*Econometrica*, Vol. 67, No. 1 (January, 1999), 1–19

# THE DEMOCRATIC POLITICAL ECONOMY OF PROGRESSIVE INCOME TAXATION

BY JOHN E. ROEMER[1]

Why do both left and right political parties typically propose progressive income taxation schemes in political competition? Analysis of this problem has been hindered by the two-dimensionality of the issue space. To give parties a choice over a domain that contains both progressive and regressive income tax policies requires an issue space that is at least two-dimensional. Nash equilibrium in pure strategies of the standard two-party game, whose players have complete preferences over a two-dimensional policy space, generically fails to exist.

I introduce a new equilibrium concept for political games, based on the fact of factional conflict within parties. Each party is supposed to consist of reformists, militants, and opportunists: each faction has a complete preference order on policy space, but together they can only agree on a partial order. Nash equilibria of the two-party game, where the policy space consists of all quadratic income tax functions, and each party is represented by its partial order, exist, and it is shown that, in such equilibria, both parties propose progressive income taxation.

KEYWORDS: Progressive taxation, politico-economic equilibrium.

## 1. INTRODUCTION

WHY DO BOTH LEFT AND RIGHT parties tend to propose progressive income tax policies in democratic political competition? Some authors (e.g., Young (1994)) have used arguments of fairness, but such arguments are surely not in the spirit of political economy, in which players (in this case, parties) are primarily assumed to be self-interested. Marhuenda and Ortuño (1995) note that the "the literature ... is still very inconclusive on the connection between progressive taxation and voting." Snyder and Kramer (1988) analyze the problem, and reach the right conclusion, but only under an unjustifiably strong constraint, namely, that parties may propose only tax functions that are ideal for some voter. Cukierman and Meltzer (1991) study the question when a Condorcet winner exists among quadratic income tax schemes, when voters have preferences over income and leisure, but succeed in demonstrating such existence only under unreasonable conditions.[2] Moreover, it is only in a "Downsian" framework that players propose the Condorcet winner, if there is one, in equilibrium.

Most formal political analysts of party competition, until recently, have assumed that both parties are Downsian—that their objective is to maximize the probability of winning office. On the other hand, a long tradition, dating from

[1] I am grateful to three referees and a co-editor for careful comments that significantly improved the exposition in the paper. Seminar audiences, at various universities and conferences, have also been most helpful.

[2] I do not wish to imply that the analysis which follows dominates that of Cukierman and Meltzer, for they work with a class of utility functions that include leisure as an argument, while I do not.

Michels (1915) to Lipset (1959) to Przeworski and Sprague (1986), takes it as given that parties represent, perhaps imperfectly, different constituencies among the population, such as economic classes. There has, in the past fifteen years, been a growing formal literature on competition between partisan parties (ones that have policy preferences)—see, e.g., Wittman (1983), Calvert (1985), Alesina and Rosenthal (1995), and Roemer (1997). But almost all of that analysis has assumed that the issue space is uni-dimensional (e.g., voting on a proportional income tax, or an affine tax scheme for which the budget balances). To study why parties adopt progressive income tax rules requires an issue space which is at least two-dimensional, that is, one in which both progressive and regressive income tax policies can be represented.

In this study, I assume that parties represent, imperfectly, different constituencies, or economic classes. Society's problem is to choose an income tax regime. Since I wish to study when that regime will be progressive, I work with the family of quadratic income taxes, where *after-tax income* of an individual takes the form $aw^2 + bw + c$, where $w$ is the individual's income. Using a balanced budget constraint, and assuming that taxes are purely redistributive, we may view the domain of policies as two dimensional, deriving $c$ as a function of $a$ and $b$, and hence regard a tax policy as an ordered pair $(a, b)$. I shall assume there are two parties, and each must propose a policy, consisting of a tax policy $(a, b)$, in the political contest.

The problem immediately encountered when working with this set-up is that, except for singular cases, a Nash equilibrium (in pure strategies) does not exist in the political game between parties—because the strategy sets are two-dimensional. This is, indeed, the source of the "inconclusiveness" referred to by Marhuenda and Ortuño above. There are two standard moves the analyst can make when confronted with the nonexistence of Nash equilibrium in pure strategies—first, to allow mixed strategies, or second, to change the game into a stage game, and use some variant of perfect equilibrium (e.g., Stackelberg or Rubinstein). But, in the case of competing political parties, I find neither of these moves appealing. I do not think we can easily interpret political parties as playing mixed strategies. And the stage-game model is appropriate only when there is a clear first mover. Some, including myself, have attempted to argue that a political contest takes place between a challenger and an incumbent, and either the challenger (Roemer (in press)) or the incumbent (Bernhardt and Ingberman (1985)) is the natural leader. These arguments, however, are inconclusive. When there is no compelling argument for one player's moving first, it is more appropriate to model the contest as one with simultaneous moves.

The heart of the present paper consists in a new equilibrium concept in such political games. In a word, I shall retain the notion of Nash equilibrium in a simultaneous-move game between the players (parties), but replace their preferences with incomplete preferences, with respect to which Nash equilibria shall exist. The proposal is historically motivated, being based upon evidence of the nature of inner-party struggle over the line (policy), which derives from twentieth-century European political history.

Let us suppose that the interest group (say, economic class) a party purports to represent has well-defined policy preferences, represented by a von Neumann-Morgenstern utility function $v_L$ defined on the space of policies, whose generic element is $t$. (In our context, $t$ stands for a policy $(a, b)$ and $L$ stands for "Left," one of the constituencies.) Parties will be assumed to be uncertain about the distribution of voter types, so that, if the Left and Right parties propose policies $t_L$ and $t_R$ respectively, there is only a probability, $\pi(t_L, t_R)$, that Left wins by majority vote. (Parties agree on the priors, and therefore calculate the same function $\pi$.) The expected utility of the Left constituency, under this pair of proposals, is therefore

$$(1.0) \qquad \Pi^L(t_L, t_R) = \pi(t_L, t_R)v_L(t_L) + (1 - \pi(t_L, t_R))v_L(t_R).$$

The recent formal, non-Downsian literature to which I referred above models political equilibrium as Nash equilibrium between parties $L$ and $R$ with preferences $\Pi^L$ and the analogous $\Pi^R$, where $v_L$ is replaced by a function $v_R$ in expression (1.0).

My institutional assumption, in contrast, is that there are three groups of actors within each party: facing a proposal $t_R$ by the Right, the *reformists* within the Left party wish to choose $t_L$ to maximize $\Pi^L(t_L, t_R)$; the *militants* in Left—who want the party to adhere as closely as possible to its principles—wish to choose $t_L$ to maximize simply $v_L(t_L)$, or equivalently, the after-tax income of Left's constituency, and the *opportunists* in Left, who desire only to win office, wish to choose $t_L$ to maximize $\pi(t_L, t_R)$. There are reformists, militants, and opportunists in the Right party who wish to maximize analogous functions.

Party histories are replete with descriptions of militants and opportunists, if not using that nomenclature—see, for example, Przeworski and Sprague's (1986) history of European social democracy. Despite the names I have given these three collective actors, it would be incorrect to view them as, necessarily, not having the interests of the constituents in mind. In the history of European social-democratic parties, an on-going debate took place concerning the purpose of participating in elections: many maintained that that purpose was not to win office, but to educate the working-class, to develop their consciousness. These *militants* believed the party should stick to its principles (here interpreted as its ideal point), and not compromise for the sake of winning office: they were revolutionary anti-reformists. Maximizing the expected utility $\Pi^L(t_L, t_R)$ is definitely a *reformist* approach, in which the decision maker cares about the *expected* utility of its constituents, given the uncertainty surrounding elections. The (revolutionary) militant, in contrast, is not particularly interested in how well the constituency will fare under the policy chosen by the current election, but rather whether the process of the election will build class consciousness or further cement an ideology. On the other hand, the actors whom I've called opportunist were not necessarily venal politicians, merely using the party as a stepping stone to a political career, although some doubtless were. They may, more generally, have believed that by holding office, the party would be able more effectively to develop the consciousness/ideology of its constituency. Winning, after all, would provide a bully pulpit from which the victorious party could address the population daily. For such *opportunists*, the objective $\Pi^L(t_L, t_R)$ was also a myopic one.

How does a party, whose activists consist of these three factions, make a decision about its policy in the electoral contest? My proposal is that it must reach *inner-party unanimity* to do so. Facing a proposal $t_R$ by the opposition, the Left will choose policy $t_L$ over an alternative $t'_L$ only if all three inner-party factions agree, that is, only if:

(1.1) $$\Pi^L(t_L, t_R) \geq \Pi^L(t'_L, t_R),$$

(1.2) $$\pi(t_L, t_R) \geq \pi(t'_L, t_R), \quad \text{and}$$

(1.3) $$v_L(t_L) \geq v_L(t'_L),$$

with at least one strict inequality. These three inequalities are, respectively, the decision criteria of the reformists, the opportunists, and the militants. Note that (1.2) and (1.3) imply (1.1), so the reformists play no active role in this conception. Requiring that all three of these inequalities (or, equivalently, inequalities (1.2) and (1.3)) hold induces a *quasi-order* for the Left on $T \times T$, where $T$ is the policy space (the details are presented in the text below), and I define $(t_L, t_R)$ to be a *party unanimity Nash equilibrium* (PUNE) if it is a Nash equilibrium, where both parties maximize with respect to their respective quasi-orders.

Consider "classical" reformist parties, whose preferences are defined by $\Pi^L(t_L, t_R)$ and the analogous $\Pi^R(t_L, t_R)$. These preferences are complete, that is, define orders on $T \times T$. Because of the two-dimensionality of the policy space, Nash equilibrium in pure strategies with respect to these preferences—call this *reformist Nash equilibrium* (RNE)—does not, as I said, generally exist. We shall note below that RNE is a refinement of PUNE. While reformist Nash equilibria fail to exist with multi-dimensional policy spaces, it turns out that party-unanimity Nash equilibria do often exist—in fact, in the application studied in this paper, there are many of them.

The main theorem states that, in all PUNE, except one singular case, both Left and Right parties play progressive tax strategies. Furthermore, it immediately follows that, in any refinement of PUNE, both parties play progressive tax strategies, so that if the three factions within the parties learn to compromise to some extent—which, formally, will mean that the party comes to adopt a quasi-order on $T \times T$ which includes the one induced by (1.1)–(1.3)—then all political equilibria with respect to that new quasi-order also entail both parties' playing progressive tax strategies.

Section 2 defines the politico-economic environment, Section 3 defines equilibrium, and Section 4 proves the main results. The entire argument can be made by use of the simple geometry of cones in two-space, and I choose the geometric exposition over a more formal, analytical one, for its transparency.

## 2. THE MODEL

Each individual wishes to maximize her after-tax income. An individual is characterized by her income, $w$, or her *type*. Individuals supply labor inelastically, as they derive no welfare from leisure. The distribution of citizen types is given by a probability measure $\mathbf{F}$ on $[0,1]$. Thus, maximum income is normalized at one.[3]

A *tax policy* is a triple $(a,b,c)$, where the after-tax income of an individual with income $w$ is $aw^2+bw+c$. Taxes are purely redistributive, so the balanced-budget condition is

$$\text{(2.1a)} \qquad \int (aw^2+bw+c)\, d\mathbf{F}(w) = \mu,$$

where $\mu = \int w\, d\mathbf{F}$ is mean income, which implies

$$\text{(2.1b)} \qquad c = -a\mu_2 - b\mu + \mu, \qquad \text{where}$$

$$\text{(2.1c)} \qquad \mu_2 \equiv \int w^2\, d\mathbf{F}.$$

Thus after-tax income is

$$a(w^2-\mu_2)+b(w-\mu)+\mu.$$

Define type $w$'s *ordinal* utility function on tax policies as

$$u(a,b,w) = a(w^2-\mu_2)+b(w-\mu)+\mu.$$

Henceforth, we understand that tax policies are two-dimensional, denoted $(a,b)$.

Let $(a,b)$ and $(a',b')$ be two tax policies and define

$$\Delta a \equiv a - a', \quad \Delta b \equiv b - b'.$$

Then a voter of type $w$ is indifferent between policies $(a,b)$ and $(a',b')$ if and only if she enjoys the same after-tax income in both, that is, if and only if

$$\text{(2.2a)} \qquad \Delta a(w^2-\mu_2)+\Delta b(w-\mu) = 0.$$

Define the function

$$\text{(2.3)} \qquad \varphi(w) \equiv \frac{w^2-\mu_2}{w-\mu}, \quad \text{for} \quad w \neq \mu.$$

It follows from (2.2a) that, for $w \neq \mu$, voter $w$ is indifferent between the two policies iff

$$\text{(2.2b)} \qquad \Delta a\, \varphi(w) + \Delta b = 0.$$

[3] My nomenclature: the population consists of *citizens*, all of whom are taxed. Citizens may choose whether or not to vote. Hence the set of *voters* may differ from the set of citizens.

This tells us that, viewing the domain of policies $(a, b)$ as $\Re^2$, the indifference curves of type $w$ are *straight lines* of slope $-\varphi(w)$. The indifference curves of type $w = \mu$ are vertical straight lines.

We have thus far restricted policies only by a budget-balancing constraint. We further make the following assumption.

ASSUMPTION A1: (i) $\forall w$, $a(w^2 - \mu_2) + b(w - \mu) + \mu \geq 0$; (ii) $\forall w$, $2aw + b \geq 0$.

A1 (i) says that every individual's after-tax income must be nonnegative; A1 (ii) says that after-tax income must be a nondecreasing function of income. Thus (i) is an individual budget constraint, and (ii) is an incentive compatibility constraint. The reader may check that the set of policies satisfying A1 is the triangle $\mathcal{T} = OUV$ illustrated in Figure 1.

Policy $O = (0, 0)$ is total confiscation of income and redistribution to the mean, and policy $T = (0, 1)$ is laissez-faire (no taxation).

Define a policy $(a, b)$ to be *progressive* iff it generates an after-tax income function which is concave in (pre-tax) income. This is equivalent to having $a \leq 0$. Thus, progressive policies are precisely those in $\Delta OUT$ in Figure 1. (Strictly) progressive policies are ones for which the assessed tax is a strictly convex function of income.

Let us study the function $\varphi$, illustrated in Figure 2. The salient facts are:

- $\varphi$ is increasing, asymptotically to infinity, in the interval $[0, \mu)$;
- $\varphi$ is increasing, beginning at the value $-\infty$, in the interval $(\mu, 1]$;

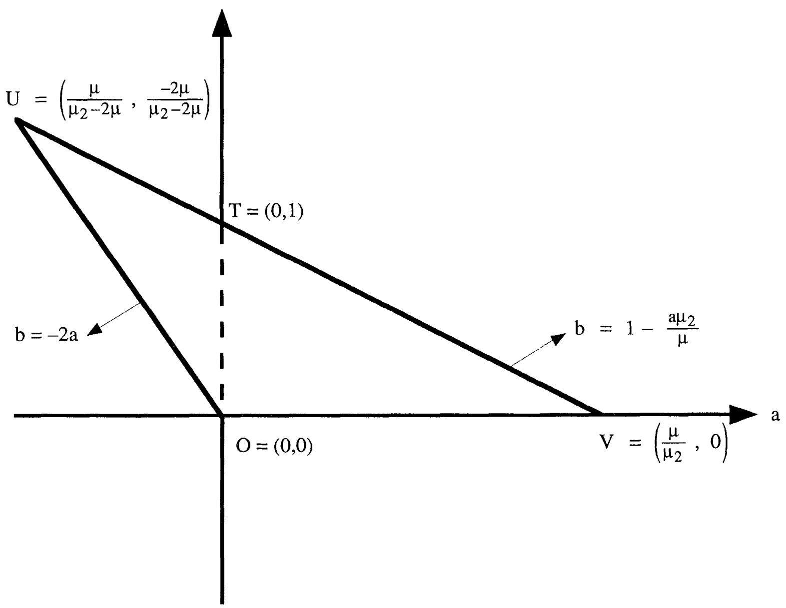


FIGURE 1

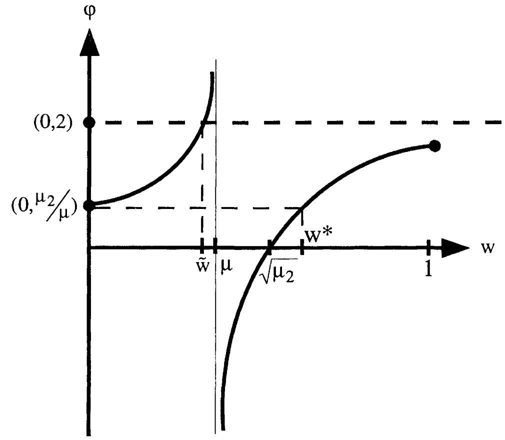


FIGURE 2

- $\varphi(w) = 0$ iff $w = \sqrt{\mu_2}$;
- $\varphi(1) > \varphi(0)$ and $\varphi(1) < 2$.

The only one of these facts I shall derive is the last one. Suppose, to the contrary, that $\varphi(1) \geq 2$. Then, by definition of $\varphi$, $\mu_2 - 1 \leq 2\mu - 2$, and so $\mu_2 \leq 2\mu - 1$. But $\mu^2 < \mu_2$ ($F$ has positive variance); hence $\mu^2 < 2\mu - 1$, which means $(\mu - 1)^2 < 0$, an impossibility.

Because $\varphi(1) > \varphi(0)$, we may define a type $w^*$ as that unique type such that $w^* > \mu$ and $\varphi(w^*) = \mu_2/\mu$. Because $\varphi(1) < 2$, we may define a type $\tilde{w}$ as that unique type such that $\varphi(\tilde{w}) = 2$. $w^*$ and $\tilde{w}$ are illustrated in Figure 2.

We next study the indifference curves and ideal points of the various types. An individual $w$ prefers policy $(a, b)$ to $(a', b')$ if and only if

$$(2.3) \qquad \Delta a(w^2 - \mu_2) + \Delta b(w - \mu) > 0.$$

It follows that:

- if $w < \mu$, then $\Delta a\, \varphi(w) + \Delta b$ decreases as utility increases;
- if $\mu < w$,then $\Delta a\, \varphi(w) + \Delta b$ increases as utility increases.

Using this observation, the following lemma immediately follows.

LEMMA 1: (i) *O is the ideal point of* $\{w \mid w < \tilde{w}\} = W_1$; (ii) *U is the ideal point of* $\{w \mid \tilde{w} < w < \mu\} = W_2$; (iii) *U is the ideal point of* $\{w \mid \mu < w < \sqrt{\mu_2}\} = W_3$; (iv) *U is the ideal point of* $\{w \mid \sqrt{\mu_2} < w < w^*\} = W_4$; (v) *V is the ideal point of* $\{w \mid w > w^*\} = W_5$.

PROOF (Refer to Figure 1):

(i) $w \in W_1$ have indifference curves which are less steep than *OU*, since $\varphi(w) < 2$, and utility increases in the south-west direction, since $w < \mu$.

(ii) $w \in W_2$ have indifference curves which are steeper than *OU*, and utility increases in the south-west direction.

(iii) $w \in W_3$ have positively sloped indifference curves, and utility increases in the north-west direction.

(iv) $w \in W_4$ have negatively sloped indifference curves that are less steep than $UV$, which increase in the north-east direction.

(v) $w \in W_5$ have negatively sloped indifference curves, steeper than $UV$, and utility increases in the north-east direction.

We next introduce political parties. There are two, called Left and Right. For simplicity, we assume that Left "represents" a particular citizen type $w_L$ and Right represents a citizen type $w_R$. We make the following assumption.

ASSUMPTION A2: (i) $w_L < \tilde{w}$, *and* (ii) $w_R > w^*$.

Thus Left (Right) represents a relatively poor (rich) citizen.

When I say that Left "represents" a citizen type $w_L$, I mean that Left's militants have the (ordinal) preferences of $w_L$, and that Left's reformists are endowed with a von Neumann-Morgenstern utility function $v_L$ on tax policies $(a, b)$ that is consistent with $w_L$'s (ordinal) after-tax income preferences—that is, $v_L$ is an ordinal transform of $u(\cdot, w_L)$. Similarly, Right's reformists are endowed with a von Neumann-Morgenstern utility function $v_R$ on tax policies consistent with $w_R$'s ordinal income preferences.

According to Lemma 1, Left's ideal policy is the point $O$, and Right's is the point $V$ (see Figure 1). Types in the region $(\tilde{w}, w^*)$ are the "moderates," with ideal point $U$. Note, as well, that A2 implies that $\varphi(w_L)$ and $\varphi(w_R)$ are both positive, and so both those types have negatively sloped indifference curves.

We next introduce uncertainty. Parties are uncertain about the distribution of *voter* income. They know the probability measure $\mathbf{F}$ of citizen income, but some citizens may not vote, which introduces uncertainty about the distribution of voter income.[4]

We represent this uncertainty as follows. There is a continuum of states $s \in [0, 1]$, and in state $s$, the probability measure of voter income is $\mathbf{F}_s$ on $[0, 1]$. Both parties share the prior that $s$ is distributed according to a probability measure $\mathbf{H}$ on $[0, 1]$. We make the following assumption.

ASSUMPTION A3: *The measures* $\mathbf{F}$, $\mathbf{H}$, *and* $\mathbf{F}_s$, *for all* $s$, *are all equivalent*[5] *to Lebesgue measure on* $[0, 1]$. *The cumulative density function of* $\mathbf{F}_s$, *denoted* $F_s$, *is a continuous function of* $s$ *as well.*

[4]This is but one convenient way of interpreting the uncertainty. One might alternatively say that voters do not necessarily have preferences which correspond to their particular incomes ... they may, for example, have preferences associated with the income they think they will have in the future, and parties are uncertain about these beliefs.

[5]That is, they assign positive measure to an event if and only if the event has positive Lebesgue measure.

Suppose Left (Right) proposes a policy $t_L = (a, b)$ ($t_R = (a', b')$). A citizen $w$ prefers $t_L$ to $t_R$ if and only if (2.3) holds. Define $W(t_L, t_R)$ as the set of types for which (2.3) holds. Then Left defeats Right (by majority vote) exactly in those states $s$ such that

$$\mathbf{F}_s(W(t_L, t_R)) > \tfrac{1}{2}. \tag{2.4}$$

We follow the usual convention that if a voter is indifferent between $t_L$ and $t_R$, she votes randomly. By the continuum assumption, it follows that $t_L$ and $t_R$ tie in state $s$ iff $\mathbf{F}_s(W(t_L, t_R)) = \frac{1}{2}$.

Let $S(t_L, t_R)$ be the set of $s$ such that (2.4) holds. It follows that, from the parties' viewpoints, Left defeats Right with probability

$$\pi(t_L, t_R) = \mathbf{H}(S(t_L, t_R)). \tag{2.5}$$

We have now defined the preferences of all three factions of each party. For Left, for example, these are represented by the functions $\Pi^L$ (reformists), $\pi$ (opportunists), and $u(\cdot, w_L)$ (militants).

## 3. THE EQUILIBRIUM CONCEPT

The standard approach, as I mentioned earlier, is to define political equilibrium as a Nash equilibrium between players with payoff functions $\Pi^L$ and $\Pi^R$, each of whose strategy space is $\mathcal{T}$—what I have called in Section 1, reformist Nash equilibrium. As RNE generally fail to exist, I propose that an inner-party struggle determines the policy of the party, and show that, under that specification, Nash equilibria in pure strategies of the political game exist.

Think of Left as evaluating whether to deviate from $t_L$ to $t'_L$ when Right is playing $t_R$. As I described earlier, the militants in Left wish to maximize $u(\cdot, w_L)$, and the opportunists wish to maximize $\pi(\cdot, t_R)$. Unanimity between these two groups is required for Left to deviate from $t_L$ to $t'_L$. The agreement of the "reformists" will follow automatically.

Stated more precisely, Left would entertain a deviation from $t_L$ to $t'_L$ only if

$$(\pi(t'_L, t_R), u(t'_L, w_L)) \geq (\pi(t_L, t_R), u(t_L, w_L)), \tag{3.1a}$$

with the convention on vector ordering given in this footnote.[6]

Similarly, at $(t_L, t_R)$, Right would entertain a deviation from $t_R$ to $t'_R$ only if

$$(1 - \pi(t_L, t'_R), u(t'_R, w_R)) \geq (1 - \pi(t_L, t_R), u(t_R, w_R)). \tag{3.1b}$$

DEFINITION: A policy pair $(t_L, t_R)$ is a *party unanimity Nash equilibrium* (PUNE) if and only if there is no $t'_L \in \mathcal{T}$ at which (3.1a) holds, and there is no $t'_R \in \mathcal{T}$ at which (3.1b) holds.

[6] $(x_1, x_2) \geq (y_1, y_2)$ if and only if $x_i \geqq y_i$ and for some $i$, $x_i > y_i$. $(x_1, x_2) \geqq (y_1, y_2)$ if and only if $x_i \geqq y_i$.

The parties, under this specification, possess incomplete preference orders on the space $\mathcal{T}\times\mathcal{T}$. Consider the quasi-order $\succsim_L$ on $\mathcal{T}\times\mathcal{T}$ defined by

$$(3.2a) \qquad (t_L, t_R) \succsim_L (t'_L, t'_R) \Leftrightarrow (\pi(t_L, t_R), u(t_L, w_L)) \geqq (\pi(t'_L, t'_R), u(t'_L, w_L)),$$

and the quasi-order $\succsim_R$ defined by

$$(3.2b) \qquad (t_L, t_R) \succsim_R (t'_L, t'_R) \Leftrightarrow (1-\pi(t_L, t_R), u(t_R, w_R)) \geqq (1-\pi(t'_L, t'_R), u(t'_R, w_R)).$$

Then a PUNE is a Nash equilibrium where Left (Right) maximizes with respect to $\succsim_L$ ($\succsim_R$).

We introduce next a refinement of PUNE:

DEFINITION: $(t_L, t_R)$ is a *strong party unanimity Nash equilibrium* iff $(t_L, t_R)$ is a PUNE and it is false that $t_L = (0,0)$ and $\pi(t_L, t_R) = 1$.

"Strongness" is a (very) weak refinement of PUNE. If we eliminate nonstrong PUNE from consideration, we are saying something about the capacity of Right's inner-party factions to compromise in the face of almost sure calamity. Recall that policy $(a,b) = (0,0)$ is complete leveling of all incomes to the mean. If, in a PUNE, $t_L = (0,0)$ and $\pi(t_L, t_R) = 1$, then facing the prospect of Left's winning almost surely with the leveling policy $(0,0)$, Right does not deviate. Strongness says that, facing such a prospect, Right's opportunist and reformist factions will be able to persuade the militant faction to deviate to a policy with a positive probability of defeating $(0,0)$.

Let us make three remarks.

REMARK 1: If $(t_L, t_R)$ is an RNE then it is a PUNE. (I.e., RNE is a refinement of PUNE.)

PROOF: Let $(t_L, t_R)$ be an RNE. Suppose there exists a deviation for one party, say Left—call the deviation $t'_L$—that both the militants and the opportunists weakly prefer, and one strongly prefers, to $t_L$. Then $t'_L$ must increase the reformists' payoff function, as well (that is, $\Pi^L(t_L, t_R) > \Pi^L(t_L, t_R)$). This contradicts the fact that $(t_L, t_R)$ is an RNE. *Q.E.D.*

REMARK 2: PUNE is an *ordinal* concept: it depends only on the ordinal preferences of the opportunists and the militants. In contrast, RNE is a cardinal concept: it depends on the *particular* von Neumann-Morgenstern utility functions that the reformist factions have.

REMARK 3: A comment about Condorcet winners. Suppose it is the case that

$$\forall s \quad \mathbf{F}_s(W_2 \cup W_3 \cup W_4) > \tfrac{1}{2}.$$

(Refer to Lemma 1 for notation.) Then $W_2 \cup W_3 \cup W_4$ is a majority coalition in all states, all of whose members have ideal point $U \in \mathcal{T}$. It follows that $U$ is a

Condorcet winner in $\mathscr{T}$: it defeats all policies except itself, for sure. Thus, it would be a Downsian equilibrium for both parties to propose $U$. (That is, both parties' proposing $U$ would constitute a Nash equilibrium of the game between two parties possessing the single internal faction of opportunists.) It would as well be a PUNE for both parties to propose $U$, since neither opportunist faction would be willing to deviate: at $(U, U)$, each party has probability one-half of victory, and under any unilateral deviation, the deviating party has probability zero of victory.[7] Nevertheless, in a PUNE, it might be the case that neither party proposes $U$.

It may be worthwhile to summarize the informational set-up of the model, which is related to Remark 2 above. *Voters* are endowed only with preferences over after-tax income, not with preferences over income lotteries. They know that the distribution of *citizen* types is **F**, and hence face no relevant uncertainty when voting over tax policies. Militants, as well, do not face a decision problem under uncertainty—they behave just like particular voters ($w_L$ and $w_R$). Opportunists want to maximize the probability of victory, a known function on the domain $\mathscr{T}\times\mathscr{T}$. Only the reformists face a decision problem under uncertainty: they must evaluate income lotteries for their respective constituents. These lotteries occur because, although the distribution of income (**F**) among *citizens* is known, the distribution of income among *voters* is uncertain ($\{\mathbf{F}_s\}$).

The parsimonious reader might well inquire why I do not dispense with the reformists, as they are not necessary for the equilibrium concept of PUNE, as I noted earlier. The answer is two-fold: because, in reality, I think that reformists exist, and because the concept of reformist Nash equilibrium is well-established in the literature, although not under that name.

## 4. ANALYSIS OF PARTY COMPETITION

Let us look at a typical strategy pair that might arise in the game between our two parties. Denote henceforth by $L$ the policy $(a, b)$ announced by Left, and by $R$ the policy $(a', b')$ announced by Right. Examine Figure 3, which reproduces the domain $\mathscr{T}$, with the two hypothetical proposals, $L$ and $R$. Assume that $w_L$ prefers $L$ to $R$, and $w_R$ prefers $R$ to $L$.

In Figure 3, the slope $m$, of $LR$ is positive. We know from Section 2 that type $\hat{w}$ is indifferent between $L$ and $R$ iff $\varphi(\hat{w}) = -m$. But this determines a unique $\hat{w}$, illustrated in Figure 4, which is an amended reproduction of Figure 2. It now follows, from examination of (2.3), that the set of types preferring $L$ to $R$ is $[0, \hat{w})$, and the set of types preferring $R$ to $L$ is $(\hat{w}, 1]$. These sets are illustrated by shading in Figure 4.

[7] We cannot generally assert, however, that a Downsian equilibrium is a PUNE. Let $(t_L, t_R)$ be a Downsian equilibrium. It follows that $\pi(t_L, t_R) = 0.5$. Suppose $t'_L$ is another policy such that $\pi(t'_L, t_R) = 0.5$. It may be that the Left militants prefer $t'_L$ to $t_L$, in which case $(t_L, t_R)$ is not a PUNE.

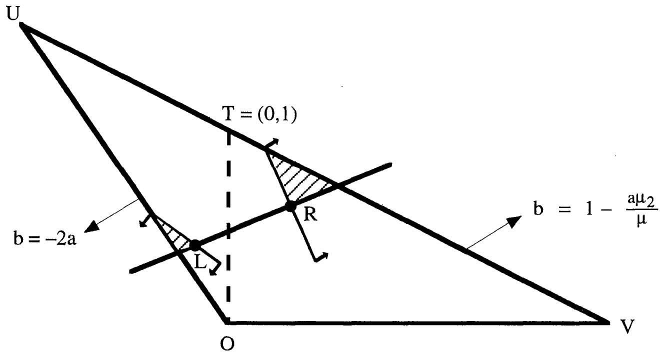


FIGURE 3

In Figure 3, I have also drawn an indifference curve for voter $w_L$ through the point $L$, and an indifference for voter $w_R$ through the point $R$. The arrows indicate the directions of increasing utility.

Now imagine, in the situation of Figure 3, that Left is considering deviating from $L$, locally, to a point in the adjacent shaded triangle, while Right is fixed at $R$. Any such deviation is preferred by Left's militants, since the shaded triangle lies on the preferred side of $w_L$'s indifference curve at $L$. Any such deviation also would reduce the steepness of the line $LR$; but according to Figure 4, that means it would *increase* the set of types who prefer the Left policy to $R$, and hence must (weakly) increase the probability that Left defeats $R$. Hence, $(L, R)$ cannot be a PUNE: both militant and opportunist factions of Left would agree to deviate from $L$ into the shaded region.

Let us define the set of policies below $w_L$'s indifference curve at $L$ and above the line $LR$ as the *cone of attractive policies for Left at L*. Similar analysis shows

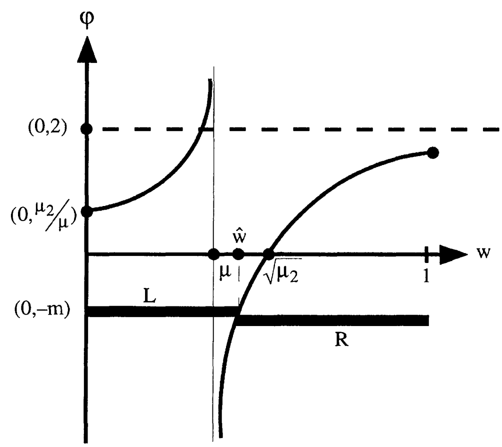


FIGURE 4

that the shaded cone at $R$ is the cone of attractive policies for Right at $R$—i.e., the set of policies to which the factions in Right would deviate at $R$, given $L$ fixed. A pair $(L, R)$ *is a PUNE only if the intersection of each cone of attractive policies with* $\mathcal{T}$ *is empty.*[8]

We shall use this technique in the rest of the paper. The analysis is simple by virtue of the pleasant fact that voter indifference curves in policy space are straight lines, and so it is easy to identify the cone of attractive policies for a party at a pair of proposals.

Before proceeding further, we state a useful general principle:

LEMMA 2: *Let* $(L, R)$ *be a policy pair in* $\mathcal{T} \times \mathcal{T}$. *A local deviation to one side of LR from L by Left increases Left's probability of victory if and only if a local deviation by Right from R on the same side of LR increases Right's probability of victory.*

PROOF: The key observation is that when Left or Right deviates from $L$ or $R$ on the *same side* of $LR$, while the other party remains fixed, they change the slope of $LR$ in *opposite ways*: e.g., if Left's deviation renders $LR$ steeper, then Right's deviation renders $LR$ less steep. The lemma now follows from consideration of what happens to the sizes of the coalitions of types favoring each party (see, for example, Figure 4). *Q.E.D.*

We also state an obvious fact:

LEMMA 3: *If* $(L, R)$ *is a PUNE, then* $w_L$ *weakly prefers L to R and* $w_R$ *weakly prefers R to L.*

PROOF: Suppose, to the contrary, that $w_L$ strictly prefers $R$ to $L$. Then Left can deviate by moving along the line $LR$ from $L$ towards $R$: this leaves the probability of victory unchanged, and increases the welfare of $w_L$, so both opportunists and militants will support the deviation. *Q.E.D.*

Our first task is to prove existence of strong PUNE.[9] To this end, we make the following assumption.

ASSUMPTION A4: *There is a set S of states, with* $\mathbf{H}(S) > 0$, *such that*

$$s \in S \Rightarrow F_s(\tilde{w}) < 1/2.$$

In particular, if A4 did not hold, the Left could win with probability one against any Right proposal by proposing its ideal point, $O$! It follows that A4 is a

[8] This is only an "only if" statement. It is possible that, locally, neither party wants to deviate, but that there is a distant deviation that is attractive for one party.

[9] It is trivial to note that $L = O$ and $R = V$ is always a PUNE, for the militants in each party will refuse to deviate from their ideal points. But in general, this pair of proposals is not a strong PUNE (i.e., Left may win with probability one).

*necessary* condition for existence of a strong PUNE. Theorem 1 says it is, as well, sufficient.

THEOREM 1: *If A1, A2, A3, and A4 hold, then there exist strong PUNE.*

PROOF: (i) Choose $L \in OU$ and $R \in UT$, such that the slope, $m$, of $LR$ is negative and close in absolute value to 2, as illustrated in Figure 5: in particular, choose $L$ to be close to $O$. The type $\hat{w}$ who is indifferent between $L$ and $R$ has $\varphi(\hat{w}) = -m$; it follows from Figure 2 that $\hat{w} \in (\tilde{w}, \mu)$. Hence, the coalition who vote Left, at these proposals, is $[0, \hat{w})$.

(ii) For any small $\varepsilon > 0$, we can choose $(L, R)$ so that the slope of $LR$ is sufficiently close to $-2$ that the coalition who vote Left is precisely $[0, \tilde{w} + \varepsilon)$. By A4 and A3 it follows that there is a set of states of positive **H**-measure such that $F_s(\tilde{w} + \varepsilon) < 1/2$. Therefore $\pi(L, R) < 1$.

(iii) From Figure 2, it follows that Left can increase the size of the Left coalition, and hence weakly increase the probability of victory, by deviating from $L$ to the left of $LR$, for such a move increases the absolute value of the slope of $LR$. By Lemma 2, Right can weakly increase its probability of victory by deviating from $R$ to the left of $LR$. Moreover, a deviation by Left from $L$ to the right of $LR$, and below $w_L$'s indifference curve through $L$, will decrease the size of the Left coalition, since $L$ is close to $O$ (again, examine Figure 2). Since $\pi(L,R) < 1$, it will therefore decrease $\pi$, since $\mathbf{F}_s$ and **H** are equivalent to Lebesgue measure. It follows that Left's cone of attraction at $L$ is the shaded region in Figure 5, and further, that there are no deviations in $\mathcal{T}$ that are attractive for Left. The last clause uses the fact that $L$ is close to $O$, so the points to the right of $LR$ and below $w_L$'s indifference curve are all close to $O$.

(iv) To calculate Right's cone of attractive policies at $R$, note that if Right deviates from $R$ to any point $R'$ on the segment $RV$, the slope of the line $LR'$ is either negative and greater in absolute value than the slope of $LR$, or positive and less in absolute value than $\mu_2/\mu$. It follows (consult Figure 2) that the type

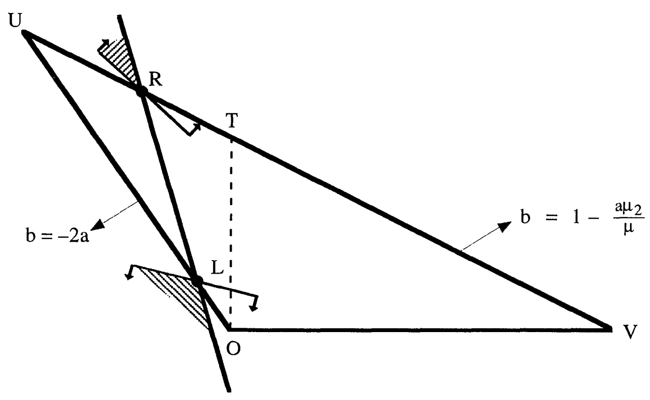


FIGURE 5

$\hat{w}$ who is indifferent between $L$ and $R'$ lies always in the interval $(\tilde{w}, w^*)$ and hence the Right coalition will be $(\hat{w}, 1]$, which is always smaller than it is at $R$. Without loss of generality, these are the only deviations from $R$ that are above $w_R$'s indifference curve at $R$ and to the right of $LR$ that we need consider (for if there is any attractive deviation in that region, there must be one on the segment $RV$). It follows that Right's cone of attraction at $R$ is the shaded region in Figure 5, and there are no attractive deviations for Right in $\mathscr{T}$.

(v) It follows that $(L, R)$ is a PUNE.

(vi) It immediately follows that $(L, R)$ is a strong PUNE. *Q.E.D.*

We now introduce the only distributional assumption of our analysis:

ASSUMPTION A5: $F_s(\mu) > 1/2$ *almost surely*.

We may think of A5 as reasonable if median income is less than mean income, for it says that in almost all states, more than half the voters have income less than the mean.

We now state our main result:

THEOREM 2: *If A1, A2, A3, and A5 hold, then in all strong PUNE, both Left and Right play progressive policies.*

In other words, the distributional assumption A5 and the (weak) refinement notion enable us to deduce the ubiquity of progressive tax proposals in political competition between Left and Right.

Some readers might find the introduction of the "strong PUNE" refinement, and hence the statement of Theorem 2, inelegant. We can, in fact, avoid that terminology, if we wish, as follows:

COROLLARY 1: *Let A1, A2, A3, and A5 hold. Then, in any PUNE, with probability one, a progressive policy wins the election.*

PROOF: Theorem 2 says that the only PUNE in which a party does not propose a progressive policy is one in which Left proposes $(0, 0)$ and wins with probability one. But $(0, 0)$ is a progressive policy. The claim immediately follows.[10]

Note that, from Remark 1 of Section 3, it follows that, if an RNE exists, then, with probability one, a progressive policy wins the election.

PROOF OF THEOREM 2:

*Step 1: There is no strong PUNE where Right plays a regressive policy.*

Suppose Right plays a policy $R$ in triangle $OTV$ (see Figure 1), but not on the line $OT$ (i.e., $R$ is regressive). Let Left play $L = O$. I claim Left wins with

[10] This circumlocution for avoiding the necessity of introducing the strong PUNE concept is due to Klaus Nehring.

probability one. For the slope of $LR$ is positive, and so the type, $w$, who is indifferent between $L$ and $R$, has $\varphi(w)$ negative, which, by Figure 2, means $w > \mu$; thus, by A5, Left wins with probability one. It follows that *any* PUNE where Right plays $R$ must have Left playing $O$, because $O$ is ideal for both Left's militants and opportunists. The conclusion follows.

Therefore, the remainder of the proof shows that Left never plays regressive in a strong PUNE.

*Step 2: There is no PUNE where either party plays a policy that is interior in $\mathscr{T}$.*

1. If $(L, R)$ is a PUNE and $L \in$ interior $\mathscr{T}$, then slope $LR = -\varphi(w_L)$.

Suppose, to the contrary, that slope $LR \neq -\varphi(w_L)$. Let $l$ be $w_L$'s indifference line containing $L$. Then $l$ and $LR$ do not coincide. The militants of Left strictly prefer any policy on line $LR$ and below $l$ (i.e., on the $O$ side of $l$), and Left's opportunists are indifferent to such a move (moving along the line $LR$ leaves the probability of victory constant). Hence $(L, R)$ is not a PUNE.

2. If $L \in$ interior $\mathscr{T}$, $R \in$ boundary $\mathscr{T}$, and slope $LR = -\varphi(w_L)$, then $(L, R)$ is not a PUNE.

Consult Figure 6, which illustrates two possible cases. Because $\varphi(w_R) > \mu_2/\mu$, $w_R$'s indifference curves are steeper than $UV$. Because $\varphi(w_R) < 2$, $w_R$'s indifference curves are less steep than $OU$. It must be that a deviation by Right from $R$ to the *right* of $LR$ weakly increases Right's probability of winning, by Lemma 2: for if Left could weakly increase its probability by deviating to the *left* of $LR$, then both its militant and opportunist factions would agree to do so, and $(L, R)$ would not be a PUNE. Therefore the cones of attraction for Right, in both cases illustrated in Figure 6, intersect $\mathscr{T}$'s interior, and so neither is a PUNE.

The third possibility, that $R \in OV$, can be similarly disposed of.

3. If $L, R \in$ interior $\mathscr{T}$ and slope $LR = -\varphi(w_L)$, $(L, R)$ is not a PUNE.

This step is easier than step 2, and is left to the reader.

4. We have now disposed of the possibility of a PUNE where $L$ is interior. Exactly symmetric arguments show there can be no PUNE where $R \in$ interior $\mathscr{T}$.

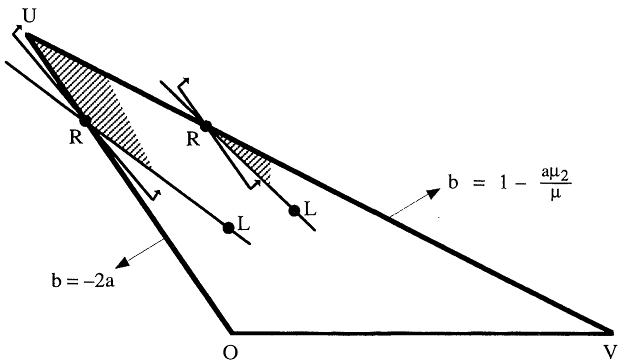


FIGURE 6

*Step 3. There is no strong PUNE where Left plays a policy L on the segment OV of $\mathcal{T}$, $L \neq O$.* (See Figure 1.)

If $(L, R)$ is a strong PUNE and $L \in OV$, then $R \in OU \cup UT$ (these are the only possibilities, by previous steps). Suppose $R \in OU$. Left's militants would like to move from $L$ along $LO$ towards $O$; it therefore must be that such a move would decrease $\pi$. Therefore, a move to the right of $L$ along $LV$ increases $\pi$. By Lemma 2, Right increases its probability of victory by moving above $LR$ (that is, on the segment $RU$) as well. But moving from $R$ along $RU$ also increases the welfare of Right's militants. This contradicts the assumption that $(L, R)$ is a PUNE.

The possibility that $R \in UT$ is similarly disposed of.

*Step 4. There is no strong PUNE where $L \in TV$.*

Again, $R \in OU \cup UT$. It is easily checked that, if $L$ moves towards $R$ along the line $LR$, Left's militants are better off than at $L$. But such a move leaves $\pi$ unchanged. Hence $(L, R)$ is not a PUNE.

Steps 3 and 4 exhaust the possibilities for a PUNE where at least Left plays a regressive strategy, for the only regressive boundary strategies are on the segments $TV$ and $OV$ of $\mathcal{T}$. *Q.E.D.*

## 5. CALIBRATION

According to the 1990 US census, mean household income in the US was \$30,900 and the standard deviation of household income was \$34,000.[11] Let us take maximum household income, for all effective purposes, to have been \$200,000 in 1990. Then, normalizing maximum income at unity, we compute that $\mu = 0.1545$ and $\mu_2 = 0.05277$. The empirical function $\varphi$ is graphed in Figure 7.

We calculate that $\tilde{w} = \$27{,}500$ and $w^* = \$68{,}300$. Our Assumption A2 requires that Left represent a voter whose income is not greater than \$27,500, and Right a voter whose income is not less than \$68,300. Given that mean income is \$30,900, these are reasonable assumptions—if not for the Democrats and Republicans in the US, then, perhaps, for Labour and the Conservatives in Great Britain.

Our Assumptions A5 and A4 say that it is always the case that at least one-half the voters have an income less than \$30,900, but there is positive probability that fewer than one-half the voters will have income less than \$27,500.

## 6. CONCLUSION

Let me summarize the model. There is a population with a distribution of income, who must vote on a redistributive tax policy, which is limited to be some quadratic function of income. Voters supply labor inelastically—so income is fixed, for each voter. There are two parties, one purporting to represent a

[11] I thank my colleague Marianne Page for computing these statistics.

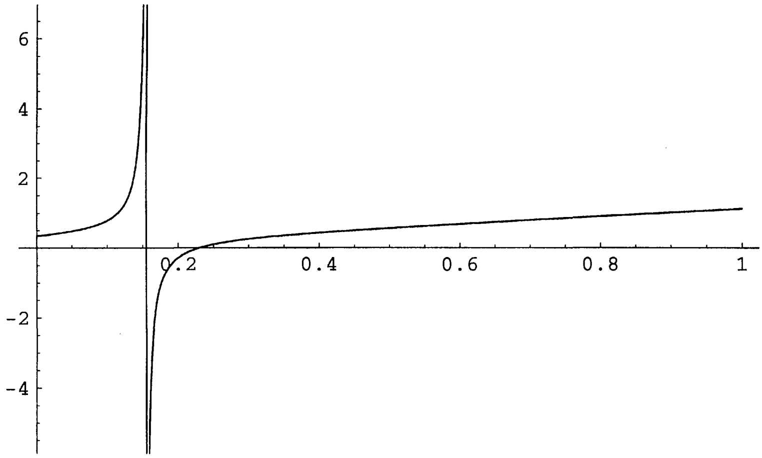


FIGURE 7.—The function $\varphi$ for the United States.

relatively poor voter, and one a relatively rich voter. Within each party, there are three factions: reformists, militants, and opportunists. Parties are unsure about the distribution of voter types on election day, so that, if two policies are before the voters, they can assign only a probability that one policy defeats the other.

Call the policy space $\mathscr{T}$. Each faction of each party has (complete) preferences on $\mathscr{T}\times\mathscr{T}$, that is, on pairs of policies that it and its opposition can propose. All three factions in a party must agree for the *party* to prefer one element in $\mathscr{T}\times\mathscr{T}$ to another: this leads to each party's having a preference quasi-order on $\mathscr{T}\times\mathscr{T}$ —it is the intersection of the preference orders of its three factions. A political equilibrium is a Nash equilibrium where each party maximizes with respect to its quasi-order.

The main result says that if most voters, in all states, have income less than mean income, then in any political equilibrium, both parties propose progressive tax policies.

The "unanimity" quasi-order is the weakest quasi-order (i.e., corresponds to the smallest binary relation in $(\mathscr{T}\times\mathscr{T})^2$) that can hold, given the inner-party struggle among the three factions, as it corresponds to the most demanding condition for inner-party agreement. It thus generates the largest possible set of (Nash) political equilibria. If the factions learn to compromise—which is to say that the party's preferences come to be represented by a quasi-order that *contains* the unanimity quasi-order—the political equilibria will be a subset of the ones here shown to exist, and hence will still consist in progressive policies.

Perhaps the assumption one would most like to weaken is the inelasticity of labor supply. Do the results remain true if individuals experience disutility from labor? One can prove the following. Suppose that individuals have preferences over income and leisure, such that the labor supply elasticity with respect to the wage is uniformly (for all incomes and all individuals) less than some number $\delta > 0$. Consider, now, a sequence of economies, letting $\delta$ approach zero. For $\delta$ sufficiently small, strong PUNE exist, and in all of them, the policies are arbitrarily close to being progressive. (The proof goes by showing that, as $\delta$ gets small, the feasible set of policies converges to the triangle $OUV$ and the indifference curves of individuals become arbitrarily close to being straight lines. We then use the results established above to get the limit result.) Is there a stronger result, saying that both parties will propose progressive policies for $\delta$ not close to zero? I expect not. If the labor supply elasticity of the high-wage agents is sufficiently large, it does not seem to me that progressive policies would necessarily be advocated by either party.

*Department of Economics, University of California, Davis CA 95616, U.S.A.; jeroemer@ucdavis.edu; http: //polar.ucdavis.edu / ~ roemer*

*Manuscript received March, 1997; final revision received January, 1998.*

## REFERENCES

ALESINA, A., AND H. ROSENTHAL (1995): *Partisan Politics, Divided Government, and the Economy*. New York: Cambridge University Press.

BERNHARDT, M. D., AND D. INGBERMAN (1985): "Candidate Reputations and the 'Incumbency Effect'," *Journal of Public Economics*, 27, 47–67.

CALVERT, R. (1985): "Robustness of the Multi-Dimensional Voting Model: Candidates' Motivations, Uncertainty, and Convergence," *American Journal of Political Science*, 29, 69–95.

CUKIERMAN, A., AND A. MELTZER (1991): "A Political Theory of Progressive Income Taxation," in *Political Economy*, by A. Melzer, A. Cukierman, and S. F. Richard. New York: Oxford University Press.

LIPSET, S. M. (1959): *Political Man: The Social Bases of Politics*. Baltimore: Johns Hopkins University Press.

MARHUENDA, F., AND I. ORTUÑO-ORTIN (1995): "Majority Voting and Progressivity," *Investigaciones Economicas*, 19, 469–473.

MICHELS, R. (1915): *Political Parties*. New York: The Free Press.

PRZEWORSKI, A., AND J. SPRAGUE (1986): *Paper Stones*. Chicago: University of Chicago Press.

ROEMER, J. E. (in press): "Why the Poor do not Expropriate the Rich: An Old Argument in New Garb," forthcoming in *Journal of Public Economics*.

——— (1997): "Political-Economic Equilibrium when Parties Represent Constituents: The Unidimensional Case," *Social Choice and Welfare*, 14, 479–502.

SNYDER, J., AND G. KRAMER (1988): "Fairness, Self-Interest, and the Politics of the Progressive Income Tax," *Journal of Public Economics*, 36, 197–230.

WITTMAN, D. (1983): "Candidate Motivation: A Synthesis of Alternatives," *American Political Science Review*, 77, 142–157.

YOUNG, P. (1994): *Equity*. Princeton: Princeton University Press.