, s$ , the cross-partial derivatives  $\partial z_i/\partial n_j$  (for  $i \neq j$ ) will be close to zero for small  $\gamma$ . Therefore, the condition of lemma 6 for global utility maximization will be satisfied for small  $\gamma$ . Therefore,  $V(\gamma)$ , the solution of the relaxed problem is also the full solution. Hence, the full solution displays no bunching for small  $\gamma$ . Therefore, we have proved the following:

**Theorem 2** *For  $\gamma$  close enough to zero, the optimal tax system displays no bunching.*

There are four notable consequences of Theorem 2. First, this no bunching theorem for small redistributive tastes also clearly applies to the standard Mirrlees (1971) one dimensional problem. In contrast to the multi-dimensional case, it can be demonstrated using the first order condition for optimality and without invoking advanced functional analysis results. To the best of our knowledge, this result does not seem to have been noticed in the extensive literature on the one-dimensional problem.

Second, it is easy to show that the multi-dimensional screening problem for the monopolist is formally equivalent to our optimal tax problem in the case of a Rawlsian objective where the government maximizes the utility of the worse-off couple  $(\underline{n}_p, \underline{n}_s)$ . In that case, it is equivalent for the government to maximize taxes subject to a minimum utility level constraint and taxes are then redistributed lump sum (as there are no income effects). In that case,  $\gamma = \infty$  and the social marginal welfare weight  $g(n_p, n_s)$  becomes a Dirac distribution with all mass at the bottom point  $(\underline{n}_p, \underline{n}_s)$ . The bunching result from Armstrong (1996) clearly applies in that case as well. This means that, as  $\gamma$  increases, we should expect bunching to appear. Exploring with numerical simulations below how large  $\gamma$  needs to be for bunching to appear is left for future research.

Third, we have shown in the previous subsection that when  $\Psi$  is quadratic, the optimal tax system is separable. In that case, the cross derivatives  $\partial z_i / \partial n_j, i \neq j$  are zero and therefore  $\partial z_i / \partial n_i \geq 0, i = p, s$  ensures that there is no bunching exactly as in the one dimensional case. We know from the one dimensional case that this happens for a wide set of parameters. Starting from  $\Psi$  quadratic, a small perturbation on  $\Psi$  will also create only a small deviation on the solution  $V$  of the relaxed problem. As a result, the solution  $V$  will also display no bunching. This shows that there should be a wide set of cases with significant redistributive tastes where the solution displays no bunching.

Finally, our no bunching result and the negative jointness result we derived in the previous subsection carry over to the standard industrial organization model used in Armstrong (1996) or Rochet and Choné (1998) if monopoly profit maximization is replaced by social welfare maximization as long as redistributive tastes are small enough. Social welfare maximization is of less direct interest in Industrial Organization than in optimal taxation. However, it is interesting to note that environments more competitive than monopoly pricing can also generate solutions with no bunching in multi-dimensional pricing problems (see Armstrong and Vickers 2001 for a recent analysis in that direction).

## 4 Model Extensions

### 4.1 Endogenous Marriage Decisions

We have demonstrated that negative jointness is optimal assuming that marriage is unresponsive to taxes. Because any form of joint tax treatment for married couples affects the incentives to marry, it is relevant to consider the case of endogenous marriage. Indeed, a classic argument

for individual taxation is that tax systems should be neutral with respect to marriage decisions. Studies estimating the effect of income taxes on marriage tend to find statistically significant but modest effects (e.g. Alm and Whittington, 1999; Eissa and Hoynes, 2000).

We now present an argument that our results survive endogenous marriage in the context of the binary model of Section 2 (the argument can easily be extended to the continuous model). Suppose that the economy is populated by individuals of type  $p$  (characterized as our primary earners by ability  $n$ ) or type  $s$  (characterized as our secondary earners by a fixed work cost  $q$ ), and that individuals can choose either to be single or to be married. We start from the optimal *separable* tax system so that marriage decisions are initially undistorted, and introduce a little bit of negative jointness as illustrated in Figure 3. As described above, this reform creates a positive direct welfare effect, while the fiscal effects from labor supply responses cancel out. In the case of endogenous marriage, there is an additional behavioral response because incentives to marry have changed. At low primary earnings, marriage has become less attractive for two-earner couples and more attractive for one-earner couples while, at high primary earnings, marriage incentives are changing in the opposite directions. However, the marriage responses created these changes have no first-order effect on utility (standard envelope theorem) and no first-order effect on government revenue, because we are starting from a separable tax schedule whereby marital status has no tax consequences. Hence, the negative jointness reform is still desirable and Proposition 3 remains valid. This implies that negative jointness should also be part of an optimal incentive scheme, although the presence of marriage distortions will tend to reduce the optimal degree of negative jointness.

## 4.2 The Collective Labor Supply Approach

A growing literature challenges the unitary approach adopted in this paper, arguing that the family should be viewed as consisting of members with conflicting interests engaging in bargaining over household resources (see Lundberg and Pollak, 1996, for a survey of this literature). Empirical studies have supported this hypothesis. For example, the influential study of Lundberg et al. (1997) showed that a policy reform which transferred a child allowance from the father to the mother significantly increased spending on the wife and children in the family.

Following the seminal contributions by Chiappori (1988, 1992), the collective labor supply model has become especially popular. This approach does not model a particular bargaining process—only Pareto efficiency is assumed—and it encompasses the unitary model as a special case. In the collective model, the within-family decision process amounts to maximizing a

weighted sum of individual utilities, where the weights may depend on factors such as innate characteristics, relative incomes, and on whom receives government transfers. It is natural to distinguish between two cases depending on the government's view on intra-family distribution. In one case, policy makers respect family sovereignty. In this case, it is easy to see that changes in intra-household distribution have no consequences for social welfare, implying that all our optimal tax results continue to apply.

In the alternative case, policy makers disagree with intra-household distribution. The findings by Lundberg et al. (1997) suggest that the government can modify within-family consumption allocation at no fiscal cost simply by transferring the benefits from one spouse to the other keeping total family income constant. As shown in the formal analysis of Kroft (2007), by transferring enough resources across spouses, the government is able to restore a fair allocation within the family.<sup>32</sup> Moreover, this within-family redistribution is not associated with an efficiency loss, because it has no fiscal cost *and* because within-family bargaining is Pareto efficient. Once within-family distributional issues are fully resolved at no efficiency cost, we are essentially back to the problem of redistribution across families which we have analyzed in this paper. Hence, the collective labor supply approach introduces a new intra-family dimension to the redistribution problem which is very interesting and calls for more work, but which appears to be independent of the inter-family redistribution problem considered in this paper.

## 5 Conclusion

This paper explored the optimal income tax treatment of couples allowing for fully general joint income tax systems. To make progress on this difficult problem, we focused on unitary models of family decision making, and assumed no income effects on labor supply and separability in the disutility of work for husbands and wives. We considered models where the secondary earner's labor supply is either binary (only extensive response) or continuous (only intensive response). Assuming independent abilities across spouses, our central theoretical result is that, if the social marginal utility of income is convex in income, then the optimal tax function has a negative cross-partial derivative everywhere, implying that the tax rate on one person is decreasing in the earnings of the spouse. Numerical simulations showed that this negative jointness result survives positive assortative matching and may even be reinforced.

The intuition for our results can be understood as follows. Redistribution from couples with

---

<sup>32</sup>For example, the credit reform studied in Lundberg et al. (1997) did not affect family budget constraints but yet had an impact on the consumption allocation within families.



high primary earnings to couples with low primary earnings follows the logic of the Mirrlees (1971) model. Indeed, the marginal tax rate on primary earners at each earnings level, averaging over their different spouses, is identical to the marginal tax rate obtained in the Mirrlees model. At a given level of primary earnings, the government values redistribution from couples with high secondary earnings to couples with low secondary earnings, and this requires a positive second-earner tax. But the value of redistributing in favor of couples with low secondary earnings diminishes as primary earnings increases, because secondary earnings become less important for family utility. Hence, the optimal second-earner tax is decreasing in primary earnings, and tends to zero as primary earnings go to infinity.

The negative jointness result may seem surprising at first glance, and at odds with the actual practice in countries using joint taxation. However, we have argued that the current practice of many European countries—such as the United Kingdom—combining an individual income tax with a family-based and means-tested welfare system creates negative jointness. In families with low primary-earner incomes, secondary earners face high tax rates due to transfer phase-out, whereas in families with medium or high primary-earner incomes, secondary earners face low tax rates because the income tax is individual.

It is interesting to note that our result, in the binary model, that the second-earner participation tax is always positive stands in contrast to Saez (2002) who showed that, for *unmarried* individuals, the presence of participation responses tend to make EITC-schemes featuring negative tax rates at the bottom desirable. We conjecture a generalization of our model allowing for participation responses for the primary earner would imply negative tax rates at the bottom for primary earners along with positive tax rates on their spouses.

It would be very interesting to extend the numerical simulations to carefully calibrated models which are closer to the real world in terms of labor supply responses and the joint distribution of spouse abilities. Such simulations would allow us to assess the quantitative importance of the negative jointness result and make it possible to assess quantitatively whether the current practice in many OECD countries of imposing family based transfer programs along with individually based income taxes is close to optimal. We leave such important extensions for future work.

On the theoretical side, we have shown that with smooth and concave social welfare functions, the solution of the multi-dimensional screening problem is regular with no bunching for a wide set of parameters. This stands in sharp contrast to previous results in the Industrial Organization literature showing that, in the case of monopoly profit maximization, the solution displays

bunching generically. With competitive environments instead of monopoly profit maximization, multi-dimensional screening problems can sometimes generate solutions with no bunching (see e.g., Armstrong and Vickers (2001, 2006) and Armstrong (2006)). Using similar techniques as the ones developed here, it might be possible to obtain qualitative properties of the optimal solution in some of those cases. Finally, although our model has focused on the case of couples taxation, it could be easily extended to other settings with multi-dimensional characteristics where the separability assumptions we have made can be applied. An example could be health and ability where health status is indirectly revealed by health expenditures while ability is revealed by earnings. Such a model could possibly be used to analyze how individual health care expenditures should be refunded by the government as a function of earnings.

## A Appendix

### A.1 Proof of Proposition 1: Optimal Tax Formulas in the Binary Model

The government maximizes

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

subject to the budget constraint

$$\begin{aligned} & \int_{\underline{n}}^{\bar{n}} \int_0^{V_1(n)-V_0(n)} [z_1(n) + w - nh(z_1(n)/n) - V_1(n)]p(q|n)f(n)dqdn + \\ & \int_{\underline{n}}^{\bar{n}} \int_{V_1(n)-V_0(n)}^{\infty} [z_0(n) - nh(z_0(n)/n) - V_0(n)]p(q|n)f(n)dqdn \geq 0, \end{aligned}$$

and the constraints arising from the couples utility maximization:  $\dot{V}_l(n) = -h(z_l(n)/n) + (z_l(n)/n)h'(z_l(n)/n)$  for  $l = 0, 1$ . Let us denote by  $\lambda$ ,  $\mu_0(n)$ ,  $\mu_1(n)$ , the multipliers associated. The transversality conditions are  $\mu_0(\underline{n}) = \mu_1(\underline{n}) = \mu_0(\bar{n}) = \mu_1(\bar{n}) = 0$ . We abbreviate  $h(z_1(n)/n)$  into  $h_1$ , etc.

The first order conditions with respect to  $z_0(n)$  and  $z_1(n)$  are

$$\begin{aligned} \mu_0 \cdot \frac{z_0}{n^2} h_0'' + \lambda \cdot (1 - h_0') \cdot (1 - P(\bar{q}|n)) \cdot f(n) &= 0, \\ \mu_1 \cdot \frac{z_1}{n^2} h_1'' + \lambda \cdot (1 - h_1') \cdot P(\bar{q}|n) \cdot f(n) &= 0. \end{aligned}$$

The first order conditions with respect to  $V_0(n)$  and  $V_1(n)$  are

$$\begin{aligned} -\dot{\mu}_0 &= \int_{V_1-V_0}^{\infty} \Psi'(V_0(n))p(q|n)f(n)dq - \lambda(1 - P(\bar{q}|n))f(n) - \lambda[T_1 - T_0]p(\bar{q}|n)f(n), \\ -\dot{\mu}_1 &= \int_0^{V_1-V_0} \Psi'(V_1(n) - q)p(q|n)f(n)dq - \lambda P(\bar{q}|n)f(n) + \lambda[T_1 - T_0]p(\bar{q}|n)f(n). \end{aligned}$$

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:

$$\begin{aligned} -\frac{\mu_0(n)}{\lambda} &= \int_n^{\bar{n}} \{ [1 - g_0(n')] (1 - P(\bar{q}|n')) f(n') + [T_1 - T_0] p(\bar{q}|n') f(n') \} dn', \\ -\frac{\mu_1(n)}{\lambda} &= \int_n^{\bar{n}} \{ [1 - g_1(n'_p)] P(\bar{q}|n') f(n') - [T_1 - T_0] p(\bar{q}|n') f(n') \} dn'. \end{aligned}$$

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 (10) and (11) in Proposition 1.

The transversality conditions imply that  $T'_1 = T'_0 = 0$  at the end points  $\underline{n}$  and  $\bar{n}$ .

Kleven et al. 2006 show that  $z_0$  and  $z_1$  weakly increasing in  $n$  is a necessary and sufficient condition for implementability (exactly as in the one dimensional Mirrlees model). If (10) and (11) generate decreasing ranges for  $z_0$  or  $z_1$  then there is bunching and the formula do not apply on the bunching portions.

## A.2 Proof of Proposition 4: Optimal Tax Formulas in the Continuous Model

We start by forming the integrated Hamiltonian:

$$\begin{aligned} H &= \int \int_D \Psi(V(n_p, n_s)) f(n_p, n_s) dn_p dn_s \\ &+ \int \int_D \lambda [z_p + z_s - n_p h_p(z_p/n_p) - n_s h_s(z_s/n_s) - V] f(n_p, n_s) dn_p dn_s \\ &+ \int \int_D [\partial_p V + h_p - (z_p/n_p) h'_p] \mu_p(n_p, n_s) dn_p dn_s \\ &+ \int \int_D [\partial_s V + h_s - (z_s/n_s) h'_s] \mu_s(n_p, n_s) dn_p dn_s, \end{aligned}$$

where  $\lambda$  is the scalar budget constraint multiplier and  $\mu_p$  and  $\mu_s$  are scalar functions of  $(n_p, n_s)$ . To simplify the problem, it is useful to use the divergence theorem from multi-variable calculus (as in Mirrlees, 1976)

$$\int \int_D (\partial_p V \mu_p + \partial_s V \mu_s) dn_p dn_s + \int \int_D V \left( \frac{\partial \mu_p}{\partial n_p} + \frac{\partial \mu_s}{\partial n_s} \right) dn_p dn_s = \int_{\partial D} V(\mu \cdot ds),$$

where  $\mu = (\mu_p, \mu_s)$  and  $ds$  denotes the normal outward vector along  $\partial D$ , the boundary of  $D$ . Using the above expression, we may rewrite the Hamiltonian to

$$\begin{aligned} H &= \int \int_D \Psi(V(n_p, n_s)) f(n_p, n_s) dn_p dn_s \\ &+ \int \int_D \lambda [z_p + z_s - n_p h_p(z_p/n_p) - n_s h_s(z_s/n_s) - V] f(n_p, n_s) dn_p dn_s \\ &+ \int \int_D [h_p - (z_p/n_p) h'_p] \mu_p(n_p, n_s) dn_p dn_s \\ &+ \int \int_D [h_s - (z_s/n_s) h'_s] \mu_s(n_p, n_s) dn_p dn_s \\ &- \int \int_D V \left( \frac{\partial \mu_p}{\partial n_p} + \frac{\partial \mu_s}{\partial n_s} \right) dn_p dn_s + \int_{\partial D} V(\mu \cdot ds). \end{aligned}$$

The transversality condition is that  $\mu \cdot ds = 0$  on the boundary  $\partial D$ . In words, the scalar product of the normal vector  $ds$  to the boundary of  $D$  and  $\mu$  must be zero at all points along the boundary  $\partial D$ . If  $D = [\underline{n}_p, \bar{n}_p] \times [\underline{n}_s, \bar{n}_s]$ , then  $\mu_p = 0$  for  $n_p = \underline{n}_p, \bar{n}_p$  and  $\mu_s = 0$  for  $n_s = \underline{n}_s, \bar{n}_s$ .

The first-order conditions in  $z_p$  and  $z_s$  are:

$$\frac{\partial H}{\partial z_p} = \lambda [1 - h'_p(z_p/n_p)] f(n_p, n_s) - \frac{z_p}{n_p} h''_p(z_p/n_p) \cdot \frac{\mu_p}{n_p} = 0, \quad (33)$$

$$\frac{\partial H}{\partial z_s} = \lambda [1 - h'_s(z_s/n_s)] f(n_p, n_s) - \frac{z_s}{n_s} h''_s(z_s/n_s) \cdot \frac{\mu_s}{n_s} = 0. \quad (34)$$

After routine rewriting and introducing the elasticity of earnings with respect to  $1 - T'_p$ , denoted by  $\varepsilon_p$ , for the primary earner, the first-order condition in  $z_p$  at  $(n_p, n_s)$  becomes

$$\frac{T'_p}{1 - T'_p} = \frac{1}{\varepsilon_p} \cdot \frac{1}{n_p f(n_p, n_s)} \cdot \frac{\mu_p}{\lambda}. \quad (35)$$

Similarly, the first-order condition in  $z_p$  at  $(n_p, n_s)$  is

$$\frac{T'_s}{1 - T'_s} = \frac{1}{\varepsilon_s} \cdot \frac{1}{n_s f(n_p, n_s)} \cdot \frac{\mu_s}{\lambda}. \quad (36)$$

The first-order condition in  $V$  at  $(n_p, n_s)$  gives the divergence equation

$$\frac{\partial \mu_p}{\partial n_p} + \frac{\partial \mu_s}{\partial n_s} = [\Psi'(\cdot) - \lambda] f(n_p, n_s). \quad (37)$$

By defining  $t_p = \mu_p/(\lambda \cdot f_s)$  and  $t_s = \mu_s/(\lambda \cdot f_p)$  and  $g(n_p, n_s) = \Psi'(\cdot)/\lambda$ , we rewrite the first-order conditions above so as to obtain the conditions (16), (17), and (18) in Proposition 4.

### A.3 Establishing that $x \rightarrow (1 - h'_j(x))/(xh''_j(x))$ is decreasing at any $x = z_j(n)/n$

The second order condition in  $z_j$  of the government maximization problem is:

$$\frac{\partial^2 H}{(\partial z_j)^2} = -\lambda \frac{1}{n_j} h''_j(z_j/n_j) f(n) - \frac{1}{n_j} h''_j(z_j/n_p) \cdot \frac{\mu_j}{n_j} - \frac{z_j}{n_j^2} h'''_j(z_j/n_j) \cdot \frac{\mu_j}{n_j} \leq 0$$

Using the first-order conditions (33) and (34) to substitute for  $\mu_j$ ,  $j = p, s$ , we have:

$$-h''_j(x_j) x_j - (1 - h'_j(x_j)) \left( 1 + \frac{h'''_j(x_j)}{h''_j(x_j)} x_j \right) \leq 0$$

where  $x_j = z_j/n_j$  and  $j = p, s$ . This inequality is equivalent to the derivative of  $x \rightarrow (1 - h'_j(x))/(xh''_j(x))$  being negative.  $\square$

Note that if  $x \rightarrow (1 - h'_j(x))/(xh''_j(x))$  is increasing in some ranges, then at the optimum,  $z_j/n_j$  cannot fall in those ranges. Mirrlees (1971) shows that assuming that  $x \rightarrow (1 - h'(x))/(xh''(x))$  is decreasing ensures that the optimum solution of the one dimensional problem is such that  $z(n)$  is *continuous* in  $n$ .

#### A.4 Proof of $T''_{ps} \leq 0$ in Theorem 1

**Lemma 7** *The second-order conditions of the household imply:  $\partial z_p / \partial n_p > 0$ ,  $\partial z_s / \partial n_s > 0$ ,  $\text{sign}(\partial T'_p / \partial n_s) = \text{sign}(\partial T'_s / \partial n_p) = \text{sign}(T''_{ps})$ .*

**Proof:**

The first-order conditions of the household equal  $1 - T'_p - h'_p(z_p/n_p) = 0$  and  $1 - T'_s - h'_s(z_s/n_s) = 0$ . Hence the second order conditions take the form:

$$T''_{pp} + \frac{1}{n_p} h''_p(z_p/n_p) > 0, \quad (38)$$

$$T''_{ss} + \frac{1}{n_s} h''_s(z_s/n_s) > 0, \quad (39)$$

$$[T''_{pp} + \frac{1}{n_p} h''_p(z_p/n_p)] \cdot [T''_{ss} + \frac{1}{n_s} h''_s(z_s/n_s)] - (T''_{ps})^2 > 0. \quad (40)$$

Total differentiation of the first-order conditions with respect to  $n_p$  gives

$$-\left[ T''_{pp} + \frac{1}{n_p} h''_p(z_p/n_p) \right] \frac{\partial z_p}{\partial n_p} - T''_{ps} \frac{\partial z_s}{\partial n_p} + \frac{z_p}{n_p^2} h''_p(z_p/n_p) = 0 \quad (41)$$

$$-T''_{ps} \frac{\partial z_p}{\partial n_p} - T''_{ss} \frac{\partial z_s}{\partial n_p} - \frac{1}{n_s} h''_s(z_s/n_s) \frac{\partial z_s}{\partial n_p} = 0 \quad (42)$$

The last equation implies

$$\frac{\partial z_s}{\partial n_p} = -\frac{T''_{ps}}{T''_{ss} + h''_s(z_s/n_s)/n_p} \cdot \frac{\partial z_p}{\partial n_p}. \quad (43)$$

Insert this in (41):

$$\frac{\partial z_p}{\partial n_p} = \frac{\frac{z_p}{n_p^2} h''_p(z_p/n_p)}{T''_{pp} + \frac{1}{n_p} h''_p(z_p/n_p) - (T''_{ps})^2 / (T''_{ss} + \frac{1}{n_s} h''_s(z_s/n_s))} > 0,$$

where the inequality follows from the second-order condition (40). It now follows from (39) and (43) that  $\text{sign}(\partial z_s / \partial n_p) = -\text{sign}(T''_{ps})$ . From  $1 - T'_s = h'_s(z_s/n_s)$ , we have  $\text{sign}(\partial z_s / \partial n_p) = -\text{sign}(\partial T'_s / \partial n_p)$ . The symmetric equation (inverting  $s$  and  $p$ ) follows in the same way.  $\square$

#### A.5 Proof of Lemma 5

$V(\gamma)$  is the unique solution of the (strictly) convex minimization problem  $\min_{V \in K} \Phi(V, \gamma)$ . Therefore the first order conditions of this convex problem are necessary and sufficient to characterize the solution  $V(\gamma)$ . The first order conditions of the Hamiltonian problem is the following nonlinear Elliptic Partial Differential Equation:

$$-\varepsilon \nabla \cdot \begin{pmatrix} n_p \left( [(1+\varepsilon)\partial_p V]^{-\frac{1}{1+\varepsilon}} - 1 \right) f(n) \\ n_s \left( [(1+\varepsilon)\partial_s V]^{-\frac{1}{1+\varepsilon}} - 1 \right) f(n) \end{pmatrix} + \left( \frac{V^{-\gamma}}{\lambda} - 1 \right) f(n) = 0, \quad (44)$$

with Neumann boundary conditions on  $\partial D$ :

$$\begin{pmatrix} n_p \left( [(1+\varepsilon)\partial_p V]^{-\frac{1}{1+\varepsilon}} - 1 \right) f(n) \\ n_s \left( [(1+\varepsilon)\partial_s V]^{-\frac{1}{1+\varepsilon}} - 1 \right) f(n) \end{pmatrix} \cdot \mathbf{u} = 0, \quad (45)$$

where  $\nabla \cdot$  denotes the divergence operator and  $\mathbf{u}$  is the normal unit vector on  $\partial D$ .  $\lambda$  is the Lagrange multiplier associated to the government budget constraint  $B(V) \geq 0$  and is such that  $\lambda = \int_D V^{-\gamma} f(n) dn$ . Thus, the budget constraint binds at the optimum:

$$\int_D \left[ \sum_{i=p,s} n_i \left( (1+\varepsilon)\partial_i V \right)^{\frac{\varepsilon}{1+\varepsilon}} - \varepsilon \sum_{i=p,s} n_i \partial_i V - V \right] f(n) dn = 0. \quad (46)$$

When  $\gamma = 0$ ,  $V = (n_p + n_s)/(1+\varepsilon)$  and  $\lambda = 1$  is the trivial laissez-faire solution that satisfies (44), (45), and (46)

Abstractly, the PDE (44), (45), and (46) can be written as:

$$\Xi(V(\gamma), \gamma) = 0, \quad (47)$$

where  $\Xi$  is a functional defined for  $V \in K$  and parameter  $\gamma$ . If  $D_V \Xi(V(0), 0)$  exists and is an invertible linear operator, then we can apply the implicit function theorem and obtain that  $\gamma \rightarrow V(\gamma)$  is differentiable at  $\gamma = 0$  with a derivative  $U = D_\gamma V \in H^1(D)$  which satisfies:

$$D_V \Xi(V(0), 0)U + D_\gamma \Xi(V(0), 0) = 0. \quad (48)$$

Note that this equation corresponds exactly to differentiating (47) with respect to  $\gamma$  (at  $\gamma = 0$ ) and applying the standard chain-rule for differentiation. Using standard differentiation rules, equation (48) can be written as the *linear* elliptic PDE:

$$\varepsilon \nabla \cdot \begin{pmatrix} n_p f(n) (\partial_p U) \\ n_s f(n) (\partial_s U) \end{pmatrix} = \left( \log V(0) - \int_D \log V(0) f(n) dn \right) f(n), \quad (49)$$

with Neumann boundary conditions on  $\partial D$ :

$$\begin{pmatrix} n_p f(n) (\partial_p U) \\ n_s f(n) (\partial_s U) \end{pmatrix} \cdot \mathbf{u} = 0, \quad (50)$$

and the linearized budget constraint:

$$\int_D U f(n) dn = 0. \quad (51)$$

The linear elliptic PDE problem (49) and (50) is a standard problem of the form  $\nabla \cdot (P(n)\nabla U) = K(n)$ , with  $P(n)$  diagonal two-by-two matrix with diagonal coefficients  $n_p f(n)$  and  $n_s f(n)$  bounded away from zero on  $D$ . The matrix  $P(n)$  is therefore coercive<sup>33</sup> and is smooth on  $\bar{D}$ . Furthermore, the boundary condition can be written as  $P(n)\nabla U \cdot \mathbf{u} = 0$  which is the co-normal derivative of the elliptic and coercive operator  $\nabla \cdot (P(n)\nabla)$ . Finally, the problem satisfies the integrability condition  $\int_D K(n) dn = 0$ .

Therefore, the problem (49) and (50) has a unique solution (up to constant) in  $H^1(D)$  (see e.g., Brezis, 1983). The linearized budget constraint (51) pins down the constant so that  $U$  is unique. Finally, because  $\partial D$  and  $f(n)$  are smooth, the solution  $U$  is actually smooth (at least of class  $C^2$ ) (see again Brezis, 1983) on  $\bar{D}$ .

More generally, in order to demonstrate that the operator  $D_V \Xi(V(0), 0)$  is invertible, we consider the general equation:

$$D_V \Xi(V(0), 0)U = \Theta, \quad (52)$$

which is the Elliptic PDE problem  $\varepsilon \nabla \cdot (P(n)\nabla U) = \Theta_1(n)$  on  $D$ , with Neumann boundary condition  $P(n)\nabla U \cdot \mathbf{u} = \Theta_2(n)$  on  $\partial D$ , and the linearized budget constraint  $\int_D U f(n) dn = \Theta_3$ . This generalized problem also has a unique solution in  $H^1(D)$  as long as the integrability condition  $\int_D \Theta_1(n) dn = \varepsilon \int_{\partial D} \Theta_2(n) ds$  is satisfied. Therefore,  $D_V \Xi(V(0), 0)$  is an invertible operator. Thus, we can apply the implicit function theorem and the lemma is demonstrated.  $\square$

## A.6 Proof of Lemma 6

We have  $u(c, z, n) = c - n_p (z_p/n_p)^{1+k} / (1+k) - n_s (z_s/n_s)^{1+k} / (1+k)$  where  $k = 1/\varepsilon$ . By definition of  $z(n)$ ,  $\partial_i V(n) = (k/(k+1))(z_i(n)/n_i)^{1+k}$  for  $i = p, s$ . Hence,  $\nabla V(n) = \partial_n u(c(n), z(n), n)$  where  $\partial_n u$  denotes the partial derivative of  $u(c, z, n)$  with respect to  $n$  (keeping  $c$  and  $z$  constant). Therefore,

$$\nabla V(m) = \partial_n u(c(n'), z(n'), m) + \int_0^1 A(n' + s \cdot (m - n'), m)(m - n') ds, \quad (53)$$

---

<sup>33</sup>There is a constant  $c > 0$  such that, for any vector  $v \in \mathbb{R}^2$ , and any  $n \in D$ , we have:  $vP(n)v \geq c|v|^2$ .



where  $A(r, m)$  is the two-by-two matrix derivative of  $r \rightarrow \partial_n u(c(r), z(r), m)$ . Using  $V(n) - V(n') = \int_0^1 \nabla V(n' + t(n - n')) \cdot (n - n') dt$ , integrating (53) from  $m = n'$  to  $m = n$  implies:

$$V(n) - V(n') = u(c(n'), z(n'), n) - V(n') + \int_0^1 \int_0^1 A(n' + st \cdot (n - n'), n' + t(n - n')) t(n - n')(n - n') ds dt.$$

Thus, the lemma is established if we can show that  $A(n, m)$  is a non-negative matrix for any  $n, m \in D$ .  $\partial_n u(c(n), z(n), m) = ((k/(k+1))(z_p(n)/m_p)^{1+k} \quad (k/(k+1))(z_s(n)/m_s)^{1+k})$ , hence:

$$A(n, m) = k \begin{pmatrix} \frac{z_p(n)^k}{m_p^{1+k}} \cdot \frac{\partial z_p(n)}{\partial n_p} & \frac{z_p(n)^k}{m_p^{1+k}} \cdot \frac{\partial z_p(n)}{\partial n_s} \\ \frac{z_s(n)^k}{m_s^{1+k}} \cdot \frac{\partial z_s(n)}{\partial n_p} & \frac{z_s(n)^k}{m_s^{1+k}} \cdot \frac{\partial z_s(n)}{\partial n_s} \end{pmatrix} \quad (54)$$

A sufficient condition for a two-by-two matrix  $(a_{ij})$  to be non-negative is that  $a_{11} \geq 0, a_{22} \geq 0$  and  $a_{11}a_{22} \geq (a_{12} + a_{21})^2/4$ . In the case of the matrix  $A(n, m)$  in (54), the first two conditions can be written as  $\partial z_p/\partial n_p \geq 0, \partial z_s/\partial n_s \geq 0$ . We have  $\partial_i V = (k/(k+1))(z_i(n)/n_i)^{1+k}$  for  $i = p, s$ . Using equality of the cross-partial derivatives of the function  $V(n)$ , we have  $\partial_{ps}^2 V = k \cdot [z_p(n)^k/(n_p^{1+k})] \partial z_p(n)/\partial n_s = \partial_{sp}^2 V = k \cdot [z_s(n)^k/(n_s^{1+k})] \partial z_s(n)/\partial n_p$ . Using this expression, we can rewrite condition  $a_{11}a_{22} \geq (a_{12} + a_{21})^2/4$  for matrix  $A(n, m)$  as:

$$\frac{\partial z_p(n)}{\partial n_p} \cdot \frac{\partial z_s(n)}{\partial n_s} \geq \left\{ \frac{1}{4} \left[ 1 + \left( \frac{n_p m_s}{n_s m_p} \right)^{1+k} \right] \cdot \left[ 1 + \left( \frac{n_s m_p}{n_p m_s} \right)^{1+k} \right] \right\} \cdot \frac{\partial z_p(n)}{\partial n_s} \cdot \frac{\partial z_s(n)}{\partial n_p}.$$

Because  $D$  is bounded and bounded away from zero, we can define  $K$  as the finite upper bound over  $n, m \in \bar{D}$  of the expression in curly brackets on the right-hand-side above. In that case, under the conditions of the lemma, we have, for any  $n, m \in D$ ,  $A(n, m) \geq 0$  and hence global maximization is established.

## A.7 Numerical Simulations

Simulations are performed with MATLAB software and our programs are available upon request. We select a grid for  $n$ , from  $\underline{n} = 1$  to  $\bar{n} = 4$  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:

$$\begin{aligned} & \int_0^{V_1 - V_0} \Psi'(V_1 - q) p(q) dq = \int_0^{V_1 - V_0} \frac{1}{(V_1 - q)^\gamma} \frac{\eta \cdot q^{\eta-1}}{q_{max}^\eta} dq \\ & = \frac{\eta}{q_{max}^\eta} \int_0^{V_1 - V_0} (V_1 - q)^{-\gamma} q^{\eta-1} dq = \frac{\eta \cdot V_1^{\eta-\gamma}}{q_{max}^\eta} \int_0^{1 - \frac{V_0}{V_1}} t^{\eta-1} (1-t)^{-\gamma} dt = \frac{\eta \cdot V_1^{\eta-\gamma}}{q_{max}^\eta} \cdot \beta\left(1 - \frac{V_0}{V_1}, \eta, 1-\gamma\right) \end{aligned}$$

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

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

We pick  $q_{max} = 2 \cdot w^{1+1/\eta}$  so that the fraction of spouses working is normalized in the situation with no taxes (when  $w$  or  $\eta$  change). We set  $w = 1$  in the simulations presented so that  $q_{max} = 2$ .

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. This is done with a sub-iterative routine that adapts  $T_0$  and  $T_1$  as the bottom  $\underline{n}$  until those conditions are satisfied. We then use the first order conditions (10), (11) 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.

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 non-decreasing so that the fixed point is implementable.

## References

- Alesina, Alberto and Andrea Ichino** (2007). "Gender based taxation." Harvard University Working Paper, April.
- Alm, James and Leslie A. Whittington** (1999). "For Love or Money? The Impact of Income Taxes on Marriage." *Economica* 66, 297-316.
- Armstrong, Mark** (1996). "Multiproduct Nonlinear Pricing." *Econometrica* 64, 51-75.
- Armstrong, Mark** (1999). "Price Discrimination by a Many-Product Firm." *Review of Economic Studies* 66, 151-168.
- Armstrong, Mark** (2006). "Recent Developments in the Economics of Price Discrimination", chapter 4 in *Advances in Economics and Econometrics: Theory and Applications* Ninth World Congress of the Econometric Society Volume II, Eds. Blundell, Newey, and Persson, Cambridge University Press.
- Armstrong, Mark and Rochet, Jean-Charles** (1999). "Multi-dimensional Screening: A User's Guide." *European Economic Review* 43, 959-79.
- Armstrong, Mark and John Vickers** (2001) "Competitive Price Discrimination", *RAND Journal of Economics* 32(4), 579-605.
- Armstrong, Mark and John Vickers** (2006) "Competitive Nonlinear Pricing and Bundling", Working Paper University College London.
- Basov, Suren** (2005). *Multi-Dimensional Screening*. Springer: Berlin.
- Blundell, Richard W. and Thomas MaCurdy** (1999). "Labor Supply: A Review of Alternative Approaches," in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics* vol. 3A. Elsevier Science B.V.: Amsterdam.
- Boskin, Michael and Eytan Sheshinski** (1983). "Optimal Tax Treatment of the Family: Married Couples." *Journal of Public Economics* 20, 281-297.
- Brezis, Haim** (1983). *Analyse Fonctionnelle: Théorie et Applications*. Masson: Paris.
- Brett, Craig** (2006), "Optimal Nonlinear Taxes for Families.", *International Tax and Public Finance* 14(3), 225-261.
- Chiappori, Pierre-André** (1988). "Rational Household Labor Supply." *Econometrica*, 56(1), 63-90.

- Chiappori, Pierre-André** (1992). “Collective labour supply and welfare.” *Journal of Political Economy* 100, 437-67.
- Cremer, Helmuth, Jean-Marie Lozachmeur, and Pierre Pestieau** (2007) “Income Taxation of Couples and the Tax Unit Choice”, Working Paper.
- Cremer, Helmuth, Pestieau, Pierre, Rochet, Jean-Charles** (2001). “Direct versus indirect taxation: The design of the tax structure revisited.” *International Economic Review* 42, 781-799.
- Deaton, Angus** (1992). *Understanding Consumption*. Oxford University Press: Oxford.
- Diamond, Peter** (1998). “Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates.” *American Economic Review* 88, 83-95.
- Eissa, Nada and Hillary Hoynes** (2000). “Tax and Transfer Policy, and Family Formation: Marriage and Cohabitation.” Working paper, December 2000.
- Eissa, Nada and Hillary Hoynes** (2004). “Taxes and the Labor Market Participation of Married Couples: The Earned Income Tax Credit.” *Journal of Public Economics* 88, 1931-1958.
- Grisvard, Pierre** (1985). *Elliptic Problems in Nonsmooth Domains*. Pitman: Boston.
- Kleven, Henrik J., Claus Kreiner, and Emmanuel Saez** (2006). “The Optimal Income Taxation of Couples”, NBER Working Paper No. 12685.
- Kroft, Kory** (2006). “A Note on Intra-household Allocation and Optimal Income Transfers”, UC Berkeley unpublished Working Paper.
- Lundberg, Shelly J., and Robert A. Pollak** (1996). “Bargaining and Distribution in Marriage.” *Journal of Economic Perspectives*, 10(4), 139–158.
- Lundberg, Shelly J., Robert A. Pollak, and Terence J. Wales** (1997). “Do Husbands and Wives Pool Their Resources: Evidence from the United Kingdom Child Tax Credit.” *Journal of Human Resources*, 32(3), 463–480.
- McAfee, R. Preston, and McMillan, John** (1988). “Multidimensional Incentive Compatibility and Mechanism Design.” *Journal of Economic Theory* 46, 335-54.
- Mirrlees, James A.** (1971). “An Exploration in the Theory of Optimal Income Taxation.” *Review of Economic studies* 38, 175-208.
- Mirrlees, James A.** (1976). “Optimal tax theory: a synthesis.” *Journal of Public Economics*

6, 327-358.

**Mirrlees, James A.** (1986). "The Theory of Optimal Taxation," in K.J. Arrow and M.D. Intriligator (eds.), *Handbook of Mathematical Economics* vol. 3. Elsevier Science B.V.: Amsterdam.

**Pechman, Joseph A.** (1987). *Federal Tax Policy*. Brookings Institution: Washington D.C.

**Piggott, John and John Whalley** (1996). "The Tax Unit and Household Production." *Journal of Political Economy* 104(2), 398-418.

**Rochet, Jean-Charles** (1987). "A Necessary and Sufficient Condition for Rationalizability in a Quasi-linear Context," *Journal of Mathematical Economics* 16(2), 191-200.

**Rochet, Jean-Charles and Choné, Philippe** (1998). "Ironing, Sweeping, and Multi-dimensional Screening", *Econometrica* 66, 783-826.

**Rochet, Jean-Charles and Stole, Lars** (2003). "The Economics of Multidimensional Screening," in *Advances in Economics and Econometrics: Theory and Applications*, Eighth World Congress.

**Rosen, Harvey** (1977). "Is It Time to Abandon Joint Filing?" *National Tax Journal* 30, 423-428.

**Saez, Emmanuel** (2001). "Using Elasticities to Derive Optimal Income Tax Rates." *Review of Economic Studies* 68, 205-229.

**Saez, Emmanuel** (2002). "Optimal Income Transfer Programs: Intensive Versus Extensive Labor Supply Responses." *Quarterly Journal of Economics* 117, 1039-1073.

**Schroyen, Fred** (2003). "Redistributive taxation and the household: the case of individual filings." *Journal of Public Economics* 87, 2527-2547.

**Wilson, Robert B.** (1993). *Nonlinear Pricing*. Oxford University Press: Oxford.

# CESifo Working Paper Series

for full list see [www.cesifo-group.org/wp](http://www.cesifo-group.org/wp)

(address: Poschingerstr. 5, 81679 Munich, Germany, [office@cesifo.de](mailto:office@cesifo.de))

---

- 2028 Jarko Fidrmuc, Neil Foster and Johann Scharler, Labour Market Rigidities, Financial Integration and International Risk Sharing in the OECD, June 2007
- 2029 Bernardina Algieri and Thierry Bracke, Patterns of Current Account Adjustment – Insights from Past Experience, June 2007
- 2030 Robert Dur and Hein Roelfsema, Social Exchange and Common Agency in Organizations, June 2007
- 2031 Alexander Libman and Lars P. Feld, Strategic Tax Collection and Fiscal Decentralisation: The Case of Russia, June 2007
- 2032 Øystein Foros, Hans Jarle Kind and Greg Shaffer, Resale Price Maintenance and Restrictions on Dominant Firm and Industry-Wide Adoption, June 2007
- 2033 Jan K. Brueckner and Kurt Van Dender, Atomistic Congestion Tolls at Concentrated Airports? Seeking a Unified View in the Internalization Debate, June 2007
- 2034 Viet Do and Ngo Van Long, International Outsourcing under Monopolistic Competition: Winners and Losers, June 2007
- 2035 Nadia Fiorino and Roberto Ricciuti, Determinants of Direct Democracy, June 2007
- 2036 Burkhard Heer and Alfred Maussner, Inflation and Output Dynamics in a Model with Labor Market Search and Capital Accumulation, June 2007
- 2037 Konstantinos Angelopoulos, Jim Malley and Apostolis Philippopoulos, Public Education Expenditure, Growth and Welfare, June 2007
- 2038 Maarten Bosker, Steven Brakman, Harry Garretsen and Marc Schramm, Adding Geography to the New Economic Geography, June 2007
- 2039 Steffen Henzel, Oliver Hülsewig, Eric Mayer and Timo Wollmershäuser, The Price Puzzle Revisited: Can the Cost Channel Explain a Rise in Inflation after a Monetary Policy Shock?, July 2007
- 2040 Rosario Crinò, Service Offshoring and White-Collar Employment, July 2007
- 2041 Carsten Hefeker and Michael Neugart, Labor Market Regulation and the Legal System, July 2007
- 2042 Bart Cockx and Muriel Dejemepe, Is the Notification of Monitoring a Threat to the Unemployed? A Regression Discontinuity Approach, July 2007
- 2043 Alfons J. Weichenrieder, Profit Shifting in the EU: Evidence from Germany, July 2007

- 2044 Annika Alexius and Bertil Holmlund, Monetary Policy and Swedish Unemployment Fluctuations, July 2007
- 2045 Axel Dreher, Jan-Egbert Sturm and Jakob de Haan, Does High Inflation Cause Central Bankers to Lose their Job? Evidence Based on a New Data Set, July 2007
- 2046 Guglielmo Maria Caporale and Luis A. Gil-Alana, Long Run and Cyclical Dynamics in the US Stock Market, July 2007
- 2047 Alessandro Balestrino, It is a Theft but not a Crime, July 2007
- 2048 Daniel Becker and Michael Rauscher, Fiscal Competition in Space and Time: An Endogenous-Growth Approach, July 2007
- 2049 Yannis M. Ioannides, Henry G. Overman, Esteban Rossi-Hansberg and Kurt Schmidheiny, The Effect of Information and Communication Technologies on Urban Structure, July 2007
- 2050 Hans-Werner Sinn, Please Bring me the New York Times – On the European Roots of Richard Abel Musgrave, July 2007
- 2051 Gunther Schnabl and Christian Danne, A Role Model for China? Exchange Rate Flexibility and Monetary Policy in Japan, July 2007
- 2052 Joseph Plasmans, Jorge Fornero and Tomasz Michalak, A Microfounded Sectoral Model for Open Economies, July 2007
- 2053 Vesa Kannianen and Panu Poutvaara, Imperfect Transmission of Tacit Knowledge and other Barriers to Entrepreneurship, July 2007
- 2054 Marko Koethenbueger, Federal Tax-Transfer Policy and Intergovernmental Pre-Commitment, July 2007
- 2055 Hendrik Jürges and Kerstin Schneider, What Can Go Wrong Will Go Wrong: Birthday Effects and Early Tracking in the German School System, July 2007
- 2056 Bahram Pesaran and M. Hashem Pesaran, Modelling Volatilities and Conditional Correlations in Futures Markets with a Multivariate t Distribution, July 2007
- 2057 Walter H. Fisher and Christian Keuschnigg, Pension Reform and Labor Market Incentives, July 2007
- 2058 Martin Altemeyer-Bartscher, Dirk T. G. Rübhelke and Eytan Sheshinski, Policies to Internalize Reciprocal International Spillovers, July 2007
- 2059 Kurt R. Brekke, Astrid L. Grasdal and Tor Helge Holmås, Regulation and Pricing of Pharmaceuticals: Reference Pricing or Price Cap Regulation?, July 2007
- 2060 Tigran Poghosyan and Jakob de Haan, Interest Rate Linkages in EMU Countries: A Rolling Threshold Vector Error-Correction Approach, July 2007

- 2061 Robert Dur and Klaas Staal, Local Public Good Provision, Municipal Consolidation, and National Transfers, July 2007
- 2062 Helge Berger and Anika Holler, What Determines Fiscal Policy? Evidence from German States, July 2007
- 2063 Ernesto Reuben and Arno Riedl, Public Goods Provision and Sanctioning in Privileged Groups, July 2007
- 2064 Jan Hanousek, Dana Hajkova and Randall K. Filer, A Rise by Any Other Name? Sensitivity of Growth Regressions to Data Source, July 2007
- 2065 Yin-Wong Cheung and Xing Wang Qian, Hoarding of International Reserves: Mrs Machlup's Wardrobe and the Joneses, July 2007
- 2066 Sheilagh Ogilvie, 'Whatever Is, Is Right'?, Economic Institutions in Pre-Industrial Europe (Tawney Lecture 2006), August 2007
- 2067 Floriana Cerniglia and Laura Pagani, The European Union and the Member States: Which Level of Government Should Do what? An Empirical Analysis of Europeans' Preferences, August 2007
- 2068 Alessandro Balestrino and Cinzia Ciardi, Social Norms, Cognitive Dissonance and the Timing of Marriage, August 2007
- 2069 Massimo Bordignon, Exit and Voice. Yardstick versus Fiscal Competition across Governments, August 2007
- 2070 Emily Blanchard and Gerald Willmann, Political Stasis or Protectionist Rut? Policy Mechanisms for Trade Reform in a Democracy, August 2007
- 2071 Maarten Bosker and Harry Garretsen, Trade Costs, Market Access and Economic Geography: Why the Empirical Specification of Trade Costs Matters, August 2007
- 2072 Marco Runkel and Guttorm Schjelderup, The Choice of Apportionment Factors under Formula Apportionment, August 2007
- 2073 Jay Pil Choi, Tying in Two-Sided Markets with Multi-Homing, August 2007
- 2074 Marcella Nicolini, Institutions and Offshoring Decision, August 2007
- 2075 Rainer Niemann, The Impact of Tax Uncertainty on Irreversible Investment, August 2007
- 2076 Nikitas Konstantinidis, Gradualism and Uncertainty in International Union Formation, August 2007
- 2077 Maria Bas and Ivan Ledezma, Market Access and the Evolution of within Plant Productivity in Chile, August 2007



- 2078 Friedrich Breyer and Stefan Hupfeld, On the Fairness of Early Retirement Provisions, August 2007
- 2079 Scott Alan Carson, Black and White Labor Market Outcomes in the 19<sup>th</sup> Century American South, August 2007
- 2080 Christian Bauer, Paul De Grauwe and Stefan Reitz, Exchange Rates Dynamics in a Target Zone – A Heterogeneous Expectations Approach, August 2007
- 2081 Ana Rute Cardoso, Miguel Portela, Carla Sá and Fernando Alexandre, Demand for Higher Education Programs: The Impact of the Bologna Process, August 2007
- 2082 Christian Hopp and Axel Dreher, Do Differences in Institutional and Legal Environments Explain Cross-Country Variations in IPO Underpricing?, August 2007
- 2083 Hans-Werner Sinn, Pareto Optimality in the Extraction of Fossil Fuels and the Greenhouse Effect: A Note, August 2007
- 2084 Robert Fenge, Maximilian von Ehrlich and Matthias Wrede, Fiscal Competition, Convergence and Agglomeration, August 2007
- 2085 Volker Nitsch, Die Another Day: Duration in German Import Trade, August 2007
- 2086 Kam Ki Tang and Jie Zhang, Morbidity, Mortality, Health Expenditures and Annuitization, August 2007
- 2087 Hans-Werner Sinn, Public Policies against Global Warming, August 2007
- 2088 Arti Grover, International Outsourcing and the Supply Side Productivity Determinants, September 2007
- 2089 M. Alejandra Cattaneo and Stefan C. Wolter, Are the Elderly a Threat to Educational Expenditures?, September 2007
- 2090 Ted Bergstrom, Rod Garratt and Damien Sheehan-Connor, One Chance in a Million: Altruism and the Bone Marrow Registry, September 2007
- 2091 Geraldo Cerqueiro, Hans Degryse and Steven Ongena, Rules versus Discretion in Loan Rate Setting, September 2007
- 2092 Henrik Jacobsen Kleven, Claus Thustrup Kreiner and Emmanuel Saez, The Optimal Income Taxation of Couples as a Multi-Dimensional Screening Problem, September 2007