ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Bohn, Henning; Stuart, Charles

Working Paper

Voting and Nonlinear Taxes in a Stylized Representative Democracy

CESifo Working Paper, No. 1058

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Bohn, Henning; Stuart, Charles (2003) : Voting and Nonlinear Taxes in a Stylized Representative Democracy, CESifo Working Paper, No. 1058, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:
https://hdl.handle.net/10419/76645

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.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

VOTING AND NONLINEAR TAXES IN A STYLIZED REPRESENTATIVE DEMOCRACY

HENNING BOHN
CHARLES STUART

CESIFO WORKING PAPER NO. 1058
CATEGORY 1: PUBLIC FINANCE
OCTOBER 2003

An electronic version of the paper may be downloaded
- *from the SSRN website: www.SSRN.com*
- *from the CESifo website: www.CESifo.de*

CESifo *Working Paper No. 1058*

VOTING AND NONLINEAR TAXES IN A STYLIZED REPRESENTATIVE DEMOCRACY

Abstract

We derive median-voter results and study the shape of redistributional taxes when voters elect a candidate who imposes taxes to maximize own utility. Under general conditions, a median-productivity candidate is a Condorcet winner. The imposed tax function is nonlinear, may place high marginal rates on very low incomes, and may have an interval of negative marginal rates below the income of the winning candidate. Marginal rates are positive throughout, however, if non-redistributional spending or altruism toward the poor are great enough.

JEL Classification: D7, H2.

Henning Bohn
Department of Economics
University of California
Santa Barbara, CA 93106
U.S.A.
bohn@econ.ucsb.edu

Charles Stuart
Department of Economics
University of California
Santa Barbara, CA 93106
U.S.A.

We would like to thank Marcus Berliant, Avinash Dixit, Antonio Rangel, Efraim Sadka, and seminar participants at Stanford University, the Econometric Society Summer Meetings, and CESifo for helpful comments.

INTRODUCTION

A difficulty in understanding how democratic forces shape tax and transfer policies is that such policies are inherently *nonlinear*, which causes cycling in models of direct voting over policy.[1] Voting cycles can sometimes be eliminated by restricting policy to linear tax functions, but this approach obviously precludes study of nonlinearities.[2] To study redistributional tax policies under majority rule without imposing linearity, we assume instead that policy is determined in a stylized representative democracy in which voters elect one of two candidates. The winning candidate is not bound by campaign promises and, once in office, implements his or her own most preferred tax function.[3]

A key assumption is that individuals differ only in productivity. This helps eliminate cycles because elections are between individuals, who are ranked along the single dimension of productivity, and not between policy functions, which are infinite-dimensional. As in Mirrlees (1971), policy is required only to satisfy incentive-compatibility and budget constraints.

Röell (1996) provided a median-voter result for the special case in which utility is quasi-linear (no income effects) and a minimum-utility constraint does not bind (no individual has zero consumption). She stated that "whether [a median-voter result] obtains in a fully general setting with quasi-linear preferences remains an open question" (p.12). We prove a median-voter result under a general utility function, whether or not a minimum-utility constraint binds.

[1] For instance, in an electorate of three voters with different incomes, any redistributional tax policy can be upset by an alternative policy that would effectively take resources from one of the voters and give the resources to the other two.

[2] Examples of the approach are in Romer (1975), Roberts (1979), and Meltzer and Richard (1981).

[3] In an important unpublished paper, Röell (1996) used the same approach to study nonlinear taxation. Osborne and Slivinski (1996) and Besley and Coate (1997) also consider a setting in which an election winner sets policy to maximize own utility.

To understand the result, consider an election between two candidates with different productivities. If elected, each would maximize own utility by choosing a tax function with a low (possibly negative) tax at the own optimal income, and possibly higher taxes at other incomes.[4] Under regularity conditions described below, there is a crossover point between the two candidates' productivities such that individuals with productivity above the crossover vote for the candidate with greater productivity and individuals with productivity below the crossover vote for the candidate with lower productivity. This property implies that the candidate who receives the votes of median-productivity voters wins under majority rule. It follows that a candidate with median productivity is a *Condorcet winner*, meaning a candidate that beats any other candidate with a different productivity.

We also describe the tax function an election winner imposes. To extract taxes from individuals with higher productivities, the tax function is strictly increasing above the winner's income. To extract taxes from individuals with *lower* productivities, the function may impose greater taxes than the winner pays at incomes just below the winner's, that is, marginal rates may be *negative* just below the winner's income. Such an outcome seems counterfactual. We show that the tax function can have positive marginal rates at all incomes if government funds non-redistributional spending or if individuals are altruistic.

An empirical observation is that the phase-out of welfare subsidies often leads to high effective marginal tax rates at the low end of the income distribution.[5] An outcome with relatively high marginal rates at low incomes can easily result from two forces in the model here: the winner wishes to extract taxes from individuals with lower productivities, but cannot squeeze positive taxes from individuals with zero

[4] We use *taxes* to mean *net taxes*, that is total taxes paid to government minus total transfers received from government.

[5] See Browning and Johnson (1979), Dickert et al. (1995), Keane and Moffitt (1998).

income so the tax function's intercept cannot be positive. As a consequence, the tax function may increase steeply just above zero income and then becomes flatter (and possibly slope downward) as income rises toward the winner's income.

The winner's tax function is always nonlinear with a kink at the winner's own income, and is strictly better than a linear tax for the winner. Thus there is an interval of individuals close to the election winner in productivity who also are better off than they would be under a linear tax. This together with the median-voter results provides a sense in which the analysis accords with Director's Law: if one of the candidates has close to median productivity, the resulting policy is better than a linear tax for middle-productivity individuals.[6]

Section I describes the model. Section II uses a numerical example to provide intuition and guide the formal analysis. In section III, we derive the policy a winning candidate would impose, characterizing the candidate's own optimal income level and the resulting utility levels of all individuals in the electorate. In section IV, we study election outcomes and provide median-voter results. In section V, we build on the derivation of the winner's optimal policy to characterize the shape of the tax function. Section VI concludes.

I. MODEL

Tax-paying units ("individuals") differ only in productivity (x); we therefore refer to an individual with productivity x as simply "individual x." Productivity has distribution function $F(x)$ with continuous density f and finite mean over support $[x_-, x_+]$ with $x_+ > x_- \geq 0$; x_+ may be infinite. We assume $f > 0$ over (x_-, x_+).

[6] See Stigler (1970), who argued that the middle classes sometimes have relatively great political clout and also benefit disproportionately from many government spending programs. Recent analyses of spending programs consistent with Director's Law are in Gouveia (1997) and Dixit and Londregan (1998).

Below, "all x" means "all $x \in [x_-, x_+]$."

Each individual has the same strictly increasing, differentiable, weakly concave utility u defined over consumption $c \geq 0$ and leisure $0 \leq l \leq 1$. Each maximizes utility by choosing how much to labor to supply, $n \equiv 1 - l$, which determines gross income, $y \equiv nx$, and hence consumption, $c \equiv y - T$, where T is the positive or negative tax the individual pays. To ensure that labor supply is always strictly less than one, we assume $\lim_{l->0} u_l(c,l)/u_c(c,l) = \infty$ for any $c > 0$, where subscripts denote partial derivatives throughout.

Before making labor-leisure choices, the individuals elect one of two exogenously given candidates by majority rule. We model government policy as set by the election winner, denoted x_e. Because cycling results from direct voting over tax functions, we assume that candidates cannot commit to set specific tax functions if elected. The winning candidate therefore designs policy to maximize own utility.

This stylized representative democracy abstracts from several aspects of political life. Notably, we allow candidates to be two *arbitrary* individuals in $[x_-, x_+]$ to prove median-voter results, but we do not study the politics of how the candidates are selected. Also, we do not study how policy is set by governments drawn from more than one (possibly overlapping) jurisdiction.[7]

Candidates' productivities and hence the policies they would impose are revealed to voters during the electoral process, so each voter knows the utility he/she would receive if either candidate were elected. Each individual votes for the candidate who would provide greater utility.

There are three limits on the power to tax. First and as in Mirrlees (1971), government can observe individuals' incomes but not their productivities; hence it

[7] Given the assumption that individuals differ *only* in productivity, allowing for many jurisdictions should matter little: if candidate productivities are the same across jurisdictions, all election winners would wish to impose the same tax function.

imposes a tax function T^* that specifies the positive or negative net tax paid by an individual as a function of the individual's gross income y. (Different election winners x_e would impose different T^*'s, but we suppress this dependence when studying the taxes imposed by a generic winner.) Second and as in Meltzer and Richard (1981), x_e is subject to the same rules as other individuals and must pay taxes according to T^*. Third and as in Röell, an individual cannot be forced to work in order to pay taxes. Thus x_e's policy can give no individual utility lower than that of a non-worker with zero taxes, or $U(x) \geq u(0,1)$ for all x, where $U(x)$ is the utility individual x obtains under the tax function chosen by x_e.[8] The latter implies that an individual's taxes cannot exceed the individual's income.

A *non-altruistic* election winner with moderate or high productivity may impose a tax function that extracts substantial revenue from low-productivity individuals, leaving them with utility close or equal to $u(0,1)$. Such an outcome is hard to square with observed welfare payments to low-income individuals. To allow for altruism and see how it affects the tax function, we assume the election winner guarantees a minimum consumption level $\alpha \geq 0$ and hence a minimum utility level $u(\alpha,1)$ to all citizens; this altruism is Rawlsian (1971) in that it reflects concern with those who are worst off. Both the possibility of altruism and the requirement that utility be bounded below by $u(0,1)$ are captured by imposing the minimum-utility constraint

$$U(x) \geq u(\alpha,1), \text{ for all } x, \text{ where } u(\alpha,1) > -\infty. \tag{1}$$

This treats α as a preference parameter common to all candidates.[9]

[8] Such a constraint would be superfluous in Mirrlees (1971) and Meltzer-Richard (1981), where the optimal or equilibrium policy is to transfer *to* low-productivity individuals.

[9] If u is unbounded and $\alpha = 0$, then $u(\alpha,1) = -\infty$ is possible, violating the assumption $u(\alpha,1) > -\infty$. In such cases we redefine c to be consumption financed *from taxed income,* and assume that individuals also obtain a fixed positive amount of consumption from home production. These assumptions make $u(0,1)$ finite and the analysis below goes through with the redefined variable c.

The election winner's choice of a tax function is subject to a government budget constraint $\int_{x_-}^{x_+} T^*(Y(x))dF(x) \geq G$, where $G \geq 0$ is an exogenous level of non-redistributional spending and Y is an individual's income when the tax function is T^*. We assume tax functions T^* exist that strictly satisfy the government budget. Specifically, we assume that α and G satisfy $G < \hat{G}(\alpha)$, where $\hat{G}(\alpha)$ is the supremum of levels of G that can be financed given tax functions that satisfy (1) for given α. Because α and G are given parameters, we suppress dependence of the problem on them when no ambiguity arises. In section V, we study how changes in α and G affect the shape of T^*.

Following Mirrlees (1971) and Seade (1982), we make the *agent monotonicity* assumption that $nu_l(c, 1-n)/u_c(c, 1-n)$ is strictly increasing in n for all $c > 0$. This will ensure that an individual with greater productivity never finds it optimal to earn less income than an individual with lower productivity. Sufficient conditions for agent monotonicity are that consumption is normal or that utility is separable.

The tax function T^* chosen by x_e induces profiles of utility $\{U(x)\}_x$ and income $\{Y(x)\}_x$ for all individuals. To characterize T^*, we follow Mirrlees (1971, 1986) by reformulating x_e's utility maximization as a control problem. Specifically, the problem of choosing T^* is equivalent to the problem of choosing a utility profile $\{U(x)\}_x$ and an income profile $\{Y(x)\}_x$ subject to incentive-compatibility constraints requiring that any individual x be induced to work $n(x) = Y(x)/x$ in order to earn the assigned income $Y(x)$.

To express the incentive-compatibility constraints as functions of the profiles $\{U(x)\}_x$ and $\{Y(x)\}_x$, denote the consumption level that provides utility U at work effort n by $c^*(U, n)$; this function is defined by the identity $U = u(c^*, 1-n)$. Incentive-compatibility then requires

$$U(x) \geq u(c^*(U(z), Y(z)/z), 1 - Y(z)/x), \text{ for all } x \text{ and } z \text{ satisfying } x \geq Y(z), \quad (2)$$

which says that income and utility profiles must be such that an individual with true productivity x does not prefer the income-and-tax package of any other individual (z) to his own. The constraints apply only for $x \geq Y(z)$ because individual x cannot earn an income $Y(z) > x$ within the time constraint $n \leq 1$.

From Mirrlees (1971, 1986), the incentive-compatibility constraints (2) are equivalent to the combination of a differential equation

$$U(x) - U(x_-) = \int_{x_-}^{x} \omega(U(z), Y(z), z)dz, \tag{3}$$

$$\text{where } \omega(U, Y, z) \equiv u_l(c^*(U(z), Y(z)/z), 1 - Y(z)/z) \cdot Y(z)/z^2,$$

plus requirements that $Y(x)$ be monotone non-decreasing and $Y(x) < x$.[10] We therefore replace (2) by (3), define the income derivative $\psi(x) \equiv dY/dx$, and ensure monotonicity by imposing[11]

$$\psi(x) \geq 0, \text{ for all } x. \tag{4}$$

Our assumption that $\lim_{l->0} u_l(c, l)/u_c(c, l) = \infty$ for any $c > 0$ ensures $Y(x) < x$.

To express the government budget constraint in terms of the profiles $\{U(x)\}_x$ and $\{Y(x)\}_x$, let

$$T(U, Y, x) \equiv Y - c^*(U, Y/x) \tag{5}$$

denote the revenue that is extracted from an individual with utility U and income Y who has productivity x. The government budget constraint is then

$$\int_{x_-}^{x_+} T(U(x), Y(x), x)dF(x) \geq G. \tag{6}$$

[10] Specifically, theorem 1 in Mirrlees (1971) shows that (2) implies a non-decreasing income profile and $Y(x) < x$. Lemma 6.1 in Mirrlees (1986) proves that (2) implies (3) for a lower bound of 0 instead of x_-; taking the difference $U(x) - U(x_-)$ yields (3) for any $x_- \geq 0$. That (3) implies (2) is a special case of Mirrlees (1986) lemma 6.3, as discussed in Mirrlees (1986, p. 1237).

[11] That is, we follow Brito and Oakland (1977) and Ebert (1992) in using a "second-order" approach. We assume that $\psi(x)$ is piecewise continuous, so $dY(x)/dx$ may be undefined at a finite number of points.

Thus x_e's problem is to maximize $U(x_e)$ subject to (1), (3), (4), and (6) by choice of $\{U(x), Y(x), \psi(x)\}_x$. The solution to x_e's problem is a profile $\{U(x), Y(x), \psi(x)\}_x$.

II. AN EXAMPLE

This section presents a numerical example that illustrates the model's main features and results. The example follows Diamond's (1998) simplification of Mirrlees (1971) in assuming affine utility with a constant elasticity of labor supply: $u(c, l) = c + v_0\{1 - (1 - l)^{1+v_1}\}/(1 + v_1)$, where v_0 is a constant and v_1 is the wage elasticity of labor supply. In the example, we take $v_0 = 5$, $v_1 = 1$, and assume that $F(x)$ is uniform on $[x_-, x_+] = [.2, 1.8]$.[12] Denote the median productivity x_M; the assumed distribution has $x_M = 1$. We assume initially that $G = 0$ and $\alpha = 0$.

Consider an election between x_M and an alternative candidate $x_A = .84$, at the 40th percentile. Figure 1 shows the tax functions x_M and x_A would set if elected. Because $G = 0$ and the government budget binds, some individuals gain net income ($T < 0$) and others lose net income ($T > 0$) from either function.

As in the figure, either candidate if elected would impose a negative tax on him/herself and would transfer resources from others by imposing a V-shaped tax function with a minimum at the candidate's own utility-maximizing income, generically denoted y_e. The steepness of the V balances two forces. A steeper V tends to extract greater resources from others by raising the taxes they pay, decreasing the tax the winner pays. But because the winner's own income level is tax-favored and taxes are based on income and not productivity, other individuals have an incentive to change their labor supplies to earn the tax-favored income y_e. Thus a steeper V also expands the set of individuals who earn y_e, raising the tax the winner pays.

An individual's incentive to earn the tax-favored income y_e decreases with

[12] Although affine utility violates the assumption $\lim_{l->0} u_l(c, l)/u_c(c, l) = \infty$, the parameter values we choose ensure that leisure is bounded away from zero so $Y(x) < x$.

the difference between the individual's productivity and x_e. Specifically, there is an interval around x_e, generically denoted $[x_1, x_2]$ below, such that x_e's optimal policy induces all individuals in $[x_1, x_2]$ to earn y_e. This bunching of incomes gives a mass-point in the income distribution at y_e and causes the profile of taxes set by any candidate to be flat on $[x_1, x_2]$. The $[x_1, x_2]$-intervals for x_M and x_A are illustrated in figure 2; interval endpoints are slightly higher for x_M than for x_A.

The minimum-utility constraint (1) limits government's ability to extract taxes from individuals with low market incomes. This explains why the two tax functions rise from the origin, peak, and then fall at incomes below y_e.

Figure 3 shows the utility profiles induced by x_M and x_A. The profiles cross at $x \cong 0.926$: individuals $x \gtrsim .926$ receive greater utility if x_M is elected; and individuals $x \lesssim .926$ receive greater utility if x_A is elected, except that individuals with minimum productivity $x = .2$ receive utility $u(0, 1)$ from both candidates. Thus all individuals with greater-than-median productivity (half the population) plus some individuals with lower-than-median productivity would vote for x_M, so x_M would win against x_A. Analogously, x_M would win against any other candidate.

The tax functions in figure 1 are notably nonlinear. For instance, x_M's function gives x_M a large negative tax and imposes positive taxes on lower-productivity individuals. This would not happen if taxes were constrained to be linear (figure 4), because x_M would then have to give large negative taxes to low-productivity individuals in order to obtain a small negative tax him/herself.[13]

Figure 5 illustrates how government spending G affects x_M's tax function. If $G = 0$, individuals with incomes near $Y(x_M)$ receive negative taxes while individuals with income near $Y(x_-)$ pay taxes, giving the tax function a negatively sloped segment below the winner's income. As G rises from zero toward $\hat{G}$ (about 0.0772 in

[13] If taxes are constrained to be linear, our model yields the same tax function as Meltzer and Richard (1981), so figure 4 provides a comparison of our model and Meltzer-Richard's.

the example), the election winner must increase taxes on middle and upper income individuals because the minimum utility constraint rules out increased taxes on individuals with low productivity. This tilts the V-shaped spike so the negatively sloped segment vanishes. As long as $G < \hat{G}$, however, the equilibrium tax function always has a kink at y_e. Thus taxes are always nonlinear but may increase monotonically with income.

Figure 6 illustrates how the altruistic consumption floor α affects the tax function. An increase in α has two effects. First, it lowers the intercept of the function (which is $-\alpha$), pulling the function down at low incomes. Second, it raises the government's need for revenue, pulling the function up at higher incomes. Both effects tend to make the tax function slope upward and may make taxes increase monotonically with income.

III. AN ELECTION WINNER'S POLICY

The first step in a general analysis is to solve x_e's problem, deriving the winning candidate's optimal own income y_e and tax function T^*. A complication is that x_e's objective $U(x_e)$ is also a point on the utility profile and hence enters a subset of the incentive-compatibility constraints. We deal with this by solving a modified version of x_e's problem with analytically convenient properties that will turn out to have the same solution as the original problem.

In the modified problem, we give x_e the additional option of choosing a separate own income y_e and tax payment T_e (and hence an own consumption $y_e - T_e$), but we subject x_e to additional incentive compatibility constraints requiring that no individual x capable of earning y_e (that is, with $x \geq y_e$) would receive greater utility from earning y_e and paying taxes T_e than from accepting the assigned utility $U(x)$.

These additional constraints are

$$U(x) \geq U_e(x) \text{ for } x \geq y_e, \tag{7}$$

where $U_e(x) \equiv u(y_e - T_e, 1 - y_e/x)$ is the utility x would receive from earning y_e and paying taxes T_e. Formally, x_e'**s modified problem** is to

$$\begin{gathered}\text{maximize } u(y_e - T_e, 1 - y_e/x_e) \\ \text{subject to (1), (3), (4), (6), and (7),}\end{gathered}$$

by choice of $(y_e, T_e, \{U(x), Y(x), \psi(x)\}_x)$.

The modified problem can be solved conveniently in two stages. The first stage is a control problem that determines the set of feasible (y_e, T_e) pairs, and the second is a simple utility maximization that determines the optimal (y_e, T_e) pair. To evaluate whether a pair (y_e, T_e) is feasible, we find the maximum revenue $R(y_e, T_e)$ that can be raised conditional on (y_e, T_e) by solving the **conditional revenue maximization problem (CRM)**:

$$\begin{gathered}\text{maximize} \int_{x_-}^{x_+} T(U(x), Y(x), x) dF(x) \\ \text{subject to (1), (3), (4), and (7),}\end{gathered} \tag{CRM}$$

by choice of $\{U(x), Y(x), \psi(x)\}_x$. A pair (y_e, T_e) is feasible if and only if $R(y_e, T_e) \geq G$.

Because the value of x_e does not enter the objective function or the constraint set of CRM, the set of feasible (y_e, T_e) pairs is the same for all potential election winners. This feature is indispensable in deriving voting theorems in the next section.

The conditioning variables y_e and T_e enter CRM only through $U_e(x)$ in the inequality constraint (7), and the form of the solution differs depending on whether (7) binds. To determine the (y_e, T_e) pairs for which (7) binds, we solve CRM without (7), which we refer to as the **unconditional revenue maximization problem**

(URM):

$$\text{maximize} \int_{x_-}^{x_+} T(U(x), Y(x), x) dF(x) \tag{URM}$$

$$\text{subject to (1), (3), and (4),}$$

by choice of $\{U(x), Y(x), \psi(x)\}_x$. We assume URM has a unique solution denoted $\{\hat{U}(x), \hat{Y}(x), \hat{\psi}(x)\}_x$. Note that the maximum revenue that can be raised under URM is $\hat{G}$.

For (y_e, T_e) pairs with low enough $U_e(x)$ so $\hat{U}(x) \geq U_e(x)$ for all $x \geq y_e$, the profile $\{\hat{U}(x)\}_x$ satisfies (7); thus for such pairs, the solution to URM also solves CRM. (Because $G < \hat{G}$, all (y_e, T_e) pairs for which $\{\hat{U}(x)\}_x$ satisfies (7) are feasible.)

For all other (y_e, T_e) pairs, $\hat{U}(x) < U_e(x)$ for at least one $x \geq y_e$, so $\{\hat{U}(x)\}_x$ violates (7). Solutions to URM and CRM therefore differ and (7) binds at some value(s) of x under a solution to CRM. Given a conditioning pair (y_e, T_e), (7) therefore binds if and only if $\hat{U}(x) < U_e(x)$ for at least one x. Let $\mathcal{B} \equiv \{(y_e, T_e) \mid \hat{U}(x) < U_e(x)$ for some $x \geq y_e\}$ denote the set of (y_e, T_e) pairs for which (7) binds so the solution to URM does *not* solve CRM. Also let $\bar{\mathcal{B}} = \{(y_e, T_e) \mid \hat{U}(x) \leq U_e(x)$, for some $x \geq y_e\}$ denote the closure of $\mathcal{B}$.

For any $(y_e, T_e) \in \mathcal{B}$, let x_b denote a value of x at which (7) binds so $U(x_b) = U_e(x_b)$. We show below that for any such x_b, CRM can be solved in two parts: choose $\{U(x), Y(x), \psi(x)\}_x$ for $x \leq x_b$ to maximize revenue obtained from individuals with productivities less than x_b; and choose $\{U(x), Y(x), \psi(x)\}_x$ for $x \geq x_b$ to maximize revenue obtained from individuals with productivities greater than x_b.[14]

For given (x_b, y_e, T_e), **problem CRM1** focuses on maximizing revenue ob-

[14] Constraint (7) turns out to hold with equality at x_e, but imposing $x_b = x_e$ would obscure the independence of the set of feasible (y_e, T_e) pairs on the winner's identity.

tained from individuals $x \leq x_b$:

$$\text{maximize} \int_{x_-}^{x_b} T(U(x), Y(x), x) dF(x) \tag{CRM1}$$

$$\text{subject to (1), (3), (4)}, U(x_b) = U_e(x_b), \text{ and } Y(x_b) \leq y_e,$$

by choice of $\{U(x), Y(x), \psi(x)\}_{x \leq x_b}$. Similarly, **problem CRM2** focuses on maximizing revenue obtained from individuals $x \geq x_b$:

$$\text{maximize} \int_{x_b}^{x_+} T(U(x), Y(x), x) dF(x) \tag{CRM2}$$

$$\text{subject to (1), (3), (4), } U(x_b) = U_e(x_b), \text{ and } Y(x_b) \geq y_e,$$

by choice of $\{U(x), Y(x), \psi(x)\}_{x \geq x_b}$. The condition $U(x_b) = U_e(x_b)$ imposed on both problems ensures that (7) holds with equality at x_b. Let $R_1(x_b, y_e, T_e)$ and $R_2(x_b, y_e, T_e)$ denote maximal values of the revenue integrals in CRM1 and CRM2.

Problems URM, CRM1, and CRM2 have the same objective function and differential equations. They differ only in domains and boundary conditions. Hence their Hamiltonians and Euler equations have the same form. The common Hamiltonian is

$$H(U, Y, \psi, \xi, \mu, x) = T(U, Y, x) \cdot f(x) + \omega(U, Y, x) \cdot \xi(x) + \psi(x) \cdot \mu(x),$$

where ξ and μ as the costate variables associated with U and Y. The implied Euler equations are

$$H_Y = T_Y \cdot f(x) + \omega_Y(U, Y, x) \cdot \xi(x) = -\mu_x(x), \tag{8}$$

$$H_U = T_U \cdot f(x) + \omega_U(U, Y, x) \cdot \xi(x) = -\xi_x(x), \tag{9}$$

where definition of T and the properties of c^* imply $T_Y = 1 - u_l/(u_c x)$ and $T_U = -1/u_c$.

We sidestep issues of existence and uniqueness of a solution to CRM with:[15]

Assumption CON: For all $(y_e, T_e) \in \bar{\mathcal{B}}$ and $x_b \in \{x \mid \hat{U}(x) \leq U_e(x)\}$, CRM1 and CRM2 each have a solution. For any solution, functions ξ and μ satisfying (8) and (9) exist, and $H(U, Y, \psi, \xi, \mu, x)$ is strictly concave in (U, Y) for all x. The functions U and Y are continuous in x, ψ is piecewise continuous in x, and ξ and μ are continuous and piecewise continuously differentiable in x.

Strict concavity of the Hamiltonian in (U, Y) under CON ensures uniqueness. (By a unique solution, we mean a solution with unique profiles $\{U(x), Y(x)\}_x$ and a profile $\{\psi(x)\}_x$ that is unique except at points of discontinuity.)

The following proposition characterizes solutions to CRM and shows that if $(y_e, T_e) \in \mathcal{B}$ so the incentive compatibility constraint (7) binds, then income-bunching on an interval $[x_1, x_2]$ is generic and CRM can be solved by separately maximizing revenue obtained from those with productivities lower than x_b and revenue obtained from those with productivities higher than x_b. These two optimizations give the segments of the tax function below and above y_e (see figure 1), where x_b is any value in the interval $[x_1, x_2]$ of individuals that earn y_e (see flat segments of tax profiles in figure 2). Proofs of propositions are in an appendix:

Proposition 1 (conditional revenue maximization) *If CON holds, CRM has a unique solution for any (y_e, T_e), and:*

1. *For any $(y_e, T_e) \in \mathcal{B}$:*

[15] Existence would be straightforward with a discrete number of productivities but is technically complicated when productivity has a continuous distribution. Mirrlees (1986, p. 1235) states that conditions like CON are "obscure" in that they restrict third partial derivatives of u, but are unavoidable in variational problems of this type. Note that the previously assumed existence and uniqueness of $\{\hat{U}(x), \hat{Y}(x), \hat{\psi}(x)\}_x$ is implied by CON, because URM is equivalent to CRM2 with $x_b = x_-$, $y_e = 0$, and $T_e = -\alpha$.

(a) *There is an interval* $[x_1, x_2]$ *with* $U(x) = U_e(x)$ *and* $Y(x) = y_e$ *for* $x \in [x_1, x_2]$*, and* $U(x) > U_e(x)$ *for* $x \notin [x_1, x_2]$*. If* $[x_1, x_2]$ *includes any* $x \in (x_-, x_+)$*, then* $x_1 < x_2$*; otherwise, either* $x_1 = x_2 = x_-$ *or* $x_1 = x_2 = x_+$*.*

(b) *For any* $x_b \in [x_1, x_2]$*, the solution to CRM1 on* $[x_-, x_b]$ *together with the solution to CRM2 on* $[x_b, x_+]$ *also solve CRM. Moreover,* $R(y_e, T_e) = R_1(x_b, y_e, T_e) + R_2(x_b, y_e, T_e)$ *for all* $x_b \in [x_1, x_2]$*.*

(c) *Revenue* $R(y_e, T_e)$ *is continuous and differentiable in* (y_e, T_e)*, and strictly increasing in* T_e*. The function* T_e^* *defined by* $R(y_e, T_e^*(y_e)) = G$ *is differentiable.*

2. *For* $(y_e, T_e) \notin \mathcal{B}$*,* $\{\hat{U}(x), \hat{Y}(x), \hat{\psi}(x)\}_x$ *solves CRM and* $R(y_e, T_e) = \hat{G}$*.*

The function T_e^* described in part 1c of the proposition gives x_e's lowest feasible own tax payment as a function of y_e. In this way, x_e's income choice y_e uniquely determines x_e's leisure $1 - y_e/x_e$, minimum feasible tax payment $T_e^*(y_e)$, and consumption $y_e - T_e^*(y_e)$.

It is convenient to represent x_e's optimization in income-consumption space where consumption is a good but income is a "bad" because greater income means less leisure. The set

$$C_e = \{(y_e, c_e) \geq 0 | c_e \leq y_e - T_e^*(y_e)\} \tag{10}$$

is the set of x_e's feasible choices. Note that T_e^* is independent of x_e, so the same feasible set C_e is available to all potential election winners. Income and leisure are negatively related by the identity $l = 1 - y/x$, so the weak concavity of u implies that an individual's indifference curves in income-consumption space are convex, continuous lines $\{(y, c) \geq 0 | u = u(\ c, 1 - y/x)\}$, with slope $u_l/(xu_c) > 0$. The optimal income choice y_e is found by maximizing $u(y_e - T_e^*(y_e), 1 - y_e/x_e)$, illustrated in figure 7 as finding the highest indifference curve tangent to the upper boundary of C_e. As in the figure, this boundary is not necessarily concave,

so there may be multiple optimal y_e. Denote x_e's set of optimal income values $y_e^*(x_e) \equiv \{0 \leq y_e \leq x_e \mid y_e = \arg\max u(y_e - T_e^*(y_e), 1 - y_e/x_e)\}$.

The following proposition characterizes the solution to x_e's modified problem and verifies that the solution also solves x_e's original problem:

Proposition 2 (solution to x_e's problem) *Assume CON. For any $x_e \in [x_-, x_+]$:*

1. *A pair (y_e, T_e) and a set of profiles $\{U(x), Y(x), \psi(x)\}_x$ solve x_e's modified problem if and only if $\{U(x), Y(x), \psi(x)\}_x$ solves CRM for $y_e \in y_e^*(x_e)$ and $T_e = T_e^*(y_e)$.*

2. *x_e's modified problem has at least one solution. All solutions satisfy $Y(x_e) = y_e$ and $U(x_e) = U_e(x_e) = u(y_e - T_e, 1 - y_e/x_e)$. Moreover, $[x_1, x_2]$ is non-empty, $x_e \in [x_1, x_2]$, and $(y_e, T_e^*(y_e)) \in \mathcal{B}$. For $y_e > 0$, optimal values of y_e satisfy the first-order condition $\partial T_e^*/\partial y_e = 1 - u_l/(u_c x_e)$.*

3. *Profiles $\{U(x), Y(x), \psi(x)\}_x$ solve x_e's original problem if and only if $(Y(x_e), T_e^*(Y(x_e)), \{U(x), Y(x), \psi(x)\}_x)$ solve x_e's modified problem.*

Because the values of x_1 and x_2 are conditional on y_e, we sometimes write $x_1 = x_1(y_e)$ and $x_2 = x_2(y_e)$ below.

Seade (1982) points out that agent monotonicity implies that an individual with greater productivity chooses a greater income. Applied to election winners, this says that $y_e^*(x_e)$ is strictly increasing:[16]

Proposition 3 (income monotonicity (Seade, 1982)) *Let $x_L < x_H$ be two candidates and let $y_L \in y_e^*(x_H)$ and $y_H \in y_e^*(x_H)$. Then $y_L < y_H$ as long as $y_H > 0$; otherwise $y_L = y_H = 0$.*

[16] Seade (1982) does not contain a proof. A formal proof in the context of our model is available from the authors.

The idea is illustrated in figure 8. At any point (y, c) in income-consumption space with $y > 0$, agent monotonicity may be written $(y/x_L)[u_l(c, 1-y/x_L)/u_c(c, 1-y/x_L)] > (y/x_H)[u_l(c, 1-y/x_H)/u_c(c, 1-y/x_H)]$; dividing both sides by y shows that at (y, c), the slope of x_L's indifference curve exceeds the slope of x_H's indifference curve. Thus increased productivity flattens the indifference curve at (y, c). All potential election winners face the same feasible set C_e, so flatter indifference curves shift the optimum (tangency) toward greater y_e, and optimal income choices satisfy $y_L < y_H$ as long as $y_H > 0$. Similarly $y_H = 0$ implies $y_L = y_H = 0$; this occurs if x_L and x_H are so low that indifference curves at $y_e = 0$ are steeper than the boundary of C_e.

Note from propositions 1 and 2 that x_e's problem has exactly as many distinct solutions as there are elements in $y_e^*(x_e)$. Because y_e^* is strictly monotone by proposition 3, $y_e^*(x_e)$ is single-valued except at isolated x_e-values, so x_e's problem has a unique solution for almost all election winners. To account for the possibility that one of the candidates may have multiple optimal y_e, we express the policy of an election winner with any $y_e \in y_e^*(x_e)$ as a function of y_e, denoting the associated income and utility profiles $\{U(x \mid y_e), Y(x \mid y_e)\}_x$.[17]

The net income-tax function T^* implied by $\{U(x \mid y_e), Y(x \mid y_e)\}_x$ can then be inferred from (5), the definition of T. Specifically, for any income level $y \in [Y(x_- \mid y_e), Y(x_+ \mid y_e)]$ and any x for which $Y(x \mid y_e) = y$, the income tax is $T^*(y \mid y_e) = T(U(x \mid y_e), Y(x \mid y_e), x)$.[18]

To clarify the relationship between T^* and T_e^*, the value $T^*(y \mid y_e)$ is the income tax an individual with income y would pay if the election winner earns income y_e, whereas $T_e^*(y_e)$ is the minimum tax an election winner with income y_e can arrange

[17] From proposition 3, a given value $y_e > 0$ is optimal for at most one value of x_e, so the election winner's productivity can always be inferred from y_e.

[18] To extend T^* to the entire real line, it suffices to set taxes prohibitively high (e.g. $T^*(y) = y - \alpha$) for $y \notin [Y(x_-|y_e), Y(x_+|y_e)]$.

for him/herself by optimally choosing an income tax function T^*. It can be shown that T_e^* is the lower envelope of all tax functions T^*, and that the tax function T^* set by a given election winner has a downward spike at $y = y_e$, where T^* touches T_e^*.

IV. ELECTIONS

Elections between candidates with the same productivity or between candidates who would choose the same y_e are uninformative so we study elections between pairs of candidates $x_L < x_H$ with $y_L < y_H$. For now, take $y_L \in y_e^*(x_L)$ and $y_H \in y_e^*(x_H)$ as given; later in the section we discuss candidates' optimal income choices when $y_e^*(x_L)$ and/or $y_e^*(x_H)$ are multi-valued.

Induced Utility Profiles

We establish that under two regularity conditions, the utility profiles induced by x_L and x_H cross only once at a productivity denoted $x_\times$ with $x_L < x_\times < x_H$, that individuals $x > x_\times$ receive greater utility from x_H, and that individuals $x < x_\times$ either receive greater utility from x_L or else receive utility $u(\alpha, 1)$ from both candidates.

A graphical intuition is as follows. When x_e picks T^*, the individuals in the economy are presented with a set of (y, c) choices denoted $C(y_e) \equiv \{(y, c) \geq 0 | c \leq y - T^*(y \mid y_e)\}$. The downward spike in T^* at $y = y_e$ implies that the boundary of $C(y_e)$ has an upward spike at $(y_e, y_e - T_e^*(y_e))$, where $C(y_e)$ touches the boundary of C_e (see figure 9). Individuals $x \in [x_1(y_e), x_2(y_e)]$ then find a corner solution at income y_e optimal, individuals $x < x_1(y_e)$ optimally earn incomes $y < y_e$, and individuals $x > x_2(y_e)$ optimally earn incomes $y > y_e$.

As y_e increases differentially from y_L to y_H, the spike point $(y_e, y_e - T_e^*(y_e))$ moves monotonically to the right along the boundary of the common feasible set C_e, illustrated in the figure by the shift from L to H. This shifts the boundary of $C(y_e)$

outward for individuals with $y > y_e$ (arrow 1 in the figure), making them strictly better off, and *inward* for individuals with $y < y_e$ (arrow 2), making them strictly worse off as long as $U(x) > u(\alpha, 1)$.

Among the set of individuals $x \in [x_1(y_H), x_2(y_L)]$ who earn $y = y_e$ for any $y_e \in [y_L, y_H]$, agent monotonicity implies that there is a crossover individual $x_\times$ whose indifference curve goes through L and H.[19] Because agent monotonicity implies that indifference curves become flatter with productivity, it must be that individuals with productivities greater than $x_\times$ receive greater utility from x_H and individuals with productivities less than $x_\times$ receive greater utility from x_L.

The regularity conditions we impose to show this single-crossing property ensure that the boundary of $C(y_e)$ is uniquely defined and changes smoothly with y_e. The boundary of $C(y_e)$ is determined by the Euler equations for CRM1 and CRM2. The first regularity condition is therefore CON, which ensures that unique solutions to CRM1 and CRM2 exist; CON also fills a role like that of second-order conditions in comparative static exercises. The second regularity condition is a "no-bunching" assumption that ensures that the solutions to CRM1 and CRM2 vary smoothly with y_e. To formalize the assumption, let $X_0(y_e) \equiv \{x \mid Y(x \mid y_e) = 0\}$ denote the set of non-workers given y_e. The no-bunching assumption is then:[20]

Assumption NB: $Y(x \mid y_e)$ increases strictly on $[x_-, x_1(y_e)) \backslash X_0(y_e)$ and on $(x_2(y_e), x_+]$ for all y_e.

Assumption NB says that the income profile that solves x_e's problem does not have flat spots ("bunching") at any productivity except those that lead to incomes of zero or y_e.

[19] This assumes y_L and y_H are close. If they are not close, $[x_1(y_H), x_2(y_L)]$ may be empty and the crossover individual may have a tangency on a leg of a spike. Proposition 4 below allows for such cases.

[20] This type of assumption is common in the literature that follows Mirrlees (1971).

A complication is that there may be a set of individuals at the lower end of the productivity distribution who would obtain utility $u(\alpha, 1)$ from both candidates. To treat this, let $X_{\min}(y_e) = \{x \mid U(x \mid y_e) = u(\alpha, 1)\}$ denote the set of individuals who would obtain utility $u(\alpha, 1)$ from a candidate with income y_e.[21]

Assumptions CON and NB are sufficient for:

Proposition 4 (single crossing of utility profiles) *Consider candidates $x_L < x_H$. Assume CON, NB, and take $y_L \in y_e^*(x_L)$ and $y_H \in y_e^*(x_H)$ as given with $y_L < y_H$. Then utility profiles $\{U(x \mid y_H)\}_x$ and $\{U(x \mid y_L)\}_x$ cross at a unique point $x_\times \in (x_L, x_H)$, and:*

1. *Individuals $x > x_\times$ have $U(x \mid y_H) > U(x \mid y_L)$;*
2. *Individuals $x < x_\times$ with $x \notin X_{\min}(y_L)$ have $U(x \mid y_L) > U(x \mid y_H)$;*
3. *Individuals $x = x_\times$ have $U(x \mid y_L) = U(x \mid y_H)$; and*
4. *Individuals in $X_{\min}(y_L)$ have $U(x \mid y_L) = U(x \mid y_H) = u(\alpha, 1)$.*

Note that because $X_{\min}(y_H)$ consists of $X_{\min}(y_L)$ plus the set of x for whom $U(x \mid y_L) > U(x \mid y_H) = u(\alpha, 1)$, individuals in $X_{\min}(y_H)$ who are not in $X_{\min}(y_L)$ receive strictly higher utility from x_L.

A key assumption underlying proposition 4 is that x_e is able to impose an *unrestricted* tax function. Ad hoc restrictions on the tax function (for instance, that it be quadratic or piecewise linear) may cause utility profiles to cross several times, leading to voting cycles. Intuitively, this is because proposition 4 relies on incentive constraints that restrict utility differences of individuals close in productivity, with

[21] Because U and Y are increasing, $X_{\min}(y_e)$ and $X_0(y_e)$ are at the lower end of the income distribution. If $X_{\min}(y_e)$ has positive measure then $X_{\min}(y_e) = X_0(y_e)$, but $X_0(y_e)$ may have positive measure even if $X_{\min}(y_e)$ has measure zero.

individuals treated monotonically better the closer they are to the election winner. This monotone link may be disrupted by restrictions on the functional form of T^*.

Median-Voter Results

Proposition 4 forms the basis for median-voter results. If $X_{\min}(y_L)$ has measure zero and the mapping y_e^* is single-valued at x_L and x_H, the logic is simple. The proposition implies that all individuals with productivity greater than $x_\times$ have $U(x \mid y_L) > U(x \mid y_H)$ and hence vote for x_H. If $x_\times < x_M$, a majority that includes median-productivity individuals therefore votes for x_H. Similarly all individuals with productivity less than $x_\times$ who are not in $X_{\min}(y_L)$ have $U(x \mid y_H) > U(x \mid y_L)$. Because $X_{\min}(y_L)$ has measure zero, this implies that almost all individuals with productivity less than $x_\times$ vote for x_L. If $x_M < x_\times$, a majority that includes x_M therefore votes for x_L. Thus median-productivity individuals always vote with the winning majority. (In the non-generic case in which $x_M = x_\times$, the election ends in a tie.)

There are two possible complications that require additional assumptions. First, individuals in $X_{\min}(y_L)$ have $U(x \mid y_L) = U(x \mid y_H)$ and are indifferent, and assumptions about how such individuals vote matter if (and only if) $X_{\min}(y_L)$ has positive measure. Three alternative reasonable assumptions might be made: individuals in $X_{\min}(y_L)$ vote for x_L, who is closer to the individual's own productivity and income (*voting by closeness*);[22] they abstain; or they randomize. Randomization is formally similar to abstention under simple assumptions about how randomization occurs, so we consider only voting by closeness and abstention.

[22] In the spirit of Benabou and Ok (2001), a preference for x_L by a nonworker with current utility $u(\alpha, 1)$ would be expected in an extended model in which future productivity is random and there is a positive probability the individual will work and have utility greater than $u(\alpha, 1)$ during some part of the winner's term of office.

When voters in $X_{\min}(y_L)$ vote by closeness, proposition 4 again implies that the winning majority contains all individuals with median productivity. When voters in $X_{\min}(y_L)$ abstain, the statement of the median-voter result changes: all voters with median productivity *among those who do not abstain* vote with the winning majority. Voters in $X_{\min}(y_L)$ and $x_\times$ abstain but the latter have measure zero, so voters with median productivity in $[x_-, x_+]\backslash X_{\min}(y_L)$ always vote with the majority.

The second complication arises if the mapping y_e^* is not single-valued at x_L or x_H, because the specific values chosen in $y_e^*(x_L)$ or $y_e^*(x_H)$ may then affect voting. The complication is non-generic because y_e^* is single-valued except at isolated points. In cases where more than one income-value gives a candidate the same utility, we assume the candidate chooses y_e to maximize the own vote share. This is natural because winning gives higher utility than losing.

For x_H, proposition 4 implies that choosing $y_H = \min\{y_e^*(x_H)\}$ yields the lowest crossing point $x_\times$ and therefore maximizes x_H's vote share against any income chosen by x_L. Similarly $y_L = \max\{y_e^*(x_L)\}$ yields the highest crossing point for x_L and therefore maximizes x_L's vote share against x_H, provided $X_{\min}(y_L)$ has measure zero or individuals vote by closeness. If $X_{\min}(y_L)$ has positive measure and individuals in $X_{\min}(y_L)$ abstain, however, an income choice less than $\max\{y_e^*(x_L)\}$ may maximize x_L's vote share because the number of abstentions may then rise with y_L. In this case, determining the winner requires inspecting voting outcomes for all elements of $y_e^*(x_L)$ against $\min\{y_e^*(x_H)\}$.

Thus:

Proposition 5 (median-voter theorem) *Consider candidates $x_L < x_H$. Assume CON, NB, $y_H = \min\{y_e^*(x_H)\} > 0$, and that x_L chooses the $y_L \in y_e^*(x_L)$ that maximizes x_L's vote share against x_H:*

1. *If $X_{\min}(y_L)$ has measure zero, then the candidate who provides greater utility to*

median-productivity individuals wins.

2. *If $X_{\min}(y_L)$ has positive measure and*

   (a) *if indifferent individuals vote based on closeness in productivity, then the candidate who provides greater utility to median-productivity individuals wins, or*

   (b) *if indifferent individuals abstain, then the candidate who provides greater utility to the median of $[x_-, x_+] \backslash X_{\min}(y_L)$ wins.*

Parts 1 and 2a of proposition 5 immediately imply:

Proposition 6 (Condorcet winner) *Assume CON, NB, and $y_M \equiv \min\{y_e^*(x_M)\} > 0$. Then x_M is the Condorcet winner if $X_{\min}(y_M)$ has measure zero, or if $X_{\min}(y_M)$ has positive measure and indifferent individuals vote based on closeness in productivity.*

When $X_{\min}(y_M)$ has positive measure and indifferent voters abstain (case 2b of proposition 5), the existence of a Condorcet winner is not guaranteed because the set of individuals who vote then depends on the specific candidate pair.[23]

V. THE SHAPE OF THE ELECTION WINNER'S TAX FUNCTION

To describe the shape of the tax function T^* set by a generic election winner x_e with income $y_e \in y_e^*(x_e)$, we derive the marginal tax schedule $dT^*(y)/dy$.

[23] By a simple fixed-point argument, there is always a smallest productivity $x_m \in (x_M, x_+)$ that is median in the set $[x_-, x_+] \backslash X_{\min}(\min\{y_e^*(x_m)\})$. Because $x_L < x_m$ may draw individuals in $X_{\min}(\min\{y_e^*(x_m)\})$ to the polls, x_L may win against x_m, and because x_L is below the median in $[x_-, x_+] \backslash X_{\min}(\min\{y_e^*(x_L)\})$, x_L may lose against an $x_l \in (x_L, x_m)$ who in turn may lose against x_m, forming a cycle. If x_m wins against all $x_L < x_m$, however, then x_m is a Condorcet winner. This occurs if no x_L induces enough individuals in $X_{\min}(\min\{y_e^*(x_m)\})$ to vote.

From Section III, the income tax satisfies $T^*(Y(x)) = T(U(x), Y(x), x)$ for all x. On intervals where Y has a differentiable inverse Y^{-1}, we may substitute $x = Y^{-1}(y)$ into this expression for T^*, differentiate with respect to y, and impose incentive compatibility (3) to obtain $dT^*(y)/dy = T_Y(U(Y^{-1}(y)), y, Y^{-1}(y))$.[24] For brevity, define $\tau(x) \equiv T_Y(U(x), Y(x), x)$, so the latter is

$$dT^*(y)/dy = \tau(Y^{-1}(y)). \tag{11}$$

The following proposition characterizes the marginal tax schedule as $\tau(Y^{-1}(y))$. As part of the proposition, we show that (11) holds at all points where Y^{-1} is well-defined regardless of differentiability. By NB, Y^{-1} is well-defined at all incomes except $y = y_e$ and possibly $y = 0$. We use the solution to CRM1 conditional on $(x_e, y_e, T_e^*(y_e))$ to evaluate $\tau(x)$ for $x < x_1$, and use the solution to CRM2 conditional on $(x_e, y_e, T_e^*(y_e))$ to evaluate $\tau(x)$ for $x > x_2$:

Proposition 7 (the net-income-tax function chosen by x_e) *Assume CON and NB and consider $x_e \in (x_-, x_+)$ with $y_e > 0$.*[25] *Then T^* is continuously differentiable on $[Y(x_-), Y(x_+)]$ except at $y = y_e$, with $dT^*(y)/dy = \tau(Y^{-1}(y))$ except at y_e and possibly at zero.*[26] *Moreover:*

1. *T^* is always nonlinear in that the marginal tax rate jumps upward at y_e. Specifically: $\tau(x_1) = lim_{y \uparrow y_e} dT^*(y)/dy < \tau(x_2) = lim_{y \downarrow y_e} dT^*(y)/dy$.*

2. *T^* is strictly increasing for $y > y_e$. If x_+ is finite, $dT^*(Y(x_+))/dy = \tau(x_+) = 0$.*

3. *T^* has three possible shapes for $y < y_e$:*

[24] More heuristically, at any point where T^* is differentiable, an individual who maximizes $u(y - T^*(y), 1 - y/x)$ by choice of y satisfies the first-order condition $dT^*(y)/dy = 1 - u_l/(u_c x)$. From section III, the right-hand side of the condition equals T_Y.

[25] Cases with $x_e = x_-$, $x_e = x_+$, and $y_e = 0$ are straightforward but give degenerate results in which $[x_0, x_1) \backslash X_0$, $[x_1, x_2]$, or $(x_2, x_+]$ are empty or contain a single point.

[26] If $X_0(y_e)$ is an interval, $Y^{-1}(0)$ is not defined and $dT^*(0)/dy = \tau(\max X_0)$.

(a) T^* may be strictly decreasing. A sufficient condition for this is that the minimum-utility constraint (1) does not bind. If $X_0(y_e)$ has measure zero, then $dT^*(Y(x_-))/dy = \tau(x_-) = 0$.

(b) T^* may increase strictly from $Y(x_-)$ to a local maximum at $Y(x_\tau)$, then decrease strictly to y_e. This occurs if and only if $\tau(x_\tau) = 0$ for some $x_\tau \in (x_-, x_1) \backslash X_0(y_e)$.

(c) T^* may be strictly increasing.

Part 1 says that the tax function has a kink at the election winner's own income level, with T^* having a lower slope from the right than from the left. Thus the tax function is always nonlinear.

Because the kinked function T^* is strictly better than a linear tax for the winner and the associated utility profile U is continuous, there is an interval of productivities around x_e that also are better off under T^* than under a linear tax. From the median-voter results of the previous section, if one of the candidates has productivity close to the median, the resulting redistributional tax policy is better than a linear tax for a range of middle-productivity individuals. This result is in accord with Director's Law.

Results in part 2 describe taxes at incomes above y_e and resemble results in Seade (1977, 1982) and Röell (1996): income taxes rise with income; and the marginal tax rate is zero at the maximum productivity x_+ if $x_+ < \infty$.

Results in part 3 state that the tax function T^* may have three shapes at incomes below y_e. If the minimum-utility constraint (1) does not bind (case a) then T^* decreases monotonically over $[Y(x_-), y_e)$ as the election winner extracts greater taxes from individuals the lower their incomes. Such an outcome is counterfactual. For marginal rates not to be negative on $[Y(x_-), y_e)$, the minimum utility constraint must bind. If (1) binds, T^* may still have a counterfactual downward sloping segment

(case b), or it may have positive marginal rates throughout (case c); in both cases, marginal rates can be high at very low incomes (see figures 1, 5, and 6).

Whether case (b) or (c) describes the tax function depends on the revenue requirement G and the level of the altruistic consumption floor α. With no revenue requirement or altruism ($G = 0$ and $\alpha = 0$), it can be shown that $T^*(y_e) < 0$ whereas $T^*(Y(x_-)) \geq 0$. This says that the election winner pays lower taxes than an individual with productivity x_-, so T^* necessarily has a downward sloping segment (case b). If G and α are high enough, however, marginal tax rates are positive throughout:[27]

Proposition 8 (the role of G and α in determining the slope of T^*) *Assume CON and NB and consider solutions to x_e's problem with $G < \hat{G}(\alpha)$. Then:*

1. *For G in a neighborhood of $\hat{G}(\alpha)$ for given α, the tax function T^* is strictly increasing on $[Y(x_-), Y(x_+)]$.*

2. *For α in a neighborhood of $\hat{G}^{-1}(G)$ for given G, the tax function T^* is strictly increasing on $[Y(x_-), Y(x_+)]$.*

The proof studies parametric variations in G and α. The proposition shows that the tendencies illustrated in figures 5 and 6 are generic.

VI. CONCLUSION

We study a simple form of representative democracy in which voters elect a single candidate who then imposes redistributional taxes on all citizens. We provide conditions for the winning candidate to receive the votes of median-productivity individuals, so a median-productivity candidate is a Condorcet winner. We also show that the tax function imposed by the election winner is always nonlinear, may have

[27] Röell (1996) assumes $G = 0$ and $\alpha = 0$, which implies that the tax function necessarily has negative marginal rates at some incomes.

positive marginal rates at all incomes if government faces non-redistributional spending requirements or if individuals are altruistic, and may impose high marginal tax rates at very low incomes.

Two assumptions are crucial for the median-voter results. The first and no doubt strongest is *one-dimensional heterogeneity* of candidates; if candidates (the objects over which voting occurs) differ along more than one dimension, voting cycles arise generally. The advantage of assuming such one-dimensional heterogeneity is that it allows analysis of how economic forces influence the shape of the tax function. The alternative approach of ruling out cycling by imposing linearity on the tax function precludes such analysis.

The second crucial assumption is that *no* restrictions are placed on the functional form of the tax function. Although a median-voter result was obtained by Meltzer and Richard (1981) by restricting the tax function to be linear, other restrictions can upset a median-voter outcome.

REFERENCES

[1] Bénabou, Roland and Efe Ok, "Social Mobility and the Demand for Redistribution: The Poum Hypothesis," *Quarterly Journal of Economics* 116. May 2001, 447-87.

[2] Besley, Timothy and Stephen Coate, "A Model of Representative Democracy," *Quarterly Journal of Economics* 112 (1), Feb. 1997, 85-114.

[3] Brito, Dagobert and William Oakland, "Some Properties of the Optimal Income Tax," *International Economic Review* 18 (2), June 1977, 407-23.

[4] Browning, Edgar and William Johnson, *The Distribution of the Tax Burden*, Washington, D.C.: American Enterprise Institute, 1979.

[5] Diamond, Peter, "Optimal Income Taxation: An Example with U-Shaped Pattern of Optimal Marginal Tax Rates," *American Economic Review* 88 (1), March 1998, 83-95.

[6] Dickert, Stacy, Scott Houser, and John Karl Scholz, "The Earned Income Tax Credit and Transfer Programs: A Study of Labor Market and Program Participation," in *Tax Policy and the Economy*, NBER, 1995, 1-50.

[7] Dixit, Avinash, and John Londregan, "Ideology, Tactics, and Efficiency in Redistributive Politics," *Quarterly Journal of Economics* 113 (2), May 1998, 497-529.

[8] Ebert, Udo, "A Reexamination of the Optimal Nonlinear Income Tax," *Journal of Public Economics* 49, 1992, 47-73.

[9] Gouveia, Miguel, "Majority Rule and the Public Provision of a Private Good," *Public Choice* 93 (3-4) Dec. 1997, 221-44.

[10] Keane, Michael and Robert Moffitt, "A Structural Model of Multiple Welfare Program Participation and Labor Supply," *International Econ. Review* 39, 1998, 553-89.

[11] Meltzer, Allan, and Scott Richard, "A Rational Theory of the Size of Government," *Journal of Political Economy* 89 (5), Oct.1981, 914-927.

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

[13] ———, "The Theory of Optimal Taxation," in: K.J. Arrow and M.D. Intriligator, *Handbook of Mathematical Economics, vol.3*, Amsterdam: North-Holland, 1986, 1197-1249.

[14] Osborne, Martin and Al Slevinski, "A Model of Political Competition with Citizen-Candidates," *Quarterly Journal of Economics* 111 (1), Feb. 1986, 65-96.

[15] Rawls, John, *A Theory of Justice*, Cambridge, Mass.: Harvard University Press, 1971

[16] Roberts, Kevin, "Voting over Income Tax Schedules," *Journal of Public Economics* 8, Dec. 1977, 329-40.

[17] Röell, Ailsa, "Voting over Nonlinear Income Tax Schedules," unpublished paper, 1996.

[18] Romer, Thomas, "Individual Welfare, Majority Voting, and the Properties of a Linear Income Tax," *Journal of Public Economics* 4, Feb. 1975, 163-85.

[19] Seade, Jesus, "On the Shape of Optimal Tax Schedules," *Journal of Public Economics* 7, 1977, 203-235.

[20] ———, "On the Sign of the Optimum Marginal Income Tax," *Review of Economic Studies* 49, 1982, 637-643.

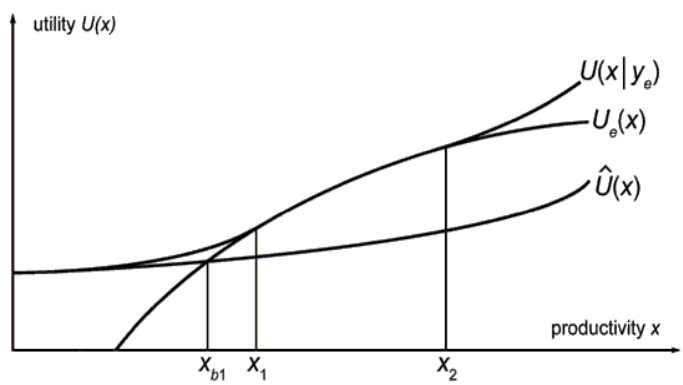
[21] Stigler, George, "Director's Law of Public Income Redistribution," *Journal of Law and Economics* 13, Apr. 1970, 1-10.

APPENDIX: PROOFS

Proposition 1 Overview: For $(y_e, T_e) \notin \mathcal{B}$, we show that solutions to URM solve CRM, establishing part 2 of the proposition. For $(y_e, T_e) \in \mathcal{B}$, we first define $x_{b1} \equiv \min\{x \mid U_e(x) \geq \hat{U}(x)\}$ (see figure A1). Let $_{(i)}$ subscripts mark solutions to CRMi for $i = 1, 2$. We construct the interval $[x_1, x_2]$ by solving CRM2 conditional on (x_{b1}, y_e, T_e), defining $x_2 \equiv \sup\{x \mid Y_{(2)}(x) = y_e\}$, then solving CRM1 conditional on (x_2, y_e, T_e) and defining $x_1 \equiv \min\{x \mid Y_{(1)}(x) = y_e\}$. The interval $[x_1, x_2]$ has constant income $Y(x) = y_e$ and hence constant utility $U(x) = U_e(x)$ as in part 1a. We combine the solutions to CRM1 and CRM2 conditional on (x_2, y_e, T_e) as

$$\{U_b(x), Y_b(x)\}_x \equiv \{U_{(1)}(x), Y_{(1)}(x)\}_{x \leq x_2} \cup \{U_{(2)}(x), Y_{(2)}(x)\}_{x \geq x_2}.$$

The remainder of parts 1a-b are established by showing that $\{U_b(x), Y_b(x)\}_x$ solves CRM, and that the combined solutions to CRM1 and CRM2 conditional on (x_b, y_e, T_e) for any $x_b \in [x_1, x_2]$ also equal $\{U_b(x), Y_b(x)\}_x$ and hence yield the same revenue. The properties of R in part 1c follow from parametric variation/value function arguments. Given R, the existence and differentiability of T_e^* follows from the implicit function theorem.



Details: CON and the Mangasarian sufficiency theorem imply that: CRM1

and CRM2 have solutions with unique profiles $\{U(x), Y(x)\}_x$; the conditions of the Maximum Principle are sufficient for optimality; and Y is continuous. Because Y is unique, its derivative ψ to is uniquely defined except at points of discontinuity. This uniqueness implies the uniqueness of the solution to CRM.

For $(y_e, T_e) \notin \mathcal{B}$, we have $\hat{U}(x) \geq U_e(x)$ for all x so the solution to URM satisfies (7) and hence is optimal for CRM. The profiles $\{\hat{U}(x), \hat{Y}(x)\}_x$ are unique by CON because URM is equivalent to CRM2 conditional on $(x_b = x_-, y_e = 0, T_e = -\alpha)$.[28] The definition of $\hat{G}$ implies $R(y_e, T_e) = \hat{G}$ for $(y_e, T_e) \notin \mathcal{B}$.

For $(y_e, T_e) \in \mathcal{B}$, the solution to URM does not satisfy (7). Because CRM1 and CRM2 conditional on (x_b, y_e, T_e) for any $x_b \in [x_1, x_2]$ satisfy (7) by construction, the combined solutions to CRM1 on $[x_-, x_b]$ and CRM2 on $[x_b, x_+]$ must differ from the solution to URM.

Because URM does not satisfy (7), the sets $\{x \mid U_e(x) > \hat{U}(x)\}$ and hence $\{x \mid U_e(x) \geq \hat{U}(x)\}$ are nonempty so $x_{b1} = \min\{x \mid U_e(x) \geq \hat{U}(x)\}$ exists. Consider CRM2 conditional on (x_{b1}, y_e, T_e). Its solution profiles $\{U_{(2)}(x), Y_{(2)}(x)\}_x$ are unique and $U_{(2)}(x_{b1}) = U_e(x_{b1})$, so $Y_{(2)}(x_{b1}) = y_e$. Thus $\{x : Y_{(2)}(x) = y_e\}$ is non-empty so x_2 is well-defined.

The special case with $x_{b1} = x_2$ is considered below; for now, assume $x_{b1} < x_2$. It cannot be that $x_2 = \infty$ is optimal (when $x_+ = \infty$) because higher revenue could then be extracted from high-productivity individuals by assigning them incomes above y_e. Specifically, define

$$\tau(x|y_e, T_e) \equiv 1 - \frac{u_1(y_e - T_e, 1 - y_e/x)}{u_c(y_e - T_e, 1 - y_e/x)x}.$$

Agent monotonicity implies that $\tau(x|y_e, T_e)$ is strictly increasing in x and converges to one as $x \to \infty$. Thus there is an $x_\tau < \infty$ such that $\tau(x|y_e, T_e) > 0$ for all $x > x_\tau$. From

[28] In CRM2 conditional on $(x_-, 0, -\alpha)$, the constraint $Y(x_b) \geq y_e$ holds trivially and the constraint $U(x_b) = U_e(x_b)$ reduces to $U(x_-) = u(\alpha, 1)$. The Euler equation (9) for URM implies $\xi(x_-) < 0$ so (1) binds, that is, $U(x_-) = u(\alpha, 1)$.

(5) and $T_Y(U_e(x), y_e, x) = \tau(x|y_e, T_e)$, an upward variation in $Y(x)$ for all $x \in (x_\tau, x_+]$ holding $U(x)$ constant then raises T, so a policy with $Y(x) = y_e$ for all $x \geq x_{b1}$, that is, with $x_2 = \infty$, cannot be optimal.

From $x_2 < \infty$, it follows that $x_2 = \max\{x : Y_{(2)}(x) = y_e\}$, so $Y_{(2)}(x_2) = y_e$ and $U_{(2)}(x_2) = U_e(x_2)$.

Consider CRM1 conditional on (x_2, y_e, T_e) and note that $x_1 = \min\{x : Y_{(1)}(x) = y_e\}$. CRM1 conditional on (x_2, y_e, T_e) and CRM2 conditional on (x_{b1}, y_e, T_e) have matching utility and income values at x_2, or $U_{(1)}(x_2) = U_{(2)}(x_2) = U_e(x_2)$ and $Y_{(1)}(x_2) = Y_{(2)}(x_2) = y_e$, so the solutions to both CRM1 and CRM2 can hold simultaneously and $\{U_b(x), Y_b(x)\}_x$ is well-defined at x_2.

The special case with $x_1 = x_2$ (and $x_{b1} < x_2$) is also be considered below; for now, assume $x_- \leq x_1 < x_2 \leq x_+$. By construction, the combined profiles $\{U_b(x), Y_b(x)\}_x$ have $U_b(x) = U_e(x)$ and $Y_b(x) = y_e$ for $x \in [x_1, x_2]$, and $Y_b(x) \neq y_e$ for $x \notin [x_1, x_2]$. Because $U_b(x) = U_e(x)$ implies $Y_b(x) = y_e$ and $U_b(x) \geq U_e(x)$ always holds, it must be that $U_b(x) > U_e(x)$ whenever $Y_b(x) \neq y_e$, that is, for $x \notin [x_1, x_2]$.

CON implies that the unique solution profiles of CRM1 and CRM2 conditional on (x_b, y_e, T_e) for any $x_b \in [x_1, x_2]$ are the same as $\{U_b(x), Y_b(x)\}_x$, and hence yield the same revenue $R_b \equiv R(y_e, T_e)$. To prove the optimality of $\{U_b(x), Y_b(x)\}_x$, suppose to the contrary that an alternative profile $\{U_a(x), Y_a(x)\}_x$ satisfied the constraints of CRM and yielded revenue $R_a > R_b$. There are two possibilities:

(a) If $Y_a(x_a) = y_e$ for any $x_a \in [x_-, x_+]$ then the combined solutions to CRM1 and CRM2 conditional on (x_a, y_e, T_e), denoted $\{U_a^*(x), Y_a^*(x)\}_x$, would yield revenue $R_a^* \geq R_a > R_b$. Because any $x_b \in [x_1, x_2]$ yields revenue R_b, it must be that $x_a \notin [x_1, x_2]$. Hence either $\min\{x : Y_a^*(x) = y_e\} > x_2$ or $\max\{x : Y_a^*(x) = y_e\} < x_1$. Consider $\min\{x : Y_a^*(x) = y_e\} > x_2$. CRM2 conditional on (x_2, y_e, T_e) satisfies the transversality condition $\xi_{(2)}(x_+) = 0$ and the Euler equation $\xi_{(2),x}(x) = -H_U > 0$ for $x \geq x_2$, so $\xi_{(2)}(x_2) < 0$. For $x > x_2$ we have $Y_{(2)}(x) > y_e$ so $\psi(x) > 0$ and hence

$\mu_{(2)}(x) = 0$ in a neighborhood above x_2. Thus $\mu_{(2),x}(x_2) = 0$. In (8), $\mu_{(2),x}(x_2) = 0$ and $\xi_{(2)}(x_2) < 0$ imply $\tau(x_2|y_e, T_e) > 0$. Because $\tau(x|y_e, T_e)$ is strictly increasing in x, $\tau(x|y_e, T_e) > 0$ for $x \in [x_2, x_+]$. Again using (5), a compensated upward variation in Y_a^* on $\{x : Y_a^*(x) = y_e\} \subset [x_2, x_+]$ would then increase revenue so $\{U_a^*(x), Y_a^*(x)\}_x$ is not revenue maximizing, a contradiction. Analogous logic applies if $x_1 > \max\{x : Y_a^*(x) = y_e\}$.

(b) If $Y_a(x) \neq y_e$ for all x then the continuity of Y_a implies that either $Y_a(x) < y_e$ or $Y_a(x) > y_e$ for all x. If $Y_a(x) < y_e$ then $\{U_a(x), Y_a(x)\}$ is feasible for CRM1 conditional on (x_+, y_e, T_e) so $R_a \leq R_1(x_+, y_e, T_e)$. Similarly if $Y_a(x) > y_e$ then $\{U_a(x), Y_a(x)\}$ is feasible for CRM2 conditional on (x_-, y_e, T_e) so $R_a \leq R_2(x_-, y_e, T_e)$. The feasible sets for CRM1 conditional on (x_+, y_e, T_e) and for CRM2 conditional on (x_-, y_e, T_e) are in the feasible set for CRM, so $R_b \geq \max\{R_1(x, y_e, T_e), R_2(x_-, y_e, T_e) \geq R_a$, also a contradiction. Thus $\{U_b(x), Y_b(x)\}_x$ is optimal.

We now consider the two special cases $x_{b1} = x_2$ and $x_{b1} < x_1 = x_2$, and show that these correspond to the boundary cases with $x_- = x_1 = x_2$ and $x_+ = x_1 = x_2$ in part 1a:

(1) Suppose $x_{b1} = x_2$: URM restricted to any $[x_-, x_b]$ satisfies all sufficient optimality conditions for CRM1 conditional on (x_b, y_e, T_e) except $Y_{(1)}(x_b) \leq y_e$ and $U_e(x_b) = U_{(1)}(x_b)$. If $x_- < x_{b1} = x_2$, then $U_e(x_{b1}) = \hat{U}(x_{b1})$ and $\hat{Y}(x_{b1}) = y_e$, so $\{\hat{U}(x), \hat{Y}(x)\}$ on $[x_-, x_{b1}]$ is the unique solution to CRM1 conditional on (x_{b1}, y_e, T_e). The combined solutions of CRM1 on $[x_-, x_{b1}]$ and CRM2 on $[x_{b1}, x_+]$ then satisfy the sufficient conditions for a solution to URM on $[x_-, x_+]$, contradicting $(y_e, T_e) \in \mathcal{B}$.[29]

[29] For the combined solutions of CRM1 and CRM2 to satisfy all sufficient conditions for a solution to URM, solutions must match at x_{b1}, that is, $Y_{(1)}(x_{b1}) = Y_{(2)}(x_{b1})$, $U_{(1)}(x_{b1}) = U_{(2)}(x_{b1})$, $\mu_{(1)}(x_{b1}) = \mu_{(2)}(x_{b1})$, and $\xi_{(1)}(x_{b1}) = \xi_{(2)}(x_{b1})$. The first two hold by construction. From $Y_{(1)}(x) < y_e$ for $x < x_{b1}$ we have $\mu_{(1)}(x) = 0$ in a neighborhood of x_{b1}, and hence from (8), $H_{(1),Y}(x_{b1}) = 0$. Similarly, $Y_{(2)}(x) > y_e$ for $x > x_2$ implies $\mu_{(2)}(x) = 0$ in a neighborhood of x_2 and hence $H_{(2),Y}(x_2) = 0$. Thus $H_{(1),Y}(x_{b1}) = H_{(2),Y}(x_{b1})$, which implies $\xi_{(1)}(x_{b1}) = \xi_{(2)}(x_{b1})$.

Thus $x_2 = x_-$. This is a boundary case in which the constraint $U_{(2)}(x) \geq U_e(x)$ binds only at x_-, and the solution to CRM2 conditional on (x_-, y_e, T_e) is the unique solution to CRM. In this case, $x_1 = x_-$ so $x_- = x_1 = x_2$.

(2) Suppose $x_1 = x_2$ with $x_{b1} < x_2$: If $x_2 < x_+$, then $U_e(x_2) = \hat{U}(x_2)$ and $\hat{Y}(x_2) = y_e$, so $\{\hat{U}(x), \hat{Y}(x)\}$ on $[x_2, x_+]$ is the unique solution to CRM2 conditional on (x_2, y_e, T_e). The combined solutions of CRM1 on $[x_-, x_2]$ and CRM2 on $[x_2, x_+]$ then satisfy the sufficient conditions for a solution to URM on $[x_-, x_+]$, contradicting $(y_e, T_e) \in \mathcal{B}$. Thus $x_1 = x_2 = x_+ < \infty$. This is a boundary case in which the constraint $U_{(1)}(x) \geq U_e(x)$ binds only at x_+, and the solution to CRM1 conditional on (x_+, y_e, T_e) is the unique solution to CRM.

Standard value-function results imply that R_1 and R_2 are continuous and differentiable with respect to $(y_e, T_e) \in \mathcal{B}$. To sign $\partial R/\partial T_e$ when $x_1 < x_2$, choose any $x_b \in (x_1, x_2)$. Because solutions to CRM1 and CRM2 are unique, so is $R = R_1 + R_2$, and

$$\begin{aligned} \partial R/\partial T_e &= \partial R_1(x_b, y_e, T_e)/\partial T_e + \partial R_2/\partial T_e(x_b, y_e, T_e) && (12) \\ &= [\xi_{(1)}(x_b) - \xi_{(2)}(x_b)] \cdot u_c(y_e - T_e, 1 - y_e/x_b). && (13) \end{aligned}$$

Subtract (9) for CRM2 conditional on (x_1, y_e, T_e) from (9) for CRM1 conditional on (x_2, y_e, T_e) to obtain

$$\xi_{(1),x}(x) - \xi_{(2),x}(x) = -\omega_U(U_e(x), y_e, x) \cdot (\xi_{(1)}(x) - \xi_{(2)}(x)), \text{ for } x \in [x_1, x_2]. \quad (14)$$

This is a homogenous linear differential equation so $\xi_{(1)} - \xi_{(2)}$ cannot change sign on $[x_1, x_2]$. Taking similar differences of Euler equations (8) and integrating over $[x_1, x_b]$ with $x_b \in [x_1, x_2]$ yields

$$\begin{aligned} I(x_b) &\equiv \int_{x_1}^{x_b} \omega_Y(U_e(x), y_e, x) \cdot [\xi_{(1)}(x) - \xi_{(2)}(x)]dx && (15) \\ &= \mu_{(2)}(x_b) + \mu_{(1)}(x_1) - \mu_{(2)}(x_1) - \mu_{(1)}(x_b). && (16) \end{aligned}$$

Because $\psi(x) \geq 0$ implies $\mu_{(i)}(x) \leq 0$ and the optimality conditions of CRM1 and CRM2 imply $\mu_{(1)}(x_1) = 0$ and $\mu_{(2)}(x_2) = 0$, it follows that $I(x_2) \geq 0$ and hence $\xi_{(1)}(x_b) - \xi_{(2)}(x_b) \geq 0$. If $\xi_{(1)}(x_b) = \xi_{(2)}(x_b)$ then (14) implies $\xi_{(1)}(x) = \xi_{(2)}(x)$ for $x \in [x_1, x_2]$ so $I(x_2) = 0$, whence $\mu_{(2)}(x_1) = \mu_{(1)}(x_2) = 0$. From (8), $\tau(x_1|y_e, T_e) = \tau(x_2|y_e, T_e)$ and because τ is strictly increasing, $x_1 = x_2$, contradicting $x_1 < x_2$. Thus

$$\xi_{(1)}(x_b) - \xi_{(2)}(x_b) > 0, \tag{17}$$

so $\partial R/\partial T_e > 0$.

In the special case with $x_- = x_1 = x_2$, the transversality condition $\xi_{(2)}(x_+) = 0$ and (9) imply $\xi_{(2)}(x_-) < 0$, whence $\partial R/\partial T_e = -\xi_{(2)}(x_-) \cdot u_c > 0$. In the special case with $x_1 = x_2 = x_+$, the solution to CRM1 conditional on (x_+, y_e, T_e) has $\xi_{(1)}(x_+) \geq 0$ and cannot satisfy the sufficient conditions for a solution to URM, so $\xi_{(1)}(x_+) \neq 0$ whence $\partial R/\partial T_e = \xi_{(1)}(x_+) \cdot u_c > 0$.

Proposition 2 1. From the definition of R, a given (y_e, T_e) satisfies the feasibility constraint $R(y_e, T_e) \geq G$, if and only if profiles $\{U(x), Y(x), \psi(x)\}_x$ exist that satisfy the constraints of x_e's modified problem (equations (1), (3), (4), (6), and (7)). Hence $(y_e, T_e, \{U(x), Y(x), \psi(x)\}_x)$ solves x_e's modified problem if and only if (y_e, T_e) maximizes $u(y_e - T_e, 1 - y_e/x_e)$ subject to $R(y_e, T_e) \geq G$. Values $T_e < T_e^*(y_e)$ are infeasible and values $T_e > T_e^*(y_e)$ are suboptimal because x_e could raise $u(y_e - T_e, 1 - y_e/x_e)$ by reducing T_e. Thus $T_e = T_e^*(y_e)$ in any solution to x_e's modified problem, so $R(y_e, T_e) = G$. The latter implies that profiles $\{U(x), Y(x), \psi(x)\}_x$ induce maximum revenue and therefore solve CRM. Given $T_e = T_e^*(y_e)$, maximizing $u(y_e - T_e, 1 - y_e/x_e)$ means that y_e maximizes $u(y_e - T_e^*(y_e), 1 - y_e/x_e)$, as claimed. Non-negative leisure implies $0 \leq y_e \leq x_e$.

2. A solution exists because y_e maximizes a continuous function on the compact set $[0, x_e]$; the standard first-order necessary condition for an interior optimum is $\partial T_e^*/\partial y_e = (1 - \frac{u_l}{u_c x_e})$. (The assumption $\lim_{l \to 0} u_l/u_c = \infty$ rules out a corner solution

at $y_e = x_e$.) Solutions to x_e's modified problem satisfy $U(x_e) \geq U_e(x_e)$ because (7) holds for $x = x_e$. If $U(x_e) > U_e(x_e)$ then x_e could raise own utility by choosing $(Y(x_e), T(x_e))$ instead of (y_e, T_e), contradicting optimality. Thus $U(x_e) = U_e(x_e)$. Because $R(y_e, T_e^*(y_e)) = G < \hat{G}$, it must be that $(y_e, T_e^*(y_e)) \in \mathcal{B}$. From proposition 1, part 1a, $[x_1, x_2]$ is non-empty, $x_e \in [x_1, x_2]$, and $Y(x_e) = y_e$.

3. Any profiles $\{U(x), Y(x), \psi(x)\}_x$ that are feasible for x_e's original problem are feasible for the modified problem by taking $y_e = Y(x_e)$ and $T_e = T(U(x_e), y_e, x_e)$. Therefore the utility x_e obtains under the original problem is less than or equal to the utility x_e obtains in any solution to the modified problem. From part 2, any solution to the modified problem satisfies $Y(x_e) = y_e$ and $U(x_e) = U_e(x_e) = u(y_e - T_e, 1 - y_e/x_e)$, so $c^*(U(x_e), Y(x_e)/x_e) = y_e - T_e$, and (2) implies (7). All solutions to the modified problem therefore have profiles in the feasible set of the original problem, so any solution to the modified problem solves the original problem. Moreover, x_e's utility is equal under the modified and original problems, so any solution to x_e's original problem solves the modified problem.

Proposition 4 We compare $\{U(x \mid y_L)\}_x$ and $\{U(x \mid y_H)\}_x$ for given x by integrating $\partial U(x \mid y_e)/\partial y_e$ over $[y_L, y_H]$. At points x_e where y_e^* is single-valued, proposition 3 also implies that y_e^* has an inverse, denoted $x_e^*(y_e)$, that is single-valued, continuous, and strictly increasing. At points x_e where y_e^* is multi-valued, the Maximum Theorem implies that y_e^* is compact-valued so $\min\{y_e^*(x_e)\}$ and $\max\{y_e^*(x_e)\}$ exist; for $y_e \in [\min\{y_e^*(x_e)\}, \max\{y_e^*(x_e)]$, either $x_e^*(y_e) = x_e$ or else $x_e^*(y_e)$ is undefined.

We first consider subintervals of $[y_L, y_H]$ over which x_e^* is strictly increasing, then deal with subintervals over which x_e^* is constant or undefined; because each of the latter subintervals is associated with an x_e at which y_e^* is multi-valued, the latter subintervals have the form $[\min\{y_e^*(x_e)\}, \max\{y_e^*(x_e)\}]$.

Consider $y_e \in (y_L, y_H)$ at which x_e^* is strictly increasing. From proposition 1,

the solution to CRM conditional on $(y_e, T_e^*(y_e))$ has $U(x) = U_e(x)$ and $Y(x) = y_e$ on an interval $[x_1, x_2]$. Because $x_1 \leq x_e \leq x_2$ from proposition 2, part 2, and $x_- \leq x_e^*(y_L) < x_e = x_e^*(y_e) < x_e^*(y_H) \leq x_+$, proposition 1 (part 1a) implies $x_1 < x_2$.

We consider in sequence given x-values in the intervals $[x_1, x_2]$, $[x_2, x_+]$, and $[x_-, x_1]$:

(1) For $x \in [x_1, x_2]$, where $U(x \mid y_e) = U_e(x)$:

$$\begin{aligned} \partial U(x \mid y_e)/\partial y_e &= \partial U_e(x)/\partial y_e = [u_c - u_l/x] - u_c \cdot \partial T_e^*/\partial y_e \\ &= u_c \cdot [\tau(x|y_e, T_e) - \partial T_e^*/\partial y_e] = u_c \cdot [\tau(x|y_e, T_e) - \tau(x_e|y_e, T_e)], \end{aligned} \tag{18}$$

where the last step uses the first-order condition in proposition 2. Because the envelope theorem implies $\partial U(x_e \mid y_e)/\partial y_e = 0$ at $x_e = x_e^*(y_e)$ and agent monotonicity implies that $\tau(x|y_e, T_e)$ and hence $\partial U(x \mid y_e)/\partial y_e$ are strictly increasing in x, it follows that $\partial U(x \mid y_e)/\partial y_e < 0$ for $x \in [x_1, x_e)$ and $\partial U(x \mid y_e)/\partial y_e > 0$ for $x \in (x_e, x_2]$.

For use below, we need to show that

$$\partial U(x_1 \mid y_e)/\partial y_e < 0 \tag{19}$$

$$\partial U(x_2 \mid y_e)/\partial y_e > 0. \tag{20}$$

From above, these follow if $x_e \neq x_1$ and $x_e \neq x_2$. Differentiate the government budget $R(y_e, T_e^*(y_e)) = G$ totally to get $\partial R(T_e, y_e)/\partial y_e + \partial R(T_e, y_e)/\partial T_e \cdot \partial T_e^*(x_e)/\partial y_e = 0$ and insert $\partial R(T_e, y_e)/\partial y_e = \partial U_e(x)/\partial y_e[\xi_2(x_e) - \xi_1(x_e)] - \mu_{(1)}(x_e) + \mu_{(2)}(x_e)$, which is the value-function result analogous to (12), as well as $\partial R(T_e, y_e)/\partial T_e$ from (12) and $\partial T_e^*(x_e)/\partial y_e = \tau(x_e|y_e, T_e)$ from proposition 2, then apply (18) and reduce to $\mu_{(1)}(x_e) = \mu_{(2)}(x_e)$. If $x_e = x_1$ then the optimality conditions for CRM1 imply $\mu_{(1)}(x_1) = \mu_{(2)}(x_1) = 0$. For $x < x_1$, the construction of x_1 implies $Y(x) < y_e$ so $\mu_{(1)}(x) = 0$. Thus $\mu_{(1)}(x) = 0$ for $x \in [x_-, x_1]$, which implies $\mu_{(1),x}(x_1) = 0$ and hence from (8) that $H_{(1),Y} = 0$ at x_1. Then (17) and (8) imply $H_{(2),Y} = -\mu_{(2),x}(x_1) < 0$. Because $\mu_{(2)}(x_1) = 0$, it must be that $\mu_{(2)}(x) > 0$ for some $x > x_1$, contradicting the

optimality condition $\mu_{(2)}(x) \leq 0$. Thus $x_e \neq x_1$, which implies (19). The proof that $x_e \neq x_2$, which implies (20), is analogous.

(2) For $x \in [x_2, x_+]$, NB implies $\mu_x(x) = 0$ in (8) so $H_Y = 0$, where subscripts $_{(2)}$ for CRM2 are suppressed. Because $H_{YY} < 0$ by CON, the income level $Y \equiv y(U, \xi, x)$ that satisfies $H_Y(x, U, Y, \psi, \xi, 0) = 0$ is a well-defined function of (U, ξ, x) that is differentiable in (U, ξ). From (3) and (9), $U(x \mid y_e)$ and $\xi(x \mid y_e)$ then satisfy the differential equations

$$\begin{aligned} U_x(x \mid y_e) &= \omega(U(x \mid y_e), y(U(x \mid y_e), \xi(x \mid y_e), x), x) \\ \xi_x(x \mid y_e) &= -T_U(U(x \mid y_e), y(U(x \mid y_e), \xi(x \mid y_e), x), x) f(x) \\ &\quad -\xi(x \mid y_e) \cdot \omega_U(U(x \mid y_e), y(U(x \mid y_e), \xi(x \mid y_e), x), x). \end{aligned} \tag{21}$$

System (21) is saddle-path stable because the characteristic matrix has zero trace and, from CON, a negative determinant. Hence, a solution to (21) (which exists by CON) is uniquely determined by the two boundary conditions $U(x_2 \mid y_e) = U_e(x_2)$, and $\xi(x_+ \mid y_e) = 0$ if $x_+ < \infty$ or $\lim_{x \to \infty} \xi(x \mid y_e) = 0$ if $x_+ = \infty$. For given y_e, (21) and $Y = y(U, \xi, x)$ uniquely determine $\{U(x \mid y_e), Y(x \mid y_e)\}_x$ on $[x_2, x_+]$. Saddle-path stability implies that an increase in y_e that raises $U_e(x_2)$ must increase $U(x \mid y_e)$ for $x \geq x_2$. The derivative $U'(x) \equiv \partial U(x \mid y_e)/\partial y_e$ exists for $x \geq x_2$ and is found in two steps. First, differentiate (21) with respect to y_e to obtain[30]

$$\begin{aligned} \partial U'(x)/\partial x &= \gamma(U, \xi, x) \cdot U'(x) + \gamma_{U'\xi}(U, \xi, x) \cdot \xi'(x) \\ \partial \xi'(x)/\partial x &= \gamma_{\xi' U}(U, \xi, x) \cdot U'(x) - \gamma(U, \xi, x) \cdot \xi'(x), \end{aligned} \tag{22}$$

where $\xi'(x) \equiv \partial \xi(x \mid y_e)/\partial y_e$, $\gamma(U, \xi, x) = \omega_U - \omega_Y \cdot H_{YU}/H_{YY}$, $\gamma_{U'\xi}(U, \xi, x) = -\omega_{YY} > 0$ and $\gamma_{\xi' U}(U, \xi, x) = -[H_{YY} \cdot H_{UU} - H_{YU} \cdot H_{YU}]/H_{YY} > 0$. Second, solve (22) subject to the boundary conditions that $U'(x_2)$ is given and $\xi'(x_+) = 0$. (These are

[30] The notation $U'(x)$ highlights that these are ordinary differential equations in x at given y_e. Also, derivatives are one-sided at set boundaries here and below.

the derivatives of the boundary conditions for (21).) Because (21) has a characteristic matrix with zero trace and a negative determinant, it is saddle-path stable. This with $U'(x_2) > 0$ from (20) implies $U'(x) > 0$, or $\partial U(x \mid y_e)/\partial y_e > 0$ for $x \in [x_2, x_+]$.

(3) For $x \in [x_-, x_1]$, let $x_0(y_e) \equiv \inf\{x \geq x_- : Y(x \mid y_e) > 0\}$ denote the lower bound of values of x at which earnings are positive. CRM1 applies so we suppress subscripts $_{(1)}$ below. As in case (2) above, $U(x \mid y_e)$ and $\xi(x \mid y_e)$ satisfy (21) and (22) for $x \in [x_0, x_1]$. One boundary condition on (21) is $U(x_1 \mid y_e) = U_e(x_1)$. Because CRM1 is constrained by the inequalities $U(x_- \mid y_e) \geq u(\alpha, 1)$ and $Y(x_- \mid y_e) \geq 0$, each of which may hold with inequality or equality and each with zero or non-zero shadow values ($\xi(x_- \mid y_e) \leq 0, \mu(x_- \mid y_e) \leq 0$), the second boundary condition depends on which of the four inequalities hold as equalities. For any y_e, one or more of the following four cases apply:

(a) $x_0 = x_-$ and $\xi(x_- \mid y_e) = 0$: In this case (21) applies on $[x_-, x_1]$ and the second boundary condition is $\xi(x_- \mid y_e) = 0$. Also, $Y(x \mid y_e) > 0$ and $U(x \mid y_e) > u(\alpha, 1)$ for $x > x_-$, so $X_{\min}(y_e)$ has measure zero.

(b) $x_0 = x_-$ and $U(x_- \mid y_e) = u(\alpha, 1)$: In this case, (21) applies on $[x_-, x_1]$ with second boundary condition $U(x_- \mid y_e) = u(\alpha, 1)$. Again $Y(x \mid y_e) > 0$ and $U(x \mid y_e) > u(\alpha, 1)$ for $x > x_-$, with $X_{\min}(y_e) = \{x_-\}$.

(c) $Y(x_- \mid y_e) = 0$ and $\xi(x_- \mid y_e) = 0$: In this case, $x_0 \geq x_-$ is endogenous and $\{U'(x), \xi'(x)\}_x$ on $[x_0, x_1]$ is described by (21) with the second boundary condition given by requirements that $(U(x_0 \mid y_e), \xi(x_0 \mid y_e))$ satisfy $H_Y(U(x_0 \mid y_e), 0, \psi, \xi(x_0 \mid y_e), 0, x_0) = 0$ and that $\xi(x_0 \mid y_e)$ be determined by the differential equation $-\xi_x(x \mid y_e) = H_U(U(x_0 \mid y_e), 0, \psi, \xi(x \mid y_e), \mu, x)$. The latter has solution $\xi(x_0 \mid y_e) = F(x_0)/u_c(c^*(U(x_0 \mid y_e), 0), 1)$. Also, $U(x \mid y_e) = U(x_- \mid y_e)$ for $x \leq x_0$, so $X_{\min}(y_e) = \oslash$ if and only if $U(x_- \mid y_e) > u(\alpha, 1)$, and $X_{\min}(y_e) = [x_-, x_0]$ if and only if $U(x_-) = u(\alpha, 1)$.

(d) $Y(x_- \mid y_e) = 0$ and $U(x_-) = u(\alpha, 1)$: In this case, $x_0 \geq x_-$ is also

endogenous and (21) applies on $[x_0, x_1]$ with the second boundary condition given by requirements that $(U(x_0 \mid y_e), \xi(x_0 \mid y_e))$ satisfy $U(x_0 \mid y_e) = u(\alpha, 1)$ and $H_Y(U(x_0 \mid y_e), 0, \psi, \xi(x_0 \mid y_e), 0, x_0) = 0$. From $Y(x \mid y_e) = 0$ and $U(x \mid y_e) = U(x_- \mid y_e) = u(\alpha, 1)$ for $x \in [x_-, x_0]$, it follows that $X_{\min}(y_e) = [x_-, x_0]$.

Denote the sets of y_e-values for which CRM1 has a solution falling into case $j \in \{a, b, c, d\}$ above by $\mathcal{Y}(j) \equiv \{y_e |$ (21) has a solution in case j for given $U(x_1 \mid y_e)\}$. Because (21) is saddle-path stable, solutions are continuous in the boundary value $U(x_1 \mid y_e)$, and because the cases are defined by equality conditions, the sets $\mathcal{Y}(j)$ are closed.

For $y_e \in \mathcal{Y}(a)$, $\xi(x_- \mid y_e) = 0$ implies the second boundary condition $\xi'(x_-) = 0$. By logic analogous to that in case (2) above, $U'(x)$ exists for $x \leq x_1$ and is the solution to (22) with boundary conditions that $U'(x_1)$ is given and $\xi'(x_-) = 0$. Saddle-path stability and (19) then imply $U'(x) < 0$ for $x \in [x_-, x_1]$.

For $y_e \in \mathcal{Y}(b)$, $U(x_- \mid y_e) = u(\alpha, 1)$ implies the second boundary condition $U'(x_-) = 0$. As above, $U'(x)$ exists for $x \leq x_1$; in this case $U'(x)$ is the solution to (22) with boundary conditions that $U'(x_1)$ is given and $U'(x_-) = 0$. Saddle-path stability and (19) then imply $U'(x) < 0$ for $x \in (x_-, x_1]$.

For $y_e \in \mathcal{Y}(c)$, either $x_0 = x_-$, which implies $y_e \in \mathcal{Y}(a)$ so all results for $\mathcal{Y}(a)$ apply, or $x_0 = x_0(y_e) > x_-$. In the latter case, it is straightforward to show that x_0 is continuous in y_e and that $U(x \mid y_e)$ and $\xi(x \mid y_e)$ are differentiable in y_e on $[x_0, x_1]$. Given differentiability, $\{U'(x), \xi'(x)\}_x$ on $[x_0, x_1]$ is uniquely determined by (22) with boundary conditions that $U'(x_1)$ is given and $\xi'(x_0) = F(x_0)(-\frac{u_{cc}(c^*(U(x_0|y_e),0),1)}{u_c(c^*(U(x_0|y_e),0),1)^3}) \cdot U'(x_0)$ (which is the derivative of $\xi(x_0 \mid y_e) = F(x_0)/u_c(c^*(U(x_0 \mid y_e), 0), 1)$). Saddle-path stability and (19) then imply $U'(x) < 0$ for $x \in [x_0, x_1]$, so $U'(x_0) < 0$. This with $U'(x) = U'(x_-) < 0$ for $x \in [x_-, x_0]$ implies $U'(x) < 0$ for $x \in [x_-, x_0]$.

For $y_e \in \mathcal{Y}(d)$, one can show that x_0 is continuous in y_e and that $U(x \mid y_e)$ and $\xi(x \mid y_e)$ are differentiable in y_e on $[x_0, x_1]$. Because $U(x \mid y_e) = u(\alpha, 1)$ for $x \leq x_0$,

it follows that $U'(x) = 0$ for $x \in [x_-, x_0]$. Then $\{U'(x), \xi'(x)\}_x$ on $[x_0, x_1]$ is uniquely determined by (22) with boundary conditions that $U'(x_1)$ is given and $U'(x_0) = 0$. Saddle-path stability and (19) then imply $U'(x) < 0$ for $x \in (x_0, x_1]$ and $\xi'(x) < 0$ for $x \in [x_0, x_1]$. Moreover, $\partial Y(x_0 \mid y_e)/\partial y_e = (-H_{YY})^{-1}\omega_Y \xi'(x_0) < 0$ implies that x_0 increases strictly with y_e.

Now consider the ordering of the sets $\mathcal{Y}(j)$, $j = 1, 2, 3, 4$. Because $U'(x_-) < 0$ for $y_e \in \mathcal{Y}(a) \cup \mathcal{Y}(c)$ and $U(x_- \mid y_e) = u(\alpha, 1)$ for $y_e \in \mathcal{Y}(b) \cup \mathcal{Y}(d)$, it must be that all points in $\mathcal{Y}(a) \cup \mathcal{Y}(c)$ are less than all points in $\mathcal{Y}(b) \cup \mathcal{Y}(d)$ except that if both are nonempty, then $\mathcal{Y}(a) \cup \mathcal{Y}(c)$ and $\mathcal{Y}(b) \cup \mathcal{Y}(d)$ meet at the single point $\max\{y_e \in \mathcal{Y}(a) \cup \mathcal{Y}(c)\} = \min\{y_e \in \mathcal{Y}(b) \cup \mathcal{Y}(d)\}$. Moreover, $y_e \in \mathcal{Y}(c) \backslash \mathcal{Y}(d)$ implies $U(x_0 \mid y_e) > u(\alpha, 1)$ and hence $X_{\min}(y_e) = \oslash$, which implies that $X_{\min}(y_e)$ has positive measure only if $y_e \in \mathcal{Y}(d)$. Because x_0 is strictly increasing for $y_e \in \mathcal{Y}(d)$, it must be that all points in $\mathcal{Y}(b)$ are less than all points in $\mathcal{Y}(d)$ except that if both are nonempty, then $\mathcal{Y}(b)$ and $\mathcal{Y}(d)$ meet at the single point $\max\{y_e \in \mathcal{Y}(b)\} = \min\{y_e \in \mathcal{Y}(d)\}$. Because (i) $X_{\min}(y_e) = \oslash$ for $y_e < \max\{y_e \in \mathcal{Y}(a) \cup \mathcal{Y}(c)\}$, (ii) $X_{\min}(y_e) = \{x_-\}$ for $y_e \in \mathcal{Y}(b)$, and (iii) x_0 is strictly increasing in y_e on $\mathcal{Y}(d)$, it follows that $X_{\min}(y_1) \subset X_{\min}(y_2)$ for any $y_1 \leq y_2$.

Summarizing derivatives for $x \in [x_-, x_1]$: Because (i) $U'(x) < 0$ for $x \in [x_-, x_1]$ and any $y_e < \max\{y_e \in \mathcal{Y}(a) \cup \mathcal{Y}(c)\}$, (ii) $U'(x) < 0$ for $x \in (x_-, x_1]$ and any $y_e \in \mathcal{Y}(b)$, and (iii) $U'(x) < 0$ for $x \in (x_0, x_1]$ and any $y_e \in \mathcal{Y}(d)$, it follows for all y_e that $U'(x) < 0$ for $x \in [x_-, x_1] \backslash X_{\min}(y_e)$ and $U(x \mid y_e) = u(\alpha, 1)$ for $x \in X_{\min}(y_e)$.

Now consider any interval $[\min\{y_e^*(x_e)\}, \max\{y_e^*(x_e)\}]$ for $x_e \in [x_e^*(y_L), x_e^*(y_H)]$ with multi-valued y_e^*. Let $\tilde{U} \equiv U(x_e \mid \max\{y_e^*(x_e)\}) = U(x_e \mid \min\{y_e^*(x_e)\})$ be x_e's utility and let $\tilde{U} = u(y_e - T_e, 1 - y_e/x_e)$ implicitly define utility-compensating taxes $T_e = \tilde{T}(y_e)$. Let $\{\tilde{U}(x \mid y_e)\}_x$ denote utility profiles that solve CRM conditional on $(y_e, \tilde{T}(y_e))$. By construction, $\tilde{U}(x_e \mid y_e) = \tilde{U}$ for $y_e \in [\min\{y_e^*(x_e)\}, \max\{y_e^*(x_e)\}]$, so $\partial \tilde{U}(x_e \mid y_e)/\partial y_e = 0$. For profiles $\{\tilde{U}(x \mid y_e)\}_x$, derivations analogous to those in

(1)-(3) imply $\partial\tilde{U}(x \mid y_e)/\partial y_e \geq 0$ for $x > x_e$ and $\partial\tilde{U}(x \mid y_e)/\partial y_e \leq 0$ for $x < x_e$. Note that $\{\tilde{U}(x \mid y_e)\}_x = \{U(x \mid y_e)\}_x$ for $y_e \in y_e^*(x_e)$. For any $y_1, y_2 \in y_e^*(x_e)$ with $y_1 < y_2$, we have $U(x \mid y_2) - U(x \mid y_1) = \tilde{U}(x \mid y_2) - \tilde{U}(x \mid y_1) = \int_{y_1}^{y_2} \partial\tilde{U}(x \mid y_e)/\partial y_e dy_e$ for all x, so $U(x_e \mid y_2) - U(x_e \mid y_1) = 0$, $U(x \mid y_2) - U(x \mid y_1) \geq 0$ for $x > x_e$, and $U(x \mid y_2) - U(x \mid y_1) \leq 0$ for $x < x_e$. (The latter implies that $X_{\min}(y_1) \subset X_{\min}(y_2)$ also holds for any $y_1 \leq y_2$ taken from multi-valued $y_e^*(x_e)$.)

To compute $U(x \mid y_H) - U(x \mid y_L)$, integrate $\partial U(x \mid y_e)/\partial y_e$ over all subintervals of $[y_L, y_H]$ over which x_e^* is strictly increasing, integrate $\partial\tilde{U}(x \mid y_e)/\partial y_e$ over all subintervals of the form $[\min\{y_e^*(x_e)\}, \max\{y_e^*(x_e)\}]$ as in the preceding paragraph, and sum. Because $[y_L, y_H]$ contains at least one non-degenerate subinterval over which x_e^* is strictly increasing, we have $U(x \mid y_H) - U(x \mid y_L) > 0$ for $x \geq x_H$, $U(x \mid y_H) - U(x \mid y_L) = 0$ for $x \in X_{\min}(y_L)$, and $U(x \mid y_H) - U(x \mid y_L) < 0$ for $x \in [x_-, x_1] \backslash X_{\min}(y_L)$.

Because $U(x \mid y_H) - U(x \mid y_L)$ is continuous in x, the mean-value theorem ensures that there is at least one crossing point $x_\times \in (x_L, x_H)$ with $U(x \mid y_H) - U(x \mid y_L) = 0$. The uniqueness of $x_\times$ in $[x_L, x_H]$ follows from agent monotonicity.

Proposition 5 Voting decisions of individuals with $x \notin X_{\min}(y_L)$ follow from proposition 4 (parts 1-3). (Because $y_H > 0$, proposition 3 implies $y_L < y_H$.)

If $X_{\min}(y_L)$ has zero measure or if all $x \in X_{\min}(y_L)$ vote for x_L, then $U(x_M \mid y_L) > U(x_M \mid y_H)$ implies that $x_\times < x_M$, all voters in $(x_M, x_+]$ prefer x_L over x_H, and $(x_M, x_+]$ is a majority. The reverse applies if $U(x_M \mid y_L) < U(x_M \mid y_H)$.

If $X_{\min}(y_L)$ has positive measure and all $x \in X_{\min}(y_L)$ abstain, let x_m denote the median of $[x_-, x_+] \backslash X_{\min}(y_L)$. Then $U(x_m \mid y_L) > U(x_m \mid y_H)$ implies $x_\times < x_m$, so all voters in $(x_m, x_+]$ prefer x_L over x_H, and this is a majority. The reverse applies if $U(x_m \mid y_L) < U(x_m \mid y_H)$.

Proposition 6 Immediate from proposition 5 (parts 1 and 2a). Whenever x_M is a candidate, $x_\times \neq x_M$ rules out a tie.

Proposition 7 From (5), $T^*(Y(x)) = T(U(x), Y(x), x)$ determines T^* for all $y \in [Y(x_-), Y(x_+)]$. The continuity of U, Y, and T implies continuity of T^*. From CON, Y has a piecewise continuous derivative ψ, so $dT^*(Y(x))/dy \cdot \psi(x) = T_U(U, Y, x) \cdot dU/dx + T_Y(U, Y, x) \cdot \psi(x) + T_x$ except at points x where ψ is discontinuous. From (3) and (5), $T_U(U, Y, x) \cdot dU/dx + T_x = -c_U^* \cdot \omega(U, Y, x) + c_n^* \cdot Y/x^2 = (-1/u_c) \cdot u_l Y/x^2 + (u_l/u_c) \cdot Y/x^2 = 0$, so $dT^*(Y(x))/dy \cdot \psi(x) = T_Y(U, Y, x) \cdot \psi(x)$.

From NB, Y^{-1}exists for all $y \in [Y(x_-), Y(x_+)]$ except possibly $y = 0$ and $y = y_e$. Moreover, $\psi(Y^{-1}(y)) > 0$ wherever Y^{-1} is defined, and ψ is continuous, so $dT^*(Y(x))/dy = T_Y(U, Y, x) = \tau(x)$ and hence $dT^*(y)/dy = T_Y(U, Y, x) = \tau(Y^{-1}(y))$. For any $y_d \in [Y(x_-), Y(x_+)] \backslash \{0,\ y_e\}$ at which $\psi(Y^{-1}(y_d))$ is discontinuous, $dT^*(y)/dy = \tau(Y^{-1}(y))$ in a neighborhood of y_d (excluding y_d), so $dT^*(y_d)/dy = lim_{y \to y_d} dT^*(y)/dy = \tau(Y^{-1}(y_d))$ is well-defined. Thus (11) holds for all $y \in [Y(x_-), Y(x_+)] \backslash \{0,\ y_e\}$. If $0 \in [Y(x_-), Y(x_+)]$, then $dT^*(0)/dy \equiv lim_{y \to 0} dT^*(y)/dy = \tau(\max X_0)$ is well-defined. Therefore T^* is continuously differentiable on $[Y(x_-), Y(x_+)] \backslash \{y_e\}$. (Because $Y^{-1}(0)$ is undefined if X_0 is an interval, (11) may not hold at $y = 0$.)

1. On $[x_1, x_2]$, $\tau(x) = \tau(x|y_e, T_e)$, which is strictly increasing in x from agent monotonicity. Thus $\tau(x_2) > \tau(x_1)$. The one-sided limits follow from the continuity of $dT^*(y)/dy = \tau(Y^{-1}(y))$ for $y < y_e$ in a neighborhood of y_e and $y > y_e$ in a neighborhood of y_e.

2. Monotonicity of Y implies $Y(x) \geq y_e$. Because $(y_e, T_e) \in \mathcal{B}$ by proposition 2, part 1b of proposition 1 requires that $\{U(x), Y(x)\}_x$ solve CRM2 on $[x_e, x_+]$ and CRM1 on $[x_-, x_e]$. Integrating the Euler equation (9):

$$\xi_{(2)}(x) = -\int_x^{x_+} \frac{\eta(x, z)}{u_c(c^*(U(z), Y(z)/z), 1 - Y(z)/z)} dF(z) + \xi_{(2)}(x_+), \qquad (23)$$

where $\eta(x, z) \equiv exp\{\int_x^z \omega_U(U(\tilde{z}), Y(\tilde{z}), \tilde{z}) d\tilde{z}\} > 0$. The transversality condition for

CRM2 is $\xi_{(2)}(x_+) = 0$, which implies $\xi_{(2)}(x) < 0$ on $[x_e, x_+)$. In (8) for CRM2, NB implies $\mu_x(x) = 0$ for $x \geq x_2$, so $\tau(x) = 1 - \frac{u_l}{u_c x} = -[\omega_Y(U, Y, x)/f(x)]\xi_{(2)}(x)$. It is straightforward to show that agent monotonicity implies $\omega_Y > 0$. Because $f > 0$ and $\xi_{(2)}(x) < 0$, it follows that $\tau(x) > 0$ on $[x_2, x_+)$, and because $\xi_{(2)}(x_+) = 0$, it follows that $\tau(x_+) = 0$ if $x_+ < \infty$.

3. Monotonicity of Y implies $Y(x) \leq y_e$. Integrating the Euler equation (9):

$$\xi_{(1)}(x) = \int_{x_-}^{x} \frac{\eta(x,z)}{u_c(c^*(U(z), Y(z)/z), 1 - Y(z)/z)} dF(z) + \xi_{(1)}(x_-). \tag{24}$$

Because (1) holds if and only if $U(x_-) \geq u(\alpha, 1)$, transversality conditions for CRM1 are $[U(x_-) - u(\alpha, 1)] \cdot \xi_{(1)}(x_-) = 0$ and $\xi_{(1)}(x_-) \leq 0$. Because $\eta/u_c > 0$, $\xi_{(1)}$ is strictly increasing. Thus there are three possibilities: $\xi_{(1)} < 0$ for all $x < x_1$; $\xi_{(1)}$ switches sign from negative to positive at a point $x_\tau \in (x_-, x_1)$ where $\xi_{(1)}(x_\tau) = 0$; or $\xi_{(1)} \geq 0$ for all $x < x_1$. As above, (8) and NB imply that $\xi_{(1)}(x)$ and $\tau(x)$ have opposite signs on $[x_0, x_1]$. If $\xi_{(1)} < 0$ for $x < x_1$ then $\tau(x) > 0$ on $[x_0, x_1)$ so T^* is strictly increasing; this is case (c). If $\xi_{(1)}(x_\tau) = 0$ for some $x_\tau \in (x_-, x_1)$, the shape of T^* depends on x_τ and x_0. If $x_\tau > x_0$, (8) and NB imply $\tau(x_\tau) = 0$, $\tau(x) > 0$ on (x_0, x_τ), and $\tau(x) < 0$ on (x_τ, x_1); this is the inverted U-shaped case (b). Finally, if $x_\tau < x_0$ or if $\xi_{(1)} \geq 0$ for $x < x_1$, (8) and NB imply $\tau(x) < 0$ on $[x_0, x_1)$, so T^* is strictly increasing; this is case (a). If (1) does not bind ($\xi_{(1)}(x_-) = 0$) then (24) implies $\xi_{(1)} > 0$ on $(x_-, x_1]$.

Proposition 8 The Euler equations for URM imply that the costate variable associated with $\hat{U}$, denoted $\hat{\xi}$, satisfies (23) for all x. Because $\eta/u_c > 0$ and $\hat{\xi}(x_+) = 0$, it follows that $\hat{\xi}(x) < 0$ for $x < x_+$. By (8), $\hat{\tau}(x) > 0$ for $x < x_+$ and hence for $x \in [x_-, x_e]$ in the solution to URM. As $\{U(x), Y(x)\} \to \{\hat{U}(x), \hat{Y}(x)\}_x$, we have $\tau(x) \to \hat{\tau}(x)$ pointwise. Since $[x_-, x_e]$ is compact, the convergence $\tau(x) \to \hat{\tau}(x) > 0$ is uniform on $[x_-, x_e]$. Hence there is a neighborhood of $(\hat{G}(\alpha), \alpha)$ such that $\tau(x) > 0$ for all $[x_-, x_e]$. Because $\tau(x) \geq \tau(x_e) > 0$ on $[x_e, x_2]$ and $\tau(x) > 0$ for $x_2 \in [x_2, x_+)$, it follows that $\tau(x) > 0$ for $x \in [x_-, x_+)$.

FIG. 1.--Net income tax functions.

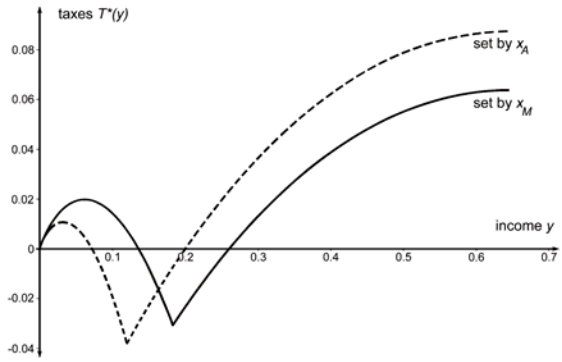


FIG. 2.--Profiles of taxes.

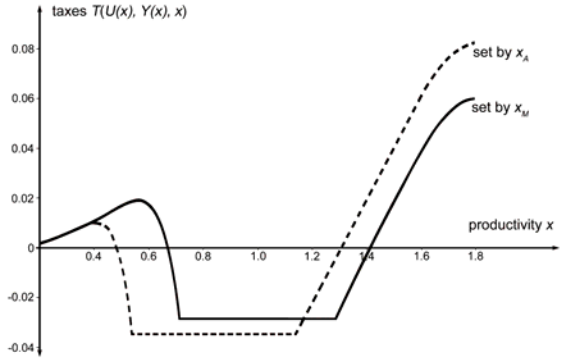


FIG. 3.--Induced utility profiles.

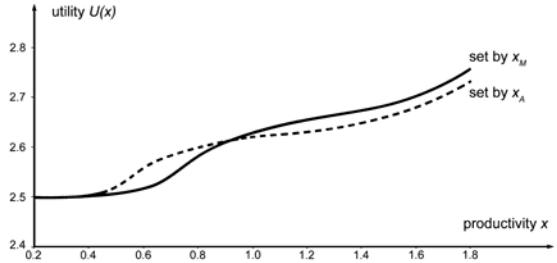


FIG. 4.--Comparison of linear and non-linear taxes set by x_M.

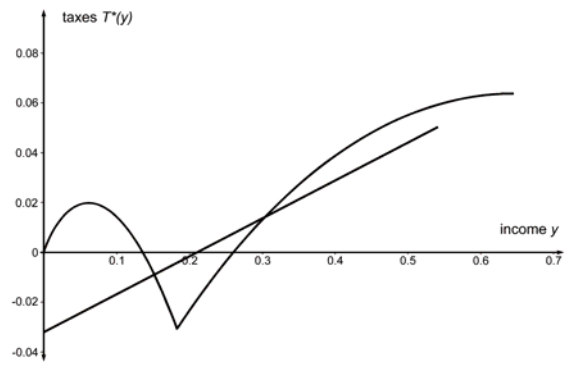


FIG. 5.--Tax function set by x_M at different G.

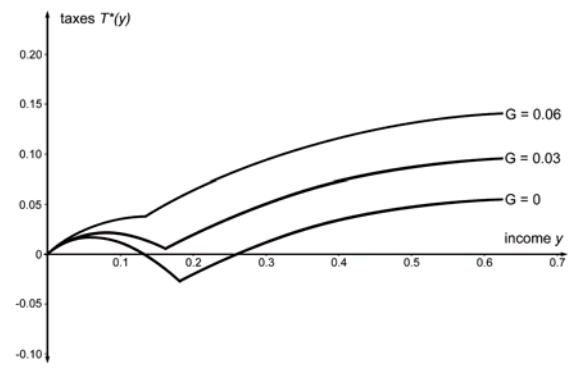


FIG. 6.--Tax function set by x_M at different α.

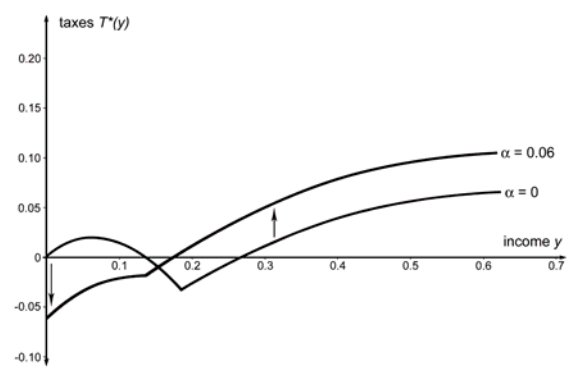


FIG. 7.--The election winner's income choice.

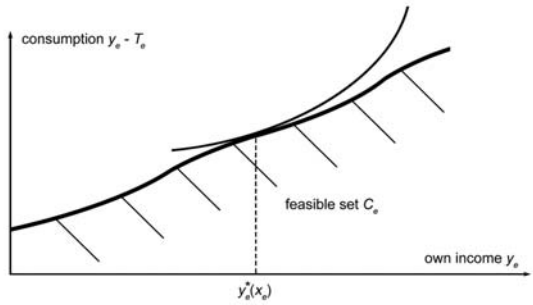


FIG. 8.--Income choices of x_L and x_H.

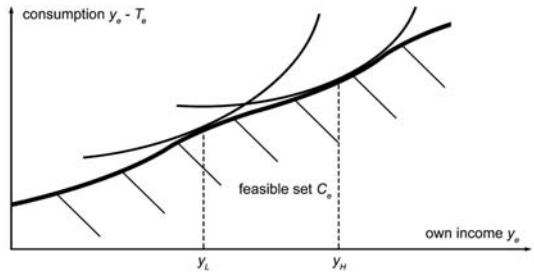
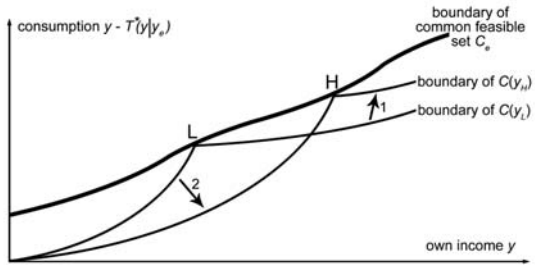


FIG. 9.--How individual income and consumption vary with y_e.



CESifo Working Paper Series

(for full list see www.cesifo.de)

---

994 Louis N. Christofides and Chen Peng, Contract Duration and Indexation in a Period of Real and Nominal Uncertainty, July 2003

995 M. Hashem Pesaran, Til Schuermann, Björn-Jakob Treutler, and Scott M. Weiner, Macroeconomic Dynamics and Credit Risk: A Global Perspective, July 2003

996 Massimo Bordignon and Sandro Brusco, On Enhanced Cooperation, July 2003

997 David F. Bradford, Addressing the Transfer-Pricing Problem in an Origin-Basis X Tax, July 2003

998 Daniel Gros, Who Needs Foreign Banks?, July 2003

999 Wolfram Merzyn and Heinrich W. Ursprung, Voter Support for Privatizing Education: Evidence on Self-Interest and Ideology, July 2003

1000 Jo Thori Lind, Fractionalization and the Size of Government, July 2003

1001 Daniel Friedman and Donald Wittman, Litigation with Symmetric Bargaining and Two-Sided Incomplete Information, July 2003

1002 Matthew Clarke and Sardar M. N. Islam, Health Adjusted GDP (HAGDP) Measures of the Relationship Between Economic Growth, Health Outcomes and Social Welfare, July 2003

1003 Volker Grossmann, Contest for Attention in a Quality-Ladder Model of Endogenous Growth, August 2003

1004 Marcel Gérard and Joan Martens Weiner, Cross-Border Loss Offset and Formulary Apportionment: How do they affect multijurisdictional firm investment spending and interjurisdictional tax competition ?, August 2003

1005 Burkhard Heer, Nonsuperneutrality of Money in the Sidrauski Model with Heterogeous Agents, August 2003

1006 V. Anton Muscatelli, Piergiovanna Natale, and Patrizio Tirelli, A Simple and Flexible Alternative to the Stability and Growth Pact Deficit Ceilings. Is it at hand?, August 2003

1007 Reto Foellmi and Josef Zweimüller, Inequality and Economic Growth: European Versus U.S. Experiences, August 2003

1008 James S. Costain and Michael Reiter, Business Cycles, Unemployment Insurance, and the Calibration of Matching Models, August 2003

1009 Marco Runkel, Optimal Contest Design when the Designer's Payoff Depends on Competitive Balance, August 2003

1010 Donald O. Parsons, Torben Tranaes and Helene Bie Lilleør, Voluntary Public Unemployment Insurance, August 2003

1011 Rüdiger Pethig and Andreas Wagener, Profit Tax Competition and Formula Apportionment, August 2003

1012 Johan Willner, Privatisation and Public Ownership in Finland, August 2003

1013 Seppo Kari and Jouko Ylä-Liedenpohja, Taxation and Valuation of International Real Investments, August 2003

1014 James Heckman, Rosa Matzkin and Lars Nesheim, Simulation and Estimation of Hedonic Models, August 2003

1015 Biswa N. Bhattacharyay, Towards a Macro-Prudential Leading Indicators Framework for Monitoring Financial Vulnerability, August 2003

1016 J. Stephen Ferris and Stanley L. Winer, Searching for Keynes: With Application to Canada, 1870-2000, August 2003

1017 Massimo Bordignon, Luca Colombo and Umberto Galmarini, Fiscal Federalism and Endogenous Lobbies' Formation, August 2003

1018 Annette Alstadsæter, The Dual Income Tax and Firms' Income Shifting through the Choice of Organizational Form and Real Capital Investments, August 2003

1019 Peter Fredriksson and Bertil Holmlund, Optimal Unemployment Insurance Design: Time Limits, Monitoring, or Workfare?, August 2003

1020 Kashif S. Mansori, Following in their Footsteps: Comparing Interest Parity Conditions in Central European Economies to the Euro Countries, August 2003

1021 Christoph Borgmann and Matthias Heidler, Demographics and Volatile Social Security Wealth: Political Risks of Benefit Rule Changes in Germany, August 2003

1022 Kjell Erik Lommerud, Bjørn Sandvik and Odd Rune Staume, Good Jobs, Bad Jobs and Redistribution, August 2003

1023 Patrick Karl O'Brien, The Governance of Globalization: The Political Economy of Anglo-American Hegemony, 1793-2003, September 2003

1024 Antonio Ciccone and Giovanni Peri, Skills' Substitutability and Technological Progress: U.S. States 1950-1990, September 2003

1025 Bjørn Sandvik, Optimal Taxation and Normalisations, September 2003

1026 Massimo Bordignon and Gilberto Turati, Bailing Out Expectations and Health Expenditure in Italy, September 2003

1027 José A. Herce, Namkee Ahn, Ricard Génova, and Joaquín Pereira, Bio-Demographic and Health Aspects of Ageing in the EU, September 2003

1028 John Komlos and Marieluise Baur, From the Tallest to (One of) the Fattest: The Enigmatic Fate of the American Population in the 20th Century, September 2003

1029 Stefan Napel and Mika Widgrén, Bargaining and Distribution of Power in the EU's Conciliation Committee, September 2003

1030 Kai Li and Dale J. Poirier, Relationship Between Maternal Behavior During Pregnancy, Birth Outcome, and Early Childhood Development: An Exploratory Study, September 2003

1031 Ivar Ekeland, James J. Heckman, and Lars Nesheim, Identifcation and Estimation of Hedonic Models, September 2003

1032 Kjetil Bjorvatn and Alexander W. Cappelen, Decentralization and the Fate of Minorities, September 2003

1033 Lars-Erik Borge and Jørn Rattsø, The Relationships Between Costs and User Charges: The Case of a Norwegian Utility Service, September 2003

1034 Maureen Were and Nancy N. Nafula, An Assessment of the Impact of HIV/AIDS on Economic Growth: The Case of Kenya, September 2003

1035 A. Lans Bovenberg, Tax Policy and Labor Market Performance, September 2003

1036 Peter Birch Sørensen, Neutral Taxation of Shareholder Income: A Norwegian Tax Reform Proposal, September 2003

1037 Roberta Dessi and Sheilagh Ogilvie, Social Capital and Collusion: The Case of Merchant Guilds, September 2003

1038 Alessandra Casarico and Carlo Devillanova, Capital-skill Complementarity and the Redistributive Effects of Social Security Reform, September 2003

1039 Assaf Razin and Efraim Sadka, Privatizing Social Security Under Balanced-Budget Constraints: A Political-Economy Approach, September 2003

1040 Michele Moretto, Paolo M. Panteghini, and Carlo Scarpa, Investment Size and Firm's Value under Profit Sharing Regulation, September 2003

1041 A. Lans Bovenberg and Peter Birch Sørensen, Improving the Equity-Efficiency Trade-off: Mandatory Savings Accounts for Social Insurance, September 2003

1042 Bas van Aarle, Harry Garretsen, and Florence Huart, Transatlantic Monetary and Fiscal Policy Interaction, September 2003

1043 Jerome L. Stein, Stochastic Optimal Control Modeling of Debt Crises, September 2003

1044 Thomas Stratmann, Tainted Money? Contribution Limits and the Effectiveness of Campaign Spending, September 2003

1045 Marianna Grimaldi and Paul De Grauwe, Bubbling and Crashing Exchange Rates, September 2003

1046 Assar Lindbeck and Dennis J. Snower, The Firm as a Pool of Factor Complementarities, September 2003

1047 Volker Grossmann, Firm Size and Diversification: Asymmetric Multiproduct Firms under Cournot Competition, September 2003

1048 Dan Anderberg, Insiders, Outsiders, and the Underground Economy, October 2003

1049 Jose Apesteguia, Steffen Huck and Jörg Oechssler, Imitation – Theory and Experimental Evidence, October 2003

1050 G. Abío, G. Mahieu and C. Patxot, On the Optimality of PAYG Pension Systems in an Endogenous Fertility Setting, October 2003

1051 Carlos Fonseca Marinheiro, Output Smoothing in EMU and OECD: Can We Forego Government Contribution? A Risk Sharing Approach, October 2003

1052 Olivier Bargain and Nicolas Moreau, Is the Collective Model of Labor Supply Useful for Tax Policy Analysis? A Simulation Exercise, October 2003

1053 Michael Artis, Is there a European Business Cycle?, October 2003

1054 Martin R. West and Ludger Wößmann, Which School Systems Sort Weaker Students into Smaller Classes? International Evidence, October 2003

1055 Annette Alstadsaeter, Income Tax, Consumption Value of Education, and the Choice of Educational Type, October 2003

1056 Ansgar Belke and Ralph Setzer, Exchange Rate Volatility and Employment Growth: Empirical Evidence from the CEE Economies, October 2003

1057 Carsten Hefeker, Structural Reforms and the Enlargement of Monetary Union, October 2003

1058 Henning Bohn and Charles Stuart, Voting and Nonlinear Taxes in a Stylized Representative Democracy, October 2003