



ELSEVIER

Journal of Public Economics 87 (2003) 2069–2103

JOURNAL OF
PUBLIC
ECONOMICS

www.elsevier.com/locate/econbase

The political economy of redistribution under asymmetric information

Sanjit Dhami*

Department of Economics, University of Newcastle Upon Tyne, Newcastle NE1 7RU, UK

Received 15 March 2000; received in revised form 28 August 2001; accepted 2 February 2002

Abstract

This paper examines the political economy of redistribution when voters have asymmetric information about the redistributive preferences of politicians and the latter cannot make credible policy commitments. The candidates in each party are endogenously selected by a process of Nash Bargaining between the competing factions. In equilibrium, there is ‘partial convergence’ of redistributive policies, support for ‘Director’s Law’, the possibility of ‘policy reversals’ across the parties, and ‘inter-term tax variability’ (political budget cycles) during the tenure of a politician. The effect of inequality on the magnitude of the redistributive activity depends in important ways on the incentives and constraints facing politicians.

© 2002 Elsevier B.V. All rights reserved.

Keywords: Signaling; Inequality; Redistribution; Political budget cycles

JEL classification: D72 (Economic Models of Political Processes); D78 (Positive Analysis of Policy-Making and Implementation); D82 (Asymmetric and Private Information); H23 (Redistributive Effects)

1. Introduction

The normative theory of redistribution characterizes optimal redistribution as the outcome of a trade-off between equity and efficiency that maximizes some

*Tel.: +44-191-222-6831; fax: +44-191-222-6548.

E-mail addresses: sanjit.dhami@ncl.ac.uk (S. Dhami), <http://www.staff.ncl.ac.uk/sanjit.dhami> (S. Dhami).

well-defined social welfare function; see for example Atkinson and Stiglitz (1980). On the other hand, positive theories view redistribution as the outcome of some well-defined political process. There are three broad strands in the latter literature.

In the ‘direct voting models’ exemplified in Romer (1975) and Roberts (1977), voters vote directly over alternative income tax schedules to determine equilibrium redistribution. Under certain conditions, the median voter makes the decisive policy choice, hence, changes in the spatial location of the median voter are critical to an understanding of redistributive policy. In the literature on ‘pork barrel politics’ on the other hand, partisan redistributive differences among political parties, and their attempts to gain strategic electoral support, determine the redistributive outcome; see for example Lindbeck and Weibull (1987, 1993), and Dixit and Londregan (1995, 1998). Strategic transfers, or pork, given to targeted voter groups translates into greater electoral support. Since the rich and the poor have a strong electoral and economic affinity with the Right and the Left party respectively, party loyalties are most likely to be loosened for middle income voters, who consequently receive the largest strategic transfers, an outcome often referred to as ‘Director’s Law’. In the third strand in the literature, Roemer (1998, 1999) models the redistributive process as the outcome of internal party struggle among the various factions within political parties. Restricting attention to those policies which are not weakly Pareto preferred by any of the factions, Roemer shows that equilibrium tax systems are progressive. Furthermore, political parties might subdue their redistributive tendencies to seek support along other (non-redistributive) policy dimensions.

Although it facilitates and develops a deeper understanding of the process of redistribution, the existing literature limits itself to the joint assumptions of ‘full-information’ and ‘full-commitment’. These joint assumptions constitute a fairly strong restriction; its relaxation is one of the central motivations of this paper.

The assertion that the ideological preferences of political candidates, especially prior to their taking up office, is public information, does not seem to be supported in public debates. Despite the role played by candidate selection processes in weeding out certain kinds of candidates, most available evidence points to at least some degree of ‘residual uncertainty’ about the preferences/ideology of politicians. In a special election report, Patrick Wintour (*The Guardian*, 9 Feb. 2001) writes ‘Mr Blair claims his first administration has laid the ghosts of its own past. The party need no longer define itself by proving it is not the old Labour party. The next few weeks, he appears to be saying, as the second term program unfolds will finally reveal the Blairite government in its true colours’. Even former Presidents can be fooled about the ideological preferences of the current incumbent. For instance, President Jimmy Carter made the following widely reported comments (*The Guardian*, 26 July 2001) about the current incumbent: ‘I thought he would be a moderate leader, but he has been very strictly conforming to

some of the more conservative members of his administration'. Indeed, there is a general perception that politicians choose to transmit (or hide) their true colors only through their implementation of actual policies.

Politicians rarely, if ever, commit to a detailed blueprint of their redistributive policy, prior to the elections. Campaign promises, in their interpretation as cheap-talk, might not always be informative.¹ The retraction of President Bush from the Kyoto Accord is just the latest in a series of instances where public perception based on campaign speeches is proved to be wrong (*The Economist*, 7 April 2001). Columnist Ed Quillen (*Denver Post*, 22 May 1990) pokes fun at the notoriously vague nature of redistributive campaign promises, thus: 'Read His Lips: 'Know' New Taxes'. Indeed Besley and Case (1995) provide indirect evidence to support the assertion that mechanisms to achieve full-commitment, for instance, reputational considerations or the ability of long-lived political parties to control short-lived politicians, are not always perfect.

This paper differs from most of the existing literature in two further respects. Firstly, voters and politicians alike care about equality as in Dixit and Londregan (1998), however, the existing literature assumes otherwise. Secondly, the candidate in each party is endogenously selected in a model that reflects the spirit of Roemer (1998, 1999), namely, that the selection of redistributive policy in a party is the outcome of internal party struggle among competing factions.

The description of the model in this paper is as follows. All individuals are privately informed about their redistributive ideology; politicians can signal their redistributive ideology to voters via their tax policy. There are two political parties, the 'Left' and the 'Right' and within each party there are two competing factions, the 'opportunists' and the 'militants'; the former care relatively more about winning the elections while the latter care relatively more about the party platform. The candidate in each party is selected as the outcome of 'Generalized Nash Bargaining' among the two factions; both parties choose their candidates simultaneously. The chosen candidate in each party is privately informed of her redistributive ideology, which can be either extreme or moderate; the latter type is more popular with the voters.

The chosen candidates contest the first election, following which the winner implements the first period redistributive policy. There is a second election at the end of the first period and the winner implements second period redistributive policy. The game lasts two periods, thus, the second period politician is a

¹For cheap talk to be valuable, different types of candidates in an election must have different preferences over the actions (voting) of the voters. However, since all politicians have identical preferences over that action, i.e. they would like voters to vote for them, campaign promises are unlikely to be informationally valuable, see Crawford and Sobel (1982). This would, however, need to be qualified if long lived political parties could control short lived politicians.

lame-duck;² since lame-duck politicians have no future to worry about, they implement their most desired policies. Crucially however, first period redistributive policy can signal the politician's redistributive type to the voters (retrospective voting) who can then infer the redistributive type of the lame-duck politician. Since the 'moderate types' in each party are relatively more popular than the 'extreme types', and political office confers ego-rents as well as an opportunity to implement a more desired policy, they must prevent the latter from hijacking their platforms. In order to credibly signal their type, the moderate types need to appear even more moderate than they actually are, to ensure that the resulting policy platform is prohibitively expensive for the extreme types in their party, to hijack. However, if a moderate type wins the second elections, then, as a lame-duck politician, she will implement her most preferred tax rate.

The increased moderation of the moderate types creates 'partial convergence' of the redistributive platforms implemented by the two parties towards the middle. Middle income voters then find the implemented redistributive policy to be closer to their preferred outcome; this is a version of 'Director's Law'. Depending on the intensity of the extreme type's desire to misrepresent her type, it is possible that the moderate types in each party move too far into the middle, and across each other, creating 'policy reversals' between the two parties. The distortion in the redistributive policy of the moderate type creates 'inter-term tax variability' (political budget cycles) over the two periods. The pattern of these political budget cycles is party-specific and is in line with the empirical evidence in Besley and Case (1995), namely, Left (resp. Right) party incumbents undertake relatively smaller (resp. larger) redistribution in their first term as compared to the second.

The results on partial convergence and Director's Law also arise under full-information and full-commitment, however, they are driven by other considerations, such as the promise of greater pork to voter-groups with greater 'political clout'. The results on policy reversals and on (inter-term) political budget cycles do not, however, have any close counterpart in that literature. Comparative static results show that 'ego rents', the 'political polarization' in society, and 'inter-party political polarization' are important determinants of equilibrium redistribution. The predictions of these comparative static results are consistent with the empirical findings of Shi and Svensson (2001).

Finally, most of the existing literature poses the relation between redistribution and the extent of income inequality in a 'median voter framework' (MVF) in which the median voter directly chooses policy. This framework ignores important

²This two-period model embodies constitutional institutions such as 'term limits' which are important for gubernatorial and presidential elections in the United States. Term limits are also a feature of presidential elections in South America (except for the Dominican Republic). Even in European parliamentary democracies where term limits are absent, finite lifetimes, institutional and social changes, and changes in the nature of internal party struggles have constrained most postwar prime ministers to a relatively small number of terms in office.

issues raised in a ‘representative democracy framework’ (RDF) in which the incentives and constraints facing politicians determine actual policy. The results of the two frameworks coincide only in very special cases; see for instance Cukierman and Speigel (2001).

In a MVF, Meltzer and Richard (1981) and Persson and Tabellini (1994) show that inequality reduces the income of the median voter who consequently chooses greater redistribution, however, the empirical evidence is mixed. Not a great deal is known about the effect of inequality on redistribution in an RDF. This paper shows that inequality, by altering the support bases of politicians, alters the incentives of the extreme type to hijack the moderate type’s platform. It is shown that fairly simple conditions on the slope of the population density, following an increase in inequality, can generate both positive and negative effects of inequality on the magnitude of the redistributive activity.

The paper is organized as follows. Section 2 describes the model while Section 3 derives the properties of the probability of reelection. Sections 4 and 5 respectively characterize the equilibrium under the benchmark case of full information and asymmetric information. Section 5 also derives the comparative static results and discusses some extensions. Section 6 examines the implications for redistribution when inequality increases, in an RDF. Finally, Section 7 concludes the paper. All proofs are collected in Appendix A.

2. The model

The characteristics of all individuals are drawn from the space $Y = W \otimes \Theta$, with generic element (w, δ) . $W = [\underline{w}, \bar{w}] \subset \mathfrak{R}_+$ is the interval of exogenous societal wealth (or income) with distribution function $F(w)$, and $\Theta = \{\delta_c, \delta_l\}$ represents the ideological type of an individual such that $\delta_c, \delta_l \in [0, 1]$ and $\delta_c < \delta_l$. Individuals of type δ_l prefer greater redistribution and are described as ‘liberals’ while individuals of type δ_c prefer lower redistribution and are described as ‘conservatives’.³ While w is public information, δ is private information to the individual; the prior probability that $\delta = \delta_l$ is ρ . The game lasts for two periods and in each period, the characteristics of any individual, (w, δ) , are identical. There are two political parties, the Left and the Right; party members in each party endogenously determine the respective party candidates, (w_L, δ_l) and (w_R, δ_c) , such that $w_L < w_R$ and $\delta_l, \delta_c \in \Theta$.

Definition 1. *Ideological polarization* in society is defined as $\Delta\delta = \{\delta_l - \delta_c\}$.

³It is more realistic to suppose, as in Dixit and Londregan (1998) that ideology might be correlated with income or that ideology in the Left party differs from that in the Right party. These extensions can be accommodated, at the cost of greater complexity, but are not critical to the results.

2.1. Redistributive policy

In any of the two periods, the politician who holds office engages in balanced budget redistribution using a linear progressive income tax $T = \{\alpha, t\}$, where $0 \leq t \leq 1$ is the constant marginal tax rate and $\alpha \geq 0$ is the per capita lumpsum transfer received by all individuals. The per capita taxes collected equal $t \int_w w dF(w) = t\mu$, where μ is mean income, thus, the government budget constraint is given by $\alpha = t\mu$.

2.2. Preferences

In any period, a (w_i, δ_j) individual derives utility $V_{ij}(t)$ defined by:

$$V_{ij}(t) = \theta \{C(t; w_i) + d\psi_{ij}\} - \frac{1}{2} \{t\mu - \delta_j \mu\}^2. \quad (2.1)$$

Individuals derive utility from their private consumption $C(t; w) = [1 - t]w + t\mu$ and ego-rents ψ_{ij} (which entail no resource costs) that accrue only if that individual holds office, in which case d equals unity, and zero otherwise. However, individuals derive disutility if the actual magnitude of redistribution $t\mu$, differs from their desired level, $\delta_j \mu$. The parameter $\theta > 0$ is the relative weight placed by the individual on private consumption relative to her ideological preferences. Using terminology introduced by Roemer (1998), the parameter θ^{-1} represents the ‘salience’ of the ideological issue with respect to redistribution.⁴ There is no discounting over the two periods and the intertemporal payoff of any (w_i, δ_j) individual, $E[W_{ij}]$, is additively separable in the expected payoffs in each period:

$$E[W_{ij}] = E[V_{ij}] + E[V_{ij}] \quad (2.2)$$

where E denotes the expectation operator with respect to the joint distribution of the ideological type of the politician and the probability of her winning the election.

2.3. Timing

The sequence of moves is shown in Fig. 1. Nature moves first by choosing the characteristics (w, δ) of all individuals. The Left and the Right parties simultaneously select candidates (w_L, δ_j) and (w_R, δ_k) , respectively; these candidates are unchanged for the following two periods.⁵ The winner of the first election, the ‘first

⁴Alesina et al. (2001) suggest that because of differences in social mobility, θ^{-1} is relatively higher for Europeans as compared to Americans.

⁵If the opposition party were to reselect its candidates at the end of the first period, results do not change, provided that the first period incumbent is allowed to contest the second elections.

Pre-Stage	First Period	Second Period
Candidate Selection Left party: (w_L, δ_j) Right party: (w_R, δ_k)	Winner implements first period tax policy	Lame-duck politician implements second period tax policy
	First Election	Second Election

Fig. 1. The sequence of moves.

period incumbent’, implements the first period redistributive tax, which voters use to infer the incumbent’s ideological type. At the end of the first period, the first period incumbent contests her ‘second election’ against the opposition candidate. The winner of the second elections (the lame-duck politician) implements second period redistributive tax policy.

2.4. Characteristics of the most preferred tax rate (MPTR)

In any period, the ‘most preferred tax rate’ (MPTR) of a (w_i, δ_j) individual, t_{ij}^* , is the solution to the first order condition $V'_{ij}(t_{ij}^*) = \theta C'(t_{ij}^*; w_i) - \mu^2\{t_{ij}^* - \delta_j\} = 0$; the second order condition is satisfied because $V''_{ij}(t_{ij}^*) < 0$. Solving out for t_{ij}^* :

$$t_{ij}^* = \delta_j + \frac{\theta(\mu - w_i)}{\mu^2}. \tag{2.3}$$

The condition $\delta_j \geq \mu^{-2}\theta(\mu - \bar{w})$ guarantees that $t_{ij}^* \geq 0, \forall (w, \delta) \in Y$, but is not vital for the results. The comparative static properties of the MPTR follow directly from Eq. (2.3); these are summarized without proof in Result 1.

Result 1. The MPTR is increasing in the individual’s ideological type, δ_j (ceteris paribus, ‘liberals’ prefer higher tax rates relative to ‘conservatives’), and decreasing in the individual’s wealth, w_i (richer voters prefer a lower redistributive tax rate). An increase in the salience parameter θ^{-1} reduces the MPTR if $\mu > w_i$ but increases the MPTR if $\mu < w_i$.

Recall that the Left and the Right party politicians are individuals with respective characteristics (w_L, δ_j) and (w_R, δ_k) : $w_L < w_R$. The following definition encapsulates the spatial redistributive distance between the two parties.

Definition 2. Political polarization is defined as $w_R - w_L > 0$ or equivalently as $t_{Lj}^* - t_{Rk}^* > 0$.

Using Result 1, liberal politicians in any party prefer higher taxes, i.e. $t_{Rc}^* < t_{Rl}^*$ and $t_{Lc}^* < t_{Ll}^*$ and since $w_L < w_R$, thus, $t_{Rc}^* < t_{Ll}^*$. Furthermore, if there is some minimum degree of political polarization in the sense that $w_R - w_L \geq \theta^{-1} \Delta\delta$, then $t_{Rl}^* < t_{Lc}^*$. In that case, the following inequalities hold:

$$t_{Rc}^* < t_{Rl}^* < t_{Lc}^* < t_{Ll}^*. \quad (2.4)$$

The two extreme inequalities are critical; the middle inequality despite being plausible, is less important. These inequalities show that the Liberal-Right and the Conservative-Left politician prefer moderate redistribution relative to the Conservative-Right and the Liberal-Left politicians; the following terminology is self evident.

Definition 3. Within their respective parties, the Liberal-Right and the Conservative-Left politicians are the ‘moderate types’ while the Conservative-Right and the Liberal-Left politicians are the ‘extreme types’.

2.5. Voting

Denote the expected redistributive tax rates of the Left and the Right party politicians by t_{Lj}^r and t_{Rk}^r , respectively, in the r th period, $r = 1, 2$. In the second election, a (w_i, δ_j) voter votes for the Left party over the Right party if $E[V_{ij}(t_{Lj}^2)] > E[V_{ij}(t_{Rk}^2)]$, while in the first election, she votes Left over Right if $E[W_{ij}(t_{Lj}^1, t_{Lj}^2)] > E[W_{ij}(t_{Rk}^1, t_{Rk}^2)]$. In any election, the winner is elected by a plurality rule. Suppose that in the r th election, a fraction x of the total voters in the population prefer the Left party candidate to the Right party candidate. Then, under plurality rule, the probability with which the Left party candidate (w_L, δ_j) wins the r th election is given by $\Pi_{Lj}^r = \Pr(x + \epsilon > 1/2)$ where ϵ is uniformly distributed over the interval $[-1/2, 1/2]$ and is independent of x .⁶ It is straightforward to compute that $\Pi_{Lj}^r = x$ and $\Pi_{Rk}^r = 1 - x$.

3. Re-election probability: second election

The probability of winning the first election is relevant only for the candidate selected by each party, but not for the subsequent redistributive equilibrium, thus,

⁶Suppose that the utility function was $V_{ij}(t_p) + \gamma_p$, where t_p is the redistributive tax expected of party $p = L, R$ and γ_p is a party-specific parameter not related to the tax rate. Then, a (w_i, δ_j) voter votes for the Left party over the Right if $V_{ij}(t_L) - V_{ij}(t_R) \geq \gamma_R - \gamma_L$. Defining $\epsilon = \gamma_R - \gamma_L$ gives the result in the text. The term ϵ is identical to the ‘net looks-shock’ term in Rogoff (1990) where it is also independent of x and is interpreted as the ‘weather on election day’ by Roemer (1998), ‘voter-bias’ by Wittman (1983), ‘utility difference in favour of the politician’ by Lindbeck and Weibull (1993) and ‘intrinsic preference in favour of the politician’ by Boadway and Keen (1999).

its consideration is postponed to Sections 4.1 and 5.5.1. Two remarks of a presentational nature are in order. Firstly, because, the game is symmetric, throughout the paper the first period incumbent is assumed to belong to the Left party, on account of alphabetical precedence; analogous results for a Right party incumbent are noted where appropriate. Secondly, it simplifies the analysis to assume that voters have asymmetric information about the Left party incumbent and full information about her Right party challenger. Section 5.5.2 relaxes this assumption; this only affects the voters expectation of the tax rate from the lame-duck Right party politician and does not affect the main results.

Suppose that at the end of the first period, voters expect the (lame-duck) Left and the Right party politicians to implement redistributive taxes t_{Lj}^* and t_{Rk}^* , respectively, in the second period. Let $(w_c(t_{Lj}^*, t_{Rk}^*), \delta_c)$ and $(w_l(t_{Lj}^*, t_{Rk}^*), \delta_l)$ respectively denote a conservative and a liberal voter whose MPTR is $(t_{Lj}^* + t_{Rk}^*)/2$. Using Eq. (2.3), w_c and w_l are:

$$\begin{aligned} w_c(t_{Lj}^* + t_{Rk}^*) &= \mu + \theta^{-1} \mu^2 \{ \delta_c - (t_{Lj}^*, t_{Rk}^*)/2 \} \\ w_l(t_{Lj}^* + t_{Rk}^*) &= \mu + \theta^{-1} \mu^2 \{ \delta_l - (t_{Lj}^*, t_{Rk}^*)/2 \}. \end{aligned} \tag{3.1}$$

Check that $w_c(t_{Lj}^*, t_{Rk}^*) < w_l(t_{Lj}^*, t_{Rk}^*) = w_c(t_{Lc}^*, t_{Rk}^*) < w_l(t_{Lc}^*, t_{Rk}^*)$. Lemma 1 shows that the wealth distribution can be partitioned into three intervals: ‘core supporters’ of each party who vote for that party, irrespective of their ideology and ‘swing voters’ whose (privately known) ideology is decisive in their voting choice.⁷

Lemma 1. *All (w, δ_j) voters such that $w < w_c(t_{Lj}^*, t_{Rk}^*)$ prefer the Left party while all (w, δ_j) voters such that $w_l(t_{Lj}^*, t_{Rk}^*) < w$ prefer the Right party; these voters constitute the ‘core support’ for a party. In the interval $(w_c(t_{Lj}^*, t_{Rk}^*), w_l(t_{Lj}^*, t_{Rk}^*))$, liberal voters vote Left while conservative voters vote Right; these voters constitute the ‘swing support’.*

Lemma 2. *The reelection probability of a (w_L, δ_j) politician against a (w_R, δ_k) politician is $\Pi_{Lj}^2(t_{Lj}^*, t_{Rk}^*) = F(w_c(t_{Lj}^*, t_{Rk}^*)) + \rho \{ F(w_l(t_{Lj}^*, t_{Rk}^*)) - F(w_c(t_{Lj}^*, t_{Rk}^*)) \}$.*

The expression for Π_{Rk}^2 follows from $\Pi_{Rk}^2 = 1 - \Pi_{Lj}^2$. Proposition 1 illustrates some features of Π_{Lj}^2 ; analogous properties can be checked to hold for Π_{Rk}^2 .

Proposition 1. $\Pi_{Lj}^2(t_{Lj}^*, t_{Rk}^*)$ is increasing as: (1) the prior probability, ρ , that a voter is liberal, increases, (2) the Right party candidate becomes ‘less centrist’,

⁷Unlike in the pork-barrel politics literature, the term ‘swing voters’ does not imply that these voters can be swayed by promises of pork (because of the absence of full-commitment). However, their privately known ideology can swing the outcome of the election results, ex-post.

i.e. t_{Rk}^* decreases, and (3) the Left party candidate becomes ‘more centrist’, i.e. t_{Lj}^* decreases.

Proposition 1 shows that candidates whose redistributive position is more centrist are more likely to win and an increase in the proportion of ‘swing voters’ who support the Left party increases Π_{Lj}^2 . The effects of ideological polarization $\Delta\delta$ and the salience parameter θ^{-1} , require further restrictions; Section 5.3 looks at these affects.

Under asymmetric information, the potential benefits to a politician of extreme type from hijacking a moderate type’s platform are directly related to the resulting increase in popularity among the voters; the following terminology is useful.

Definition 4. The ‘differential election probability’ between the two types in the Left and the Right party, respectively, in the r th election is defined as $\Delta\Pi_L^r = \Pi_{Lc}^r - \Pi_{Ll}^r$ and $\Delta\Pi_R^r = \Pi_{Rl}^r - \Pi_{Rc}^r$.

To minimize notation and to emphasize that the first period incumbent belongs to the Left party, write $w(t_{Lc}^*, t_{Rk}^*) = w(t_{Lc}^*)$. Using the definition of Π_{Lj}^2 in Lemma 2 it can be checked that $\Delta\Pi_L^2$ equals:

$$\Delta\Pi_L^2 = \rho\{F(w_l(t_{Lc}^*)) - F(w_c(t_{Lc}^*))\} + (1 - \rho)\{F(w_l(t_{Ll}^*)) - F(w_c(t_{Ll}^*))\}. \tag{3.2}$$

The intuition runs as follows. For the moderate and the extreme types, the ‘core support’ lies respectively in the intervals $[w_c, w_c(t_{Lc}^*)]$ and $[w_c, w_c(t_{Ll}^*)]$ while the respective ‘swing support’ lies in the intervals $[w_c(t_{Lc}^*), w_l(t_{Lc}^*)]$ and $[w_c(t_{Ll}^*), w_l(t_{Ll}^*)]$. Since the reelection probability of any type equals the fraction of voters who prefer that type, thus, if a moderate is replaced by an extreme type, a fraction $\rho\{F(w_l(t_{Lc}^*)) - F(w_c(t_{Lc}^*))\}$ of voters in the interval $[w_c(t_{Lc}^*), w_l(t_{Lc}^*)]$ and another fraction $(1 - \rho)\{F(w_l(t_{Ll}^*)) - F(w_c(t_{Ll}^*))\}$ of voters in the interval $[w_c(t_{Ll}^*), w_l(t_{Ll}^*)]$ switch to the Right party. Eq. (3.2) can be rewritten as:

$$\Delta\Pi_L^2 = \rho \int_{w_c(t_{Lc}^*)}^{w_l(t_{Lc}^*)} F'(x) dx + \{1 - \rho\} \int_{w_c(t_{Ll}^*)}^{w_l(t_{Ll}^*)} F'(x) dx > 0. \tag{3.3}$$

The sign of Eq. (3.3) follows because $F'(x) > 0$, $w_l(t_{Lc}^*) > w_c(t_{Lc}^*)$ and $w_l(t_{Ll}^*) > w_c(t_{Ll}^*)$. Thus, the moderate types, since they are ‘centrists’ relative to the extreme types in their party, are more likely to win elections. Analogously, if the first period incumbent belongs to the Right party, write $w(t_{Lj}^*, t_{Rk}^*) = w(t_{Rk}^*)$ to simplify notation and check that $\Pi_{Rk}^2 = 1 - F(w_l(t_{Rk}^*)) + (1 - \rho)\{F(w_l(t_{Rk}^*)) - F(w_c(t_{Rk}^*))\}$. The differential reelection probability in this case is:

$$\Delta \Pi_R^2 = \rho \int_{w_c(t_{Rc}^*)}^{w_l(t_{Rc}^*)} F'(x) dx + \{1 - \rho\} \int_{w_c(t_{Rl}^*)}^{w_c(t_{Rc}^*)} F'(x) dx > 0. \tag{3.4}$$

4. Full information equilibrium

Suppose that political parties can credibly transmit all relevant information about their candidate to the voters. Then, for any of the two elections, forward-looking voters correctly anticipate the post-election redistributive tax policy and since no new information about the incumbent arises on account of her tax policy, the result in Lemma 3 is obvious.

Lemma 3. *Under full information, the election probability of any politician in any of the two elections is identical i.e. $\Pi_{Lj}^1 = \Pi_{Lj}^2 = \Pi_{Lj}$ and $\Pi_{Rk}^1 = \Pi_{Rk}^2 = \Pi_{Rk}$.*

Voters know that the lame-duck politician will implement her MPTR; t_{Lj}^* if she belongs to the Left party and t_{Rk}^* if she belongs to the Right party. Since voters know the type of all politicians so $\Pi_{Lj}^2 = \Pi_{Lj}^2(t_{Lj}^*, t_{Rk}^*)$, which is independent of the first period redistributive taxes. Thus, the optimization problem of the first period (Left party) incumbent is to choose t_1 to maximize $E[W_{Lj}(t_1, t_{Lj}^*)] = V_{Lj}(t_1) + \Pi_{Lj}^2 V_{Lj}(t_{Lj}^*) + (1 - \Pi_{Lj}^2) V_{Lj}(t_{Rk}^*)$; the optimal solution is $t_1 = t_{Lj}^*$. Hence, under full information any politician implements her MPTR in both periods.

Despite $\Delta \Pi_L^1 > 0$ and $\Delta \Pi_R^1 > 0$, full information prevents the extreme type in each party from misrepresenting her type. Under full information, there are no political budget cycles over the terms of a politician, no convergence of party platforms, no policy reversals, and no support for Director’s Law. These results will change when asymmetric information is introduced in Section 5.

4.1. A simple model of candidate selection under full information

This section briefly sketches a simple model of candidate selection in the spirit of Roemer (1998, 1999). Relative to asymmetric information about the politician, the model of candidate selection is more elegant under full information mainly because (1) one can abstract from issues of voters having to guess the subsequent ‘Perfect Bayesian Equilibrium’, and (2) the election probabilities in the two elections are identical (see Lemma 3). For some candidate selection issues under asymmetric information, see Section 5.5.1. Suppose that the Left and the Right parties stand respectively on some ‘initial’ redistributive platforms, t_L and t_R : $t_L < t_R$; these platforms arise from some (unmodeled) history of the game. Each of the two parties has a set of party members who are exogenously given. The candidate selection process makes the following four assumptions:

[A1]. Equilibrium concept: Both parties simultaneously select their candidates

given the candidate chosen by the other party; the equilibrium concept is (Cournot) Nash equilibrium.

Since the game is symmetric, consider first the derivation of the Left party ‘reaction function’ $t_{Lj}(t_{Rk})$, which gives its desired redistributive tax rate for any redistributive tax rate t_{Rk} chosen by the Right party. Using Eq. (2.3) any desired redistributive policy can be mapped to the characteristics $(w_L, E[\delta])$ of the desired party candidate, where $E[\delta] = \rho\delta_l + (1 - \rho)\delta_c$ is the expected ideological position of any voter.⁸

[A2]. Factions: Party members state if their MPTR $t_{ij}^* \geq t_L$. Denote the set of party members who state $t_{ij}^* > t_L$ by **M** and the set of party members who state $t_{ij}^* < t_L$ by **O**.

It is obvious that party members in the set **M** prefer the ‘initial’ party position t_L relative to the MPTR of any party member in the set **O** because t_L is relatively closer to their own MPTR. For this reason, they are termed as ‘*militants*’. On the other hand, the MPTR of party members in **O** is relatively ‘centrist’, as compared to the ‘initial’ party position t_L , and has a higher probability of being successful in the election (see Proposition 1); hence they are termed as ‘*opportunists*’. In spirit, but not in the detail, the terms militants and opportunists correspond to the usage of these terms in Roemer (1998, 1999). In Roemer, the opportunists maximize the probability of winning the elections by advocating more centrist positions, while the militants simply want to adhere to the ‘initial’ party position.

[A3]. Bargaining among factions: The median (or representative) party members in each of the groups **M** and **O**, bargain with each other, using Generalized Nash Bargaining, to arrive at the consensus candidate for the party. If they fail to reach an agreement, then the opposition candidate wins by default.

The utility functions of the median faction members in the groups *M* and *O* are $E[W_{Mj}]$, and $E[W_{Oj}]$; the respective MPTRs are t_M^* and t_O^* ; the respective bargaining powers are $\lambda_M > 0$ and $\lambda_O > 0$: $\lambda_O + \lambda_M = 1$;⁹ and the respective disagreement payoffs are d_M and d_O .

⁸The assumption that all individuals are willing to take up the responsibility of office if chosen to represent their party is not overly restrictive. If taking up political office is costly and the pool of potential candidates is limited, then the one closest to the party’s desired position is chosen.

⁹Roemer (1998, 1999) provides a brief historical sketch of the evolution of this bargaining power. For instance the New British Labour party is often distinguished from the Old Labour Party in terms of a decrease in the bargaining power of the militants relative to the opportunists. Evidence suggests that the source of this bargaining power lies in the degree of enfranchisement, voter turnout, costs of political participation, changes in partisanship among voters, social mobility, and changes in political communication and campaign technologies; see for example Dalton and Wattenberg (2000), Müller (1999) and Epstein (1980).

[A4]. Technical assumption: Define $w_c = \mu - \theta^{-1} \mu^2 \{t_L - \delta_c\}$ and $w_l = \mu - \theta^{-1} \mu^2 \{t_L - \delta_l\}$. The measure of party members in $[w_c, w_l]$ is zero.

Using Eq. (2.3), it is publicly known that irrespective of ideological preference (1) for all party members (w_i, δ_j) : $w_i < w_c$, the MPTR $t_{ij}^* > t_L$, and (2) for all party members (w_i, δ_j) : $w_i > w_l$, the MPTR $t_{ij}^* < t_L$. Thus, the only credible announcement by party members with $w_i < w_c$ is to join the group **M**, while that for $w_i > w_l$ is to join group **O**. For all Party members $(w_i, \delta_l) \in [w_c, w_l]$, $t_{il}^* > t_L$ while for all party members $(w_i, \delta_c) \in [w_c, w_l]$, $t_{ic}^* < t_L$ hence their (unknown) ideology is critical in their decision to join a particular faction.¹⁰ However, since the measure of these party members is zero, their (possible) misrepresentation of type does not affect the location of the median voters in each of the two factions.

Under full information about the politicians and no intertemporal discounting, the respective payoffs of the median members in the factions **M** and **O** are $E[W_{Mj}] = 2\Pi_{Lj}V_{Mj}(t_{Lj}) + (2 - \Pi_{Lj})V_{Mj}(t_{Rk})$ and $E[W_{Oj}] = 2\Pi_{Lj}V_{Oj}(t_{Lj}) + (2 - \Pi_{Lj})V_{Oj}(t_{Rk})$; these expressions use the result in Lemma 3.

Lemma 4. *In the bargaining game between the two factions in the Left party, the set of feasible tax rates belongs to the interval $T = [t_O^*, t_M^*]$.*

An analogous condition to the one stated in Lemma 4 holds for the Right party. Hence a sufficient condition for $t_{Rk}^* < t_{Lj}^*$, or alternatively $w_L < w_R$ (using Result 1) is that the median member of the opportunist faction in the Left party prefers a higher redistributive tax rate relative to his counterpart in the Right party.

Given Assumption A3, the disagreement payoffs d_M and d_O are found by setting $\Pi_{Lj} = 0$ in the objective functions; thus $d_M = 2V_{Mj}(t_{Rk})$ and $d_O = 2V_{Oj}(t_{Rk})$. The Generalized Nash Bargaining solution for $t_{Lj}(t_{Rk})$ solves the following optimization problem:

$$t_{Lj} \in \arg \max \Lambda = (E[W_{Oj}] - d_O)^{\lambda_O} (E[W_{Mj}] - d_M)^{\lambda_M}$$

subject to: $t_{Lj} \in T$.

Letting primes denote derivatives, $t_{Lj}(t_{Rk})$ solves the first order condition:

$$\frac{\partial \Lambda}{\partial t_{Lj}} = \frac{\Pi'_{Lj}}{\Pi_{Lj}} + \frac{2\lambda_M V'_{Mj}(t_{Lj})}{2V_{Mj}(t_{Lj}) - V_{mj}(t_{Rk})} + \frac{2\lambda_O V'_{Oj}(t_{Lj})}{2V_{Oj}(t_{Lj}) - V_{Oj}(t_{Rk})} = 0. \tag{4.1}$$

The three terms in (4.1) show the costs and benefits of choosing a less centrist (higher t_{Lj}) position. The first term is negative ($\Pi'_{Lj} < 0$ from Proposition 1); a

¹⁰Strategic considerations can induce party members in $[w_c, w_l]$ to misrepresent their type if by doing so the redistributive position chosen by the party is closer to their MPTR. The current model is not suited to these considerations because the bargaining power of the two factions is not determined endogenously.

less centrist candidate is less likely to be elected, which is costly to both factions. The second term is positive ($V'_{Mj}(t_{Lj}) > 0$ when $t_{Lj} \in T$); the militants benefit by putting up a less centrist candidate whose MPTR is closer to theirs. The third term is negative ($V'_{Oj}(t_{Lj}) < 0$ when $t_{Lj} \in T$); the MPTR of a less centrist candidate is farther away from the one preferred by the opportunists. The net effect of the second and the third terms crucially depends on the relative bargaining powers of the two factions. The comparative static results are stated in Proposition 2.

Proposition 2. *For any t_{Rk} , the ‘desired’ redistributive position chosen by the Left party is relatively centrist (t_{Lj} lower) if λ_o increases. In general, the effect of t_{Rk} on t_{Lj} is ambiguous but as $\lambda_o \rightarrow 1$, the Left party reciprocates the choice of a more centrist Right party candidate by choosing a more centrist Left party candidate.*

Proposition 2 shows that an increase in λ_o empowers the opportunists to press for a more centrist redistributive position (lower t_{Lj}^*) that is more electable. If the Right party fields a more centrist candidate (higher t_{Rk}^*), there are two countervailing effects. Firstly, since the Right party candidate is now closer to the middle of the political spectrum, Π_{Lj} falls (see Proposition 1) which hurts both factions. Secondly, choosing a more centrist Left party candidate in response would benefit the opportunists (as $V'_{Oj}(t_{Lj}) < 0$) but harm the militants (as $V'_{Mj}(t_{Lj}) > 0$). Hence, the net effect is ambiguous but when $\lambda_o \rightarrow 1$ the militants can do little to prevent a more electable Left party candidate from being fielded; in that case, $t_{Lj}(t_{Rk})$ is downward sloping.

An interesting anecdotal illustration of Proposition 2 is the history of the British Labour party. In the three decades prior to the election of Tony Blair as prime minister, purists (militants) in the Labour party were extremely powerful (λ_M high) and the party lost election after election without compromising on its ideological convictions. However, following the decade of the 1980s there was a change in the membership of the labour party and a change in its institutional set-up that significantly increased the bargaining power of the opportunists and allowed Blair to be elected on fairly centrist policies.¹¹

The reaction function for the Right party $t_{Rk}(t_{Lj})$ can be derived analogously. In general, there is no presumption that the Nash equilibrium is unique, however, Fig. 2 shows the reaction functions for each party $t_{Rk}(t_{Lj})$ and $t_{Lj}(t_{Rk})$ in the case of a unique equilibrium outcome. These reaction functions are drawn for two possible values of the bargaining power of the opportunists in each party; a lower value, i.e.

¹¹There was a relative shift in the membership of the Labour party from Trade Union activists (militants) to middle class white collar workers (opportunists). Institutional changes initiated by Neil Kinnock and John Smith in the 1980s and the early 1990s further weakened the power of the militants in the selection of the party candidate. For a book-length treatment of these issues see Driver and Martell (1998).

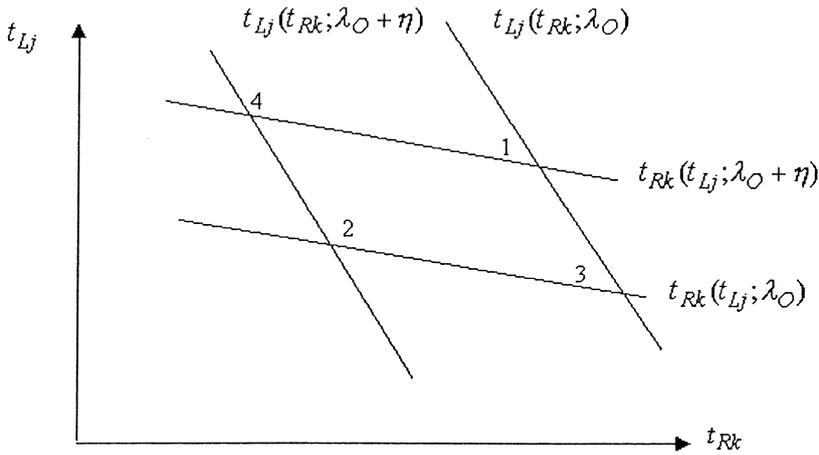


Fig. 2. Candidate selection equilibrium.

λ_O and a higher value, i.e. $\lambda_O + \eta$, where $\eta > 0$. The intersection points ‘3’ and ‘4’ represent equilibria when the bargaining power of the opportunists in each party is low and high, respectively, while intersection points ‘1’ and ‘2’ illustrate two intermediate cases.

5. Asymmetric information equilibrium

This section characterizes the Perfect Bayesian Nash Equilibria when voters have asymmetric information about the politician’s ideological type. Attention is restricted to pure strategies alone, hence, hybrid equilibria are ruled out. By placing an appropriate upper bound on political rents, all pooling equilibria can be ruled out. Thus, the focus will be on the separating equilibria; by ruling out dominated strategies a unique separating equilibrium can be characterized.

Focus initially on the second election. In forming inferences about the ideological type of the first period incumbent, voters update their beliefs, conditional on the observed first period tax policy, t . Common posterior beliefs are: $\omega = \text{Pr}(\text{Type is } \delta_c | t)$ and $1 - \omega = \text{Pr}(\text{Type is } \delta_j | t)$. In a ‘separating equilibrium’, the tax policy fully reveals the ideological type of the incumbent, while in a pooling equilibrium, it carries no information content.

5.1. Separating equilibrium

The game following the selection of the first period (Left party) incumbent is as follows. The set of players comprises of the incumbent whose characteristics are (w_L, δ_j) , where $\delta_j \in \Theta$ is private information, and the voters whose characteristics

are drawn from $Y = W \otimes \Theta$. The action set of the voters $\nu = \{0, 1\}$ denotes a binary voting choice; 1 denotes a vote ‘for’ and 0 denotes a vote ‘against’ the incumbent in the second election. The action set of the politician is a choice of the tax rate in the interval $[0, 1]$. Any voter’s strategy is defined by the mapping $[0, 1] \rightarrow \{0, 1\}$, i.e. a voting choice for all possible announcements of the tax rate by the politician. The strategy of the politician is defined by the mapping $\Theta \rightarrow [0, 1]$, i.e. a choice of the tax rate for each of the two possible types. The single period and the intertemporal payoffs are given in (2.1) and (2.2), respectively.

Let the set of possible tax rates implemented by a moderate incumbent, (w_L, δ_c) , in a separating equilibrium be denoted by S_{Lc} , with generic element t_{Lj}^1 . Thus (consistent with Bayes rule), for any $t_{Lj}^1 \in S_{Lc}$, the following posterior beliefs can support a separating equilibrium: $\omega = \Pr(\delta = \delta_c | t_{Lj}^1 \in S_{Lc}) = 1$ and zero otherwise. The set S_{Lc} must satisfy the following two conditions, labelled C1 and C2:

$$\Delta \Pi_{Lc}^2 \{V_{Lc}(t_{Lc}^*) - V_{Lc}(t_{Rk}^*) + \psi_{Lc}\} \geq V_{Lc}(t_{Lc}^*) - V_{Lc}(t_{Lj}^1) \tag{C1}$$

$$\Delta \Pi_{Ll}^2 \{V_{Ll}(t_{Ll}^*) - V_{Ll}(t_{Rk}^*) + \psi_{Ll}\} \leq V_{Ll}(t_{Ll}^*) - V_{Ll}(t_{Lj}^1). \tag{C2}$$

Condition C1 is similar to a *participation constraint*; it ensures that the future gain to the moderate type from separation (LHS of C1), exceeds the current loss (RHS of C1). Condition C2 is similar to an *incentive compatibility constraint*; it ensures that the current cost to the extreme type from misrepresenting her type (RHS of C2) exceeds the future gain (LHS of C2). Denote by S_1 and S_2 , respectively, the set of tax rates that satisfy conditions C1 and C2, and let S_2^C denote the complement of S_2 . The set of feasible separating tax rates $S_{Lc} = S_1 \cap S_2$. Lemma 5 records a useful result that facilitates the subsequent comparative static results.

Lemma 5. *Condition C2 binds in a separating equilibrium.*

The moderate type’s optimal tax policy, t_{Lc}^S , is the solution to the following problem:

$$t_{Lc}^S \in \arg \max E[W_{Lc}(t_{Lj}^1, t_{Lc}^*)] = V_{Lc}(t_{Lj}^1) + \Pi_{Lc}^2 \{V_{Lc}(t_{Lc}^*) + \psi_{Lc}\} + \{1 - \Pi_{Lc}^2\} V_{Lc}(t_{Rk}^*)$$

subject to: $t_{Lj}^1 \in S_{Lc} = S_1 \cap S_2$.

Denoting the Lagrangian multipliers on the conditions C1 and C2 by χ_1 and χ_2 , respectively, t_{Lc}^S is the solution to the following first order condition:

$$\frac{\partial V_{Lc}(t_{Lc}^S)}{\partial t_{Lj}^1} = \left[\frac{\chi_2}{1 + \chi_1} \right] \frac{\partial V_{Ll}(t_{Lc}^S)}{\partial t_{Lj}^1}. \tag{5.1}$$

The RHS of Eq. (5.1) is strictly positive unless $\chi_2 = 0$, thus, relative to the full information case, $V'_{Lc}(t_{Lc}^S) > V'_{Lc}(t_{Lc}^*) = 0$, hence, the moderate type implements an

even more moderate tax policy under asymmetric information i.e. $t_{Lc}^S < t_{Lc}^*$; the case $\chi_2 = 0$ is explored in Example 1.

Lemma 6. *If dominated strategies are ruled out, then the unique solution to the moderate type’s problem is $t_{Lc}^S = \inf S_2^C \leq t_{Lc}^*$.*

Fig. 3 provides a diagrammatic illustration of the separating equilibrium in the policy space, $\{\alpha, t\}$. The respective indifference curves of the moderate and the extreme types corresponding to the full information outcome are shown as MM and EE , the government budget constraint ($\alpha = t\mu$) is shown as the line BB through the origin and the respective MPTRs are shown as t_{Lc}^* and t_{Ll}^* . The depiction of the sets S_1 and S_2 reflects the fact that individual payoffs are decreasing in taxes that lie farther away from one’s MPTR; at the boundaries of the set S_2^C , the moderate type’s platform is sufficiently distant so as to give zero net benefits to the extreme type from misrepresenting her type. The indifference curve of the extreme type passing through $\{\alpha, t_{Lc}^S\}$ is shown by $E1E1$; the vertical distance G between EE and $E1E1$ is the current loss to the extreme type from misrepresenting her type; in equilibrium it just equals the future gain (LHS of C2).

Given the posterior beliefs, the moderate type’s tax policy immediately reveals her type to the voters, so there is no gain to the extreme type from distorting her tax policy relative to her MPTR, t_{Ll}^* , hence $t_{Ll}^S = t_{Ll}^*$. Thus, the separating equilibrium is characterized as $\{t_{Lc}^S, t_{Ll}^*, \omega, \rho\}$: $t_{Lc}^S = \inf S_2^C \leq t_{Lc}^*$. In the second period, lame duck politicians simply implement their MPTR.

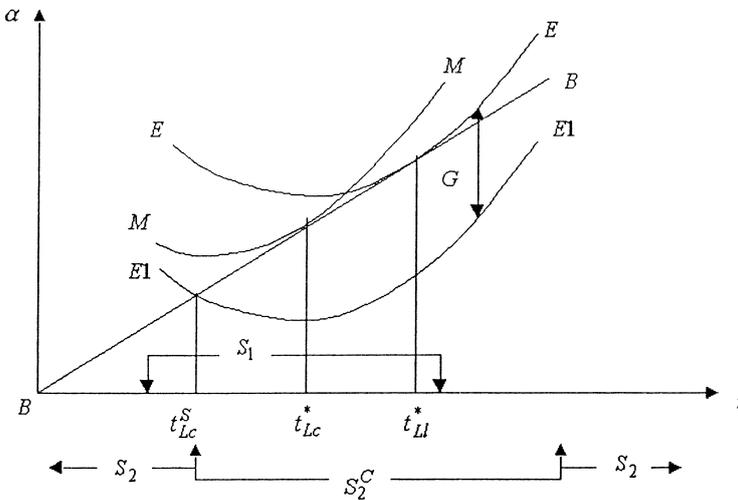


Fig. 3. Separating equilibrium: Left party incumbent.

When $\chi_2 = 0$ at $t_{Lc}^S = t_{Lc}^*$, the solutions under asymmetric and full information coincide; Example 1 considers this possibility when the distribution of population is uniform.

Example 1. Suppose that the population density is uniform over $[0, 1]$ and the Right party politician is liberal. Condition C2, when evaluated at $t_{Lc}^1 = t_{Lc}^*$ can be rewritten as:

$$4\psi_{Ll} \leq \theta \Delta\delta - \{t_{Ll}^* - t_{Rl}^*\}^2/2. \quad (5.2)$$

If (5.2) holds, then asymmetric information has no bite relative to full information. As ‘political polarization’, $t_{Ll}^* - t_{Rl}^*$ increases, the opportunity cost to the extreme type from losing the election increases while an increase in ‘ideological polarization’, $\Delta\delta$, increases $t_{Ll}^* - t_{Lc}^*$, thus imposing greater current losses on the extreme type from misrepresenting her type. Thus, political polarization increases the incentive of the extreme type to mimic the moderate type while ideological polarization reduces it. These comparative static effects are discussed more generally in Section 5.3.

Furthermore, an increase in rents reduces the likelihood that condition (5.2) will hold, because rents increase the incentive of the extreme type to hijack the moderate type’s platform. Consider the numerical example: $\theta = 2$, $t_{Ll}^* - t_{Rl}^* = 0.1$ and $\Delta\delta = 0.2$. In this case, condition (5.2) reduces to $\psi_{Ll} \leq 0.1$, hence, the solutions under asymmetric and full information are identical if political rents do not exceed 20% of the average income ($\mu = 0.5$), otherwise the two solutions are distinct.

5.1.1. First period incumbent belongs to the Right party

Recalling that the moderate Right party politician is a ‘liberal’ while the extreme type is a ‘conservative’, Fig. 4, which is analogous to Fig. 3, summarizes the equilibrium outcome and Lemma 7 states the formal result without proof. The proofs and the notation are strictly analogous to the case of the Left party incumbent.

Lemma 7. *If the incumbent belongs to the right party then ruling out dominated strategies, the moderate (liberal) type announces $t_{Rl}^S = \sup S_2^C$: $t_{Rl}^* \leq t_{Rl}^S$ while the extreme (conservative) type announces $t_{Rc}^S = t_{Rc}^*$.*

The only difference from the case of the Left party incumbent is that the moderate type in the Right party increases first period redistributive taxes relative to her MPTR.

depending on the bargaining power and the location of the factions in each party, relatively centrist candidates might be chosen (see Fig. 2). Secondly, the moderate types in each party move towards the middle to signal their type, which has no counterpart in the full information models. Furthermore, partial convergence is period and type contingent; it does not occur in the case of lame-duck or extreme type politicians. Partial convergence towards the middle ground, tilts redistributive policy towards that preferred by middle income voters and in this sense, a version of Director's Law holds, but it too is period and type contingent. This contrasts with the full-information and full-commitment literature, where Director's Law holds unconditionally; for example Dixit and Londregan (1998).

5.2.2. *Inter-term tax variability (political budget cycles)*

The extreme types in each party announce their MPTR in each period, while the moderate types distort redistributive policy away from their MPTR ($t_{RI}^* \leq t_{RI}^S$ and $t_{Lc}^S \leq t_{Lc}^*$) in the first period, but as lame duck politicians in the second period, revert back to their MPTR. The resulting pattern of inter-term variation in the redistributive tax rate (political budget cycles) is shown in Fig. 5. The pattern of political budget cycles is party-specific. It is optimal for the moderate type in the Left party to have a smaller government in her first period in office relative to the second, while the converse is true for the moderate type in the Right party.

Besley and Case (1995) use a fairly comprehensive data set on gubernatorial elections in 52 US states that had term limits, over the period 1950–1986. Their results show a pattern of party-specific political budget cycles that is identical to the one predicted in this paper, although the strength of the political budget cycle was stronger for Democrats (Left party) relative to the Republicans (Right party). In an interesting study, Shi and Svensson (2001) find that the magnitude of the political budget cycles depends on the share of informed voters in the electorate. This is also consistent with the predictions of this paper; there are no political budget cycles when voters have full information about the politician while such cycles exist under asymmetric information about the politician.

5.2.3. *Policy reversals*

While $t_{RI}^* \leq t_{RI}^S$ and $t_{Lc}^S \leq t_{Lc}^*$, there is no presumption that $t_{RI}^S < t_{Lc}^S$. Indeed, if $t_{Lc}^S < t_{RI}^S$ then a *policy reversal* takes place; 'conservative Left party politicians' could appear even more conservative than 'liberal Right party politicians'. The comparative static results in Section 5.3 show that policy reversals could arise due to a variety of reasons such as an increase in political rents, and increased 'political' and 'ideological' polarization. One is tempted to conjecture that in an appropriately modified model, if politicians could signal their ideological type through public debt, *policy reversals* might take the form of a liberal Right party politician incurring relatively more debt as compared to a conservative Left party politician; for an alternative explanation and theory see Persson and Svensson (1987).

5.2.4. Relation to other signaling models

In signaling models of government policy it is typical to assign the labels of distinct political parties, Left or Right, to the two incarnations of the same politician; see for instance the survey in Persson and Tabellini (1990). Notwithstanding this questionable interpretation, this framework also does not allow for a meaningful distinction between the candidate selection stage and the implementation of government policy. Furthermore, such interpretation can sometimes lead to unexpected comparative static effects. For instance, within the redistributive framework of this paper, if there were only one politician with two possible types, say, liberal and conservative, then separation would entail ‘policy divergence’ as the two types distance their policy away from the other in order to signal their type to the voters. Such ‘policy divergence’ seems counterfactual and contradicts the evidence on the pattern of party-specific political budget cycles and on Director’s Law. The preceding remarks also apply, at least to a degree, to the related paper by Rogoff (1990) whose focus is not on the central redistributive issues of this paper such as partial convergence, party-specific political budget cycles and the effect of inequality on the size of the government. Furthermore, his assumption of homogenous voters and the absence of political parties is perhaps less appropriate in a redistributive context.

5.3. Comparative static results

This subsection explores the comparative static effects of political rents, beliefs, ideological/political polarization, and the salience parameter, while taking as given the candidates selected by the two parties.

5.3.1. Political rents

An increase in political rents, by increasing the benefit of holding office, increases the incentive of the extreme type to masquerade as a moderate type. The latter counters this by announcing an even more moderate policy, that is too costly to mimic. Proposition 3 examines the resulting distortion in first period redistributive policy and consequently in inter-term tax variability.

Proposition 3. *An increase in ψ_{Ll} makes first period redistributive policy of the moderate type even more moderate, i.e. reduces t_{Lc}^S , but leaves t_{Lj}^* unchanged for $j = c, l$. This increases inter-term tax variability, $t_{Lc}^* - t_{Lc}^S$.*

Analogous results for a Right party incumbent are stated without proof in Corollary 1; recall that the moderate type in this case is a liberal politician.

Corollary 1. *If the first period incumbent belongs to the Right party, then an*

increase in ψ_{Rc} increases t_{Rl}^S , but leaves t_{Rk}^* unchanged, thus, inter-term tax variability, $t_{Rl}^S - t_{Rl}^*$ increases.

Under the assumptions of full-information and full-commitment, it has been shown that more electoralist parties converge more; see for example Alesina and Rosenthal (1995: 27). A similar result also holds under the converse assumptions, because Proposition 3 and Corollary 1 show that moderate politicians implement more centrist policies (greater convergence) in response to an increase in electoral benefits. Furthermore, as the magnitude of political rents increases, the moderate types might move in the direction of too much moderation, leading to ‘policy reversals’ among the two parties. The prediction about an increase in the magnitude of political budget cycles following an increase in rents, is also borne out by the empirical results in Shi and Svensson (2001).

5.3.2. Beliefs

An increase in ρ , the prior probability of a voter being of liberal type, alters the ‘swing voter support’ for each of the types; this alters the probability of election (see Lemma 2), as well as the *differential reelection probability* (see Eq. (3.2)). Differentiating Eq. (3.2) with respect to ρ :

$$\frac{\partial \Delta \Pi_L^2}{\partial \rho} = \left(\frac{\mu^2 \Delta \delta}{\theta} \right) \int_{w_l(t_{Ll}^*)}^{w_l(t_{Lc}^*)} F''(x) dx. \quad (5.3)$$

The sign of $\partial \Delta \Pi_L^2 / \partial \rho$ depends on the change in the relative ‘swing voter support’ for the two types. Since $w_l(t_{Ll}^*) < w_l(t_{Lc}^*)$, if $F''(x) > 0$ (resp. $F''(x) < 0$) in the interval $[w_l(t_{Ll}^*), w_l(t_{Lc}^*)]$, then $\partial \Delta \Pi_L^2 / \partial \rho$ is positive (resp. negative). The resulting change in $\Delta \Pi_L^2$ alters the incentive of the extreme type to masquerade her type; Proposition 4 formalizes the outcome.

Proposition 4. *Consider an increase in ρ . If $F'' > 0$ in the interval $[w_l(t_{Ll}^*), w_l(t_{Lc}^*)]$ then t_{Lc}^S decreases and inter-term tax variability, $t_{Lc}^* - t_{Lc}^S$, increases. However, if $F'' < 0$ in that interval, then t_{Lc}^S increases and inter-term tax variability, $t_{Lc}^* - t_{Lc}^S$, decreases.*

Proposition 4 shows that the effect of beliefs (about the unknown redistributive ideology of the voters) on redistribution depends on how the ‘swing voters’ are distributed in the population. Hence, societies in which there is a greater possibility of voters being liberal, might actually exhibit less liberal redistribution (if $F'' < 0$); a result that does not arise under full information about the politician’s ideological type.

5.3.3. Polarization: ideological and political

Suppose that ideological polarization ($\Delta\delta = \delta_l - \delta_c$) and political polarization ($t_{Lj}^* - t_{Rk}^*$) increase, respectively, on account of a decrease in δ_c and in t_{Rk}^* . The crucial effect is on the differential probability of reelection. Implicitly differentiate Eq. (3.2) with respect to $\Delta\delta$ and t_{Rk}^* , respectively, to get:

$$\frac{\partial \Delta \Pi_L^2}{\partial \Delta \delta} = \frac{\partial \Delta \Pi_L^2}{\partial (1/\delta_c)} = - \left(\frac{\mu^2}{\theta \delta_c^2} \right) \rho \int_{w_c(t_{Li}^*)}^{w_l(t_{Li}^*)} F''(x) dx \tag{5.4}$$

$$\frac{\partial \Delta \Pi_L^2}{\partial (1/t_{Rk}^*)} = \left(\frac{\mu^2}{2\theta(t_{Rk}^*)^2} \right) \left\{ \rho \int_{w_c(t_{Lc}^*)}^{w_l(t_{Lc}^*)} F''(x) dx + \{1 - \rho\} \int_{w_c(t_{Li}^*)}^{w_l(t_{Li}^*)} F''(x) dx \right\}. \tag{5.5}$$

The sign of the derivatives in (5.4) and (5.5) is completely determined by the shape of the distribution function (sign of $F''(x)$) in the domain of ‘swing voters’ $[w_c, w_l]$. The intuition is similar to the one in Section 5.3.2, however, while ideological polarization (as defined) only affects the ‘swing voters’ for the moderate type, ‘political polarization’ affects ‘swing voters’ for both types.

5.3.3.1. Ideological polarization

The change in $\Delta \Pi_L^2$ (see (5.4)) following an increase in ideological polarization, alters the incentive of the extreme type to misrepresent her type. If the Right party opponent is a conservative, then the incentive of the extreme type in the left party masquerade her type increases through a second effect. In that case, a fall in δ_c reduces t_{Rc}^* ; the resulting increase in political polarization, $t_{Li}^* - t_{Rc}^*$, increases the opportunity cost to the Left party extremist from losing the elections. If the Right party politician is a liberal, then the second effect is absent. Proposition 5 formalizes the equilibrium outcome.

Proposition 5. *Following an increase in $\Delta\delta$, if $F'' > 0$, in the interval $[w_c(t_{Lc}^*), w_l(t_{Lc}^*)]$ then t_{Lc}^S increases irrespective of the type of the Right party politician, but the affect on inter-term tax variability, $t_{Lc}^* - t_{Lc}^S$, is ambiguous. If the Right party politician is a liberal and $F'' < 0$, then t_{Lc}^S decreases and inter-term tax variability increases.*

5.3.3.2. Political polarization

The main primitives of political polarization are the relative bargaining powers of the two factions in each party (see Section 4.1), and there is some evidence that political polarization in the British context, alluded to earlier, has been falling. There are two effects of political polarization. Firstly, an increase in $t_{Lj}^* - t_{Rk}^*$ increases the opportunity cost to the (w_l, δ_l) politician of losing the election, thereby increasing her incentive to masquerade her type. The second effect arises

due to the change in $\Delta\Pi_L^2$ (see (5.5)); if $F''(x) > 0$ in the interval $[w_c, w_l]$ then $\Delta\Pi_L^2$ increases, complementing the first effect.

Proposition 6. *If $F'' > 0$ in the interval $[w_c(t_L^*), w_l(t_{Lc}^*)]$ then an increase in political polarization increases t_{Lc}^S and dampens inter-term tax variability $t_{Lc}^* - t_{Lc}^S$.*

Under full-information, since politicians announce their MPTR, an increase in political polarization polarizes tax policy to the same extent, i.e. $t_{Lj}^* - t_{Rk}^*$ increases, but there is no affect on political budget cycles. However, under asymmetric information, if $F''(x) > 0$, then an increase in political polarization reduces (1) the degree of partial convergence ($\therefore t_{Lc}^S$ increases), (2) the strength of Director’s Law and, (3) the magnitude of political budget cycles.

5.3.4. *The significance of societal attitudes to inequality*

Alesina et al. (2001) argue that due to differences in social mobility between Europeans and Americans, the former prefer greater equality relative to the latter. In very rough terms, this corresponds to a higher ‘salience parameter’, θ^{-1} , for the Europeans relative to the Americans. Without offering any general results for redistribution, Example 2 focuses on the special case of a uniform population distribution.

Example 2. Suppose that the population is distributed uniformly over $[0, 1]$, $\psi_{Ll} = 0$ and $t_{Rk}^* = t_{Rl}^*$. Using Eq. (3.2), $\Delta\Pi_L^2 = \Delta\delta/4\theta$, hence, $\partial\Delta\Pi_L^2/\partial\theta = (-\Delta\Pi_L^2/\theta) < 0$. Implicitly differentiating condition C2, gives

$$\frac{\partial t_{Lc}^S}{\partial\theta^{-1}} = \frac{1}{\theta^2} \left\{ \sqrt{\frac{\Delta\delta}{\theta}}(w_R - w_L) + \frac{(w_L - \mu)}{\mu^2} \right\}$$

while differentiating Eq. (2.3) gives $\partial t_{Lc}^*/\partial\theta^{-1} = \mu^2/\theta^2 (w_L - \mu)$. A sufficient condition for $\partial t_{Lc}^S/\partial\theta^{-1} \geq 0$ and $\partial t_{Lc}^*/\partial\theta^{-1} \geq 0$ is the empirically plausible condition that the wealth of the Left party politician be at least as great as the average, i.e. $\mu \leq w_L$. In that case, it can be checked that $\partial t_{Lc}^S/\partial\theta^{-1} > \partial t_{Lc}^*/\partial\theta^{-1} > 0$ and so inter-term tax variability dampens. Similar results hold for a Right party incumbent. These findings are summarized and restated in Result 2.

Result 2. For a uniform distribution, if $\mu \leq w_L$, then an increase in θ^{-1} increases equilibrium redistribution and dampens inter-term tax variability.

Under full-information, an increase in θ^{-1} always increases redistributive taxes if $\mu \leq w_L$. Example 2 replicates this result under asymmetric information, and so redistributive taxes in this simple example would be predicted to be higher in

Europe relative to the United States, however, political budget cycles are predicted to be more pronounced in the latter.

5.4. Pooling equilibrium

A separating equilibrium will not exist if the moderate type does not find it worthwhile to prevent the extreme type from imitating her redistributive policy. In terms of the notation introduced in Section 5.1, this occurs when $S_1 \subset S_2^C$. Denote the (identical) tax policy chosen in a pooling equilibrium by both types in the Left party as t_L^P and the probability of reelection of the Left party incumbent, in the second election, by $\Pi_L^2(t_L^P, t_{Rk}^*)$. Since the moderate type is relatively popular as compared to the extreme type, it is easy to check that $\Pi_L^2(t_{Ll}^*, t_{Rk}^*) < \Pi_L^2(t_L^P, t_{Rk}^*) < \Pi_L^2(t_{Lc}^*, t_{Rk}^*)$ where $\Pi_L^2(t_{Lc}^*, t_{Rk}^*)$ and $\Pi_L^2(t_{Ll}^*, t_{Rk}^*)$ are the reelection probabilities of a moderate and an extreme type, respectively, under a separating equilibrium. The proof is analogous to that in Section 2.5, hence it is omitted.

Definition 5. In a pooling equilibrium, the differential reelection probability is defined as $\Delta\Pi^P = \Pi_L^2(t_L^P, t_{Rk}^*) - \Pi_L^2(t_{Ll}^*, t_{Rk}^*)$.

Consistent with Bayes rule, on observing t_L^P voters do not revise their posterior beliefs, i.e. $\omega = \Pr(\delta = \delta_c | t = t_L^P) = \rho$ and zero otherwise. Conditions C3 and C4, are necessary for a pooling equilibrium:

$$\text{Type } (w_L, \delta_c): \Delta\Pi^P \{V_{Lc}(t_{Lc}^*) - V_{Lc}(t_{Rk}^*) + \psi_{Lc}\} \geq V_{Lc}(t_{Lc}^*) - V_{Lc}(t_L^P) \tag{C3}$$

$$\text{Type } (w_L, \delta_l): \Delta\Pi^P \{V_{Ll}(t_{Ll}^*) - V_{Ll}(t_{Rk}^*) + \psi_{Ll}\} \geq V_{Ll}(t_{Ll}^*) - V_{Ll}(t_L^P). \tag{C4}$$

In conditions C3 and C4, the RHS is the current loss from participating in a pooling equilibrium while the LHS is the expected future gain; in a successful pooling equilibrium, the latter is relatively larger. Condition C4 can be rewritten as:

$$\Gamma = \frac{V_{Ll}(t_{Ll}^*) - V_{Ll}(t_L^P)}{\Delta\Pi^P} - \{V_{Ll}(t_{Ll}^*) - V_{Ll}(t_{Rk}^*)\} \leq \psi_{Ll}.$$

Proposition 7 shows that the pooling equilibrium can be ruled out if rents, ψ_{Ll} , are appropriately bounded.

Proposition 7. *If $\Gamma \leq \psi_{Ll}$ then the pooling equilibrium survives the ‘intuitive criteria’. But if ψ_{Ll} is bounded above by Γ then the pooling equilibrium can be ruled out.*

5.5. Some extensions

5.5.1. Candidate selection under asymmetric information

In selecting the party candidate under asymmetric information, the factions in each party must have common beliefs about: (1) the reelection probabilities of all types in each of the following two periods, (2) the type of PBNE, separating or pooling, that is expected to prevail in the first period, and (3) the distortion in redistributive policy relative to the MPTR as a result of signaling. Although, the analysis of candidate selection is not as elegant as under full information, results similar to those in Proposition 2 can be shown to hold in this case as well.¹² In particular: (1) the selected candidate depends on the relative bargaining powers of the two factions in the party, and (2) without further restrictions, one cannot predict if a party chooses a more centrist candidate in response to a centrist candidate chosen by the other party.

5.5.2. Private information about the challenger

Introducing private information about the challenger does not alter the main characteristics of the solution. Under asymmetric information about the Right party challenger, the analysis goes through by redefining the critical voters w_c and w_l who constitute the endpoints of the interval of *swing voters*. Instead of their MPTR being $(t_{Lj}^* + t_R^*)/2$, redefine this to equal

$$\frac{(t_{Lj}^*)^2 - E[(t_R^*)^2]}{2\{t_{Lj}^* - E[t_R^*]\}}$$

where $E[t_R^*] = \rho t_{Rl}^* + (1 - \rho)t_{Rc}^*$. This recalculation only affects the probability of reelection but nevertheless $\Delta \Pi_L^2 > 0$ and $\Delta \Pi_R^2 > 0$ continue to hold. Hence, at the cost of greater complexity, one can still show that the results on partial-convergence, Director's Law and the pattern of political budget cycles are unaffected. However, the comparative static results become more complicated and need to take account of the effect on the redistributive policy of the Right party challenger. While this does not effect the result on rents, the other comparative static results are somewhat amended.¹³ Candidate selection issues also become more complicated because voters and the factions within parties now also need to guess the type of PBNE that will hold in the Right party should it win the election.

¹²Proofs available from the author on request.

¹³It is not difficult to check that results do not change at all, if ideological polarization $\Delta\delta$ and beliefs ρ are party-specific.

6. Inequality and redistribution

At an empirical level, the relation between inequality and the size of the government is mixed; see for instance the survey in Persson and Tabellini (2000). However, one basic strand continues to run through the theoretical literature on this relationship, namely, that redistributive policy is chosen in a median voter framework (MVF) under full-information; for example Meltzer and Richard (1981), Persson and Tabellini (1994, 2000). In the MVF, following changes to inequality, equilibrium redistributive policy changes as a result of the spatial downward movement in the position of the median voter, who chooses a more redistributive policy. Relative to the existing literature, the analysis in this section is conducted in a representative democracy framework (RDF) under asymmetric-information. Thus, the incentives and constraints faced by politicians play a central role. Since income inequality can be a contentious concept, its usage in the paper is defined as follows.

Definition 6. An increase in inequality is defined as a *mean preserving spread* (MPS) in the wealth distribution. Formally, it corresponds to an increase in the parameter σ in the distribution function $F(w, \sigma)$ from σ_1 to σ_2 .

If the MPS takes place in some ‘localized interval’ $[w_m, w_n]$ such that its intersection with the interval of swing voters $[w_c, w_l]$ is empty, then $\Delta\Pi_L^2$ and the conditions C1 and C2 are unaffected and so there is no change in equilibrium redistribution. Thus, in what follows, changes in income inequality will be ‘global changes’ in the sense that $[w_c, w_l] \subset [w_m, w_n]$. Letting subscripts denote the partial derivatives of $F(w, \sigma)$ and partially differentiating the expression for $\Delta\Pi_L^2$ in Eq. (3.3) with respect to σ :

$$\frac{\partial \Delta\Pi_L^2}{\partial \sigma} = \rho \int_{\sigma_1}^{\sigma_2} \int_{w_c(t_{Lc}^*)}^{w_l(t_{Lc}^*)} F_{12}(x, y) \, dx \, dy + \{1 - \rho\} \int_{\sigma_1}^{\sigma_2} \int_{w_c(t_{Li}^*)}^{w_l(t_{Li}^*)} F_{12}(x, y) \, dx \, dy. \tag{6.1}$$

Denote the *differential probability of election* conditional on σ by $\Delta\Pi_L^2(\sigma)$. Since $\sigma_2 > \sigma_1$, and $w_l > w_c$, Result 3 follows directly from Eq. (6.1).

Result 3. $\Delta\Pi_L^2(\sigma_2) - \Delta\Pi_L^2(\sigma_1)$ is positive (resp. negative) if $\forall w \in [w_c(t_{Li}^*), w_l(t_{Lc}^*)]$, $F_{12}(w; \sigma)$ is positive (resp. negative).

The intuition is obvious from the definition of $\Delta\Pi_{Lj}^2$ in Eq. (3.2). Following an increase in inequality, if $F_1(w, \sigma)$ increases in the interval $[w_c, w_l]$, then the moderate type receives relatively greater ‘swing support’ while changes in the ‘core support’ in the interval $[w, w_c(t_{Li}^*)]$ are common to both types and so cancel out, hence $\Delta\Pi_{Lj}^2$ increases. The converse takes place if $F_1(w, \sigma)$ decreases. Fig. 6

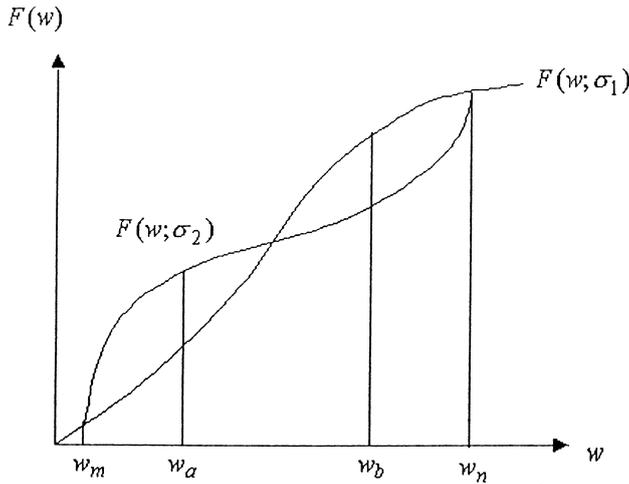


Fig. 6. Income inequality and redistribution.

illustrates the various cases following a MPS. In the interval $[w_c, w_l] \subset (w_a, w_b)$, $F_{12} < 0$, thus, using Result 3 it follows that $\Delta \Pi_L^2(\sigma_2) < \Delta \Pi_L^2(\sigma_1)$. Analogously, in the interval $[w_c, w_l] \subset [w_m, w_n] / [w_a, w_b]$, $F_{12} > 0$, thus, $\Delta \Pi_L^2(\sigma_2) > \Delta \Pi_L^2(\sigma_1)$.

Proposition 8. *If $\forall w \in [w_c(t_{Ll}^*), w_l(t_{Lc}^*)]$, $F_{12}(w; \sigma)$ is positive (resp. negative) then inequality reduces (resp. increases) the size of the government, $\alpha = t_{Lc}^S \mu$, and inter-term tax variability $t_{Lc}^* - t_{Lc}^S$ increases (resp. decreases).*

Using Eq. (3.4) it can be checked that Proposition 8 also holds if the first period incumbent belongs to the Right party. Proposition 8 provides a simple condition ($F_{12}(w; \sigma) > 0$ in the domain of *swing voters*) for the size of the government to decrease in response to an increase in income inequality. This contrasts with the typical result derived in the MVF under full-information, namely, that inequality always increases the size of the government. Furthermore, if changes in inequality were to affect the representative positions within the factions of militant and opportunist party members, then the candidate selection model would predict changes to the selected candidate in each party; this remains an interesting issue for further research.

7. Conclusions

This paper reconsiders redistributive policy when voters and politicians have asymmetric information about the redistributive preferences of each other, and the

latter cannot make credible commitments. The politicians are selected endogenously using a simple model of candidate selection. The equilibrium redistributive outcome is influenced by political rents, political and ideological polarization, and the salience of the redistributive issue.

Some standard results in the full-information and the full-commitment literature continue to hold with qualifications that are politician and period-specific; these include partial convergence and Director's Law. However, other results such as *policy reversals* and *inter-term political budget cycles* do not have a counterpart in the earlier literature. Finally, the effect of inequality on the size of the government depends crucially on the incentives and constraints faced by politicians, unlike in the literature that directly appeals to the median voter framework.

There are several possible directions for future research, some of which have been partly explored in the literature. These include the roles of party reputations in sustaining particular redistributive policies and the interaction of redistributive choice with other policy variables. Although, the candidate selection model in this paper is fairly simple, it illustrates the difficulty in deriving general results. The selection of candidates in a general redistributive context, especially in response to the changing distribution of income and to social mobility remains a fascinating area for research.

Acknowledgements

For insightful comments and discussions, I am greatly indebted to Daniel Seidmann, Jack Mintz, Tim Barmby, Francesco Giovannoni, Craig Brett, Gareth Myles, Abhinay Muthoo, Mick Keen, Michael Smart, Anthony Shorrocks and Melvyn Coles. I am also grateful for the very useful comments made during presentations of the paper at Toronto, Essex, Exeter, Newcastle and Dundee.

Appendix A

Proof of Lemma 1

In the second election, a (w_i, δ_j) voter with MPTR t_{ij}^* votes Left over Right if $V_{ij}(t_{Lj}^*) > V_{ij}(t_{Rk}^*) \Leftrightarrow (t_{Lj}^* - t_{Rk}^*) \mu^2 \{t_{ij}^* - (t_{Lj}^* + t_{Rk}^*)/2\} > 0$. Hence, the voting choice depends on the relative magnitudes of t_{ij}^* and $(t_{Lj}^* + t_{Rk}^*)/2$. Denoting by $V_{cc}(t)$ and $V_{il}(t)$, the respective utility functions of critical voters $[w_c, \delta_c]$ and $[w_l, \delta_l]$ (see Eq. (3.1)) check that (1) $V_{cc}(t_{Lj}^*) = V_{cc}(t_{Rk}^*)$, and $V_{il}(t_{Lj}^*) = V_{il}(t_{Rk}^*)$. (2) For all voters $[w_i, \delta_j]$, $j = c, l$: $w_i < w_c$, $V_{ij}(t_{Lj}^*) > V_{ij}(t_{Rk}^*)$. (3) For all voters $[w_i, \delta_j]$, $j = c, l$: $w_i > w_l$, $V_{ij}(t_{Lj}^*) < V_{ij}(t_{Rk}^*)$. (4) For all voters $[w_i, \delta_l]$: $w_i \in [w_c, w_l]$, $V_{il}(t_{Lj}^*) > V_{il}(t_{Rk}^*)$. (5) For

all voters $[w_i, \delta_c]$: $w_i \in [w_c, w_l]$, $V_{ic}(t_{Lj}^*) < V_{ic}(t_{Rk}^*)$. The derivation of these conditions completes the proof. \square

Proof of Lemma 2

In the interval $[w_c, w_l]$, the (unknown) ideology of the voters, which is decisive in their voting choice, follows a Bernoulli distribution ξ . With probability $0 \leq \rho \leq 1$, the voter is liberal and votes Left ('success' denoted by '1') and with probability $1 - \rho$, the voter is conservative and votes Right ('failure' denoted by '0'). Formally, $\xi(y|\rho) = \rho^y(1 - \rho)^{1-y}$, if $y = 0, 1$ and $\xi(y|\rho) = 0$ otherwise.

The random variable Y defined as the number of successes in n independent repetitions (because ideological types are independent across voters) of the Bernoulli experiment follows the Binomial distribution

$$B(y|n, \rho) = \binom{n}{y} \rho^y (1 - \rho)^{n-y}, \text{ if } y = 0, 1 \text{ and } B(y|n, \rho) = 0 \text{ otherwise.}$$

The mean and variance of this distribution are given, respectively, by $\varpi = n\rho$ and $\sigma^2 = n\rho(1 - \rho)$.

The proportion of voters who vote Left in the interval $[w_c, w_l]$ equals the proportion of liberal voters in this interval; this proportion is defined as $\lim_{n \rightarrow \infty} (Y/n)$ (because the interval $[w_c, w_l]$ is continuous). For any $\epsilon > 0$:

$$\Pr(|Y/n - \rho| \geq \epsilon) = \Pr(|Y - n\rho| \geq n\epsilon) = \Pr(|Y - \varpi| \geq \sigma k),$$

where $k = \epsilon \sqrt{n/\rho(1 - \rho)}$.

Using Chebyshev's inequality, $\Pr(|Y - \varpi| \geq \sigma k) \leq \rho(1 - \rho)/n\epsilon^2$, so $\Pr(|Y/n - \rho| \geq \epsilon) \leq \rho(1 - \rho)/n\epsilon^2$. Thus, it follows that $\lim_{n \rightarrow \infty} \Pr(|Y/n - \rho| \geq \epsilon) = 0$ and $\lim_{n \rightarrow \infty} \Pr(|Y/n - \rho| < \epsilon) = 1$. Therefore, the expected proportion of liberal voters in the interval $[w_c, w_l]$ equals ρ . Lemma 1 shows that all voters with $w < w_c$ vote Left and all voters with $w > w_l$ vote Right, hence $\Pi_{Lj}(t_{Lj}^*, t_{Rk}^*)$ equals $F(w_c(t_{Lj}^*, t_{Rk}^*)) + \rho\{F(w_l(t_{Lj}^*, t_{Rk}^*)) - F(w_c(t_{Lj}^*, t_{Rk}^*))\}$. \square

Proof of Proposition 1

Implicitly differentiating $\Pi_{Lj}^2(t_{Lj}^*, t_{Rk}^*)$ and using Eq. (3.1), the relevant comparative static results are:

$$\frac{\partial \Pi_{Lj}}{\partial \rho} = \int_{w_c(t_{Lj}^*, t_{Rk}^*)}^{w_l(t_{Lj}^*, t_{Rk}^*)} F'(x) dx > 0,$$

and

$$\frac{\partial \Pi_{Lj}^2}{\partial t_{Rk}} = \frac{\partial \Pi_{Lj}^2}{\partial t_{Lj}} = \left(\frac{-\mu^2}{2\theta} \right) \{ (1 - \rho)F'(w_c(t_{Lj}^*, t_{Rk}^*)) + \rho F'(w_l(t_{Lj}^*, t_{Rk}^*)) \} < 0.$$

The signs of these three derivatives complete the proof. □

Proof of Lemma 3

In the first election, a (w_i, δ_j) voter votes for the Left party over the Right if $V_{ij}(t_{Lj}^*) + \Pi_{Lj}^2 V_{ij}(t_{Lj}^*) + (1 - \Pi_{Lj}^2) V_{ij}(t_{Rk}^*) > V_{ij}(t_{Rk}^*) + (1 - \Pi_{Lj}^2) V_{ij}(t_{Rk}^*) + \Pi_{Lj}^2 V_{ij}(t_{Lj}^*)$, or equivalently, if $V_{ij}(t_{Lj}^*) > V_{ij}(t_{Rk}^*)$, which is identical to her voting decision in the second election. Hence, the voting decisions and the election probabilities are identical for all politicians in the two elections. □

Proof of Lemma 4

The first step is to specify the set \mathbf{X} of possible physical agreements, with generic element t_{Lj} , on the ‘desired’ redistributive tax rate; these simply belong to the interval $[0, 1]$. The disagreement outcome is t_{Rk}^* because if the two factions cannot agree, then the Right party candidate wins by default. Defining \mathfrak{R} as the set of real numbers, the player’s utility functions are $E[W_{Oj}]: \mathbf{X} \cup \{t_{Rk}^*\} \rightarrow \mathfrak{R}$, and $E[W_{Mj}]: \mathbf{X} \cup \{t_{Rk}^*\} \rightarrow \mathfrak{R}$. Define Ω as the set of utility pairs obtained through agreement, $\Omega = \{E[W_{Oj}], E[W_{Mj}]: t_{Lj} \in \mathbf{X}\}$. In order to derive the GNB solution, one needs two technical conditions;¹⁴ described as A1 and A2.

[A1] The Pareto frontier Ω^P of Ω is the graph of a concave function $E[W_{Mj}] = F(E[W_{Oj}])$ whose domain is a closed interval $\mathbf{I} \subset \mathfrak{R}$ and $\exists E[W_{Oj}] \in \mathbf{I}$ such that $E[W_{Oj}] > d_o$ and $F(E[W_{Oj}]) > d_m$.

[A2] The set Ω^W of weakly Pareto efficient utility pairs is closed.

By the definition of the MPTR and the properties of the utility function, it follows that in the intervals $[0, t_{Oj}^*)$, and $(t_{Mj}^*, 1]$ $E[W_{Oj}]$ and $E[W_{Mj}]$ are both respectively increasing and decreasing in t_{Lj} while over the interval $[t_{Oj}^*, t_{Mj}^*]$ $E[W_{Oj}]$ is decreasing while $E[W_{Mj}]$ is increasing in t_{Lj} . Furthermore, $[t_{Oj}^*, t_{Mj}^*]$ is a closed interval, hence condition A1 is satisfied if $t_{Lj} \in [t_{Oj}^*, t_{Mj}^*]$. It is straightforward to check that condition A2 is automatically satisfied. Hence the feasible interval for the purposes of Generalized Nash Bargaining is $T = [t_{Oj}^*, t_{Mj}^*]$. □

Proof of Proposition 2

The second order condition (soc) with respect to t_{Lj} is

$$\frac{\partial^2 \Lambda}{\partial (t_{Lj})^2} = \frac{\partial}{\partial t_{Lj}} \left[\frac{\partial \Pi'_{Lj}}{\Pi_{Lj}} \right] + \{\text{a negative term}\}.$$

¹⁴The relevant details and procedure follows Muthoo (1999) closely.

Thus, a sufficient condition for the soc to hold is $\partial/\partial t_{Lj}[\partial\Pi'_{Lj}/\Pi_{Lj}] \leq 0$ which requires that $\Pi''_{Lj} \leq 0$. Implicitly differentiate Eq. (4.1) with respect to λ_o :

$$\frac{\partial t_{Lj}}{\partial \lambda_o} = \left(-\frac{\partial^2 \Lambda}{\partial (t_{Lj})^2} \right)^{-1} \left\{ \frac{2V'_{Oj}(t_{Lj})}{2V_{Oj}(t_{Lj}) - V_{Oj}(t_{Rk})} - \frac{2V'_{Mj}(t_{Lj})}{2V_{Mj}(t_{Lj}) - V_{mj}(t_{Rk})} \right\} < 0.$$

The sign follows because $V'_{Mj}(t_{Lj}) > 0$ and $V'_{Oj}(t_{Lj}) < 0$ when $t_{Lj} \in T$. Implicitly differentiate Eq. (4.1) with respect to t_{Rk} :

$$\begin{aligned} \frac{\partial t_{Lj}}{\partial t_{Rk}} = & \left(-\frac{\partial^2 \Lambda}{\partial (t_{Lj})^2} \right)^{-1} \left\{ \frac{\partial}{\partial t_{Rk}} \left(\frac{\partial \Pi'_{Lj}}{\Pi_{Lj}} \right) + \frac{2\lambda_M V'_{Mj}(t_{Lj}) V'_{Mj}(t_{Rk})}{\{2V_{Mj}(t_{Lj}) - V_{mj}(t_{Rk})\}^2} \right. \\ & \left. + \frac{2\lambda_o V'_{Oj}(t_{Lj}) V'_{Oj}(t_{Rk})}{\{2V_{Oj}(t_{Lj}) - V_{Oj}(t_{Rk})\}^2} \right\}. \end{aligned}$$

In the braces, the first term is negative

$$\left(\cdot \frac{\partial}{\partial t_{Rk}} \left(\frac{\partial \Pi_{Lj} / \partial t_{Rk}}{\Pi_{Lj}} \right) = \frac{\partial}{\partial t_{Lj}} \left(\frac{\partial \Pi_{Lj} / \partial t_{Lj}}{\Pi_{Lj}} \right) < 0 \right),$$

the second is positive ($\cdot: V'_{mj}(t_{Lj}) > 0$), while the third is negative ($\cdot: V'_{Oj}(t_{Lj}) < 0$). Hence, the overall sign of the derivative is ambiguous. However, if $\lambda_o \rightarrow 1 \Leftrightarrow \lambda_M \rightarrow 0$, then the second term disappears and the overall sign of the derivative is negative. \square

Proof of Lemma 5

Suppose to the contrary that condition C2 does not bind at some t^1_{Lc} but does so at some $t' \neq t^1_{Lc}$. By definition, $V_{Lj}(t') > V_{Lj}(t^1_{Lc})$, and hence $t^1_{Lc} < t' \leq t^*_{Lc} < t^*_{Ll}$ ($\cdot: V' > 0$). Thus, $V_{Lc}(t') > V_{Lc}(t)$ which ensures that (a) the participation constraint holds at t' , and (2) the expected payoff of the moderate type increases. This contradicts the original choice of t^1_{Lc} . \square

Proof of Lemma 6

The discussion following Eq. (5.1) has already shown that $t^S_{Lc} \leq t^*_{Lc}$. By definition, the extreme type cannot mimic any $t \in S_2$ and the payoff of the moderate type $V_{Lc}(t)$ is decreasing in $|t - t^*_{Lc}|$, thus $t^S_{Lc} = \inf S^C_2$ strictly dominates any $t^S_{Lc} < \inf S^C_2$ for the moderate type.¹⁵ \square

¹⁵If the extreme type chooses to mimick the moderate type when indifferent between mimicking and not mimicking, then the solution is $t^S_{Lc} = \inf S^C_2 - \epsilon$, where $\epsilon > 0$ is arbitrarily small.

Proof of Proposition 3

From Eq. (2.3), t_{Lj}^* is independent of ψ_{Lj} . Since condition C2 binds (see Lemma 5), implicitly differentiating it, one gets

$$\frac{\partial t_{Lc}^S}{\partial \psi_{Ll}} = \frac{-\Delta \Pi_L^2}{\mu^2 \{t_{Lj}^* - t_{Lc}^S\}} < 0.$$

Since t_{Lj}^* is unchanged but t_{Lc}^S decreases, thus, inter-term tax variability $t_{Lc}^* - t_{Lc}^S$ increases. \square

Proof of Proposition 4

Implicitly differentiating condition C2:

$$\frac{\partial t_{Lc}^S}{\partial \rho} = \frac{-(\partial \Delta \Pi_L^2 / \partial \rho) \{(\mu^2 / 2) \{t_{Ll}^* - t_{Rkj}^*\}^2 + \psi_{Ll}\}}{\mu^2 \{t_{Ll}^* - t_{Lc}^S\}}.$$

The claim about t_{Lc}^S now follows by using Eq. (5.3) while the claim on $t_{Lc}^* - t_{Lc}^S$ follows by noting that t_{Lc}^* is not affected. \square

Proof of Proposition 5

Implicitly differentiating condition C2 one gets:

$$\frac{\partial t_{Lc}^S}{\partial \Delta \delta} = \frac{(\partial \Delta \Pi_L^2 / \partial \Delta \delta) \{(\mu^2 / 2) \{t_{Ll}^* - t_{Rkj}^*\}^2 + \psi_{Ll}\} + (\partial t_{Rk}^* / \partial (1/\delta_c)) (\mu^2 / 2) \Delta \Pi_L^2 \{t_{Ll}^* - t_{Rkj}^*\}}{-\mu^2 \{t_{Ll}^* - t_{Lc}^S\}}.$$

The claim about t_{Lc}^S now follows by using Eq. (5.4) and by noting that $\partial t_{Rk}^* / \partial (1/\delta_c) < 0$, while $\partial t_{Rl}^* / \partial (1/\delta_c) = 0$. The affect on inter-term tax variability, $t_{Lc}^* - t_{Lc}^S$, now follows by noting that the fall in δ_c decreases t_{ic}^* , $i = L, R$ (using Eq. (2.3)). \square

Proof of Proposition 6

Implicitly differentiating condition C2:

$$\frac{\partial t_{Lc}^S}{\partial (1/t_{Rk}^*)} = \frac{(1/t_{Rk}^*)^2 \Delta \Pi_L^2 \{t_{Ll}^* - t_{Rkj}^*\} + (\partial \Delta \Pi_L^2 / \partial (1/t_{Rk}^*)) \{ \{t_{Ll}^* - t_{Rkj}^*\}^2 + (2\psi_{Ll} / \mu^2) \}}{2 \{t_{Ll}^* - t_{Lc}^S\}} > 0.$$

Eq. (5.5) shows that if $F'' \geq 0$ then $\partial \Delta \Pi_L^2 / \partial t_{Rk}^* \geq 0$. The claim about t_{Lc}^S now

follows directly, while the claim on $t_{Lc}^* - t_{Lc}^S$ follows because t_{Lc}^* is not affected. \square

Proof of Proposition 7

A very brief sketch of the proof follows; for more details see Cho and Kreps (1987). If $\Gamma \leq \psi_{Lj}$, then the ‘intuitive criteria’ does not rule out the pooling equilibrium because $S^1 \subset S_2^C$, hence there does not exist any tax policy in the complement of the set S^1 that ‘equilibrium dominates’ those in the interior of the set S^1 for a moderate incumbent. However, if $\psi_{Lj} \leq \Gamma$, then the failure of condition C4 to hold rules out the pooling equilibrium. \square

Proof of Proposition 8

Implicitly differentiating condition C2:

$$\frac{\partial t_{Lc}^S}{\partial \sigma} = \frac{-(\partial \Delta II_L^2 / \partial \sigma) \{(\mu^2/2)\{t_{Ll}^* - t_{Rk}^*\}^2 + \psi_{Lj}\}}{\mu^2 \{t_{Ll}^* - t_{Lc}^S\}}.$$

The sign of $(\partial t_{Lc}^S / \partial \sigma) = -$ sign of $(\partial \Delta II_L^2 / \partial \sigma)$. The proof now follows by using Result 3 and noting that t_{Lj}^* does not change. \square

References

- Alesina, A., Di Tella, R., MacCulloch, R., 2001. Inequality and happiness: are Europeans and Americans different?. NBER, Working paper W8198.
- Alesina, A., Rosenthal, H., 1995. Partisan Politics, Divided Government and the Economy. Cambridge University Press, New York.
- Atkinson, A.B., Stiglitz, J.E., 1980. Lectures on Public Economics. McGraw-Hill, London.
- Besley, T., Case, A., 1995. Does electoral accountability affect economic policy choices? Evidence from gubernatorial term limits. Quarterly Journal of Economics 110, 769–798.
- Boadway, R., Keen, M., 1999. Redistribution. In: Atkinson, A.B., Bourguignon, F. (Eds.), Handbook of Income Distribution. North Holland, New York.
- Calvert, R., 1985. Robustness of the multidimensional voting model: candidate’s motivations, uncertainty, and convergence. American Journal of Political Science 29, 69–95.
- Cho, I.-K., Kreps, D., 1987. Signaling games and stable equilibria. Quarterly Journal of Economics 102, 179–222.
- Crawford, V., Sobel, J., 1982. Strategic information transmission. Econometrica 50, 1431–1451.
- Cukierman, A., Spiegel, Y., 2001. When is the median voter paradigm a reasonable guide for policy choices in a representative democracy?. Tel Aviv University, Mimeo.
- Dalton, R.J., Wattenberg, M.P. (Eds.), 2000. Parties Without Partisans: Political Change In Advanced Industrial Democracies. Oxford University Press, Oxford.
- Dixit, A., Londregan, J., 1998. Ideology, tactics and efficiency in redistributive politics. Quarterly Journal of Economics 124, 497–529.

- Dixit, A., Londregan, J., 1995. Redistributive politics and economic efficiency. *American Political Science Review* 89, 856–866.
- Driver, S., Martell, L., 1998. *New Labour: Politics After Thatcherism*. Polity Press, Cambridge.
- Epstein, L.D., 1980. *Political Parties in Western Democracies*. Transaction Books, New Jersey.
- Lindbeck, A., Weibull, J., 1993. A model of political equilibrium in a representative democracy. *Journal of Public Economics* 51, 195–209.
- Lindbeck, A., Weibull, J., 1987. Balanced budget redistribution as the outcome of political competition. *Public Choice* 52, 273–297.
- Meltzer, A.H., Richard, S.F., 1981. A rational theory of the size of the government. *Journal of Political Economy* 89, 914–927.
- Müller, W.C., 1999. *Policy, Office, or Votes*. Cambridge University Press, Cambridge.
- Muthoo, A., 1999. *Bargaining Theory With Applications*. Cambridge University Press, Cambridge.
- Persson, T., Tabellini, G., 2000. *Political Economics: Explaining Economic Policy*. MIT Press, Cambridge.
- Persson, T., Tabellini, G., 1994. Is inequality harmful to growth? *American Economic Review* 84, 600–621.
- Persson, T., Tabellini, G., 1990. *Macroeconomic Policy, Credibility and Politics*. Harwood Academic, London.
- Persson, T., Svensson, L.E.O., 1987. Why a stubborn conservative would run a deficit: policy with time inconsistent preferences. *Quarterly Journal of Economics* 104, 325–345.
- Roemer, J.E., 1999. The democratic political economy of progressive income taxation. *Econometrica* 67, 1–19.
- Roemer, J.E., 1998. Why the poor do not expropriate the rich: an old argument in new garb. *Journal of Public Economics* 70, 399–424.
- Roberts, K.W.S., 1977. Voting over income tax schedules. *Journal of Public Economics* 8, 329–340.
- Rogoff, K., 1990. Equilibrium political budget cycles. *American Economic Review* 80, 21–36.
- Romer, T., 1975. Individual welfare, majority voting and the properties of a linear income tax. *Journal of Public Economics* 4, 163–168.
- Shi, M., Svensson, J., 2001. *Political budget cycles: do they differ between developed and developing countries*. Harvard University, Mimeo.
- Wittman, D., 1983. Candidate selection: a synthesis of alternatives. *American Political Science Review* 77, 142–157.