Let's Pay Lower Taxes: Revenue Extraction by Median Voters

Henning Bohn and Charles Stuart*

May 8, 2012

Abstract

People prefer paying lower taxes. Lower taxes at one income can be financed by higher taxes at other incomes, which we term revenue extraction. We study revenue extraction when taxpayers elect representatives who set incentive-compatible tax policy and a minimum-utility constraint limits what can be taken from the poor. Revenue extraction by median-income voters is then a Condorcet outcome, and resulting policy broadly resembles U.S. policy: taxes are progressive, the poor receive subsidies but face high effective marginal tax rates, and high-income taxpayers pay most taxes.

*Correspondence to bohn@econ.ucsb.edu, stuart@econ.ucsb.edu; Department of Economics, University of California, Santa Barbara, CA 93106. Earlier title: "Voting and Nonlinear Taxes in a Stylized Representative Democracy." We thank Marcus Berliant, Avinash Dixit, Antonio Rangel, Efraim Sadka, and seminar participants at Stanford University, the Econometric Society Summer Meetings, and CESifo. JEL codes D7, H2. Keywords: revenue extraction, median-voter theorem, Condorcet winner, Director's Law, non-linear tax, voting, Mirrlees model, representative democracy, income tax.

I. INTRODUCTION

Tax revenue in the U.S. comes mainly from people with high income. In 2007, for instance, the top 40 percent of U.S. households paid 85.4 percent of total Federal taxes and 98.7 percent of Federal individual income taxes—see table 1. We describe such a pattern as a median-voter outcome in which middle-income voters gain by imposing high taxes on high-income taxpayers and using the extracted revenue to keep taxes low at middle incomes.[1]

Table 1—Income and Federal Tax Shares by Household Income Quintile, 2007

	lowest	2nd	middle	4th	highest	lower 60%
Pretax Income	4.0	8.4	13.1	19.3	55.9	25.5
All Federal Taxes	0.8	4.4	9.2	16.5	68.9	14.4
Individual Income Taxes	-3.0	-0.3	4.6	12.7	86.0	1.3

Source: Congressional Budget Office (2010).

To model how voting pressure from the middle can lead to revenue extraction, we make two stylized assumptions:

1. Individuals differ only by productivity, so voting is over a single dimension. This helps avoid cycling, which would obscure tendencies toward a median-voter outcome.[2]

[1] The argument that the middle classes benefit disproportionately from some government spending programs because of political clout is sometimes known as "Director's Law" (Stigler, 1970). Analyses of programs consistent with Director's Law are in Gouveia (1997) and Dixit and Londregan (1998).

[2] Linearity is sometimes imposed to reduce dimensionality and rule out cycling (Romer, 1975; Roberts, 1979; Meltzer and Richard, 1981; Krusell and Ríos-Rull, 1999) but this does not work here because revenue extraction is fundamentally non-linear.

2. To capture electoral competition between different productivities, we assume voters elect one of two exogenously given candidates, each of whom if elected would set policy to maximize the utility of individuals with a single known productivity.[3] Having elections between candidates representing two single productivity values also helps avoid cycling, which would sometimes occur if politicians were modeled more generally as maximizing functions of arbitrary coalitions of individuals with different productivities. For brevity, a *candidate's* productivity, income, taxes, or utility here means the productivity, income, taxes, or utility of the *constituents* the candidate represents.

Policy is an incentive-compatible income-tax function without linearity or other shape restrictions, as in Mirrlees (1971). Net revenue must cover given spending on public goods. Forced labor is not allowed and there is a limit to how little a taxpayer can be left to consume, which means policy must also satisfy minimum-utility constraints.[4] Subject to these constraints, the election winner gains from a tax system that extracts revenue from others to give a low tax at the own income.

When voters compare two candidates, they see the income tax functions the candidates would set and vote for the candidate whose tax function would provide greater utility. We provide conditions under which all individuals with productivity

[3]This follows the representative-democracy literature (Osborne and Slivinski, 1996; Besley and Coate, 1997). Without modeling the underlying reputational process, we take as a starting point that trust in a politician's position matters and changing positions is costly, so politicians never deviate from maximizing the utility of the individuals they represent. An alternative (possibly less empirically plausible) interpretation would be to assume that a politician's personal productivity is the same as the productivity of the politician's constituents, so that once in office, the politician can be counted on to maximize constituents' utilities.

[4]Röell (1996) provides a median-voter result with a non-binding minimum-utility constraint when utility is linear in income and government has no revenue requirements. In a different context, Martimort (2001) applies optimal taxation in a political-economy model.

below a crossover vote for the candidate with lower productivity, and all individuals above the crossover vote for the candidate with higher productivity. This implies that median-productivity voters are always on the winning side. Thus a median-productivity candidate (one who would maximize the utility of median-productivity individuals) is a Condorcet winner, and there is broad voting pressure toward the median-voter outcome.[5]

The tax function a candidate with median or close-to-median productivity would set has broad features of the U.S. tax system:

1. Maximum revenue is extracted from higher-income individuals. This is consistent with data in table 1.

2. Taxes are progressive in the middle of the income distribution, consistent with the graduated brackets of the income tax.

3. Revenue extraction from lower-income individuals is restricted by the minimum-utility constraint. Under reasonable conditions, low-income individuals receive welfare (negative overall taxes) but face high marginal taxes. That welfare recipients face high marginal taxes because of means-testing is well-documented.[6]

The analysis is consistent with the positions taken by both major U.S. political parties in recent years. The current Democratic administration has pushed increasing taxes for those earning above about $250,000/year, and the previous Republican administration pushed middle-class tax cuts. That roughly half of households have

[5] As is common, we take the candidates and their positions as exogenously given and leave unmodelled the processes by which individuals become candidates and get on the ballot. That a candidate must attract the votes of median-productivity individuals to win also means candidates may benefit from becoming advocates for middle voters in these unmodelled processes.

[6] e.g. Browning and Johnson (1979), Dickert et al. (1995), Keane and Moffitt (1998).

zero Federal income-tax liability[7] is also consistent with a tax system in which middle-income voters are themselves "tax-preferred."

Section II describes the model. Section III describes how revenue-extracting income taxes are optimal for an election winner. Section IV studies elections and provides median-voter results. Section V studies the shape of the winner's tax function in more detail. Proofs are in an appendix (included here for reviewers, not for publication).

II. MODEL

Individuals have identical preferences over consumption $c \geq 0$ and leisure $0 \leq l \leq 1$ but differ in productivity x. Productivity has distribution $F(x)$ on an interval $[x_-, x_+]$ with $0 \leq x_- < x_+ \leq \infty$, where F has finite mean and continuous density f with $f > 0$ on (x_-, x_+). Individual income is $y \equiv nx$ where $n \equiv 1 - l$ is labor. The government can tax income but not leisure or productivity. An individual consumes $c \equiv y - T(y)$ where $T(y)$ is net income taxes.

Income taxes T are determined by the winner of a majority-rule election between two exogenously given candidates. The winner sets policy to maximize the utility of individuals with productivity denoted x_e; the two candidates have different values of x_e. Individuals first vote for the candidate whose policy would provide greater utility; then with T set by the winner, they choose income y (or equivalently labor supply n) to maximize utility $u(c, l) = u(y - T(y), 1 - \frac{y}{x})$. We make standard assumptions on u, including agent-monotonicity.[8]

[7] In 2009, 51 percent of households had zero or negative Federal tax liability (Joint Committee on Taxation, 2011).

[8] Namely, u is strictly increasing, strictly concave, and three times differentiable. To ensure $l > 0$, we assume $\lim_{l \to 0} u_l(c, l)/u_c(c, l) = \infty$ for any $c > 0$, where subscripts denote partial derivatives. We make the agent-monotonicity assumption that $nu_l(c, 1 - n)/u_c(c, 1 - n)$ increases strictly in n for all $c > 0$. This ensures that individual income choices do not decrease with productivity. Agent

As in Mirrlees (1971), income taxes T are obtained from the solution to a control problem with incentive and other constraints. The controls are a profile $\{U(x), Y(x)\}$ of functions on $[x_-, x_+]$ that specify distributions of utility and income across individuals.[9] This profile determines profiles of: (i) labor $n = \frac{Y}{x}$; (ii) consumption $c = c^*(U, n)$, where $c^*(U, n)$ defined by $U = u(c^*, 1 - n)$ is the consumption that provides utility U given labor supply n; and (iii) taxes $t(U, Y, x) \equiv Y - c^*(U, \frac{Y}{x})$. Taxes t imply income taxes $T(Y(x)) \equiv t(U(x), Y(x), x)$ for all incomes in $[Y(x_-), Y(x_+)]$.

Mirrlees (1971, 1986) shows that the incentive constraints are captured by the combination of differential equations

$$U(\tilde{x}) - U(x_-) = \int_{x_-}^{\tilde{x}} \omega(U(x), Y(x), x)dx \tag{1}$$

for $\tilde{x} \geq x_-$ where $\omega(U, Y, x) \equiv u_l(c^*(U(x), \frac{Y(x)}{x}), 1 - \frac{Y(x)}{x})\frac{Y(x)}{x^2} \geq 0$, plus the requirement that $Y(x)$ be non-decreasing.[10] To ensure this income-monotonicity, we use the income derivative $\psi(x) \equiv \frac{dY}{dx}$ as a control, impose

$$\psi(x) \geq 0 \tag{2}$$

for all x, and treat $Y(x)$ as a state variable.[11] The incentive constraints (1) with $\omega \geq 0$ imply that $U(x)$ is also non-decreasing.

Tax revenue must cover exogenous public-good spending $G \geq 0$ so the government budget constraint is

$$\int_{x_-}^{x_+} t(U(x), Y(x), x)dF(x) \geq G, \tag{3}$$

monotonicity holds if consumption is normal or utility is separable.

[9] A profile here means a function or collection of functions on $[x_-, x_+]$. We omit braces when referring to single-function profiles. If the functions are defined only on a subset $S \subset [x_-, x_+]$, the profile is denoted $\{\}_S$.

[10] An additional requirement from Mirrlees is $Y(x) < x$, but this holds under assumptions made on u in footnote 8.

[11] The approach follows Brito and Oakland (1977) and Ebert (1992).

where the integral is total tax revenue.

Policy must also satisfy the minimum-utility constraint

$$U(x_-) \geq u(\alpha, 1), \tag{4}$$

where $\alpha \geq 0$ and $u(\alpha, 1) > -\infty$.[12] If $\alpha = 0$, (4) says the government cannot make anyone worse off than a person who consumes zero and does not work; this assumes a person always has the option of dropping out of the organized economy, not working and consuming nothing. The minimum-utility constraint is needed to rule out forced labor: without (4) when $\alpha = 0$, government could leave individuals with utility below $u(0, 1)$, which requires $l < 1$ and would mean people are forced to work in return for zero consumption.[13] The (reasonable) case with a positive consumption floor α allows dropouts to earn and consume positive income outside the organized economy.[14]

An election winner $x_e \in [x_-, x_+]$ maximizes $U(x_e)$ subject to incentive (1, 2), budget (3), and minimum-utility (4) constraints by choice of $\{U(x), Y(x), \psi(x)\}$. To ensure that a solution exists, we assume α and G satisfy $G < \hat{R}(\alpha)$ where $\hat{R}(\alpha)$ is the maximum revenue that can be raised with incentive-compatible tax functions satisfying the minimum-utility constraint for given α; formally $\hat{R}(\alpha)$ is the value of the revenue integral in:

$$\text{maximize} \int_{x_-}^{x_+} t(U(x), Y(x), x) dF(x) \tag{$\widehat{\text{RM}}$}$$

$$\text{subject to (1), (2), and (4)}$$

[12] Because $U(x)$ is non-decreasing, (4) ensures $U(x) \geq u(\alpha, 1)$ for all x. Thus if the minimum-utility constraint binds, it binds for those with lowest productivity.

[13] The minimum-utility constraint introduces a fundamental no-forced-labor limit on what can be extracted from the poor, and provides an essential balancing pressure here on governments looking to extract revenue. Such a constraint is not usually included in Mirrlees-type models in which the policy objective is to help the poor.

[14] An alternative interpretation is that all individuals have altruistic preferences defined over the minimal consumption in society, and all prefer a consumption floor of α over any other floor.

by choice of $\{U(x), Y(x), \psi(x)\}$ for given α. The Hamiltonian for $\widehat{\mathrm{RM}}$ is

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

where ξ and μ are the costates for U and Y. We assume H is strictly concave in (U, Y) and that a solution to $\widehat{\mathrm{RM}}$ exists with continuous U and Y, piecewise continuous ψ, and continuous and piecewise continuously differentiable ξ and μ.[15] Under these assumptions, the profile that solves $\widehat{\mathrm{RM}}$, $\{\hat{U}(x), \hat{Y}(x)\}$, is unique.[16]

For any candidate x_e, there is a utility profile that would result if the candidate were to win. For any pair of candidates, the winner is the candidate who would give higher utility to a majority.

III. REVENUE-EXTRACTING TAX FUNCTIONS

A generic winner's constituents end up earning some income $y_e \geq 0$ and paying taxes $T(y_e)$, so they end up with utility $u(y_e - T(y_e),\ y_e/x_e)$. Given y_e, the winner maximizes this utility by imposing a tax function that extracts maximal revenue from all taxpayers to make $T(y_e)$ as low as possible. This makes income y_e maximally tax-preferred.

To characterize the income tax functions that extract maximal revenue and make $T(y_e)$ as low as possible, consider an arbitrary y_e and a specific value T_e of the tax payment at y_e. The tax payment T_e is feasible if and only if, given (y_e, T_e),

[15]These assumptions ensure that the conditions of the Maximum Principle describe an optimum. We do not look for conditions on primitives that imply the assumptions, or that imply assumptions CON and LB below. Simple conditions on primitives are elusive: Mirrlees (1986, p. 1235) states that "obscure" existence conditions that restrict third-order partials are unavoidable in variational problems of this type.

[16]If ψ has points of discontinuity, there is a trivial multiplicity in $\psi(x)$ because ψ can be altered at such points without altering $\{U(x), Y(x)\}$. Hence we call a solution unique if it has a unique profile $\{U(x), Y(x)\}$.

the maximum revenue $R(y_e, T_e)$ that can be extracted from all taxpayers by varying $\{U(x), Y(x), \psi(x)\}$ satisfies $R(y_e, T_e) \geq G$.

Because others with sufficient productivity to earn y_e cannot be stopped from also earning y_e and paying taxes T_e, the profile underlying $R(y_e, T_e)$ must give each individual $x \geq y_e$ at least the utility $U_e(x) \equiv u(y_e - T_e, 1 - y_e/x)$ that x would get from earning y_e and paying T_e:

$$U(x) \geq U_e(x). \tag{5}$$

If y_e is tax-preferred, a set of individuals may opt to earn y_e, that is, different x may "bunch" at income y_e. Formally, $R(y_e, T_e)$ is then the maximized value of the revenue integral in:

$$\text{maximize} \int_{x_-}^{x_+} t(U(x), Y(x), x) dF(x) \tag{RM}$$
$$\text{subject to (1), (2), (4), and (5)}$$

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

Some (y_e, T_e) pairs are uninteresting in that $U_e(x) \leq \hat{U}(x)$ for all x, so no one would be better off picking (y_e, T_e) than they would be under $\widehat{\text{RM}}$. For such pairs, it will turn out that: the solution to $\widehat{\text{RM}}$ also solves RM so $R(y_e, T_e) = \hat{R}(\alpha) > G$, which says revenue exceeds required spending on public goods; and no winner would ever pick such a pair.

For all other pairs, $U_e(x) > \hat{U}(x)$ for some x so someone would be better off picking (y_e, T_e) than under $\widehat{\text{RM}}$. The set of these pairs is $\mathcal{P} \equiv \{(y_e, T_e) \mid U_e(x) > \hat{U}(x)$ for some $x\}$. For $(y_e, T_e) \in \mathcal{P}$, solutions to RM and $\widehat{\text{RM}}$ differ so there is at least one value x at which (5) binds. Let x_b denote such an x, so $U(x_b) = U_e(x_b)$. Given x_b, proposition 1 below shows that RM can be solved by separately: maximizing revenue from individuals with productivities below x_b and incomes below y_e; and maximizing revenue from individuals with productivities above x_b and incomes above y_e.

The two maximizations are conditional on (x_b, y_e, T_e) and have ranges of integration and control below (B) and above (A) productivity x_b:

$$\text{maximize} \int_{x_-}^{x_b} t(U(x), Y(x), x) dF(x) \tag{RM$_B$}$$

$$\text{subject to (1), (2), (4), } U(x_b) = U_e(x_b), \text{ and } Y(x_b) \leq y_e$$

by choice of $\{U(x), Y(x), \psi(x)\}_{x \leq x_b}$; and

$$\text{maximize} \int_{x_b}^{x_+} t(U(x), Y(x), x) dF(x) \tag{RM$_A$}$$

$$\text{subject to (1), (2), (4), } U(x_b) = U_e(x_b), \text{ and } Y(x_b) \geq y_e$$

by choice of $\{U(x), Y(x), \psi(x)\}_{x \geq x_b}$. Let $R_B(x_b, y_e, T_e)$ and $R_A(x_b, y_e, T_e)$ denote maximal values of the revenue integrals in the respective problems. Taxpayers $x \leq x_b$ earn income $Y(x) \leq Y(x_b) \leq y_e$ and taxpayers $x \geq x_b$ earn income $Y(x) \geq Y(x_b) \geq y_e$, so the two revenue maximizations yield separate segments of the income tax function for incomes below y_e and for incomes above y_e.[17]

Figure 1 illustrates. The minimum-utility constraint anchors income taxes at $-\alpha$.[18] The top tax function, from the solution to $\widehat{\text{RM}}$, is the upper boundary of $\mathcal{P}$. Consider a winner who earns income y_e, who would pay taxes $\widehat{T}_e$ at this income under $\widehat{\text{RM}}$. A small reduction in the tax from $\widehat{T}_e$ to T'_e would raise the winner's utility. Such a reduction pulls down the entire tax function, but pulls it down particularly at income y_e: conditioning on (y_e, T_e) with $T_e < \widehat{T}_e$ gives an incentive-compatible tax function in

[17] The two inequalities suggest that x_b earns $Y(x_b) = y_e$. This is correct if $x_- < x_b < x_+$ (see proposition 1), but not necessarily if x_b is at the boundary of $[x_-, x_+]$. Intuitively, if y_e is very low, solutions to RM may have $x_b = x_-$ and $Y(x_-) > y_e$, and if y_e is very high, solutions may have $x_b = x_+$ and $Y(x_+) < y_e$. We allow $Y(x_b) \neq y_e$ to account for these boundary cases.

[18] All figures consider the empirically plausible case in which the minimum-utility constraint binds ($U(x_-) = u(\alpha, 1)$), $\alpha > 0$, and some individuals earn zero income ($Y(x_-) = 0$), so $T(0) = -\alpha$. The propositions below also cover cases with $U(x_-) > u(\alpha, 1)$, $\alpha = 0$, and $Y(x_-) > 0$. The range of implied shapes of the tax function is explained in section V.

which income y_e is tax-preferred relative to all other incomes in that (y_e, T_e) delivers taxes T_e at income y_e, and imposes taxes at all other incomes to maximize revenue. It will turn out that any such tax function has a kink at y_e.

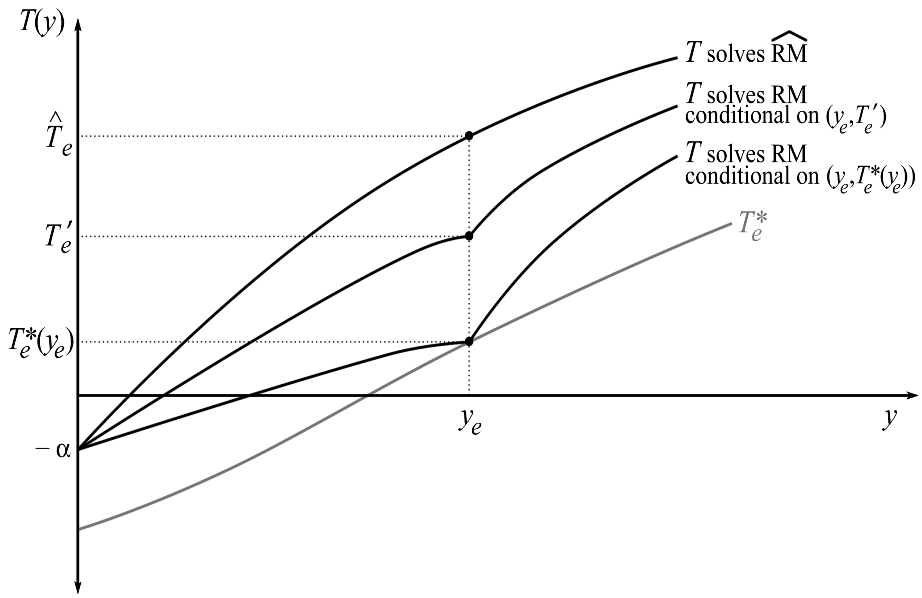

Figure 1. Revenue-Extracting Income Tax Functions

To fund G, the lowest feasible tax at income y_e is the value T_e defined by $R(y_e, T_e) = G$. If this is repeated for all values of y_e, the result is the function T_e^* that gives the lowest feasible tax at different incomes y_e. In figure 1, T_e^* is the grey curve, and the value T_e defined by $R(y_e, T_e) = G$ equals $T_e^*(y_e)$.

To deal with existence and uniqueness of solutions, we first extend the assumptions made earlier about $\widehat{\text{RM}}$ to RM_B and RM_A, as the Hamiltonians are the same:

CON: For $(y_e, T_e) \in \mathcal{P}$ and $x_b \in \{x \mid \hat{U}(x) \leq U_e(x)\}$, H is strictly concave in (U, Y), and RM_B and RM_A each have a solution with continuous U and Y, piecewise continuous ψ, and continuous and piecewise continuously differentiable ξ and μ.

Second, we impose a limited-bunching condition (LB) to ensure that solutions to RM_B and RM_A vary smoothly with (y_e, T_e). Bunching of different productivities at a common income of y_e must be allowed because y_e is tax-preferred, and bunching at zero income is sometimes implied by the minimum-utility constraint. Bunching at incomes other than zero and y_e, however, would introduce discontinuities and make it difficult to derive firm results. The regularity condition we impose limits bunching to incomes of zero and y_e:

LB: For $(y_e, T_e) \in \mathcal{P}$, $Y(x)$ increases strictly for $0 < Y(x) < y_e$ in RM_B, and increases strictly for $Y(x) > y_e$ in RM_A.

LB ensures that income tax functions are continuous except at kinks at tax-preferred incomes y_e, and also provides regularity in proofs of voting results below. Then (proofs of propositions are in an appendix):

Proposition 1 (Revenue Maximization) *RM has a unique solution for any* (y_e, T_e).

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

   (a) *There is a bunching interval* $[x_1, x_2]$ *with* $U(x) = U_e(x)$ *for* $x \in [x_1, x_2]$, *and* $U(x) > U_e(x)$ *for* $x \notin [x_1, x_2]$.

   (b) *If* $[x_1, x_2]$ *includes any* $x \in (x_-, x_+)$, *then* $x_1 < x_2$ *and* $Y(x) = y_e$ *for* $x \in [x_1, x_2]$. *If* $[x_1, x_2] \cap (x_-, x_+) = \varnothing$, *then either* $x_1 = x_2 = x_-$ *and* $Y(x_-) \geq y_e$, *or* $x_1 = x_2 = x_+$ *and* $Y(x_+) \leq y_e$.

   (c) *For any* $x_b \in [x_1, x_2]$, *the solution to* RM_B *on* $[x_-, x_b]$ *together with the solution to* RM_A *on* $[x_b, x_+]$ *solve RM, and* $R(y_e, T_e) = R_B(x_b, y_e, T_e) + R_A(x_b, y_e, T_e)$ *is the same for all* x_b.

2. *The function* T_e^* *defined by* $R(y_e, T_e^*(y_e)) = G$ *exists and is differentiable for* $y_e \in [0, x_+)$.

The function T_e^* in part 2 is the grey curve in the figures that gives the minimum feasible tax at income y_e. The *feasible set* of incomes and taxes for any winner is therefore $\{(y_e, T_e \mid T_e \geq T_e^*(y_e))\}$, with lower boundary T_e^*, as in figure 2. The winner's identity x_e does not enter the objective function or the constraint set of RM, so the set of feasible (y_e, T_e) pairs is the same for all candidates.

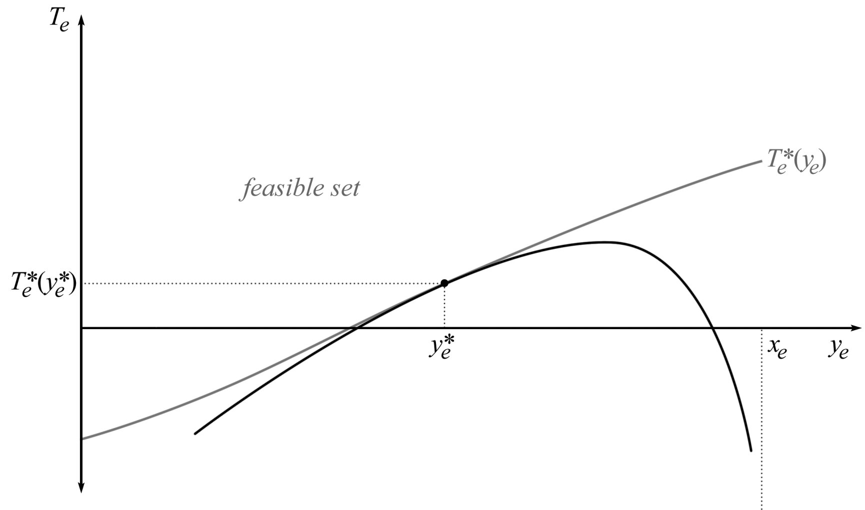

Figure 2. The Winner's Income Choice

Indifference curves in (y_e, T_e) space are inverted-U-shaped, as in figure 2. Differentiating the winner's utility $u(y_e - T_e, 1 - \frac{y_e}{x_e})$, the slope of an indifference curve is $1 - \frac{u_l}{u_c x_e} < 1$. When y_e is low so consumption is scarce and leisure plentiful, indifference curves may have positive slope, as drawn. As income increases, consumption becomes plentiful and leisure scarce so indifference curve slopes fall, become negative, and reach an asymptote at or before x_e, where leisure goes to zero. At any y_e, lower T_e means higher utility.

The winner's optimal choice is on the lowest indifference curve that touches $T_e^*(y_e)$, at income y_e^* in figure 2. Because T_e^* is continuous and $y_e \in [0, x_e]$ is bounded, there is at least one solution. Because the feasible set is not necessarily concave, however, the winner may be indifferent between multiple income values. Denote the set of optimal income values

$$\mathcal{Y}_e^*(x_e) \equiv \{0 \le y_e \le x_e \mid y_e = \arg\max u(y_e - T_e^*(y_e), 1 - y_e/x_e)\}.$$

Interior solutions ($y_e > 0$) satisfy the first-order condition

$$1 - \frac{u_l}{u_c x_e} = \partial T_e^*/\partial y_e,$$

with u_l and u_c evaluated at the tangency point $(y_e^* - T_e^*(y_e^*), 1 - y_e^*/x_e)$. As in the figure, the slope of the winner's indifference curve equals the slope of T_e^*.[19]

Proposition 2 shows that the procedure of conditioning on (y_e, T_e), maximizing revenue RM to set $T_e = T_e^*(y_e)$, and then choosing y_e to maximize $u(y_e - T_e^*(y_e), 1 - y_e/x_e)$ fully solves the winner's problem of choosing a profile $\{U(x), Y(x), \psi(x)\}$ to maximize $U(x_e)$, and also shows that any solution to the winner's problem can be interpreted as extracting maximal revenue:

Proposition 2 (Solution to the Winner's Problem) *For any $x_e \in [x_-, x_+]$:*

1. *The solution $\{U(x), Y(x), \psi(x)\}$ to RM for $(y_e, T_e^*(y_e))$ at any $y_e \in \mathcal{Y}_e^*(x_e)$ also maximizes $U(x_e)$ subject to (1,) (2), (3), and (4).*

2. *Any profile $\{U(x), Y(x), \psi(x)\}$ that maximizes $U(x_e)$ subject to (1), (2), (3), and (4) also solves RM for $(Y(x_e), T_e^*(Y(x_e))) \in \mathcal{P}$, and $Y(x_e) \in \mathcal{Y}_e^*(x_e)$.*

From part 2, the winner picks tax-preferred income $y_e = Y(x_e)$ and the pair $(y_e, T_e^*(y_e))$ lies in $\mathcal{P}$.

IV. ELECTIONS

We study elections between pairs of candidates $x_L < x_H$.

[19] If $0 \in \mathcal{Y}_e^*(x_e)$ then $1 - \frac{u_l}{u_c x_e} < \partial T_e^*/\partial y_e$ is possible. A corner solution with $y_e = x_e$ can be ruled out because $1 - \frac{u_l}{u_c x_e} \to -\infty$ as $y_e \to x_e$. Hence all solutions with $y_e > 0$ are interior.

Income Monotonicity

The productivity of a candidate's constituents determines the income the constituents would earn and hence the income-tax function the candidate would set if elected. Seade (1982) shows that agent monotonicity implies an individual with greater productivity chooses greater income. Applied to election winners, this means that $\mathcal{Y}_e^*(x_e)$ increases strictly as follows:

Proposition 3 (Income Monotonicity (Seade, 1982)) *Let $x_L < x_H$ be two candidates and let $y_L \in \mathcal{Y}_e^*(x_L)$ and $y_H \in \mathcal{Y}_e^*(x_H)$. Then $y_L < y_H$ as long as $y_H > 0$; otherwise $y_L = y_H = 0$.*

Figure 3 illustrates. All potential election winners face the same feasible set. Agent monotonicity implies that the slope of indifference curves through any point (y_e, T_e) increases strictly with x, as drawn at point a. Thus greater x_e shifts the tangency point toward greater y_e so $y_L < y_H$ except at a corner with $y_H = 0$.

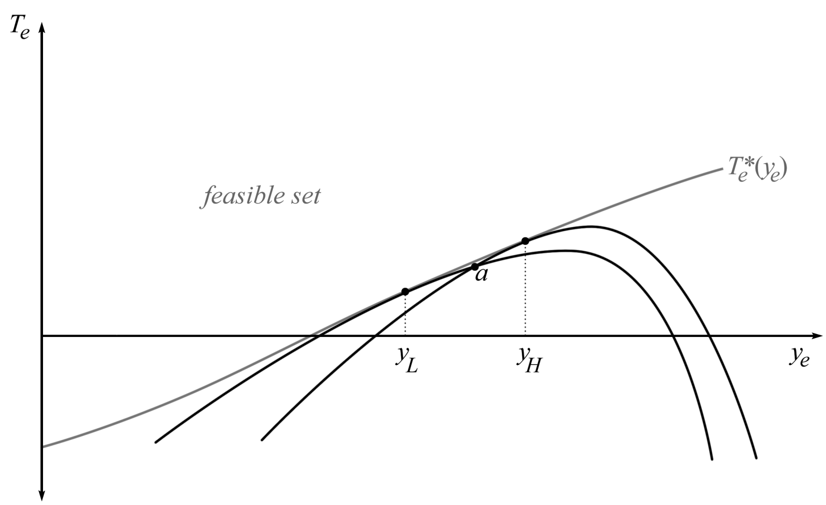

Figure 3. Income Choices of x_L and x_H

From propositions 1 and 2, the winner's problem has exactly as many distinct solutions as there are elements in $\mathcal{Y}_e^*(x_e)$. Because $\mathcal{Y}_e^*$ is increasing under proposition

3, it is single-valued except at isolated x_e-values. Thus the winner's problem has a unique solution for almost all winners. To encompass the non-generic cases in which a candidate has several optimal y_e values, we express the policy of an election winner with given $y_e \in \mathcal{Y}_e^*(x_e)$ as a function of y_e, denoting the associated income and utility profile $\{U(x \mid y_e), Y(x \mid y_e)\}$ and the implied income-tax function $T(y \mid y_e)$.

For now, we take tax-preferred incomes $y_L \in \mathcal{Y}_e^*(x_L)$ and $y_H \in \mathcal{Y}_e^*(x_H)$ as given; later we consider candidates' optimal income choices when $\mathcal{Y}_e^*(x_L)$ and/or $\mathcal{Y}_e^*(x_H)$ are multi-valued.

Single-Crossing of Utility Profiles

We now show that the utility profiles set by two candidates $x_L < x_H$ cross only once at a productivity denoted $x_\times$ with $x_L < x_\times < x_H$.

From proposition 3, candidates $x_L < x_H$ would impose tax functions $T(y \mid y_L)$ and $T(y \mid y_H)$ with different tax-preferred incomes $y_L < y_H$ as long as $y_H > 0$, as in figure 4. The tax-preferred pairs of income and taxes are at the kink points a and c; these lie on T_e^*, the grey curve in previous figures, which is suppressed in figures 4 and 5.

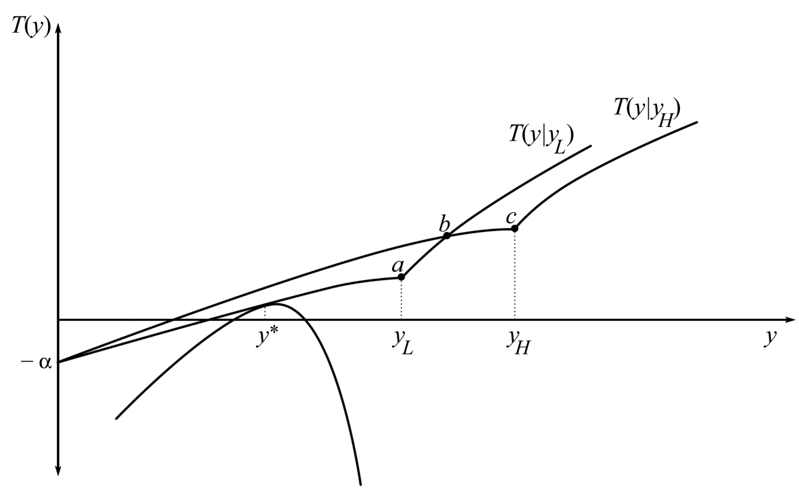

Figure 4. A Taxpayer's Opportunity Set with Candidates $x_L < x_H$.

A taxpayer compares the point of highest utility on $T(y \mid y_L)$ against the point of highest utility on $T(y \mid y_H)$, and votes for the candidate whose tax policy would give greater utility. The taxpayer's productivity x determines the position of the taxpayer's indifference curves. The taxpayer with the indifference curve drawn in figure 4 has a relatively low x. This taxpayer would have greater utility if x_L imposes policy and the taxpayer chooses to earn income $y^* = Y(x \mid y_L)$.

Agent monotonicity implies that the indifference curve at any point $(y, T(y))$ become steeper as x rises. Starting from the tangency in figure 4, this means that as productivity rises successively, indifference curve maps and peaks move to the right and the tangency moves to the right along $T(y \mid y_L)$ until it reaches the kink at a. As x rises further and indifference maps shift further to the right, a productivity denoted $x_\times$ is reached for which an indifference curve simultaneously touches somewhere on segment ab and somewhere on segment bc. A possible configuration is in figure 5. A taxpayer with productivity $x_\times$ gets the same utility from candidates x_L and x_H. Taxpayers $x < x_\times$ get strictly greater utility from candidate x_L (unless the taxpayer has very low productivity and would end

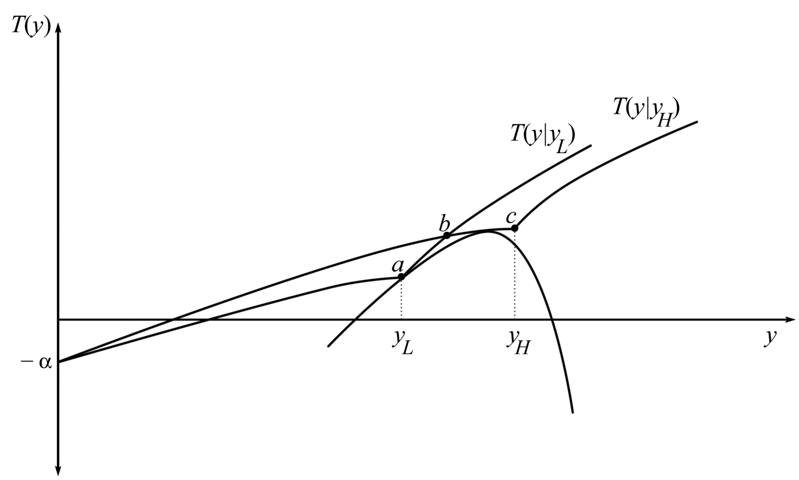

Figure 5. Taxpayer $x_\times$

up getting utility $u(\alpha, 1)$ from both candidates). By similar logic, taxpayers $x > x_\times$ get strictly higher utility from candidate x_H. That is, utility profiles cross once. Because $x_\times$'s indifference curve is steeper at each point than the indifference curve of a winner who would set y_L, and flatter than the indifference curve of a winner who would set y_H, it follows that $x_L < x_\times < x_H$.

Let $X_{u(\alpha,1)}(y_e) = \{x \mid U(x \mid y_e) = u(\alpha, 1)\}$ denote the set of individuals who would obtain the minimum utility $u(\alpha, 1)$ under a tax function with tax-preferred income y_e.[20] Then

Proposition 4 (Single Crossing of Utility Profiles) *Consider candidates $x_L < x_H$ with given $y_L \in \mathcal{Y}_e^*(x_L)$ and $y_H \in \mathcal{Y}_e^*(x_H)$, and $y_H > 0$. Then utility profiles $U(x \mid y_H)$ and $U(x \mid y_L)$ 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 $x \notin X_{u(\alpha,1)}(y_L)$ 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 $X_{u(\alpha,1)}(y_L)$ have $U(x \mid y_L) = U(x \mid y_H) = u(\alpha, 1)$.*

Two groups are indifferent between candidates x_L and x_H: those with productivity $x_\times$ (who have measure zero) and those in $X_{u(\alpha,1)}(y_L)$.[21]

[20] Because U is increasing, $X_{u(\alpha,1)}(y_e)$ is an interval that starts at x_-, or else is either empty (if $U(x \mid y_e) > u(\alpha, 1)$ for all x) or consists of the single point x_- (in the borderline case with $U(x \mid y_e) = u(\alpha, 1)$ only for $x = x_-$).

[21] As shown in the proof of proposition 4 in the appendix, the set of voters held at utility $u(\alpha, 1)$ expands as the tax-preferred income increases: $X_{u(\alpha,1)}(y_L) \subseteq X_{u(\alpha,1)}(y_H)$ for $y_L < y_H$. Thus an individual in $X_{u(\alpha,1)}(y_L)$ is also in $X_{u(\alpha,1)}(y_H)$ and receives $u(\alpha, 1)$ from both candidates. An individual in $X_{u(\alpha,1)}(y_H)$ but not in $X_{u(\alpha,1)}(y_L)$ is covered in part 2.

Median-Voter Results

Proposition 4 forms the basis for median-voter results. If $X_{u(\alpha,1)}(y_L)$ has measure zero and $\mathcal{Y}_e^*$ is single-valued at x_L and x_H, the logic is simple. Individuals $x > x_\times$ have $U(x \mid y_H) > U(x \mid y_L)$ and hence vote for x_H, and almost all individuals $x < x_\times$ have $U(x \mid y_L) > U(x \mid y_H)$ and hence vote for x_L. Let $x_M \equiv F^{-1}(1/2)$ denote the median productivity. If $x_M < x_\times$, a majority that includes median-productivity voters therefore vote for x_L, and if $x_M > x_\times$, a majority that includes median-productivity voters vote for x_H. Thus the candidate who attracts the vote of median-productivity voters wins.

If $X_{u(\alpha,1)}(y_L)$ has positive measure, assumptions about how indifferent individuals vote matter. Three alternative assumptions might be made: individuals in $X_{u(\alpha,1)}(y_L)$ 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.

If $X_{u(\alpha,1)}(y_L)$ has positive measure and voters in $X_{u(\alpha,1)}(y_L)$ vote by closeness, proposition 4 again implies that the candidate who attracts the votes of median-productivity voters wins. If $X_{u(\alpha,1)}(y_L)$ has positive measure and voters in $X_{u(\alpha,1)}(y_L)$ abstain, the statement of the median-voter result changes: the candidate who attracts median among voters *who do not abstain* wins. Summarizing:[23]

[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.

[23]A key underlying assumption is that the winner imposes an *unrestricted* tax function. Although linearity is sometimes imposed in other contexts to rule out cycling, ad hoc shape restrictions such as a finite number of brackets, or a quadratic form for T may cause utility profiles to cross several times, which can lead to cycles here. Formally, proposition 4 relies on incentive constraints that restrict

Proposition 5 (Median-Voter Theorem) *Consider candidates $x_L < x_H$ and assume $\mathcal{Y}_e^*(x_L) = \{y_L\}$ and $\mathcal{Y}_e^*(x_H) = \{y_H\}$ are single-valued with $y_H > 0$:*

1. *If $X_{u(\alpha,1)}(y_L)$ has measure zero, then the candidate who provides greater utility to median-productivity individuals wins.*[24]

2. *If $X_{u(\alpha,1)}(y_L)$ has positive measure and*

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

   (b) *indifferent individuals abstain, then the candidate who provides greater utility to the median of $[x_-, x_+] \backslash X_{u(\alpha,1)}(y_L)$ wins.*

The requirement that $y_H > 0$ excludes the trivial case in which both x_L and x_H impose $T(y \mid 0)$.

Parts 1 and 2a of Proposition 5 immediately imply:

Proposition 6 (Condorcet Winner) *Assume $\mathcal{Y}_e^*(x_M) = \{y_M\}$ is single-valued with $y_M > 0$. If either $X_{u(\alpha,1)}(y_M)$ has measure zero or indifferent individuals vote by closeness, then x_M wins against any other candidate.*

That is, x_M is the Condorcet winner if $y_M > 0$. If $y_M = 0$ then x_M is not quite a Condorcet winner because any candidate $x_e < x_M$ would also set tax-preferred income $y_e = 0$ and taxes $T(y \mid 0)$, and would tie against x_M. Because x_M would win against any x_e who would set $y_e > 0$, however, the tax function $T(y \mid 0)$ preferred by x_M is always implemented. If $X_{u(\alpha,1)}(y_M)$ has positive measure and indifferent

utility differences of individuals close in productivity, with individuals treated monotonically better the closer they are to the election winner. Shape restrictions may disrupt this monotone link.

[24]That is, if $U(x_M \mid y_L) > U(x_M \mid y_H)$ then x_L wins, and if $U(x_M \mid y_L) < U(x_M \mid y_H)$ then x_H wins. In the non-generic case in which $x_\times = x_M$ so $U(x_M \mid y_L) = U(x_M \mid y_H)$, the vote is tied.

voters abstain (as in 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.[25]

Matters are more complicated but results are similar if $\mathcal{Y}_e^*$ has multiple values at x_L, x_H, or x_M. Such cases are non-generic because $\mathcal{Y}_e^*$ is monotone and therefore single-valued except at isolated x_e. If $\mathcal{Y}_e^*(x_L)$ has multiple values, choosing $y_L = \max\{\mathcal{Y}_e^*(x_L)\}$ gives x_L the largest vote share against any y_H provided either $X_{u(\alpha,1)}(\max\{\mathcal{Y}_e^*(x_L)\})$ has measure zero or voters in $X_{u(\alpha,1)}(y_L)$ vote by closeness, because from the logic of proposition 4, higher y_L then raises $U(x \mid y_L)$ in a neighborhood of $x_\times$ and expands the interval $[x_-, x_\times]$. Similarly if $\mathcal{Y}_e^*(x_H)$ has multiple values, choosing $y_H = \min\{\mathcal{Y}_e^*(x_H)\}$ gives x_H the largest vote share against any y_L provided $y_H > 0$ (so x_H and x_L do not both set $y_e = 0$). Thus candidates have incentives to appeal to the median voter by picking high y_L and low y_H, and given these choices of y_L and y_H, the candidate who provides greater utility to median-productivity individuals wins as in proposition 5, parts 1 and 2a.[26]

[25] In detail: a simple fixed-point argument implies there is a smallest productivity $x_m \in (x_M, x_+)$ that is median in $[x_-, x_+] \backslash X_{u(\alpha,1)}(y_m)$, the set of voters who would not abstain if x_m were to set y_m. By proposition 5 (part 2b), x_m wins against any $x_e > x_m$, and no candidate other than x_m can be a Condorcet winner. Because $x_L < x_m$ may draw individuals in $X_{u(\alpha,1)}(y_m)$ to the polls, however, x_L may win against x_m and because x_L is below the median in $[x_-, x_+] \backslash X_{u(\alpha,1)}(y_L)$, x_L would lose against some $x_l \in (x_L, x_m)$, who may in turn 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 $X_{u(\alpha,1)}(y_m)$ to vote.

[26] If individuals in $X_{u(\alpha,1)}(y_L)$ abstain and $X_{u(\alpha,1)}(\max\{\mathcal{Y}_e^*(x_L)\})$ has positive measure, then $y_L = \max\{\mathcal{Y}_e^*(x_L)\}$ may not give x_L the greatest share of votes because the number of abstentions may rise with y_L. Then by choosing an income lower than $\max\{\mathcal{Y}_e^*(x_L)\}$, the lower-productivity candidate may gain the votes of some individuals who would abstain if x_L chose income $\max\{\mathcal{Y}_e^*(x_L)\}$. In this case, determining the winner requires inspecting voting outcomes for all elements of $\mathcal{Y}_e^*(x_L)$ against $\min\{\mathcal{Y}_e^*(x_H)\}$. If there is a $y_L \in \mathcal{Y}_e^*(x_L)$ such that x_L captures the median-productivity voter in

Similarly, we show in the appendix that if $\mathcal{Y}_e^*(x_M)$ has multiple values, then x_M wins with any $y_M \in \mathcal{Y}_e^*(x_M)$ against any other candidate provided $\min\{\mathcal{Y}_e^*(x_M)\}$ > 0. That is, proposition 6 does not require the assumption that $\mathcal{Y}_e^*(x_M)$ is single-valued. If $0 \in \mathcal{Y}_e^*(x_M)$ then x_M wins against any candidate who would set $y_e > 0$ so the election outcome always gives x_M maximum utility.

V. THE SHAPE OF THE WINNER'S TAX FUNCTION

We describe in more detail the marginal tax schedule $dT(y)/dy$ set by the winner. Although the analysis above suggests that election winners are likely to have median or close-to-median productivity, the analysis in this section is general and allows the winner to have any income $y_e \geq 0$. Recall that T is defined on $[Y(x_-), Y(x_+)]$.[27]

Under LB, T is differentiable and $Y^{-1}(y)$ is single-valued except at $y = y_e$ and possibly at $y = 0$. At incomes where T is differentiable, the taxpayer's first-order condition for maximizing $u(y - T(y), 1 - \frac{y}{x})$ is $\frac{dT(y)}{dy} = 1 - \frac{u_l}{u_c x}$. The derivative of the tax profile $t(U, Y, x) \equiv Y - c^*(U, \frac{Y}{x})$ with respect to Y is $\tau(x) \equiv t_Y(U(x), Y(x), x) = 1 - \frac{u_l}{u_c x}$. Therefore the marginal tax rate is

$$\frac{dT(y)}{dy} = \tau(Y^{-1}(y)) \tag{6}$$

on $[Y(x_-), Y(x_+)]$, except at $y = y_e$ and possibly at $y = 0$.

We evaluate $\tau(x)$ using the solution to RM_B for $x < x_1$ and the solution to RM_A for $x > x_2$; both are conditional on $(x_e, y_e, T_e^*(y_e))$. The Euler equation

$[x_-, x_+] \backslash X_{u(\alpha,1)}(y_L)$ when x_H picks $\min\{\mathcal{Y}_e^*(x_H)\}$, then x_L wins by choosing this y_L; if not, x_H wins.

[27] To extend the range of T to $[0, x_+)$ and preserve continuous differentiability when $Y(x_-) > 0$, one may define $T(y)$ for $y < Y(x_-)$ by $u(y - T(y), 1 - y/x_-) = U(x_-)$, which lets the tax function coincide with x_-'s indifference curve at utility $U(x_-)$.

$H_Y = \tau(x) \cdot f(x) + \omega_Y(U, Y, x) \cdot \xi(x) = 0$ implies

$$\tau(x) = -\frac{\omega_Y(U, Y, x)}{f(x)} \xi(x), \tag{7}$$

where $\omega_Y(U, Y, x)/f(x) > 0$, so $\tau(x)$ has the same sign as $-\ \xi(x)$. The proof of the following proposition derives results by studying ξ. (As with Mirrleesian taxation in other contexts, the density f also enters (7). The fine structure of how the marginal rate changes with income therefore depends on precisely how f varies with x, about which we make no assumptions.)

Proposition 7 (Net-Income-Tax Function Chosen by x_e) *For $x_e \in (x_-, x_+)$ with $y_e > 0$, T is continuously differentiable on $[Y(x_-), Y(x_+)]$ except at $y = y_e$, with $dT(y)/dy = \tau(Y^{-1}(y))$ for $y > 0$. Moreover:*

1. $\tau(x_1) = lim_{y \uparrow y_e} dT(y)/dy < \tau(x_2) = lim_{y \downarrow y_e} dT(y)/dy$.

2. *For $y > y_e$, $dT(y)/dy > 0$ except that $dT(Y(x_+))/dy = \tau(x_+) = 0$ if $x_+ < \infty$.*

3. *For $y < y_e$:*

   (a) *T may increase strictly.*

   (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 $x_\tau \in (x_-, x_1) \backslash \{x \mid Y(x \mid y_e) = 0\}$.*

   (c) *T may decrease strictly. A sufficient condition for this is that the minimum-utility constraint (4) does not bind.*

To keep the statement of proposition 7 simple, the proposition does not cover special cases in which the winner is $x_e = x_-$ or $x_e = x_+$, or in which the winner earns income $y_e = 0$. These cases are degenerate in that y_e is at the boundary of

$[Y(x_-), Y(x_+)]$ so revenue maximization is one-sided with either $R_B = 0$ or $R_A = 0$. Parts 2 and 3, respectively, still hold.[28]

To summarize this proposition and earlier material on the shape of taxes:

Part 1 says that the marginal tax jumps upward at the winner's own income, as at the kink in figure 1. The single kink in the highly stylized setting here broadly fits the bracket structure of individual income taxes in the U.S. and elsewhere, under which marginal rates rise with income.

Part 2 resembles results in Seade (1977, 1982) and Röell (1996): taxes rise with income above the winner's income, but the marginal rate is zero at the maximum productivity x_+ if x_+ is finite. Recall also from proposition 1 that taxes maximize R_A, so maximal revenue is extracted from higher-income households, consistent with data in table 1. These data indicate that the U.S. individual income tax in particular extracts revenue from taxpayers with high income.

Part 3 says that T is either increasing, inverted-U shaped, or decreasing at incomes below y_e.[29] The range of possible outcomes reflects a range of possible specific assumptions that might be made to describe different economies. For instance, in a hypothetical economy with a narrow productivity distribution so the lowest productivity is only slightly below the median and far above zero, the minimum-utility constraint may not bind so a winner in the middle would extract revenue from the

[28] The proposition also does not cover dT/dy at $y = 0$. Because $Y(x)$ is non-decreasing, $\{x \mid Y(x \mid y_e) = 0\}$ is either empty, contains only x_-, or is an interval with lower endpoint x_-. If $\{x \mid Y(x \mid y_e) = 0\}$ is empty then nobody earns $y = 0$ so the marginal tax rate is defined on $[Y(x_-), Y(x_+)]$ except at $y = y_e$. If $\{x \mid Y(x \mid y_e) = 0\} = \{x_-\}$, then $Y^{-1}(0)$ is single-valued so (6) holds at $y = 0$. If $\{x \mid Y(x \mid y_e) = 0\}$ is an interval, the definition $dT(0)/dy \equiv \tau(\max\{x \mid Y(x \mid y_e) = 0\})$ makes dT/dy continuously differentiable at $y = 0$.

[29] A result like that in Seade (1977, 1982) and Röell (1996) also holds at x_-: if the minimum-utility constraint (4) does not bind and $Y(x_-) > 0$ (a subcase of part 3c), then $dT(Y(x_-))/dy = \tau(x_-) = 0$. However, $\tau(x_-) < 0$ in the cases of parts 3a and 3b.

relatively poor (case 3c, with negative marginal taxes at all incomes below y_e). In an economy with a wide productivity range that stretches from near zero to very high, on the other hand, the minimum-utility constraint may bind (case 3a or 3b).

Revenue extraction thus has different consequences below and above y_e because of the minimum-utility constraint. Importantly, revenue extraction tends to imply high *marginal* taxes at the lowest incomes. This shows up in figures 4-5, where the tax function is anchored at $T(0) = -\alpha$ and has a kink at y_e. Together, these yield a function that tends[30] to have greater slope at very low incomes than just below y_e. These high marginal rates help the winner extract revenue from the bottom half of the income distribution by raising taxes at incomes intermediate between the lowest incomes and y_e.

Marginal taxes in most countries are positive, which is case 3a. This case requires not only a binding minimum-utility constraint but also positive G and/or α. That is, in a stripped-down model of revenue extraction with $G = \alpha = 0$, the winner would use revenue extracted from others to provide a negative net tax at the own income so $T(y_e) < 0$. Suppose $Y(x_-) = 0$, which is empirically reasonable. Then $T(0) = -\alpha = 0 > T(y_e)$, so taxes necessarily have a downward-sloping segment. As revenue requirements G or α rise, the tax function is generally pulled up at y_e, and as α rises, the intercept $T(0)$ is also pulled down, so case 3a must apply if G and/or α are high enough:

Proposition 8 (Role of G and α in Determining the Slope of T) *Consider $(G, \alpha) \geq 0$ with $G < \hat{R}(\alpha)$:*

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

[30] As noted, the exact shape of taxes depends on the shape of f.

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

The results thus suggest that a binding minimum-utility constraint and positive revenue requirements G and α are important elements in an empirically plausible model. If revenue requirements are high enough, middle-income voters may pay substantial taxes, which is the European pattern. If revenue requirements are smaller, middle-income voters may pay low taxes, which is the current pattern in the U.S.

VI. SUMMARY

We study democratic pressures to redistribute income in a static median-voter setting. To avoid cycling that would obscure tendencies toward a median-voter outcome, we assume that individuals differ only in productivity and that electoral competition is between representatives for two different single productivities. The winner sets the redistributional tax function that is optimal for the productivity the winner represents, and voters know the tax functions a candidate would set if the candidate were to win. Then median-productivity voters always vote on the winning side, and a representative for median-productivity voters would beat any other candidate. Resulting tax policy extracts maximal revenue from high-productivity voters. With a binding minimum-utility constraint that restricts what can be extracted from the poor, resulting policy coarsely describes redistributional taxation in the U.S.: taxes are progressive, the poor receive subsidies but face high effective marginal tax rates, and high-income taxpayers pay most taxes.

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] Congressional Budget Office, Shares of Federal Tax Liabilities for All Households, by Comprehensive Household Income Quintile, 1979-2007, online at http://www.cbo.gov/publications/collections/tax/2010/AverageFedTaxRates2007.pdf.

[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] Joint Committee on Taxation, Memorandum "Subject: Information on Income Tax Liability for Tax Year 2009," Apr. 29, 2011.

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

[12] Martimort, David, "Optimal Taxation and Strategic Budget Deficit Under Political Regime Switching," *Review of Economic Studies* 68 (3), July 2001, 573-92.

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

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

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

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

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

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

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

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

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

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

APPENDIX: PROOFS – NOT INTENDED FOR PUBLICATION

Proposition 1 (Revenue Maximization)

Proposition 1 is proved in a sequence of lemmas by constructing the revenue-maximizing policy conditional on (y_e, T_e), verifying that this policy is in fact optimal, and then establishing the claims in the proposition. Lemmas 1.1-1.3 first characterize solutions to $\widehat{\text{RM}}$, RM_B, and RM_A, from which the solution to RM is constructed. Solution profiles are marked naturally so $\{\hat{U}(x), \hat{Y}(x)\}$ solve $\widehat{\text{RM}}$, $\{U_B(x), Y_B(x)\}$ solve RM_B, etc. From the definition of $\mathcal{P}$, U_e lies above $\hat{U}$ for some x if $(y_e, T_e) \in \mathcal{P}$, so

$$x_s \equiv \inf\{x \in [x_-, x_+] \mid U_e(x) > \hat{U}(x)\}$$

exists and is well-defined. (Throughout, when $x_+ = \infty$, we take $[x_-, x_+]$ to mean $[x_-, \infty)$.) Figure A1 illustrates the construction for the interior case in which $x_- < x_s < x_1 < x_2 < x_+$. (The proof is general and allows for corners in which one or more of these inequalities are equalities.)

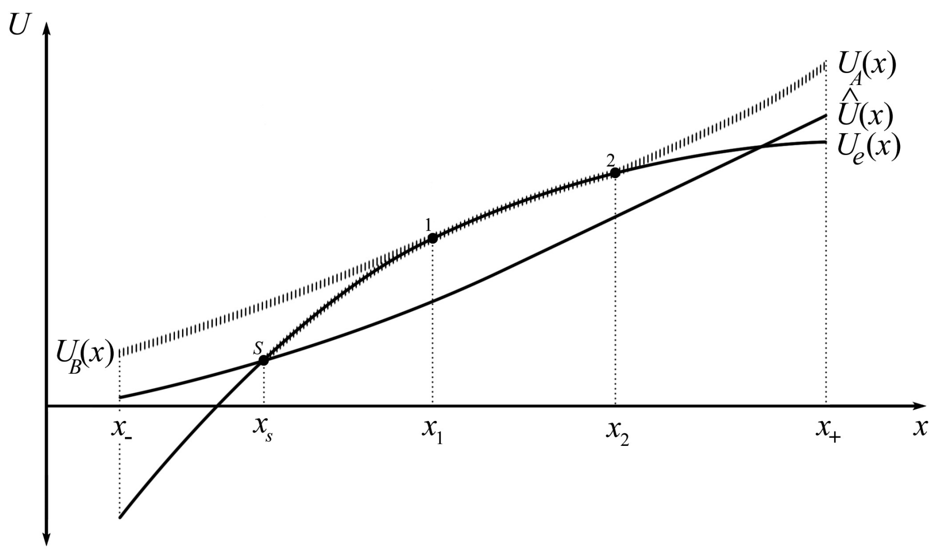

Figure A1— Construction of the Revenue-Maximizing Policy

To construct x_1 and x_2, start at x_s and consider RM_A given (x_s, y_e, T_e). In the figure, the utility profile U_A that solves this is the vertically-hatched curve that starts at point S and runs upward to the right as productivity runs toward x_+. Lemma 1.4 shows that in general,

$$x_2 \equiv \sup\{x \in [x_s, x_+] \mid U_A(x) = U_e(x)\}$$

exists under the solution to RM_A and is finite. Lemma 1.5 shows that this solution has a segment on $[x_s, x_2]$ with utility $U_A(x) = U_e(x)$ and (provided $x_2 < x_+$, as in the figure) a segment on $[x_2, x_+]$ with $U_A(x) > U_e(x)$. Given x_2, now consider RM_B given (x_2, y_e, T_e). The utility profile U_B that solves this is the vertically-hatched curve that starts from point 2 and runs down and to the left as productivity runs toward x_-. Lemma 1.6 shows that

$$x_1 \equiv \inf\{x \in [x_-, x_2] \mid U_B(x) = U_e(x)\}$$

is well-defined and that the solution to RM_B has a segment on $[x_1, x_2]$ with utility $U_B(x) = U_e(x)$ and (provided $x_1 > x_-$, as in the figure) a segment on $[x_-, x_1]$ with $U_B(x) > U_e(x)$. Thus U_B for $x \leq x_2$ matches U_A for $x \geq x_1$ from the solution to RM_A; in the figure this occurs between points 1 and 2. Lemmas 1.7-1.9 show that the profile U_{RM} obtained by combining U_B and U_A at any $x_b \in [x_1, x_2]$, which is the upper hatched curve in the figure, is the unique solution to RM. Lemmas 1.10-1.11 show existence and differentiability of T_e^*. Finally, proposition 1 is proved from the lemmas.

We suppress the dependence of U_e, x_s, x_1, and x_2 on (y_e, T_e) when considering given (y_e, T_e). Define the following boundary conditions separately for $\widehat{\mathrm{RM}}$, RM_B, and RM_A:

Boundary conditions for $\widehat{\mathrm{RM}}$: $U(x_-) = u(\alpha, 1)$, $Y(x_-) \geq 0$, $\mu(x_-) \leq 0$, and $\mu(x_-)Y(x_-) = 0$ at x_-; and $\xi(x_+) = 0$ at x_+. (Throughout, conditions at x_+ are limit conditions if $x_+ = \infty$; e.g. here $\xi(x_+) = \lim_{x\to\infty} \xi(x) = 0$.)

Boundary conditions for RM$_B$: $U(x_b) = U_e(x_b)$, $Y(x_b) \leq y_e$, $\mu(x_b) \leq 0$, and $\mu(x_b)\,[Y(x_b) - y_e] = 0$ at x_b; and $U(x_-) \geq u(\alpha, 1)$, $\xi(x_-) \leq 0$, and $\xi(x_-)$ $[U(x_-) - u(\alpha, 1)] = 0$ at x_-.

Boundary conditions for RM$_A$: $U(x_b) = U_e(x_b)$, $Y(x_b) \geq y_e$, $\mu(x_b) \leq 0$, and $\mu(x_b)\,[Y(x_b) - y_e] = 0$ at x_b; and $\xi(x_+) = 0$ at x_+.

Lemma 1.1: Solutions to $\widehat{\text{RM}}$, RM$_B$, and RM$_A$ and must satisfy the following necessary conditions for optimality: (*i*) $\psi(x) \geq 0$, $\mu(x) \leq 0$, and $\mu(x)\psi(x) = 0$ for $x \in [x_-, x_+]$; (*ii*) the Euler equations

$$\partial H/\partial Y = H_Y(U, Y, \xi, x) = t_Y(U, Y, x) \cdot f(x) + \omega_Y(U, Y, x) \cdot \xi(x) = -\mu_x(x), \quad \text{(A.1)}$$

$$\partial H/\partial U = H_U(U, Y, \xi, x) = t_U(U, Y, x) \cdot f(x) + \omega_U(U, Y, x) \cdot \xi(x) = -\xi_x(x), \quad \text{(A.2)}$$

for $x \in [x_-, x_+]$, where the definition of t and the properties of c^* imply $t_Y = 1 - u_l/(u_c x)$ and $t_U = -1/u_c$; and (*iii*) the differential equations for the state variables, which are (1) and $Y(\tilde{x}) - Y(x_-) = \int_{x_-}^{\tilde{x}} \psi(x)dx$ for $\tilde{x} \in [x_-, x_+]$; and (*iv*) the boundary conditions above.

Proof: $\widehat{\text{RM}}$, RM$_B$ given any (x_b, y_e, T_e), and RM$_A$ given any (x_b, y_e, T_e) have the same Hamiltonian H and differ only in boundary conditions. Let $\zeta(x) \geq 0$ be the Kuhn-Tucker multiplier on (2). The Maximum Principle requires that the generalized Hamiltonian $H(U, Y, \psi, \xi, \mu, x) + \zeta(x)\psi(x)$ satisfy the Euler equation $\partial(H + \zeta\psi)/\partial\psi = \mu(x) + \zeta(x) = 0$ for all x. Hence $\mu(x) = -\zeta(x) \leq 0$ for all x so (2) and the Kuhn-Tucker conditions $\zeta(x) \geq 0$ and $\zeta(x)\psi(x) = 0$ imply (*i*). Conditions (*ii*) and (*iii*) follow directly from the Maximum Principle. Because $\partial H/\partial Y$ and $\partial H/\partial U$ do not depend on μ and ψ, they can be written as functions of (U, Y, ξ, x) only. For RM$_B$ and RM$_A$, the boundary conditions in (*iv*) follow directly from the problems' constraints, noting that Kuhn-Tucker conditions apply in case of inequality constraints. For $\widehat{\text{RM}}$, the boundary conditions also follow directly from the problem's constraints, except that (4) directly implies $\hat{U}(x_-) \geq u(\alpha, 1)$, $\hat{\xi}(x_-) \leq 0$, and $\hat{\xi}(x_-)\left[\hat{U}(x_-) - u(\alpha, 1)\right] = 0$.

To show that the boundary condition $\hat{U}(x_-) = u(\alpha, 1)$ holds with equality, note from (A.2) that

$$\hat{\xi}_x(x) = -t_U(\hat{U}, \hat{Y}, x) \cdot f(x) - \omega_U(\hat{U}, \hat{Y}, x) \cdot \hat{\xi}(x).$$

Because $t_U(\hat{U}, \hat{Y}, x) \cdot f(x) > 0$ for all x, $\hat{\xi}_x(x) > 0$ whenever $\hat{\xi}(x)$ is in a neighborhood of zero. Hence the boundary condition $\hat{\xi}(x_+) = 0$ implies $\hat{\xi}(x) < 0$ for all x, which implies $\hat{\xi}(x_-) < 0$. Thus the Kuhn-Tucker conditions reduce to $\hat{U}(x_-) = u(\alpha, 1)$. QED.

Remark: Because $\mu(x) = -\zeta(x)$ for all x, we streamline the exposition below by expressing the multiplier on (2) in terms of μ, which makes ζ redundant.

Lemma 1.2: Under CON: (*i*) problems $\widehat{\text{RM}}$, RM_B, and RM_A have solutions with unique utility and income profiles $\{U(x), Y(x)\}$;[31] (*ii*) the conditions in lemma 1.1 are sufficient for optimality; (*iii*) $\psi(x)$ is uniquely defined except at points where Y is not differentiable.

Proof: From the Mangasarian sufficiency theorem (e.g. Seierstad and Sydsaeter 1987, p. 287), CON implies that $\widehat{\text{RM}}$, and RM_B and RM_A given (x_b, y_e, T_e), each have a solution with unique profiles $\{U(x), Y(x)\}$ and that the necessary conditions for optimality are also sufficient. Because Y is unique, its derivative ψ is uniquely defined except at points where Y is not differentiable. QED.

Lemma 1.3: In any solution to $\widehat{\text{RM}}$, RM_B, or RM_A, if $U(x) = U_e(x)$ and $\frac{dU}{dx} = \frac{dU_e}{dx}$, then $Y(x) = y_e$. If $U(x) = U_e(x)$ and $\frac{dU}{dx} - \frac{dU_e}{dx} \neq 0$, then $Y(x) - y_e$ has the same sign as $\frac{dU}{dx} - \frac{dU_e}{dx}$.

Proof: From (1), $\frac{dU}{dx}(x) = \omega(U(x), Y(x), x)$ and $\frac{dU_e}{dx}(x) = \omega(U_e(x), y_e, x)$. Agent monotonicity implies $\omega_Y(U, Y, x) > 0$ for $Y > 0$, so ω is invertible with respect to Y; moreover, for given x and given $U(x) = U_e(x)$, $Y(x)$ increases strictly

[31] When referring to profiles that may solve either $\widehat{\text{RM}}$, RM_B or RM_A, we omit the hats or subscripts used to mark solutions to these problems.

with $\frac{dU}{dx}$. Hence $Y(x) - y_e$ has the same sign as $\frac{dU}{dx} - \frac{dU_e}{dx}$, and $Y(x) = y_e$ if and only if $\frac{dU}{dx} = \frac{dU_e}{dx}$. QED.

Lemma 1.4: For given $(y_e, T_e) \in \mathcal{P}$, the solution to RM$_A$ given (x_s, y_e, T_e) satisfies $x_2 < \infty$.

Proof: The claim is trivial if $x_+ < \infty$ because $x_2 \leq x_+$ so suppose $x_+ = \infty$. Agent monotonicity implies that the marginal tax rate

$$t_Y(U_e(x), y_e, x) = 1 - \frac{u_l(y_e - T_e, 1 - y_e/x)}{u_c(y_e - T_e, 1 - y_e/x)x}$$

increases strictly in x. Moreover, $\frac{u_l(y_e - T_e, 1 - y_e/x)}{u_c(y_e - T_e, 1 - y_e/x)x} \to 0$ as $x \to \infty$ so there is a value $x_\tau < \infty$ such that $t_Y(U_e(x), y_e, x) > 0$ for $x \geq x_\tau$. Assume for contradiction that $x_2 = \infty$. Then from lemma 1.3, $\{U_A(x), Y_A(x)\}_{x \geq x_s} = \{U_e(x), y_e\}_{x \geq x_s}$ solves RM$_A$. Moreover, $t_Y(U_A(x), Y_A(x), x) = t_Y(U_e(x), y_e, x) > 0$ for $x > x_\tau$. Hence a marginal increase in $Y(x)$ for $x \in (x_\tau, x_+]$, holding $U(x) = U_e(x)$ constant, would increase $t(U(x), Y(x), x)$ for $x \in (x_\tau, x_+]$, satisfy the constraints of RM$_A$, and yield higher revenue. This would contradict the optimality of $\{U_e(x), y_e\}_{x \geq x_s}$, so $x_2 < \infty$. QED.

Lemma 1.5: For $(y_e, T_e) \in \mathcal{P}$, let $\{U_A(x), Y_A(x), \xi_A(x), \mu_A(x)\}_{x \geq x_s}$ denote the profile that solves RM$_A$ given (x_s, y_e, T_e). (1) If $x_s > x_-$ then $x_2 > x_s$. (2) If $x_s = x_-$ then $x_2 \geq x_s$. (3) If $x_2 > x_s$ (for any $x_s \geq x_-$) then: (*i*) $U_A(x) = U_e(x)$ and $Y_A(x) = y_e$ for $x \in [x_s, x_2]$; (*ii*) $U_A(x) > U_e(x)$ and $Y_A(x) > y_e$ for $x > x_2$; (*iii*) for any $x_b \in [x_s, x_2]$, the segment $\{U_A(x), Y_A(x), \xi_A(x), \mu_A(x)\}_{x \geq x_b}$ solves RM$_A$ given (x_b, y_e, T_e).

Proof: (1) Note that $x_s < x_+$ because $\{x \mid U_e(x) > \hat{U}(x)\} \neq \varnothing$ for $(y_e, T_e) \in \mathcal{P}$ and $U_e - \hat{U}$ is continuous in x. Hence $x_s > x_-$ implies $x_s \in (x_-, x_+)$. Continuity of U_e and $\hat{U}$ then imply $U_e(x_s) = \hat{U}(x_s)$. Because $U_e(x) > \hat{U}(x)$ in a neighborhood of $x > x_s$, differentiability of U_e and $\hat{U}$ imply $\frac{dU_e}{dx}(x_s) - \frac{d\hat{U}}{dx}(x_s) \geq 0$. Hence $\hat{Y}(x_s) \leq y_e$ from lemma 1.3. Because $x_2 \geq x_s$ by construction, $x_2 > x_s$ holds if we can rule out $x_2 = x_s$.

Suppose to the contrary that $x_2 = x_s$. Then $U_A(x) > U_e(x)$ for $x > x_s$, so $Y_A(x) > y_e$ for $x > x_s$ and $\mu_A(x_s) = 0$. Define $\tilde{y}_e \equiv \hat{Y}(x_s)$ and $\tilde{T}_e \equiv \hat{T}(\hat{Y}(x_s))$, and let $\{\tilde{U}(x), \tilde{Y}(x), \tilde{\xi}(x), \tilde{\mu}(x)\}_{x \geq x_s}$ denote the profile that solves RM_A given $(x_s, \tilde{y}_e, \tilde{T}_e)$. Because $\tilde{y}_e \leq y_e$ implies $\tilde{\mu}(x_s) = 0$, the two RM_A-problems satisfy the same boundary conditions: $\tilde{U}(x_s) = U_A(x_s)$, $\mu_A(x_s) = \tilde{\mu}(x_s) = 0$, and $\xi_A(x_s) = \tilde{\xi}(x_+) = 0$. By lemma 1.2, profiles $\{\tilde{U}(x), \tilde{Y}(x)\}_{x \geq x_s}$ and $\{U_A(x), Y_A(x)\}_{x \geq x_s}$ must be identical. However, $\tilde{Y}(x_s) = \hat{Y}(x_s)$ with $\tilde{Y}$ increasing implies $\frac{d\hat{U}}{dx}(x) \geq \frac{d}{dx}\tilde{U}_e(x)$ for $x \geq x_s$, so $\{\hat{U}(x), \hat{Y}(x)\}_{x \geq x_s}$ satisfies all constraints of RM_A given $(x_s, \tilde{y}_e, \tilde{T}_e)$, and hence $\{\hat{U}(x), \hat{Y}(x)\}$ satisfies $\hat{U}(x) \geq U_e(x)$ for all x, contradicting $(y_e, T_e) \in \mathcal{P}$. Thus $x_2 > x_s$.

(2) Trivial because the definition of x_2 implies $x_2 \geq x_s$.

(3) Given $x_s < x_2$, the constraint $U_A(x) \geq U_e(x)$ implies $U_A(x) > U_e(x)$ for $x > x_2$ and $U_A(x) = U_e(x)$ for $x \in [x_s, x_2]$. The latter implies $\frac{dU_A}{dx}(x) = \frac{dU_e}{dx}(x)$ for $x \in [x_s, x_2]$, so by lemma 1.3, $Y_A(x) = y_e$ for $x \in [x_s, x_2]$. Consider $x > x_2$ in a neighborhood of x_2. Then $U_A(x) > U_e(x)$ implies $\frac{dU_A}{dx}(x) > \frac{dU_e}{dx}(x)$, so $Y_A(x) > y_e$ by lemma 1.3. Because Y_A is increasing, $Y_A(x) > y_e$ for $x > x_2$. Inspection of the sufficient conditions (see lemma 1.1) then shows that $\{U_A(x), Y_A(x), \xi_A(x), \mu_A(x)\}_{x \geq x_b}$ solves RM_A given (x_b, y_e, T_e) for any $x_b \in [x_s, x_2]$. QED.

Lemma 1.6: For given $(y_e, T_e) \in \mathcal{P}$, let $\{U_B(x), Y_B(x), \xi_B(x), \mu_B(x)\}_{x \leq x_2}$ denote the profile that solves RM_B given (x_2, y_e, T_e). (1) If $x_- < x_2 < x_+$, then $x_1 < x_2$. (2) If $x_2 = x_+$ or $x_2 = x_-$, then $x_1 \leq x_2$. (3) If $x_1 < x_2$ (for any x_2), then (*i*) $U_B(x) = U_e(x)$ and $Y_B(x) = y_e$ for $x \in [x_1, x_2]$; (*ii*) $U_B(x) > U_e(x)$ and $Y_B(x) < y_e$ for $x < x_1$; and (*iii*) for any $x_b \in [x_1, x_2]$, the segment $\{U_B(x), Y_B(x), \xi_B(x), \mu_B(x)\}_{x \leq x_b}$ solves RM_B given (x_b, y_e, T_e).

Proof: (1) From lemma 1.5, $x_2 > x_-$ implies $x_2 > x_s$, and $x_2 > x_s$ implies $U_A(x_2) = U_e(x_2)$ and $Y_A(x_2) = y_e$. Because $x_1 \leq x_2$ by construction, $x_1 < x_2$ if we can rule out $x_1 = x_2$.

Suppose to the contrary that $x_1 = x_2$, where $x_- < x_2 < x_+$. Then the solutions to RM_B and RM_A given (x_2, y_e, T_e) can be combined as follows: By construction, $U_B(x_2) = U_A(x_2)$ and $Y_B(x_2) = Y_A(x_2)$. Because $Y_B(x) \neq y_e$ for $x < x_2$, $\mu_B(x_2) = 0$ in a neighborhood of x_2, so $\frac{d}{dx}\mu_B(x_2) = 0$. Similarly, $Y_A(x) \neq y_e$ for $x > x_2$ implies $\mu_A(x_2) = 0$ and $\frac{d}{dx}\mu_A(x_2) = 0$. From (A.1), matching values of U, Y, and μ_x at x_2 imply $\xi_A(x_2) = \xi_B(x_2)$. Because all state and costate variables match, the profile obtained by combining solutions to RM_B given (x_2, y_e, T_e) and RM_A given (x_2, y_e, T_e) are continuous at x_2 and satisfy the conditions in lemma 1.1. Moreover, because RM_B satisfies the boundary conditions for $\widehat{\mathrm{RM}}$ at x_- and RM_A satisfies the boundary conditions for $\widehat{\mathrm{RM}}$ at x_+, the combined profile satisfies the sufficient conditions for $\widehat{\mathrm{RM}}$, which contradicts the assumption $(y_e, T_e) \in \mathcal{P}$. Thus $x_1 < x_2$.

(2) Trivial because the definition of x_1 implies $x_1 \leq x_2$.

(3) Given $x_1 < x_2$, $U_B(x) = U_e(x)$ for $x \in [x_1, x_2]$ implies $\frac{dU_B}{dx}(x) = \frac{dU_e}{dx}(x)$ and hence $Y_B(x) = y_e$. Because $U_B(x) > U_e(x)$ for $x < x_1$ in a neighborhood of x_1, $\frac{dU_B}{dx}(x) < \frac{dU_e}{dx}(x)$, so $Y_B(x) < y_e$ by lemma 1.3. Because Y_B is increasing, $Y_B(x) < y_e$ for $x < x_1$. Inspection of the sufficient conditions (see lemma 1.2) shows that $\{U_B(x), Y_B(x), \xi_B(x), \mu_B(x)\}_{x \leq x_b}$ solves RM_B given (x_b, y_e, T_e) for any $x_b \in [x_1, x_2]$. QED.

Lemmas 1.5-1.6 suggest a candidate solution for RM:

Definition: Consider $(y_e, T_e) \in \mathcal{P}$. If $x_- < x_2 < x_+$, define

$$\begin{aligned}\{U_{\mathrm{RM}}(x), Y_{\mathrm{RM}}(x)\}_{x \leq x_2} &\equiv \{U_B(x), Y_B(x)\}_{x \leq x_2}, \text{ and} \\ \{U_{\mathrm{RM}}(x), Y_{\mathrm{RM}}(x)\}_{x > x_2} &\equiv \{U_A(x), Y_A(x)\}_{x > x_2};\end{aligned}$$

if $x_2 = x_-$, define $\{U_{\mathrm{RM}}(x), Y_{\mathrm{RM}}(x)\} \equiv \{U_A(x), Y_A(x)\}$; and if $x_2 = x_+$, define $\{U_{\mathrm{RM}}(x), Y_{\mathrm{RM}}(x)\} \equiv \{U_B(x), Y_B(x)\}$.[32] Also, let $R_{\mathrm{RM}}(y_e, T_e) \equiv R_A(x_2, y_e, T_e) +$

[32] The use of x_2 in this definition is without loss of generality. One could have used any $x_b \in [x_1, x_2]$ to split $[x_-, x_+]$ into segments for which solutions to RM_B and RM_A are used to define

$R_B(x_2, y_e, T_e)$ be resulting revenue.[33]

The next lemma shows that revenue from any feasible profile $\{U_0(x), Y_0(x)\}$ is bounded by revenues obtained from RM_B and RM_A combined at a suitable point x_{b0}. This is then used in lemmas 1.8 and 1.9 to show that $\{U_{\text{RM}}(x), Y_{\text{RM}}(x)\}$ generates greater revenue than all other profiles that combine segments solving RM_B and RM_A, and hence is optimal.

Lemma 1.7: Consider any $(y_e, T_e) \in \mathcal{P}$ and any profile $\{U_0(x), Y_0(x)\}$ (not necessarily continuous) that satisfies the constraints of RM given (y_e, T_e), and let R_0 be its revenue. Then for some $x_{b0} \in [x_-, x_+]$ and some $T_{e0} \le T_e$, solutions to RM_B and RM_A given (x_{b0}, y_e, T_{e0}) yield revenue $R_A(x_{b0}, y_e, T_{e0}) + R_B(x_{b0}, y_e, T_{e0}) \ge R_0$.

Proof: There are three cases: (i) Suppose $Y_0(x) \ge y_e$ for some $x \in [x_-, x_+]$. Then define $x_{b0} = \inf\{x \in [x_-, x_+] \mid Y_0(x) \ge y_e\}$ and define T_{e0} by $U_0(x_{b0}) = u(y_e - T_{e0}, 1 - \frac{y_e}{x_{b0}})$. Because $U_0(x_{b0}) \ge u(y_e - T_e, 1 - \frac{y_e}{x_{b0}})$, $T_{e0} \le T_e$. By construction: $U_0(x_{b0}) = u(y_e - T_{e0}, 1 - y_e/x)$; $Y_0(x) \le y_e$ for $x \le x_{b0}$; and $Y_0(x) \ge y_e$ for $x \ge x_{b0}$. Hence $\{U_0(x), Y_0(x)\}_{x \le x_{b0}}$ satisfies the constraints of RM_B given (x_{b0}, y_e, T_{e0}) and $\{U_0(x), Y_0(x)\}_{x > x_{b0}}$ satisfies the constraints of RM_A given (x_{b0}, y_e, T_{e0}). Because solutions to RM_B and RM_A maximize revenue on their respective domains, $R_A(x_{b0}, y_e, T_{eb}) + R_B(x_{b0}, y_e, T_{eb}) \ge R_0$.

(ii) Suppose $Y_0(x) < y_e$ for all $x \in [x_-, x_+]$, and $x_+ < \infty$. Then define $x_{b0} = x_+$, define T_{eb} by $U_0(x_+) = u(y_e - T_{eb}, 1 - \frac{y_e}{x_+})$, and note that $\{U_0(x), Y_0(x)\}$ satisfies the constraints of RM_B given (x_+, y_e, T_{e0}), so $R_B(x_+, y_e, T_{e0}) \ge R_0$. Because $R_A(x_+, y_e, T_{e0}) = 0$, $R_A(x_{b0}, y_e, T_{e0}) + R_B(x_{b0}, y_e, T_{e0}) \ge R_0$.

(iii) Suppose $Y_0(x) < y_e$ for all $x \in [x_-, x_+]$, and $x_+ = \infty$. There is a $x_\tau < \infty$ such that $t_Y(U_0(x), Y_0(x), x) > 0$ for $x \ge x_\tau$, because $\frac{u_l(c^*(U_0(x), Y_0(x)/x), 1 - Y_0(x)/x)}{u_c(c^*(U_0(x), Y_0(x)/x), 1 - Y_0(x)/x)x} \to 0$

$\{U_{\text{RM}}(x), Y_{\text{RM}}(x)\}$.

[33]The definition of R_{RM} also applies to the border cases $x_2 = x_-$ and $x_2 = x_+$ because $R_B(x_-, y_e, T_e) = 0$ and $R_A(x_+, y_e, T_e) = 0$, respectively.

as $x \to \infty$. Define $x_{b0} = x_\tau$. Define Y_{0alt} by $Y_{0alt}(x) = Y_0(x)$ for $x < x_{b0}$ and $Y_{0alt}(x) = y_e$ for $x \geq x_{b0}$. Then by arguments analogous to the proof of lemma 1.4, $\{U_0(x), Y_{0alt}(x)\}$ generates revenue $R_{0alt} > R_0$. Moreover, profile $\{U_0(x), Y_{0alt}(x)\}$ satisfies the assumptions of case (i) above, which implies $R_A(x_{b0}, y_e, T_{e0}) + R_B(x_{b0}, y_e, T_{e0}) \geq R_{0alt} > R_0$. QED

Lemma 1.8: For $(y_e, T_e) \in \mathcal{P}$: (1) $R_A(x_b, y_e, T_e) + R_B(x_b, y_e, T_e) = R_{\text{RM}}(y_e, T_e)$ for any $x_b \in [x_1, x_2]$; (2) $R_A(x_{b0}, y_e, T_e) + R_B(x_{b0}, y_e, T_e) < R_{\text{RM}}(y_e, T_e)$ for $x_{b0} \notin [x_1, x_2]$; and (3) $R_{\text{RM}}(y_e, T_{e0}) < R_{\text{RM}}(y_e, T_e)$ for $T_{e0} < T_e$.

Proof: Part 1 follows directly from lemmas 1.5-1.6, parts 3. In part 2, either $x_{b0} < x_1$ or $x_{b0} > x_2$. If $x_{b0} < x_1$,

$$\begin{aligned} R_A(x_{b0}, y_e, T_e) &= \int_{x_{b0}}^{x_+} t(U(x), Y(x), x)dF(x) \\ &= \int_{x_{b0}}^{x_2} t(U_A(x), Y_A(x), x)dF(x) + R_A(x_2, y_e, T_e) \end{aligned}$$

from lemma 1.5, part 3. Because $\{U_{\text{RM}}(x), Y_{\text{RM}}(x)\}_{x \leq x_2}$ is the unique solution to RM_B given (x_2, y_e, T_e), $R_B(x_2, y_e, T_e) > \int_{x_{b0}}^{x_2} t(U_A(x), Y_A(x), x)dF(x) + R_B(x_{b0}, y_e, T_e)$. Adding:

$$\begin{aligned} & R_A(x_{b0}, y_e, T_e) + R_B(x_{b0}, y_e, T_e) \\ < \; & R_A(x_2, y_e, T_e) + R_B(x_2, y_e, T_e) = R_{\text{RM}}(y_e, T_e). \end{aligned}$$

(3) For any $T_{e0} < T_e$, $u(y_e - T_{e0}, 1 - \frac{y_e}{x}) > u(y_e - T_e, 1 - \frac{y_e}{x})$. Hence the profiles that solve RM_B and RM_A given (x_2, y_e, T_e) satisfy the constraints of RM_B and RM_A given (x_2, y_e, T_{e0}), which implies $R_{\text{RM}}(y_e, T_e) \leq R_{\text{RM}}(y_e, T_{e0})$. Moreover, $R_{\text{RM}}(y_e, T_e) \neq R_{\text{RM}}(y_e, T_{e0})$ because the profiles given (x_2, y_e, T_e) and given (x_2, y_e, T_{e0}) differ and the optimal solution is unique by lemma 1.2. Hence $R_{\text{RM}}(y_e, T_e) < R_{\text{RM}}(y_e, T_{e0})$. QED.

Lemma 1.9: For any $(y_e, T_e) \in \mathcal{P}$, the unique solution to RM is $\{U_{\text{RM}}(x), Y_{\text{RM}}(x)\}$, and $R(y_e, T_e) = R_{\text{RM}}(y_e, T_e)$.

Proof: From lemma 1.7, revenue R_0 for any profile that satisfies the constraints of RM is bounded by $R_A(x_{b0}, y_e, T_{e0}) + R_B(x_{b0}, y_e, T_{e0}) \geq R_0$ for some x_{b0} and $T_{e0} \leq T_e$. From lemma 1.8, $R_A(x_{b0}, y_e, T_{e0}) + R_B(x_{b0}, y_e, T_{e0}) \leq R_{\text{RM}}(y_e, T_e)$, with strict inequality unless $T_{e0} = T_e$ and $x_{b0} \in [x_1, x_2]$. Hence $R_{\text{RM}}(y_e, T_e) \geq R_0$, so $\{U_{\text{RM}}(x), Y_{\text{RM}}(x)\}$ solves RM.

From lemma 1.7 and the inequalities above, a profile $\{U_0(x), Y_0(x)\}$ cannot yield $R_0 = R_{\text{RM}}(y_e, T_e)$ unless it satisfies the constraints of RM_B and RM_A given (x_b, y_e, T_e). From lemma 1.2, $\{U_{\text{RM}}(x), Y_{\text{RM}}(x)\}$ is the only profile that satisfies these constraints and attains revenue $R_{\text{RM}}(y_e, T_e)$, so the solution is unique. QED.

The next two lemmas are used to prove proposition 1 part 2. Define $\mathcal{P}_y \equiv \{y_e \mid (y_e, T_e) \in \mathcal{P} \text{ for some } T_e\}$. Also define $R(y_e, T_e) = \hat{R}(\alpha)$ for $(y_e, T_e) \notin \mathcal{P}$, so R is defined for all (y_e, T_e).

Lemma 1.10: (1) For any $y_e \in [\hat{Y}(x_-), \hat{Y}(x_+)]$, $(y_e, T_e) \in \mathcal{P}$ if and only if $T_e < \hat{T}(y_e)$; (2) for any $y_e \in \mathcal{P}_y \backslash [\hat{Y}(x_-), \hat{Y}(x_+)]$, there is a critical value $\check{T}(y_e)$ such that $(y_e, T_e) \in \mathcal{P}$ if and only if $T_e < \check{T}(y_e)$; and (3) if $x_+ < \infty$, then $\mathcal{P}_y \supseteq [0, x_+)$.

Proof: (1) Because $\hat{Y}$ is continuous and non-decreasing, there is an x_{b0} such that $y_e = \hat{Y}(x_{b0})$ for any $y_e \in [\hat{Y}(x_-), \hat{Y}(x_+)]$. For $T_e < \hat{T}(\hat{Y}(x_{b0}))$, $U_e(x_{b0}) = u(y - T_e, 1 - \frac{y_e}{x_{b0}}) > \hat{U}(x_{b0})$, so $(y_e, T_e) \in \mathcal{P}$ by the definition of $\mathcal{P}$. Conversely if $T_e \geq \hat{T}(\hat{Y}(x_{b0}))$, then $U_e(x_{b0}) \leq \hat{U}(x_{b0})$. Because $\hat{Y}$ is non-decreasing, $\hat{Y}(x) \leq y_e$ for $x < x_{b0}$ and $\hat{Y}(x) \geq y_e$ for $x > x_{b0}$. From lemma 1.3, this implies $\frac{d\hat{U}}{dx} \leq \frac{dU_e}{dx}$ for $x < x_{b0}$ and $\frac{d\hat{U}}{dx} \geq \frac{dU_e}{dx}$ for $x > x_{b0}$, so $U_e(x) \leq \hat{U}(x)$ for all x. Thus $(y_e, T_e) \notin \mathcal{P}$.

(2) If $y_e \in \mathcal{P}_y \backslash [\hat{Y}(x_-), \hat{Y}(x_+)]$ then either (i) $y_e < \hat{Y}(x_-)$, (ii) $\hat{Y}(x_+) < y_e < x_+$, or (iii) $y_e = x_+$. (i) For $y_e < \hat{Y}(x_-)$, which implies $\hat{Y}(x_-) > 0$, define $\check{T}(y_e)$ by $u(y - \check{T}(y_e), 1 - \frac{y_e}{x_-}) = \hat{U}(x_-)$, so $(y_e, \check{T}(y_e))$ provides utility $\hat{U}(x)$ at $x = x_-$. Then (as in part 1 above) $T_e < \check{T}(y_e)$ implies $U_e(x_-) > \hat{U}(x_-)$ so $(y_e, T_e) \in \mathcal{P}$, and $T_e \geq \check{T}(y_e)$ implies $U_e(x) \leq \hat{U}(x)$ for all x so $(y_e, T_e) \notin \mathcal{P}$. (ii) For $\hat{Y}(x_+) < y_e < x_+$, which implies $\hat{Y}(x_+) < x_+$, define $\check{T}(y_e)$ by $u(y - \check{T}(y_e), 1 - \frac{y_e}{x_+}) = \hat{U}(x_+)$, so $(y_e, \check{T}(y_e))$

provides utility $\hat{U}(x)$ at $x = x_+$. Then $T_e < \check{T}(y_e)$ implies $U_e(x_+) > \hat{U}(x_+)$ so $(y_e, T_e) \in \mathcal{P}$, and $T_e \geq \check{T}(y_e)$ implies $U_e(x) \leq \hat{U}(x)$ for all x so $(y_e, T_e) \notin \mathcal{P}$. (*iii*) For $y_e = x_+$, which implies $x_+ < \infty$, there are two cases. If $u(c,l) \to -\infty$ as $l \to 0$, then $u(y - \check{T}(y_e), 1 - \frac{y_e}{x_-}) \to -\infty$ as $x \to x_+$ so there is no finite $\check{T}(x_+)$, whence $\mathcal{P}_y = [0, x_+)$. And if $u(c, 0)$ is finite, $u(y - \check{T}(x_+), 0) = \hat{U}(x_+)$ is defined as in case (*ii*), so $(x_+, T_e) \in \mathcal{P}$ for $T_e < \check{T}(x_+)$, whence $\mathcal{P}_y = [0, x_+]$.

(3) By the construction of $\check{T}$ in part 2, $\mathcal{P}_y = [0, x_+)$ unless $x_+ < \infty$ and $u(c, 0) < \infty$, in which case $\mathcal{P}_y = [0, x_+]$. Either way, $\mathcal{P}_y \supseteq [0, x_+)$. QED.

Definition: Define $\hat{T}(y_e) \equiv \check{T}(y_e)$ for $y_e \in \mathcal{P}_y \backslash [\hat{Y}(x_-), \hat{Y}(x_+)]$ so $\hat{T}$ is defined for $y_e \in \mathcal{P}_y$.

Lemma 1.11: R is (*i*) strictly increasing in T_e and (*ii*) continuously differentiable in (y_e, T_e) for $(y_e, T_e) \in \mathcal{P}$.

Proof: From standard value-function results, R_B and R_A are continuous and differentiable, so $R = R_B + R_A$ is continuous and differentiable.

(*i*) To sign $\partial R/\partial T_e$ when $x_1 < x_2$, choose any fixed $x_b \in (x_1, x_2)$. We write $(x \mid y_e, T_e)$ to emphasize (when needed) that profiles depend on (y_e, T_e). Then

$$\begin{aligned} \partial R/\partial T_e &= \partial R_B(x_b, y_e, T_e)/\partial T_e + \partial R_A(x_b, y_e, T_e)/\partial T_e \\ &= [\xi_B(x_b \mid y_e, T_e) - \xi_A(x_b \mid y_e, T_e)] \cdot u_c(y_e - T_e, 1 - y_e/x_b). \qquad \text{(A.3)} \end{aligned}$$

From lemmas 1.5-1.6, parts 3, the domain of ξ_B can be extended to $[x_-, x_2]$ by considering RM_B given (x_2, y_e, T_e) and the domain of ξ_A can be extended to $[x_1, x_+]$ by considering RM_A given (x_1, y_e, T_e), and these RM_B and RM_A problems have the same utility and income profiles on $[x_1, x_2]$. (However, the costate variables differ, as shown below.)

For $x \in [x_1, x_2]$, subtract (A.2) for RM_A given (x_1, y_e, T_e) from (A.2) for RM_B given (x_2, y_e, T_e) to obtain

$$\frac{d}{dx}\xi_B(x) - \frac{d}{dx}\xi_A(x) = -\omega_U(U_e(x), y_e, x) \cdot (\xi_B(x) - \xi_A(x)), \text{ for } x \in [x_1, x_2]. \qquad \text{(A.4)}$$

Because this is a homogenous linear differential equation, $\xi_B - \xi_A$ cannot change sign on $[x_1, x_2]$. Taking similar differences of Euler equations (A.1) and integrating over $[x_1, x_b]$ for any given $x_b \in [x_1, x_2]$:

$$\begin{aligned} I(x_b) &\equiv \int_{x_1}^{x_b} \omega_Y(U_e(x), y_e, x) \cdot [\xi_B(x) - \xi_A(x)]dx && \text{(A.5)} \\ &= \mu_A(x_b) + \mu_B(x_1) - \mu_A(x_1) - \mu_B(x_b). && \text{(A.6)} \end{aligned}$$

Because $\mu_B(x), \mu_A(x) \leq 0$ from lemma 1.1 and because the optimality conditions for RM_B and RM_A imply $\mu_B(x_1) = 0$ and $\mu_A(x_2) = 0$, it follows that $I(x_2) \geq 0$ and hence $\xi_B(x_b) - \xi_A(x_b) \geq 0$.

To show $\xi_B(x_b) > \xi_A(x_b)$, suppose for contradiction that $\xi_B(x_b) = \xi_A(x_b)$. Then (A.4) implies $\xi_B(x) = \xi_A(x)$ for $x \in [x_1, x_2]$ so $I(x_2) = 0$, whence $\mu_A(x_1) = \mu_B(x_2) = 0$. Then (as in the proof of lemma 1.6) RM_B and RM_A can be combined to obtain profiles that satisfy the sufficient conditions for a solution to $\widehat{\text{RM}}$, a contradiction. Thus

$$\xi_B(x_b) - \xi_A(x_b) > 0, \tag{A.7}$$

so (A.3) implies $\partial R/\partial T_e > 0$.

In the corner case $x_1 = x_2 = x_-$, the transversality condition $\xi_A(x_+) = 0$ and (A.2) imply $\xi_A(x_-) < 0$ so $\partial R/\partial T_e = -\xi_A(x_-) \cdot u_c > 0$. In the corner case $x_1 = x_2 = x_+$, the solution to RM_B given (x_+, y_e, T_e) has $\xi_B(x_+) \geq 0$. The case $\xi_B(x_+) = 0$ can be ruled out because it would imply (by arguments as in the proof of lemma 1.1) that RM_B given (x_+, y_e, T_e) would satisfy the sufficient conditions for a solution to $\widehat{\text{RM}}$, contradicting $(y_e, T_e) \in \mathcal{P}$. Hence $\xi_B(x_+) > 0$, so $\partial R/\partial T_e = \xi_B(x_+) \cdot u_c > 0$. Therefore $\partial R/\partial T_e > 0$ for $(y_e, T_e) \in \mathcal{P}$.

(*ii*) Note that

$$\begin{aligned}\partial R/\partial y_e &= \partial R_B(x_b, y_e, T_e)/\partial y_e + \partial R_A(x_b, y_e, T_e)/\partial y_e \\ &= \mu_A(x_b \mid y_e, T_e) - \mu_B(x_b \mid y_e, T_e) \\ &\quad +\partial U_e(x_b)/\partial y_e[\xi_A(x_b \mid y_e, T_e) - \xi_B(x_b \mid y_e, T_e)],\end{aligned} \tag{A.8}$$

where $\partial U_e(x)/\partial y_e = u_c(y_e - T_e, 1 - y_e/x_b) - u_l(y_e - T_e, 1 - y_e/x_b)/x_b$ is continuous. From (A.3) and (A.8), the derivatives $\partial R/\partial T_e$ and $\partial R/\partial y_e$ are continuous if the costate variables ξ_A, ξ_B, μ_A, and μ_B are continuous in (y_e, T_e) at x_b.

Consider first the continuity of ξ_A and μ_A. On $[x_2, x_+]$, LB implies $\mu_A(x) = 0$ in (A.1) so $H_Y(U_A, Y_A, \xi_A, x) = 0$. Because $H_{YY} < 0$ by CON, income $Y \equiv y(U, \xi, x)$ is an implicit function defined by $H_Y(U, Y, \xi, x) = 0$ and is differentiable in (U, ξ). Moreover, (A.1) and (A.2) with Y replaced by $y(U_A(x \mid y_e, T_e), \xi_A(x \mid y_e, T_e), x)$ is a system of two differential equations in U_A and ξ_A, which determine $\{U_A(x \mid y_e, T_e), \xi_A(x \mid y_e, T_e)\}_{x \geq x_2}$. Because the system is saddle-path stable (with characteristic matrix having a zero trace and, using CON, a negative determinant) and has boundary conditions $U_A(x_2 \mid y_e, T_e) = u(y_e - T_e, 1 - y_e/x_2)$ and $\xi_A(x_+ \mid y_e, T_e) = 0$ that are continuous in (y_e, T_e), it follows that $\xi_A(x \mid y_e, T_e) = 0$ is continuous in (y_e, T_e) for $x \geq x_2$.

On $[x_b, x_2]$, (A.2) is a linear differential equation for ξ_A, which has a solution that is continuous in the boundary value $\xi_A(x_2 \mid y_e, T_e)$. Hence $\xi_A(x_b \mid y_e, T_e)$ is also continuous in (y_e, T_e). Because $\mu_A(x_2 \mid y_e, T_e) = 0$, integration of (A.1) over $[x_b, x_2]$ implies that $\mu_A(x_b \mid y_e, T_e)$ is a function of ξ_A and (y_e, T_e), and hence is continuous in (y_e, T_e). Thus, $\xi_A(x_b)$ and $\mu_A(x_b)$ are continuous.

The costate variables ξ_B and μ_B are continuous by analogous arguments. Hence $\partial R/\partial T_e$ and $\partial R/\partial y_e$ are continuous. QED.

Proof of proposition 1: Existence of a unique solution to RM follows from lemma 1.9 for $(y_e, T_e) \in \mathcal{P}$ and from the uniqueness of a solution to $\widehat{\text{RM}}$ for $(y_e, T_e) \notin$

$\mathcal{P}$. The claims in part 1a and 1c follow directly from lemma 1.6, parts 1-3, and lemma 1.5, part 3. If $[x_1, x_2]$ includes any $x \in (x_-, x_+)$, part 1b also follows from lemma 1.6. If $[x_1, x_2]$ does not include any $x \in (x_-, x_+)$, then either $x_1 = x_2 = x_-$ or $x_1 = x_2 = x_+$. If $x_1 = x_2 = x_-$, RM solves RM_A given (x_-, y_e, T_e), which includes $y_e \leq Y(x_-)$ as a constraint. If $x_1 = x_2 = x_+$, RM solves RM_B given (x_+, y_e, T_e), which includes $y_e \geq Y(x_-)$ as a constraint. This proves part 1b.

In part 2, existence of T_e^* requires showing that for any $y_e \in \mathcal{P}_y$, $R(y_e, T_e) = G$ for some T_e. Recall that $\check{T}$ defined in lemma 1.10 satisfies $R(y_e, \check{T}(y_e)) = \hat{R} > G$, and that R decreases strictly as T_e falls below $\hat{T}(y_e)$. Taxes for $y < y_e$ are bounded by $T(y) \leq y < x$ because $c \geq 0$, $l > 0$, and taxes for $y \geq y_e$ are bounded by $T(y) \leq T_e + y - y_e < T_e + x - y_e$; otherwise $u(y - T(y), 1 - \frac{y}{x}) \geq U_e(x) > u(y_e - T_e, 1 - \frac{y}{x})$. Thus $R(y_e, T_e) \leq \int_{x_-}^{x+} x dF(x) + (T_e - y_e)(1 - F(x_b))$ for some $F(x_b) < 1$, so $R(y_e, T_e) \to -\infty$ as $T_e \to -\infty$. Hence for any $G \geq 0$ and any $\epsilon > 0$ there is a value $\tilde{T}(y_e \mid G - \epsilon)$ so that $R(y_e, \tilde{T}(y_e \mid G - \epsilon)) \leq G - \epsilon$. By the mean-value theorem, there is a unique $T_e^*(y_e) \in (\tilde{T}(y_e \mid G - \epsilon), \check{T}(y_e))$ that satisfies $R(y_e, T_e^*(y_e)) = G$. Differentiability of T_e^* follows from the implicit function theorem because from lemma 1.11, R is continuously differentiable. QED.

Proposition 2

To relate the winner's problem (maximize $U(x_e)$ by choice of $\{U(x), Y(x), \psi(x)\}$ subject to (1)-(4)) to problem RM, define the winner's **modified problem** as maximize $U_e(x_e) = u(y_e - T_e, 1 - y_e/x_e)$ by choice of $\{U(x), Y(x), \psi(x)\}$ and values (y_e, T_e), subject to (1)-(5). The modified problem explicitly gives the winner the additional choice of (y_e, T_e), but subjects the winner to the additional constraint (5). Then:

Lemma 2.1: A pair (y_e, T_e) and a profile $\{U(x), Y(x), \psi(x)\}$ solve the modified problem if and only $y_e \in \mathcal{Y}_e^*(x_e)$, $T_e = T_e^*(y_e)$, and $\{U(x), Y(x), \psi(x)\}$ solve RM given (y_e, T_e).

Proof: (i) Suppose $(y_e, T_e, \{U(x), Y(x), \psi(x)\})$ satisfies the constraints (1)-(5) of the modified problem, and let R_e denote the revenue obtained under $\{U(x), Y(x), \psi(x)\}$. By the definition of RM, $(y_e, T_e, \{U(x), Y(x), \psi(x)\})$ satisfy the constraints of RM given (y_e, T_e), which implies $R(y_e, T_e) \geq R_e$. Because $R_e \geq G$ from (3), it follows that $R(y_e, T_e) \geq G$ so $T_e \geq T_e^*(y_e)$. By the construction of T_e^*, profile $\{U(x), Y(x), \psi(x)\}$ that satisfies (1), (2), (4), and (5) cannot satisfy (3), so values $T_e < T_e^*(y_e)$ are not feasible for the modified problem. Hence solutions to the modified problem must satisfy $T_e \geq T_e^*(y_e)$. Thus feasible (y_e, T_e) necessarily satisfy $T_e \geq T_e^*(y_e)$. Note that $T_e = T_e^*(y_e)$ is feasible for any y_e by construction.

(ii) Suppose $(y_{e0}, T_{e0}, \{U_0(x), Y_0(x), \psi_0(x)\})$ solves the modified problem. Because $u(y_e - T_e, 1 - y_e/x_e)$ decreases strictly in T_e and $T_e \geq T_e^*(y_e)$ from (i), solutions to the modified problem must satisfy $T_{e0} = T_e^*(y_{e0})$. Hence y_{e0} must maximize $u(y_e - T_e^*(y_e), 1 - y_e/x_e)$, so by the definition of $\mathcal{Y}_e^*$, $y_{e0} \in \mathcal{Y}_e^*(x_e)$. Because the solution to RM given $(y_{e0}, T_e^*(y_{e0}))$ is unique, $\{U_0(x), Y_0(x), \psi_0(x)\}$ must solve RM given (y_e, T_e).

(iii) Suppose $y_{e1} \in \mathcal{Y}_e^*(x_e)$, $T_{e1} = T_e^*(y_{e1})$, and $\{U_1(x), Y_1(x), \psi_1(x)\}$ solve RM given (y_{e1}, T_{e1}), and let $U_1 = u(y_{e1} - T_{e1}, 1 - y_{e1}/x_e)$ denote the winner's utility. Because RM given (y_{e1}, T_{e1}) satisfies (1), (2), (4), and (5) and because $R(y_{e1}, T_{e1}) = G$ satisfies (3), $(y_{e1}, T_{e1}, \{U_1(x), Y_1(x), \psi_1(x)\})$ satisfies the constraints of the modified problem. Because $y_{e1} \in \mathcal{Y}_e^*(x_e)$, $U_1 \geq u(y_e - T_e^*(y_e), 1 - y_e/x_e) \geq u(y_e - T_e, 1 - y_e/x_e)$ for all y_e and $T_e \geq T_e^*(y_e)$. Because $T_e \geq T_e^*(y_e)$ is necessary for feasibility, (y_{e1}, T_{e1}) maximizes the winner's utility. QED.

Lemma 2.2: All solutions to the modified problem satisfy $x_e \in [x_1, x_2]$ and $U(x_e) = U_e(x_e)$. If $x_e \in (x_-, x_+)$, then $Y(x_e) = y_e$. If $x_e \in \{x_-, x_+\}$, solutions with $y_e \neq Y(x_e)$ are possible but inessential in that profiles $\{U(x), Y(x), \psi(x)\}$ that solve RM given $(y_e, T_e^*(y_e))$ are identical (everywhere) to the profile that solves RM given $(Y(x_e), T_e^*(Y(x_e)))$, which have $y_e = Y(x_e)$.

Proof: Solutions to the modified problem satisfy $U(x_e) \geq U_e(x_e)$ because (5) holds at $x = x_e$. If $U(x_e) > U_e(x_e)$ then x_e could raise own utility by choosing $(Y(x_e), T(x_e))$ (which is feasible) instead of (y_e, T_e), contradicting optimality. Thus $U(x_e) = U_e(x_e)$. From proposition 1, part 1a, this implies $x_e \in [x_1, x_2]$. For $x_e \in (x_-, x_+)$, proposition 1, part 1b implies $Y(x_e) = y_e$.

To show that $y_e \neq Y(x_e)$ is possible for $x_e \in \{x_-, x_+\}$, first suppose $x_e = x_-$ and consider any $y_{e0} \in \mathcal{Y}_e^*(x_-)$. Note that RM given $(y_{e0}, T_e^*(y_{e0}))$ is solved by RM_A given $(x_-, y_{e0}, T_e^*(y_{e0}))$, where $(y_{e0}, T_e^*(y_{e0}))$ constrains RM_A only through $U(x_-) = U_e(x_-)$. If $Y(x_-) > 0$, which is possible, then for any $y_e \leq Y(x_-)$ there is a value T_e such that $u(y_e - T_e, 1 - \frac{y_e}{x_-}) = U(x_-) = u(y_{e0} - T_e^*(y_{e0}), 1 - \frac{y_{e0}}{x_-})$ at productivity x_-. Consider values (y_{e-}, T_{e-}) with these properties. By construction, RM_A given (x_-, y_{e-}, T_{e-}) is solved by the same profile $\{U(x), Y(x), \psi(x)\}$ that solves RM_A given $(x_-, y_{e0}, T_e^*(y_{e0}))$. Because RM_A given (x_-, y_{e-}, T_{e-}) attains the maximum revenue, it must be that $T_{e-} = T_e^*(y_{e-})$. Hence the profile $\{U(x), Y(x), \psi(x)\}$ that solves RM given $(y_{e0}, T_e^*(y_{e0}))$ also solves RM given $(y_{e-}, T_e^*(y_{e-}))$ for $y_{e-} < Y(x_-)$, which means $\mathcal{Y}_e^*(x_-) = [0, Y(x_-)]$. Thus $\mathcal{Y}_e^*(x_-)$ includes values $y_e < Y(x_e)$.

Second, suppose $x_e = x_+ < \infty$ and consider any $y_{e0} \in \mathcal{Y}_e^*(x_+)$. Note that the solution satisfies $Y(x_+) < x_+$ and is also the solution to RM_B given $(x_+, y_{e0}, T_e^*(y_{e0}))$. Reasoning as in the first case, RM given $(y_e, T_e^*(y_e))$ for $y_e > Y(x_+)$ is solved by the same profile $\{U(x), Y(x), \psi(x)\}$, so $\mathcal{Y}_e^*(x_+) = [Y(x_+), x_+)$ includes values $y_e > Y(x_e)$. QED.

Remark: The special cases with $y_e \neq Y(x_e)$ for $x_e \in \{x_-, x_+\}$ are noted for mathematical completeness but are economically uninteresting because the utility and income profile that obtains for any $y_e \neq Y(x_e)$ is same as the profile obtained for $y_e = Y(x_e)$. Intuitively, $(y_e, T_e^*(y_e))$ are pairs on the winner's highest indifference curve. The proof also shows that $\mathcal{Y}_e^*(x_-) = [0, Y(x_-)]$ and that, if $x_+ < \infty$, $\mathcal{Y}_e^*(x_+) = [Y(x_+), x_+)$, so there are examples with multiple solution to x_e's modified problem.

Lemma 2.3: (1) If profile $\{U(x), Y(x), \psi(x)\}$ solves the winner's (original) problem, then $(Y(x_e), T_e^*(Y(x_e)), \{U(x), Y(x), \psi(x)\})$ solves the winner's modified problem. (2) If $(y_e, T_e, \{U(x), Y(x), \psi(x)\})$ solves the modified problem, then $\{U(x), Y(x), \psi(x)\}$ solves the original problem.

Proof: Profile $\{U(x), Y(x), \psi(x)\}$ that is feasible for the original problem is feasible for the modified problem by taking $y_e = Y(x_e)$ and $T_e = T(U(x_e), y_e, x_e)$. Therefore utility $U(x_e)$ in the original problem is less than or equal to the utility the winner obtains in any solution to the modified problem. From lemma 2.2, $y_e = Y(x_e)$ and $U(x_e) = U_e(x_e)$ for $x_e \in \{x_-, x_+\}$, so all solutions to the modified problem are feasible for the original problem, which means utility the winner obtains in any solution to the modified problem equals the utility $U(x_e)$ in the original problem. For $x_e \in \{x_-, x_+\}$, lemma 2.2 implies that solutions to the modified problem with $y_e \neq Y(x_e)$ have the same profile $\{U(x), Y(x), \psi(x)\}$ as solutions with $y_e = Y(x_e)$ and attain the same utility, so utilities in solutions to the original and modified problems are again equal.

Given equal utilities, part 1 follows because $(Y(x_e), T_e^*(Y(x_e)), \{U(x), Y(x), \psi(x)\})$ is feasible for the modified problem and attains maximum utility. Part 2 follows because, by lemma 2.2, profile $\{U(x), Y(x), \psi(x)\}$ that solves the modified problem is the same for $y_e \neq Y(x_e)$ and for $y_e = Y(x_e)$, so even if $y_e \neq Y(x_e)$ in the modified problem, $\{U(x), Y(x), \psi(x)\}$ solves the original problem. QED.

Proof of proposition 2: Follows directly from lemmas 2.1 and 2.3. QED.

Remark: In describing revenue maximization generally, proposition 1 imposes no restrictions on y_e so cases with $y_e \notin [Y(x_-), Y(x_+)]$ are possible. Lemma 2.2 implies that there is no loss of generality in restricting attention to revenue maximizations with $y_e \in [Y(x_-), Y(x_+)]$.

Proposition 3

The result is stated in Seade (1982) without proof. The proof is facilitated by the following lemma, which also justifies the graphical intuition in the text.

Lemma 3.1: Agent-monotonicity implies that the slope of indifference curves through any pair (y, T) increases strictly in x for all $y > 0$.

Proof: Let $T = y - c^*(U, \frac{y}{x})$ be the tax that yield utility U. Differentiating with respect to y, the slope of the indifference curve through (y, T) is

$$\frac{\partial T}{\partial y} = 1 - \frac{1}{x}\frac{\partial c^*}{\partial y} = 1 - \frac{1}{x}\frac{u_l(y - T, 1 - \frac{y}{x})}{u_c(y - T, 1 - \frac{y}{x})} \equiv S(x \mid y, T),$$

where $c = c^*(U, \frac{y}{x}) = y - T$. Agent-monotonicity requires that $n\frac{u_l(c,1-n)}{u_c(c,1-n)}$ increase strictly with n for given c. Because $n = \frac{y}{x}$ decreases with x for all $y > 0$, $\frac{y}{x}\frac{u_l(c,1-y/x)}{u_c(c,1-y/x)} = n\frac{u_l(c,1-n)}{u(c,1-n)}$ decreases strictly with x. Hence $\frac{\partial T}{\partial y} = 1 - \frac{1}{y} n \frac{u_l(c,1-n)}{u(c,1-n)}$ increases strictly with x for any (y, T) with $y > 0$. QED.

Proof of proposition 3: (i) Let $U_L(y_e) = u(y_e - T_e^*(y_e), 1 - \frac{y_e}{x_H})$ and $U_H(y_e) = u(y_e - T_e^*(y_e), 1 - \frac{y_e}{x_H})$ denote the utility levels of x_L and x_H at income y_e. Note that

$$\begin{aligned}\frac{\partial U_i(y_e)}{\partial y_e} &= u_c(y_e - T_e^*(y_e), 1 - \frac{y_e}{x_i})\left[1 - \frac{1}{x_i}\frac{u_l(y_e - T_e^*(y_e), 1 - y_e/x_i)}{u_c(y_e - T_e^*(y_e), 1 - y_e/x_i)} - \frac{\partial T_e^*}{\partial y_e}\right] \\ &= u_c(y_e - T_e^*(y_e), 1 - \frac{y_e}{x_i})\left[S(x_i \mid y_e, T_e^*(y_e)) - \frac{\partial T_e^*}{\partial y_e}\right]\end{aligned}$$

for $i = L, H$. By lemma 3.1, $S(x \mid y_e, T_e)$ increases strictly with x for $y_e > 0$, so $\frac{\partial U_L(y_e)}{\partial y_e} < \frac{\partial U_H(y_e)}{\partial y_e}$. For $y_H \in \mathcal{Y}_e^*(x_H)$ with $y_H > 0$, the optimality condition $S(x_H \mid y_H, T_e^*(y_H)) = \frac{\partial T_e^*}{\partial y_e}$ implies $\frac{\partial U_H(y_H)}{\partial y_e} = 0$ and hence $\frac{\partial U_L(y_H)}{\partial y_e} < 0$. This implies $U_L(y_H - \epsilon) > U_L(y_H)$ for some $\epsilon > 0$, proving that $y_H \notin \mathcal{Y}_e^*(x_L)$. Thus $y_L \neq y_H$ for $y_H > 0$.

(ii) Because $y_H \in \mathcal{Y}_e^*(x_H)$, $U_H(y_e) \leq U_H(y_H)$ for all y_e, so $T_e^*(y_e) \geq y_e - c^*(U_H(y_H), \frac{y_e}{x_H})$ for all y_e. At $(y_H, T_e^*(y_H))$, x_H's indifference curve with utility $U_H(y_H)$ crosses x_L's indifference curve with utility $U_L(y_H)$. Because $S(x_H \mid y_H, T_e^*(y_H)) > S(x_L \mid y_H, T_e^*(y_H))$, $y_e - c^*(U_H(y_H), \frac{y_e}{x_H}) > y_e - c^*(U_L(y_H), \frac{y_e}{x_L})$ for $y_e > y_H$ in a neighborhood of y_H; from lemma 3.1, the latter inequality holds for $y_e > y_H$. Hence $T_e^*(y_e) \geq y_e - c^*(U_H(y_H), \frac{y_e}{x_H}) > y_e - c^*(U_L(y_H), \frac{y_e}{x_L})$, which implies $U_L(y_e) < U_L(y_H)$ and hence $y_e \notin \mathcal{Y}_e^*(x_L)$ for $y_e > y_H$. Thus $y_L \leq y_H$.

(*iii*) For $y_H > 0$, $y_L \neq y_H$ from (*i*) and $y_L \leq y_H$ from (*ii*) imply $y_L < y_H$ For $y_H = 0$, $U_H(0) = U_L(0) = u(-T_e^*(0), 1)$. Also, $\frac{y_e}{x_H} < \frac{y_e}{x_L}$ for $y_e > 0$ implies $y_e - c^*(U_H(0), \frac{y_e}{x_H}) > y_e - c^*(U_L(0), \frac{y_e}{x_L})$. As in (*ii*), $y_e \notin \mathcal{Y}_e^*(x_L)$ for $y_e > y_H$, so $y_L = y_H = 0$. QED.

Proposition 4

For given x, we compare $U(x \mid y_L)$ and $U(x \mid y_H)$ by considering $U(x \mid y_e)$ as a function of y_e and integrating $\partial U(x \mid y_e)/\partial y_e$ over $[y_L, y_H]$ to obtain $U(x \mid y_H) - U(x \mid y_L)$.[34] Because $x_L < x_H$ and $y_H > 0$, proposition 3 implies $y_L < y_H$, so the interval $[y_L, y_H]$ is nondegenerate.

If $\mathcal{Y}_e^*$ is single-valued, proposition 3 implies that $\mathcal{Y}_e^*$ has an inverse, denoted $x_e^*(y_e)$, that is single-valued, continuous, and increases strictly. However, $\mathcal{Y}_e^*$ may be multi-valued for some x_e, and $y_e \notin \mathcal{Y}_e^*(x_e)$ is possible for some $y_e \in [\inf \mathcal{Y}_e^*(x_e), \sup \mathcal{Y}_e^*(x_e)]$, so $x_e^*(y_e)$ may not exists for some y_e. To handle such missing values in $x_e^*(y_e)$, note that $\mathcal{Y}_e^*(x_e)$ is compact-valued by the Maximum Theorem, so $\min\{\mathcal{Y}_e^*(x_e)\}$ and $\max\{\mathcal{Y}_e^*(x_e)\}$ exist and $x_e^*(y_e)$ is well-defined at the boundaries of $\mathcal{Y}_e^*(x_e)$. Hence one can partition $[y_L, y_H]$ into subintervals where $x_e^*(y_e)$ is single-valued, continuous, and strictly increasing (henceforth *type-S* intervals), and subintervals of the form $[\min\{\mathcal{Y}_e^*(x_e)\}, \max\{\mathcal{Y}_e^*(x_e)]$ for x_e where $\mathcal{Y}_e^*$ is multi-valued (*type-M* intervals).

We prove results separately for each type of interval. Because type-M intervals are closed, type-S intervals can be taken as open for purposes of defining a non-

[34] Profiles conditional on y_e should be interpreted as objects generated by RM for $(y_e, T_e) = (y_e, T_e^*(y_e))$, so T_e implicitly varies with y_e. This differs from the treatment in the proof of proposition 1, where y_e and T_e varied independently. Values x_1 and x_2 also depend on $(y_e, T_e^*(y_e))$.

overlapping partition. However, boundary points can be included when studying a particular interval. To streamline the notation associated with variations in y_e, we use primes to denote partial derivatives with respect to y_e (e.g., $U'(x \mid y_e) \equiv \partial U(x \mid y_e)/\partial y_e$).

Lemma 4.1: Consider $y_e \in (y_L, y_H)$ with $x_e^{*\prime}(y_e) > 0$, so y_e lies in a type-S interval, and $x_e = x_e^*(y_e)$. The solution to the winner's problem satisfies $x_1 < x_e < x_2$.

Proof: Because $x_1 \leq x_e \leq x_2$ from lemma 2.2, and $x_- \leq x_e^*(y_L) < x_e = x_e^*(y_e) < x_e^*(y_H) \leq x_+$, proposition 1 (part 1a) implies $x_1 < x_2$. Thus one must show that $x_e \neq x_1$ and $x_e \neq x_2$.

Because $x_e \in [x_1, x_2]$, proposition 1 (parts 1a,b) and proposition 2 (part 2) imply that the solution to the winner's problem solves RM_A given $(x_e, y_e, T_e^*(y_e))$ and RM_B given $(x_e, y_e, T_e^*(y_e))$, with $U_A(x_e) = U_B(x_e) = U_e(x_e)$ and $Y_A(x_e) = Y_B(x_e) = y_e$, and the optimal y_e implies $\partial T_e^*(y_e)/\partial y_e = S(x_e \mid y_e, T_e^*(y_e))$. From (A.3) and (A.8) in the proof of lemma 1.11, where $\partial U_e(x_e)/\partial y_e = [u_c - u_l/x] = u_c S(x_e \mid y_e, T_e)$, we have

$$\begin{aligned}
\frac{\partial R}{\partial y_e} &= \mu_A(x_e \mid y_e, T_e) - \mu_B(x_e \mid y_e, T_e) + S(x_e \mid y_e, T_e) u_c [\xi_A(x_e \mid y_e, T_e) - \xi_B(x_e \mid y_e, T_e)] \\
&= \mu_A(x_e \mid y_e, T_e) - \mu_B(x_e \mid y_e, T_e) - S(x_e \mid y_e, T_e) \frac{\partial R}{\partial T_e}.
\end{aligned}$$

Differentiating $R(y_e, T_e^*(y_e)) = G$ totally, $\frac{\partial R}{\partial y_e} + \frac{\partial R}{\partial T_e} \cdot \partial T_e^*(x_e)/\partial y_e = \frac{\partial R}{\partial y_e} + \frac{\partial R}{\partial T_e} \cdot S(x_e \mid y_e, T_e) = 0$, so $\mu_A(x_e \mid y_e, T_e) = \mu_B(x_e \mid y_e, T_e)$.

Because $\xi_B(x_b) > \xi_A(x_b)$ for all $x_b \in [x_1, x_2]$, as shown in the proof of lemma 1.11, (A.1) implies $\frac{\partial}{\partial x}\mu_A(x_b) > \frac{\partial}{\partial x}\mu_B(x_b)$. Because Y_B is increasing in a neighborhood below x_1, $\mu_B(x) = 0$ so $\frac{\partial}{\partial x}\mu_B(x_1) = 0$ and $\frac{\partial}{\partial x}\mu_A(x_1) > 0$. Similarly, Y_A is increasing above x_2, which implies $\mu_A(x) = 0$ and $\frac{\partial}{\partial x}\mu_A(x_2) = 0$, so $\frac{\partial}{\partial x}\mu_B(x_2) < 0$.

To show $x_e \neq x_1$, note that $x_e = x_1$ would imply $\mu_A(x_e) = \mu_A(x_1) = 0$. Combined with $\frac{\partial}{\partial x}\mu_A(x_1) > 0$, this would imply $\mu_A(x) > 0$ for some $x > x_1$, contradicting the optimality condition $\mu_A(x) \leq 0$ in lemma 1.1. To show $x_e \neq x_2$, note that $x_e = x_2$

would imply $\mu_B(x_e) = \mu_B(x_2) = 0$. Combined with $\frac{\partial}{\partial x}\mu_B(x_2) < 0$, this would imply $\mu_B(x) > 0$ for some $x < x_2$, contradicting the optimality condition $\mu_B(x) \leq 0$ in lemma 1.1. QED.

Lemma 4.2: Consider $y_e \in (y_L, y_H)$ with $x_e^{*\prime}(y_e) > 0$, so y_e lies in a type-S interval. Then $U'(x \mid y_e) < 0$ for $x \in [x_1, x_e)$, $U'(x \mid y_e) > 0$ for $x \in (x_e, x_2]$, and $U'(x_e \mid y_e) = 0$.

Proof: Because $U(x \mid y_e) = U_e(x)$ for $x \in [x_1, x_2]$,

$$\begin{aligned} U'(x \mid y_e) &= dU_e(x)/dy_e = [u_c - u_l/x] - u_c \cdot \partial T_e^*/\partial y_e \\ &= u_c \cdot [S(x \mid y_e, T_e^*(y_e)) - \partial T_e^*/\partial y_e] = u_c \cdot [S(x \mid y_e, T_e^*(y_e)) - S(x_e \mid y_e, T_e^*(y_e)]. \end{aligned} \tag{A.9}$$

Hence $U'(x_e \mid y_e) = 0$, and the inequalities for $x < x_e$ and $x > x_e$ follow because S increases strictly in x from lemma 3.1. QED.

Lemma 4.3: Consider $y_e \in (y_L, y_H)$ with $x_e^{*\prime}(y_e) > 0$, so y_e lies in a type-S interval. Then $U'(x \mid y_e) > 0$ for $x \in [x_2, x_+]$.

Proof: Recall that $\{U(x \mid y_e), Y(x \mid y_e), \xi(x \mid y_e), \mu(x \mid y_e)\}_{x \geq x_2}$ solves RM_A for $(y_e, T_e^*(y_e))$. In (A.1), LB implies $\mu(x) = 0$ and $\mu_x(x) = 0$, so $H_Y(U, Y, \xi, x) = 0$ for all x. Because $H_{YY} < 0$ by CON, $H_Y(U, Y, \xi, x) = 0$ defines a unique income level $Y \equiv y(U, \xi, x)$, which can be used to replace $Y(x)$ in (1) and (A.2). By the implicit function theorem, y is differentiable in (U, ξ) for given x. Hence (1) and (A.2) imply that $U(x \mid y_e)$ and $\xi(x \mid y_e)$ satisfy the differential equations

$$\begin{aligned} U_x(x \mid y_e) &= \omega(U(x \mid y_e), y(U(x \mid y_e), \xi(x \mid y_e), x), x) \\ \xi_x(x \mid y_e) &= -t_U(U(x \mid y_e), y(U(x \mid y_e), \xi(x \mid y_e), x), x) f(x) \\ &\quad -\xi(x \mid y_e) \cdot \omega_U(U(x \mid y_e), y(U(x \mid y_e), \xi(x \mid y_e), x), x). \end{aligned} \tag{A.10}$$

System (A.10) is saddle-path stable because the characteristic matrix has a zero trace and, from CON, a negative determinant. Hence, a solution to (A.10) (which exists by CON) is uniquely determined by two boundary conditions, which are $U(x_2 \mid$

$y_e) = U_e(x_2)$ and $\xi(x_+ \mid y_e) = 0$. For any given y_e, (A.10) and $Y = y(U, \xi, x)$ uniquely determine $\{U(x \mid y_e), Y(x \mid y_e)\}$ on $[x_2, x_+]$. 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 (A.10) with respect to y_e to obtain[35]

$$U'_x(x \mid y_e) = \gamma(U, \xi, x) \cdot U'(x \mid y_e) + \gamma_{U'\xi}(U, \xi, x) \cdot \xi'(x \mid y_e) \quad \text{(A.11)}$$
$$\xi'_x(x \mid y_e) = \gamma_{\xi'U}(U, \xi, x) \cdot U'(x \mid y_e) - \gamma(U, \xi, x) \cdot \xi'(x \mid y_e),$$

where $\gamma(U, \xi, x) = \omega_U - \omega_Y \cdot H_{YU}/H_{YY}$, $\gamma_{U'\xi}(U, \xi, x) = -\omega_{YY} > 0$ and $\gamma_{\xi'U}(U, \xi, x) = -[H_{YY} \cdot H_{UU} - H_{YU} \cdot H_{YU}]/H_{YY} > 0$. Second, solve (A.11) subject to the boundary conditions that $U'(x_2 \mid y_e)$ is given by (A.9) and $\xi'(x_+ \mid y_e) = 0$. (These are the derivatives of the boundary conditions for (A.10).) Because (A.10) has a characteristic matrix with a zero trace and a negative determinant, it is saddle-path stable. Because $U'(x_2 \mid y_e) = dU_e(x_2)/dy_e > 0$ from lemma 4.2, saddle-path stability implies $U'(x \mid y_e) > 0$ for $x \in [x_2, x_+]$. QED.

The analysis of $U(x \mid y_e)$ and $U'(x \mid y_e)$ for $x \in [x_-, x_1]$ requires multiple case distinctions because RM_B for $(y_e, T^*_e(y_e))$ is constrained by $U(x_- \mid y_e) \geq u(\alpha, 1)$ 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 ($\xi(x_- \mid y_e) \leq 0$, $\mu(x_- \mid y_e) \leq 0$). To organize the cases, define $x_0(y_e) \equiv \inf\{x \geq x_- \mid Y(x \mid y_e) > 0\}$, which is the maximum productivity in the set of non-workers if the set is non-empty and $x_0(y_e) = x_-$ otherwise, and define the sets

$$\begin{aligned}
\Xi_a &\equiv \{y_e \in \mathcal{P}_y \mid x_0(y_e) = x_-, U(x_- \mid y_e) = u(\alpha, 1)\}, \\
\Xi_b &\equiv \{y_e \in \mathcal{P}_y \mid x_0(y_e) = x_-, \xi(x_- \mid y_e) = 0\}, \\
\Xi_c &\equiv \{y_e \in \mathcal{P}_y \mid Y(x_- \mid y_e) = 0, U(x_- \mid y_e) = u(\alpha, 1)\}, \\
\Xi_d &\equiv \{y_e \in \mathcal{P}_y \mid Y(x_- \mid y_e) = 0, \xi(x_- \mid y_e) = 0\},
\end{aligned}$$

[35]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.

where $\{U(x \mid y_e), Y(x \mid y_e), \xi(x \mid y_e)\}_{x \leq x_1}$ denotes the solution to RM_B given $(y_e, T_e^*(y_e))$. The following properties describe RM_B:

Lemma 4.4: For $y_e \in \mathcal{P}_y$, $y_e \in \Xi_j$ for some $j \in \{a, b, c, d\}$.

Proof: Because $\xi(x_- \mid y_e)[U(x_- \mid y_e) - u(\alpha, 1)] = 0$ from lemma 1.1, $\xi(x_- \mid y_e) = 0$ or $U(x_- \mid y_e) = u(\alpha, 1)$ (or both) for y_e. Because $x_0(y_e) > x_-$ implies $Y(x_- \mid y_e) = 0$, it must be that $x_0(y_e) = x_-$ or $Y(x_- \mid y_e) = 0$ (or both) for y_e. Hence $y_e \in \Xi_j$ for some $j \in \{a, b, c, d\}$. QED.

Remark: The sets Ξ_j define possible configurations of boundary conditions. Not all of them necessarily occur; that is, $\Xi_j = \varnothing$ for some j is possible. (For example, if $\alpha = 0$, $x_- > 0$, $u_c/u_l \to \infty$ as $c \to 0$, and tax rates are bounded away from 100%, then $Y(x_- \mid y_e) > 0$, so $\Xi_c = \Xi_d = \varnothing$.) The task is to prove the proposition for all possible cases.

Lemma 4.5: Consider $y_e \in (y_L, y_H)$ with $x_e^{*\prime}(y_e) > 0$, so y_e lies in a type-S interval. Then $U'(x \mid y_e) < 0$ for $x \in [x_-, x_1] \backslash X_{u(\alpha,1)}(y_e)$.

Proof: Recall that $\{U(x \mid y_e), Y(x \mid y_e), \xi(x \mid y_e), \mu(x \mid y_e)\}_{x \leq x_1}$ solves RM_B given $(y_e, T_e^*(y_e))$. As in lemma 4.3, LB implies $Y(x \mid y_e) = y(U, \xi, x)$ for $x \in [x_0(y_e), x_1]$, so $U(x \mid y_e)$ and $\xi(x \mid y_e)$ satisfy (A.10) and (A.11). One boundary condition for (A.10) is $U(x_1 \mid y_e) = U_e(x_1)$, which is invoked in all cases below. The corresponding condition for (A.11) is $U'(x_1 \mid y_e) = dU_e(x_1)/dy_e$, so $U'(x_1 \mid y_e) < 0$ from lemma 4.2. Additional boundary conditions for (A.10) and (A.11), and their implications, require case distinctions:

(a) For $y_e \in \Xi_a$, $x_0(y_e) = x_-$ implies that (A.10) holds on $[x_-, x_1]$, and $U(x_- \mid y_e) = u(\alpha, 1)$ provides the second boundary condition, so there is a unique solution. The analysis of (A.11) is analogous to the proof of lemma 4.3: $\{U'(x \mid y_e), \xi'(x \mid y_e)\}$ for $x \leq x_1$ solves (A.11), with boundary conditions that (i) $U'(x_1 \mid y_e)$ satisfies (A.9) and (ii) $U'(x_- \mid y_e) = 0$, which follows from $U(x_- \mid y_e) = u(\alpha, 1)$. Saddle-path stability (as detailed in the proof of lemma 4.3) and $U'(x_1 \mid y_e) < 0$ (from lemma 4.2) imply $U'(x \mid y_e) < 0$ for $x \in (x_-, x_1]$. Because $Y > 0$ for $x > x_-$ implies $U(x \mid y_e) > u(\alpha, 1)$, it follows that $X_{u(\alpha,1)}(y_e) = \{x_-\}$, so $U'(x \mid y_e) < 0$ for $[x_-, x_1] \backslash X_{u(\alpha,1)}(y_e)$.

(*b*) For $y_e \in \Xi_b$, cases with $U(x_- \mid y_e) = u(\alpha, 1)$ imply $y_e \in \Xi_a$ so the results in (*a*) apply. Otherwise $U(x_- \mid y_e) > u(\alpha, 1)$, which means $y_e \in \Xi_b \backslash \Xi_a$. Then $x_0(y_e) = x_-$ implies that (A.10) applies on $[x_-, x_1]$ and $\xi(x_- \mid y_e) = 0$ provides the second boundary condition, again ensuring a unique solution. Taking derivatives, $\{U'(x \mid y_e), \xi'(x \mid y_e)\}_{x \leq x_1}$ is determined by (A.11) with boundary conditions that (*i*) $U'(x_1 \mid y_e)$ satisfies (A.9) and (*ii*) $\xi'(x_- \mid y_e) = 0$. Saddle-path stability and $U'(x_1 \mid y_e) < 0$ imply $U'(x \mid y_e) < 0$ for $x \in [x_-, x_1]$. Also, $U(x_- \mid y_e) > u(\alpha, 1)$ for $y_e \in \Xi_b \backslash \Xi_a$ implies $X_{u(\alpha,1)}(y_e) = \varnothing$.

(*c*) For $y_e \in \Xi_c$, cases with $x_0(y_e) = x_-$ imply $y_e \in \Xi_a$ so the results in (*a*) apply. Otherwise $x_0(y_e) > x_-$, which means $y_e \in \Xi_c \backslash \Xi_a$. Then for $x \leq x_0 \equiv x_0(y_e)$, $Y = 0$ implies constant U, so $U(x \mid y_e) = U(x_- \mid y_e) = u(\alpha, 1)$. Because $Y > 0$ in a neighborhood of $x > x_0$, so $\psi(x) > 0$, lemma 1.1 implies $\mu(x) = 0$ whence $\mu_x(x_0) = 0$; also, $Y > 0$ implies $U(x \mid y_e) > u(\alpha, 1)$ for $x > x_0$. Thus $X_{u(\alpha,1)}(y_e) = [x_-, x_0]$. From (A.1) with $\mu_x(x_0) = 0$, $H_Y(u(\alpha, 1), 0, \xi(x_0 \mid y_e), x_0) = 0$. Because LB holds for (x_0, x_1), (A.10) holds on $[x_0, x_1]$, and the second boundary condition is the open endpoint condition at x_0 that $U(x_0 \mid y_e) = u(\alpha, 1)$ and $H_Y(u(\alpha, 1), 0, \xi(x_0 \mid y_e), x_0) = 0$. 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 $[x_0, x_1]$. Taking derivatives at x_0, $U'(x_0 \mid y_e) = 0$ and $\xi'(x_0 \mid y_e) = 0$, so $\{U'(x \mid y_e), \xi'(x \mid y_e)\}_{[x_0, x_1]}$ is determined by (A.11) with boundary conditions that (*i*) $U'(x_1 \mid y_e)$ satisfies (A.9) and (*ii*) at x_0, $U'(x_0 \mid y_e) = 0$ and $\xi'(x_0 \mid y_e) = 0$. Saddle-path stability and $U'(x_1 \mid y_e) < 0$ imply $U'(x \mid y_e) < 0$ for $x \in (x_0, x_1]$. Also, $X_{u(\alpha,1)}(y_e) = [x_-, x_0]$ because $U(x \mid y_e) = u(\alpha, 1)$ iff $x \leq x_0$. Thus $U'(x \mid y_e) < 0$ for $[x_-, x_1] \backslash X_{u(\alpha,1)}(y_e)$.

(*d*) If $y_e \in \Xi_d$, cases with $U(x_- \mid y_e) = u(\alpha, 1)$ imply $y_e \in \Xi_a \cup \Xi_c$ so results in

(a) or (c) apply. Cases with $U(x_- \mid y_e) > u(\alpha,1)$ and $x_0(y_e) = x_-$ imply $y_e \in \Xi_b \backslash \Xi_a$ so the results in (b) apply. In all other cases, $x_0(y_e) > x_-$ and $U(x_- \mid y_e) > u(\alpha,1)$, which means $y_e \in \Xi_d \backslash (\Xi_a \cup \Xi_b \cup \Xi_b)$. Then for $x \leq x_0$, $Y = 0$ implies constant U, so $U(x \mid y_e) = U(x_- \mid y_e) > u(\alpha,1)$ and $X_{u(\alpha,1)}(y_e) = \varnothing$. Because $Y > 0$ in a neighborhood of $x > x_0$, it follows that $\mu(x) = 0$ so $\mu_x(x_0) = 0$, and hence $H_Y(u(\alpha,1), 0, \xi(x_0 \mid y_e), x_0) = 0$ follows as in (c). Because LB holds for $x \in (x_0, x_1)$, (A.10) applies on $[x_0, x_1]$ with an open endpoint condition at x_0. To derive the latter, note that (A.2) for $[x_-, x_0]$ implies $\xi_x(x \mid y_e) = -H_U(U(x_0 \mid y_e), 0, \xi(x \mid y_e), x)$, which can be integrated to obtain

$$\xi(x_0 \mid y_e) = F(x_0)/u_c(c^*(U(x_0 \mid y_e), 0), 1), \tag{A.12}$$

using the boundary condition $\xi(x_- \mid y_e) = 0$. Taking derivatives at x_0,

$$\xi'(x_0 \mid y_e) = F(x_0)\left(-\frac{u_{cc}(c^*(U(x_0 \mid y_e), 0), 1)}{u_c(c^*(U(x_0 \mid y_e), 0), 1)^3}\right) \cdot U'(x_0 \mid y_e). \tag{A.13}$$

Thus $\{U'(x \mid y_e), \xi'(x \mid y_e)\}_{[x_0, x_1]}$ is determined by (A.11) with boundary conditions (i) that $U'(x_1 \mid y_e)$ satisfies (A.9) and (ii) condition (A.13) at x_0. Saddle-path stability and $U'(x_1 \mid y_e) < 0$ imply $U'(x \mid y_e) < 0$ for $x \in [x_0, x_1]$. For $x \leq x_0$, $U(x \mid y_e) = U(x_0 \mid y_e)$ implies $U'(x \mid y_e) = U'(x_0 \mid y_e)$, so $U'(x \mid y_e) < 0$ for $x \in [x_-, x_1]$.

By lemma 4.4 (part 1), cases a-d cover all possible configurations of boundary conditions, so $U'(x \mid y_e) < 0$ for $[x_-, x_1] \backslash X_{u(\alpha,1)}(y_e)$. QED.

Lemma 4.6: Consider a set $\Xi_S \subseteq [y_L, y_H]$ with $x_e^{*\prime}(y_e) > 0$ for $y_e \in \Xi_S$, which means Ξ_S is a type-S interval. Then for any y_l, $y_h \in \Xi_S$ with $y_l < y_h$, (1) $X_{u(\alpha,1)}(y_l) \subseteq X_{u(\alpha,1)}(y_h)$ and (2) $x_0(y_h) > x_0(y_l)$ whenever $\max X_{u(\alpha,1)}(y_h) > x_-$.

Proof: (1) Either (i) $y_h \in (\Xi_b \cup \Xi_d) \backslash (\Xi_a \cup \Xi_c)$ or (ii) $y_h \in \Xi_a$ or (iii) $y_h \in \Xi_c \backslash\, \Xi_a$. (i) For $y_h \in (\Xi_b \cup \Xi_d) \backslash (\Xi_a \cup \Xi_c)$, $U(x_- \mid y_h) > u(\alpha,1)$ by definition and $U'(x_- \mid y_e) < 0$ from lemma 4.5, so $U(x_- \mid y_l) > U(x_- \mid y_h) > u(\alpha,1)$. Hence

$X_{u(\alpha,1)}(y_l) = \varnothing = X_{u(\alpha,1)}(y_h)$ so $X_{u(\alpha,1)}(y_l) \subseteq X_{u(\alpha,1)}(y_h)$. (*ii*) For $y_h \in \Xi_a$, $U(x \mid y_h) > u(\alpha,1)$ for $x > x_-$ by definition and $U'(x_- \mid y_e) < 0$ from lemma 4.5, so $U(x \mid y_l) > U(x \mid y_h) > u(\alpha,1)$ for $x \in (x_-, x_1]$. Hence $X_{u(\alpha,1)}(y_l) \subseteq \{x_-\} = X_{u(\alpha,1)}(y_h)$. (*iii*) For $y_h \in \Xi_c \backslash\ \Xi_a$, $X_{u(\alpha,1)}(y_h) = [x_-, x_0(y_h)]$ with $x_0(y_h) > x_-$. Because Y satisfies $H_Y(U, Y, \xi, x) = 0$ for $x > x_0(y_h)$ and Y is continuous, $Y(x_0 \mid y_h)$ satisfies $H_Y(u(\alpha,1), Y(x_0 \mid y_h), \xi(x_0(y_h) \mid y_h), x_0) = 0$ at $x_0 = x_0(y_h)$. Taking derivatives with respect to y_e for given x_0, $\partial Y(x_0 \mid y_h)/\partial y_e = (-H_{YY})^{-1}\omega_Y \xi'(x_0 \mid y_h)$. Because $H_{YY} < 0$, $\omega_Y > 0$, and (A.13) plus $U'(x_0 \mid y_e) < 0$ imply $\xi'(x_0 \mid y_h) < 0$, $\partial Y(x_0 \mid y_h)/\partial y_e < 0$ at given x_0. Because $Y(x_0 \mid y_h) = 0$, x_0 must increase to satisfy $Y(x_0 \mid y_e) \geq 0$, so $x_0(y_e)$ increases strictly with y_e. Hence $x_0(y_l) < x_0(y_h)$ and $X_{u(\alpha,1)}(y_l) = [x_-, x_0(y_l)] \subset [x_-, x_0(y_h)] = X_{u(\alpha,1)}(y_h)$. (2) Because $\max X_{u(\alpha,1)}(y_h) > x_-$ implies $y_h \in \Xi_c \backslash\ \Xi_a$, $x_0(y_h) > x_0(y_l)$ follows from the argument in (*iii*) that $x_0(y_e)$ increases strictly in y_e. QED.

Lemma 4.7: Consider a set $\Xi_S \subseteq [y_L, y_H]$ with $x_e^{*\prime}(y_e) > 0$ for $y_e \in \Xi_S$, so again Ξ_S is a type-S interval. For any y_l, $y_h \in \Xi_S$ with $y_l < y_h$, denote $x_h = x_e^{*\prime}(y_h)$ and $x_l = x_e^{*\prime}(y_l)$. Then $U(x \mid y_h) - U(x \mid y_l) > 0$ for $x \geq x_h$, $U(x \mid y_h) - U(x \mid y_l) < 0$ for $x \in [x_-, x_l] \backslash X_{u(\alpha,1)}(y_l)$, and $U(x \mid y_h) = U(x \mid y_l) = u(\alpha,1)$ for $x \in X_{u(\alpha,1)}(y_l)$.

Proof: For any $x \in [x_-, x_+]$, $U(x \mid y_h) - U(x \mid y_l) = \int_{y_l}^{y_h} U'(x \mid y_e) dy_e$. For $x \geq x_h$, lemma 4.3 implies $U'(x \mid y_e) > 0$ for $y_e \in [y_l, y_h]$, so $U(x \mid y_h) - U(x \mid y_l) > 0$. For $x \leq x_l$, lemma 4.6 implies that $X_{u(\alpha,1)}(y_l) \subseteq X_{u(\alpha,1)}(y_h)$, so $[x_-, x_l] \backslash X_{u(\alpha,1)}(y_l)$ is the union of $[x_-, x_l] \backslash X_{u(\alpha,1)}(y_h)$ and $X_{u(\alpha,1)}(y_h) \backslash X_{u(\alpha,1)}(y_l)$, and $[x_-, x_l] \backslash X_{u(\alpha,1)}(y_h) \subseteq [x_-, x_l] \backslash X_{u(\alpha,1)}(y_l)$. For $x \in [x_-, x_l] \backslash X_{u(\alpha,1)}(y_h)$, lemma 4.5 implies $U'(x \mid y_e) < 0$ for $y_e \in [y_l, y_h]$, so $U(x \mid y_h) - U(x \mid y_l) < 0$. For $x \in X_{u(\alpha,1)}(y_h) \backslash X_{u(\alpha,1)}(y_l)$, $U(x \mid y_l) > u(\alpha,1) = U(x \mid y_h)$ by the definition of $X_{u(\alpha,1)}$, so $U(x \mid y_h) - U(x \mid y_l) < 0$. For $x \in X_{u(\alpha,1)}(y_l)$, lemma 4.6 implies $X_{u(\alpha,1)}(y_l) = X_{u(\alpha,1)}(y_l) \cap X_{u(\alpha,1)}(y_h)$, so $U(x \mid y_h) = U(x \mid y_l) = u(\alpha,1)$ by the definition of $X_{u(\alpha,1)}$. QED.

Remark: Although lemma 4.7 proves results that resemble the claims of proposition 4, a general proof of lemma 4.7 requires treatment of multi-valued $\mathcal{Y}_e^*(x_e)$ (within $[y_L, y_H]$). To do this, we use the following definitions.

Definitions: If $\mathcal{Y}_e^*(x_e) \cap [y_L, y_H]$ has multiple elements for some $x_e \in [x_L, x_H]$, define $y_{e\min}(x_e) \equiv \min\{\mathcal{Y}_e^*(x_e) \cap [y_L, y_H]\}$ and $y_{e\max}(x_e) \equiv \max\{\mathcal{Y}_e^*(x_e) \cap [y_L, y_H]\}$. Let $\Xi_M(x_e) = [y_{e\min}(x_e), y_{e\max}(x_e)]$ formally define a type-M interval. Let $\tilde{U}^*(x_e) \equiv U(x_e \mid y_e)$ where $y_e \in \Xi_M^*(x_e)$ denote x_e's maximum utility. Let $\Xi_M^*(x_e) = \Xi_M(x_e) \cap \mathcal{Y}_e^*(x_e)$ be the subset of $\Xi_M(x_e)$ for which $U(x_e \mid y_e) = \tilde{U}^*(x_e)$. For $y_e \in \Xi_M(x_e)$, define $\tilde{T}(y_e)$ by $u(y_e - \tilde{T}(y_e), 1 - \frac{y_e}{x_e}) = \tilde{U}^*(x_e)$, and let $\{\tilde{U}(x \mid y_e), \tilde{Y}(x \mid y_e), \tilde{\xi}(x \mid y_e), \tilde{\mu}(x \mid y_e)\}$ denote the solution to RM for $(y_e, \tilde{T}(y_e))$.

Remark: Because $\mathcal{Y}_e^*(x_e)$ is not necessarily an interval, some elements of $\Xi_M(x_e)$ may not maximize the winner's utility. This is a complication because if $U(x_e \mid y_e)$ varies on $\Xi_M(x_e)$, one cannot rule out $U'(x_e \mid y_e) \neq 0$. Hence the analysis of type-S intervals in lemmas 4.2-4.3, which relies on $U'(x_e \mid y_e) = 0$, does not generalize to type-M intervals. To sidestep this complication, we consider problems RM for $(y_e, \tilde{T}(y_e))$ instead of RM for $(y_e, T_e^*(y_e))$, having constructed $\tilde{T}$ such that $\tilde{U}(x_e \mid y_e)$ is constant for $y_e \in \Xi_M(x_e)$, which implies $\tilde{U}'(x_e \mid y_e) = 0$. As the following lemmas show, this will allow us to derive results about $U(x_e \mid y_e)$.

Lemma 4.8: Consider $x_e \in [x_L, x_H]$ with $y_{e\min}(x_e) < y_{e\max}(x_e)$. Then: (1) $U(x_e \mid y_e) = \tilde{U}(x_e \mid y_e)$ for $y_e \in \Xi_M^*(x_e)$; and (2) $\partial\tilde{U}(x_e \mid y_e)/\partial y_e = 0$, $\partial\tilde{U}(x \mid y_e)/\partial y_e \geq 0$ for $x \geq x_e$ and $\partial\tilde{U}(x \mid y_e)/\partial y_e \leq 0$ for $x \leq x_e$, for $y_e \in \Xi_M(x_e)$.

Proof: (1) For $y_e \in \Xi_M^*(x_e)$, $U(x_e \mid y_e) = U_e(x_e)$ implies $u(y_e - \tilde{T}(y_e), 1 - \frac{y_e}{x_e}) = \tilde{U}^*(x_e) = U(x_e \mid y_e) = u(y_e - T_e^*(y_e), 1 - \frac{y_e}{x_e})$. Hence $\tilde{T}(y_e) = T_e^*(y_e)$ and $U(x_e \mid y_e) = \tilde{U}(x_e \mid y_e)$. (2) For $y_e \in \Xi_M(x_e)$, $u(y_e - T_e^*(y_e), 1 - \frac{y_e}{x_e}) \leq \tilde{U}^*(x_e) = u(y_e - \tilde{T}(y_e), 1 - \frac{y_e}{x_e})$ implies $\tilde{T}(y_e) \leq T_e^*(y_e)$, whence $R(y_e, \tilde{T}(y_e)) \leq R(y_e, T_e^*(y_e)) = G < \hat{G}$ and $(y_e, \tilde{T}(y_e)) \in \mathcal{P}$. Thus proposition 1 ensures the existence of an interval $[x_1, x_2]$ with $\tilde{U}(x \mid y_e) = u(y_e - \tilde{T}(y_e), 1 - \frac{y_e}{x})$ and $\tilde{Y}(x \mid y_e) = y_e$ for $x \in [x_1, x_2]$. Moreover, $x_e \in [x_1, x_2]$ holds by the same arguments as the proof of lemma 2.2.

(Note, however, that the arguments for $x_e \neq x_1, x_2$ in the proof of lemma 4.1 do not have an analogy here, so $x_e = x_1$ or $x_e = x_2$ are possible.) As in the proof of lemma 4.2, $\tilde{U}'(x \mid y_e) - \tilde{U}'(x_e \mid y_e) = u_c \cdot [S(x \mid y_e, T_e^*(y_e)) - S(x_e \mid y_e, T_e^*(y_e)]$ where $\tilde{U}'(x_c \mid y_e) = 0$ by construction. Following the proofs of lemmas 4.2-4.3, $\tilde{U}'(x \mid y_e) \geq 0$ for $x > x_e$ (with strict inequality if $x_2 > x_e$), and following the proofs of lemmas 4.4-4.6, $\tilde{U}'(x \mid y_e) \leq 0$ for $x < x_e$ (with strict inequality if $x_1 < x_e$ and $\tilde{U}(x \mid y_e) > u(\alpha, 1)$). QED.

Lemma 4.9: Consider $x_e \in [x_L, x_H]$ with $y_{e\,\min}(x_e) < y_{e\,\max}(x_e)$. Then for any $y_l \in \Xi_M^*(x_e)$ and $y_h \in \Xi_M^*(x_e)$ with $y_l < y_h$: $U(x_e \mid y_h) = U(x_e \mid y_l)$; $U(x \mid y_h) \geq U(x \mid y_l)$ for $x > x_e$; $U(x \mid y_h) \leq U(x \mid y_l)$ for $x \in [x_-, x_e) \backslash X_{u(\alpha,1)}(y_l)$; and $U(x \mid y_h) = U(x \mid y_l) = u(\alpha, 1)$ for $x \in X_{u(\alpha,1)}(y_l)$.

Proof: From $y_l \in \Xi_M^*(x_e)$, $y_h \in \Xi_M^*(x_e)$, and lemma 4.8 (part 1), $\tilde{U}(x \mid y_h) - \tilde{U}(x \mid y_l) = \int_{y_l}^{y_h} \tilde{U}'(x \mid y_e) dy_e = U(x \mid y_h) - U(x \mid y_l)$. Hence $U(x \mid y_h) \geq U(x \mid y_l)$ for $x > x_e$ and $U(x \mid y_h) \leq U(x \mid y_l)$ for $x < x_e$ follow from lemma 4.8 (part 2). Also, $U(x_e \mid y_h) = U(x_e \mid y_l)$ because $y_l \in \Xi_M^*(x_e)$ and $y_h \in \Xi_M^*(x_e)$. From $X(y_l) \subseteq X(y_h)$ and the definition of $X_{u(\alpha,1)}$, $U(x \mid y_h) = U(x \mid y_l) = u(\alpha, 1)$. QED.

Proof of proposition 4: Because $x_L < x_H$, $[y_L, y_H]$ includes at least one type-S interval. If $[y_L, y_H]$ includes type-M intervals, for each of them either $y_{e\,\min}(x_e) = y_L$ or $y_{e\,\min}(x_e) = \max \Xi_S$ is the upper endpoint of a type-S interval; and either $y_{e\,\max}(x_e) = y_H$ or $y_{e\,\max}(x_e) = \min \Xi_S$ for another interval of type S. Hence $[y_L, y_H]$ decomposes into alternating type-S and type-M intervals. Pick any type-S interval and denote it $[y_l, y_h]$ so $y_L \leq y_l < y_h \leq y_H$. Then:

(a) For $x \in [x_H, x_+]$, lemmas 4.7 and 4.9 imply $U(x \mid y_H) \geq U(x \mid y_h) > U(x \mid y_l) \geq U(x \mid y_L)$.

(b) For $x \in [x_-, x_L]$, lemmas 4.7 and 4.9 imply $U(x \mid y_H) - U(x \mid y_L) \leq U(x \mid y_h) - U(x \mid y_l) \leq 0$. Within $[x_-, x_L]$: (*i*) if $x \in X_{u(\alpha,1)}(y_L)$, $X_{u(\alpha,1)}(y_L) \subseteq X_{u(\alpha,1)}(y_H)$ implies $U(x \mid y_H) = U(x \mid y_L) = u(\alpha, 1)$; (*ii*) if $x \in X_{u(\alpha,1)}(y_H) \backslash X_{u(\alpha,1)}(y_L)$, $U(x \mid$

$y_L) > U(x \mid y_H) = u(\alpha, 1)$ by the definition of $X_{u(\alpha,1)}$; (*iii*) if $x \in [x_-, x_L] \backslash X_{u(\alpha,1)}(y_H)$, then $x \in [x_-, x_L] \backslash X_{u(\alpha,1)}(y_l)$, so lemma 4.7 implies $U(x \mid y_H) \leq U(x \mid y_h) < U(x \mid y_l) \leq U(x \mid y_L)$.

(c) For $x \in (x_L, x_H)$, $U(x \mid y_H) - U(x \mid y_L)$ is continuous in x, negative at x_L and positive at x_H, so the mean-value theorem implies the existence of a 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. A unique crossing point implies $U(x \mid y_H) - U(x \mid y_L) > 0$ for $x \in (x_\times, x_H)$ and $U(x \mid y_H) - U(x \mid y_L) < 0$ for $x \in (x_L, x_\times)$.

Then proposition 4 part 1 follows from steps (a) and (c-i) above, part 2 follows from steps (b-ii, iii) and (c-ii), part 3 follows from step (c), and part 4 follows from step (b-i). QED.

Proposition 5

Proof: Voting decisions of individuals $x \notin X_{u(\alpha,1)}(y_L)$ follow from proposition 4 (parts 1-3). In detail: (1) If $X_{u(\alpha,1)}(y_L)$ has zero measure, then $U(x_M \mid y_L) > U(x_M \mid y_H)$ implies that $x_\times > x_M$ and that $(x_-, x_\times)$ is a majority. Because voters in $[x_-, x_\times) \backslash X_{u(\alpha,1)}(y_L) \supseteq (x_-, x_\times)$ prefer x_L over x_H, x_L wins. If $U(x_M \mid y_L) < U(x_M \mid y_H)$, analogous arguments imply that x_H wins. If $U(x_M \mid y_L) = U(x_M \mid y_H)$ and $x_\times = x_M$ then $(x_-, x_\times)$ and $(x_\times, x_+]$ both have measure 1/2 and the vote is tied. (2a) If indifferent individuals vote by closeness, then because $X_{u(\alpha,1)}(y_L)$ is an interval that starts at x_-, individuals $x \in X_{u(\alpha,1)}(y_L)$ vote for x_L and hence voting choices are as if $X_{u(\alpha,1)}(y_L)$ had zero measure. (2b) If $X_{u(\alpha,1)}(y_L)$ has positive measure and $x \in X_{u(\alpha,1)}(y_L)$ abstain, let x'_M denote the median of $[x_-, x_+] \backslash X_{u(\alpha,1)}(y_L)$. Then $U(x'_M \mid y_L) > U(x'_M \mid y_H)$ implies $x_\times > x'_M$, so $(x_-, x_\times)$ has greater measure than $(x_\times, x_+]$ and x_L wins. The reverse applies if $U(x'_M \mid y_L) < U(x'_M \mid y_H)$. QED.

Proposition 6

Proof: Immediate from proposition 5 (parts 1 and 2a).

Remark: As noted in the text, proposition 6 generalizes to (non-generic) cases with multi-valued $\mathcal{Y}_e^*(x_M)$. We prove this here as a lemma:

Lemma 6.1: (1) Suppose $0 \notin \mathcal{Y}_e^*(x_M)$ and either $X_{u(\alpha,1)}(y_M)$ has zero measure for $y_M = \min\{\mathcal{Y}_e^*(x_M)\}$ or indifferent individuals vote by closeness. Then x_M wins against any other candidate. (2) Suppose $0 \in \mathcal{Y}_e^*(x_M)$. Then: x_M wins against any other candidate who sets $y_e > 0$; x_M ties against candidates who sets $y_e = 0$; and regardless of opponent and election outcome, the winner's tax function maximizes x_M's utility.

Proof: (1-i) Suppose $0 \notin \mathcal{Y}_e^*(x_M)$ and $X_{u(\alpha,1)}(y_M)$ has zero measure for $y_M = \min\{\mathcal{Y}_e^*(x_M)\}$. (Note that taking the minimum $y_e \in \mathcal{Y}_e^*(x_M)$ is least restrictive because low y_e minimizes the measure of $X_{u(\alpha,1)}(y_e)$.) For opponents with given $x_e > x_M$, proposition 3 implies that $y_e > y_M > 0$ for any $y_e \in \mathcal{Y}_e^*(x_e)$. Hence proposition 4 with $x_H = x_e$, $x_L = x_M$, and $y_H \in \min\{\mathcal{Y}_e^*(x_e)\} > 0$ implies $x_\times > x_M$ and $U(x \mid y_M) > U(x \mid y_e)$ for $x \in [x_-, x_\times)\setminus X_{u(\alpha,1)}(y_M)$. Because $X_{u(\alpha,1)}(y_M)$ has zero measure, $[x_-, x_\times)\setminus X_{u(\alpha,1)}(y_M) \supseteq (x_-, x_\times)$, which is a majority; so x_M wins. For opponents with $x_e < x_M$, proposition 4 with $x_H = x_M$, $x_L = x_e$, and $y_H = y_M > 0$ implies $x_\times < x_M$ and $U(x \mid y_M) > U(x \mid y_e)$ for $x \in (x_\times, x_+]$. Because $(x_\times, x_+]$ is a majority, x_M wins.

(1-ii) Suppose $0 \notin \mathcal{Y}_e^*(x_M)$ and indifferent individuals vote by closeness. Then as in (1-i), proposition 4 with $x_e > x_M$ implies $x_\times > x_M$. Moreover, voting by closeness implies that individuals in $X_{u(\alpha,1)}(y_M)$ vote for x_M, so $x \in [x_-, x_\times)$ vote for x_M; because this is a majority, x_M wins. Also as in (1-i), proposition 4 with $x_e < x_M$, implies $x_\times < x_M$ and $U(x \mid y_M) > U(x \mid y_e)$ for $x \in (x_\times, x_+]$; because $(x_\times, x_+]$ is a majority, x_M wins.

(2) Suppose $0 \in \mathcal{Y}_e^*(x_M)$. By proposition 3, candidates who set $y_e > 0$ must have $x_e > x_M$, so proposition 4 with $x_H = x_e$, $x_L = x_M$ implies $x_\times > x_M$. Note that $X_{u(\alpha,1)}(0) = \varnothing$ because candidates who set $y_e = 0$ maximize $U(x \mid 0)$, which implies $U(x \mid 0) > u(\alpha, 1)$. Hence, $U(x \mid 0) > U(x \mid y_e)$ for $x \in [x_-, x_\times)$, which is a majority. Thus by choosing $y_M = 0$, x_M wins. Opponents who set $y_e = 0$ are trivially tied if x_M chooses $y_M = 0$ so both candidates set the same policy. If $\mathcal{Y}_e^*(x_M)$ is multi-valued and x_M chooses $y_M > 0$, $0 \in \mathcal{Y}_e^*(x_M)$ implies $x_\times = x_M$, so the vote is tied. In all cases, the winner implements $U(x \mid 0)$, which maximizes $U(x_M \mid 0)$ because $0 \in \mathcal{Y}_e^*(x_M)$. QED.

Proposition 7

From $t(U, Y, x) \equiv Y - c^*(U, \frac{Y}{x})$, $T(Y(x)) = t(U(x), Y(x), x)$ determines T for all $y \in [Y(x_-), Y(x_+)]$. Continuity of U, Y , and t implies continuity of T. From CON, Y has a piecewise continuous derivative ψ, so $\frac{d}{dy}T(Y(x)) \cdot \psi(x) = t_U(U, Y, x)\frac{dU}{dx} + t_Y(U, Y, x)\psi(x) + t_x$, except at x where ψ is discontinuous. From (1) and $t(U, Y, x) \equiv Y - c^*(U, \frac{Y}{x})$, $t_U(U, Y, x)\frac{dU}{dx} + t_x = -c_U^* \cdot \omega(U, Y, x) + c_n^* \cdot Y/x^2 = (-1/u_c) \cdot u_l Y/x^2 + (u_l/u_c) \cdot Y/x^2 = 0$, so $\frac{d}{dy}T(Y(x)) \cdot \psi(x) = t_Y(U, Y, x) \cdot \psi(x)$.

From LB, Y^{-1} exists for all $y \in [Y(x_-), Y(x_+)]$ except possibly at $y = 0$ and $y = y_e$. Moreover, $\psi(Y^{-1}(y)) > 0$ wherever Y^{-1} is defined, and ψ is continuous, so $\frac{d}{dy}T(Y(x)) = t_Y(U, Y, x) = \tau(x)$ whence $dT(y)/dy = t_Y(U, Y, x) = \tau(Y^{-1}(y))$. 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, (6) holds for all $y \in [Y(x_-), Y(x_+)] \backslash \{0,\, y_e\}$. If $0 \in [Y(x_-), Y(x_+)]$, then $dT(0)/dy \equiv lim_{y \to 0} dT(y)/dy = \tau(\max\{x \mid Y(x \mid y_e) = 0\})$ is well-defined. Thus, T is continuously differentiable on $[Y(x_-), Y(x_+)] \backslash \{y_e\}$. (Because $Y^{-1}(0)$ is undefined if $\{x \mid Y(x \mid y_e) = 0\}$ is an interval, (6) may not hold at $y = 0$.)

1. On $[x_1, x_2]$, $\tau(x) = S(x \mid y_e, T_e)$, which increases strictly 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 for $y > y_e$ in a neighborhood of y_e.

2. Monotonicity of Y implies $Y(x) \geq y_e$. Because $(y_e, T_e) \in \mathcal{P}$ by proposition 2, proposition 1 (2b) requires that $\{U(x), Y(x)\}$ solve RM$_A$ on $[x_e, x_+]$ and RM$_B$ on $[x_-, x_e]$. Integrating the Euler equation (A.2):

$$\xi_A(x) = -\int_x^{x_+} \frac{\eta(x,z)}{u_c(c^*(U(z), Y(z)/z), 1 - Y(z)/z)} dF(z) + \xi_A(x_+), \tag{A.14}$$

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

3. Monotonicity of Y implies $Y(x) \leq y_e$. Integrating the Euler equation (A.2):

$$\xi_B(x) = \int_{x_-}^{x} \frac{\eta(x,z)}{u_c(c^*(U(z), Y(z)/z), 1 - Y(z)/z)} dF(z) + \xi_B(x_-). \tag{A.15}$$

Because (4) holds if and only if $U(x_-) \geq u(\alpha, 1)$, the transversality conditions for RM$_B$ are $[U(x_-) - u(\alpha, 1)] \cdot \xi_B(x_-) = 0$ and $\xi_B(x_-) \leq 0$. Because $\eta/u_c > 0$, ξ_B is strictly increasing. Thus there are three possibilities: $\xi_B < 0$ for $x < x_1$; ξ_B switches sign from negative to positive at a point $x_\tau \in (x_-, x_1)$ where $\xi_B(x_\tau) = 0$; or $\xi_B \geq 0$ for $x < x_1$. As above, (A.1) and LB imply that $\xi_B(x)$ and $\tau(x)$ have opposite signs on $[x_0, x_1]$.

If $\xi_B < 0$ for $x < x_1$ then $\tau(x) > 0$ on $[x_0, x_1)$ so T increases strictly; this is case c. If $\xi_B(x_\tau) = 0$ for some $x_\tau \in (x_-, x_1)$, the shape of T depends on x_τ and x_0. If $x_\tau > x_0$, (A.1) and LB 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 $\xi_B \geq 0$ for

$x < x_1$, (A.1) and LB imply $\tau(x) < 0$ on $[x_0, x_1)$, so T increases strictly; this is case (a). If (4) does not bind ($\xi_B(x_-) = 0$), then (A.15) implies $\xi_B > 0$ on $(x_-, x_1]$, which is case a. Q.E.D.

Proposition 8

The Euler equations for $\widehat{\text{RM}}$ imply that the costate variable associated with $\hat{U}$, denoted $\hat{\xi}$, satisfies (A.14) for all x. Because $\eta/u_c > 0$ and $\hat{\xi}(x_+) = 0$, it follows that $\hat{\xi}(x) < 0$ for $x < x_+$. By (A.1), $\hat{\tau}(x) > 0$ for $x < x_+$ and hence for $x \in [x_-, x_e]$ in the solution to $\widehat{\text{RM}}$. As $\{U(x), Y(x)\} \rightarrow \{\hat{U}(x), \hat{Y}(x)\}$, we have $\tau(x) \rightarrow \hat{\tau}(x)$ pointwise. Because $[x_-, x_e]$ is compact, the convergence $\tau(x) \rightarrow \hat{\tau}(x) > 0$ 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 $x \geq x_2$, it follows that $\tau(x) > 0$ for all $x \in [x_-, x_+]$. Q.E.D.