Electoral Competition and Partisan Policy Feedbacks*

Carlo Prato†

March 18, 2013

Abstract

We study dynamic electoral competition with *partisan policy feedbacks* (a situation in which a policy systematically affects the electorate's future political preferences) in a public finance environment. Two parties with diverging preferences over redistribution choose socially undesirable levels of public employment because employment status systematically affects a citizen's beliefs about redistribution. We provide an explicit microfoundation for this dynamic linkage, and investigate main determinants of the resulting distortions.

Keywords: Dynamic Electoral Competition, Public Employment, Policy Feedback, Constitutions.

JEL Classification: H00, D 72, D 78.

---

*I would like to thank David Austen-Smith, Eddie Dekel, Hulya Eraslan, Bard Harstad, Salvatore Nunnari, as well as seminar and conference participants at Northwestern, EUI, Harris, PET11, ESEM13, NYU, Caltech, Georgetown, Pompeu Fabra, IAE, and Princeton for their help and comments. All mistakes are mine.

†Politics Department, Princeton University, and Edmund A. Walsh School of Foreign Service, Georgetown University.

Introduction

Policies, together with their main social and economic effects, often have indirect long term consequences on the electorate's *political attitudes*. As E. E. Schattschneider put it, "new policies create a new politics." For example, privatizing public assets might influence voters' future attitude towards redistribution (Biais and Perotti, 2002), or the funding of religious education is likely to shape the electorate's future views on a number of social policy issues. The mechanism through which policies affect voters' future political attitudes is referred to as *policy feedback* (Pierson, 1996).

In this paper, we formally study *partisan policy feedbacks,* a situation where policies influence voters' future induced preferences *over politicians* (or parties). More specifically, we study how, in a democracy, political actors strategically choose socially undesirable policies with the goal of influencing the electorate's future political preferences in their favor. We provide an explicit informational microfoundation for partisan policy feedbacks, and study the main determinants of the resulting distortions in a public finance environment.[1]

The simplest instance of a partisan policy feedback is the *Curley effect*: a politician with a strong socio-ethnic affiliation influences migration flows with the goal of increasing the relative size of his own socio-ethnic group (Glaeser and Shleifer, 2005). This mechanism is named after[2] James M. Curley, four-terms mayor of Boston in the period 1914–1950. Curley, the son of an Irish immigrant, became

[1] We choose a public finance environment because it allows to relate our theory with the existing theoretical and empirical literature in straightforward way, but we believe that it is possible to extend this framework to other policy environments.

[2] Glaeser and Shleifer argue that the same type of mechanism can explain other political failures, leading to underdevelopment and conflict in racially divided polities (for example, Coleman Young's Detroit in the period 1973-1993, or Robert Mugabe's Zimbabwe in the period 1987-present).

famous for aggressively adopting various types of populist measures that favored the inflow into Boston of poor, Catholic, Irish immigrants and drove out of the city the wealthier, Protestant, Anglo-Saxon elite. By discouraging business and reducing the city's fiscal base, his policies hampered Boston's economy. Nevertheless, by systematically and continuously reshaping the electorate is his favor, Curley was able to build a long and successful[3] political career.

This paper shows that the same logic behind the Curley effect–the desire of politicians to shape the future electoral environment in their favor–can be much broader: policies can be used to manipulate migration flows not only across space, but also *across information sets and economic interests*.[4] Moreover, this is the first paper investigating the main institutional and non–institutional determinants of the policy distortions associated with partisan policy feedbacks.

We study a dynamic model of electoral competition with endogenous occupational choice, where office is associated with policy-making power over public employment (which determines public good provision) and redistribution. Two parties have diverging preferences over the latter dimension, and compete for power by committing to a platform on the public employment/public good dimension. We

[3]In addition to his four mayoral terms, Curley also served as Governor and (twice) as Congressman of Massachusetts.

[4]Similar explanations have been proposed for instances of socially undesirable privatization and and subsidies to home ownership: Biais and Perotti (2002) to explain the wave of privatizations that occurred in various European countries during the nineties: by increasing the median voters' relative income, right wing governments implemented these policies in order to shift voters' long term political attitudes in a conservative direction. Several columnists (e.g., Becker, Stolberg and Labaton in *The New York Times*, December 20, 2008) have linked the increase in the subsidy to home ownership in the last two decades in the United States with the goal of shifting the electorate in a conservative way. Ortalo-Magne' and Prat (2011) is a first attempt into jointly investigating the economic and political consequences of home ownership subsidies in a dynamic setting.

model partisan policy feedbacks by formalizing Popkin's (1991) idea of political information as a "by-product or everyday life." More specifically, we assume[5] that a citizen's employment status (public vs private sector) influences her information about the distribution of entrepreneurial productivities: due to their indirect exposure to the private sector (i.e., their noisier information), public sector workers systematically underestimate the social cost of redistribution.

Parties have then an incentive to manipulate public employment (and thus public good provision) in order to improve their long term electoral appeal even in a Downsian environment[6] where electoral competition "pushes" platforms towards the socially optimal level. The reason is that parties optimally trade-off current electoral strength for a better future electoral environment. In this paper, we describe this dynamic trade-off and study its main determinants.

The first contribution of the paper is to show how distortions arise due to the interaction between policy feedback and two key elements. First, political actors are *differentiated*, and second, political platforms are related to implemented policies in a systematic way. In this model, differentiation comes from the fact that parties are associated to different levels of redistribution.[7]) The systematic linkage between platforms and policies comes from a commitment assumption and a *constitution.*[8] Although we choose a purely informational microfoundation for tractability,[9] the

[5]As extensively discussed in the rest of the paper, this assumption is not, by any mean, the only possible way to generate these effects.

[6]Voting follows one of the standard version of the probabilistic voting model, pioneered by Lindbeck and Weibull (1987) and Dixit and Londregan (1995) and extensively applied in Persson and Tabellini's *Political Economics* (2002).

[7]Difference is preferences, albeit natural, is not the only source of differentiation: similar results can be obtained by assuming that political actors differ in their *ability* to implement policies (Krasa and Polborn, 2009).

[8]We define a constitution as a mapping from electoral outcomes into policy-making rights

[9]After describing the model, we discuss how such modeling choice allows for a much cleaner

policy feedback can also originate directly from changes voters' economic interests induced by policies.[10]

This highlights two fundamental differences between this paper and Glaeser and Shleifer (2005). First, in this paper politicians "spread" conflict[11] from a redistributive policy dimension to a common value dimension (public good provision), without exploiting an underlying conflict in the electorate: political representation can then generate a stark form of Pareto inefficiency *absent any underlying disagreement among voters*. Second, the dynamic link between policies and voters' electoral attitudes is explicitly microfounded, rather than hard wired into preferences.

The second contribution of the paper is to study how time preference, political persistence, and voters' ideological volatility affect parties' incentive to choose socially undesirable policies. More far sighted politicians and/or voters (and, more generally, more political persistence) increase inefficiency in public good provision. This finding contrasts with what most of the recent literature in dynamic political economy would suggest.[12] In particular, the idea that more patient voters can lead

---

analysis of the parties' dynamic trade-off.

[10]For example, home owners have a direct interest in public order, and therefore more likely to support a "law and order" candidate or, more generally, to be conservative (Pattie et al., 1995; Pratt, 1986).

[11]This mechanism is similar to Morelli and Van Weelden (2011), but comes from a completely different source.

[12]Several authors have identified short termism as a primary source of political failures, especially in contexts involving redistributive politics (Acemoglu and Robinson, 2000 and 2001; Dixit and Londregan, 1995; Kundu 2007). Recent contributions to this literature have explored political failures in the various policy areas: Battaglini and Coate (2008) on public debt; Azzimonti (2011) on excessive investment taxation; Besley and Persson (2010) on the development of fiscal capacity; Aidt and Dutta, (2007) on public investment; Acemoglu et al. (2009) on labor supply distortions induced by redistribution. In all these papers, the source of the distortions lies in the inability of

to larger inefficiencies in the political process is a peculiar feature of partisan policy feedbacks.

The third contribution of the paper is to study how the type of constitution influences partisan policy feedbacks. Building on Lijphart (1999), we compare a *majoritarian* constitution to a *consensual* constitution (where parties' influence over policies is proportional to their electoral strength). Majoritarian constitutions are associated to less desirable public good provision (because of the absence of the moderating effect of bargaining), and lower redistribution. Moreover, when the horizon is finite, majoritarian constitutions display lower political polarization (that is, platform divergence), while this difference vanishes as the time horizon approaches infinity, highlighting an interesting long term neutrality.

Finally, we analyze the effect of changes in income inequality (modeled as a flatter distribution of individual productivities): while higher inequality increases polarization (McCarty, et al., 2008), its effect on public good provision is ambiguous (platforms are more extreme, but on average there is less underprovision of public goods).

1 The model

Economic environment. A polity is composed of a unit-mass continuum of citizens, and lasts T periods. In every period, each citizen chooses a *sector* (private or public) and, if in the private sector, an *occupation* (entrepreneur or worker). We normalize to zero the income received by a worker. Citizen i, if becomes an entrepreneur, earns his productivity a_{it}, drawn from a uniform with mean zero and

the policy maker (or legislative proposer) to be dynamically consistent. Since future political power is uncertain and current payoffs are fully appropriable, having more persistence in political power and/or more far sighted political actors would mitigate these distortions.

density σ_t: $a_{it} \sim U_{[-1/2\sigma_t, 1/2\sigma_t]}$. Realizations of $\sigma_t \sim \mathbb{R}^+$ are iid. A citizen's sector and employment choices depend on government policies and his own productivity.

The government provides a public good and redistributes income from self employed to workers using a lump sum tax $\tau_t \in \mathbb{R}^+$. The public sector employs[13] a share $x_t \in [0, 1]$ of the population to produce a public good using a decreasing return technology: $g_t = g(x_t)$. For simplicity, we assume $g(x) = -(x - x^*)^2/2$. We interpret x^* as the point at which the workers' marginal product of labor is equalized across sectors.[14]

The lump sum tax, imposed on entrepreneurs, finances a transfer of per capita amount b_t for the workers. A citizen then chooses to become an entrepreneur if and only if $a_{it} - \tau_t \geq b_t$. As a consequence, b_t is implicitly defined by the zero[15] of

$$b \Pr(a \leq b + \tau) - \tau \Pr(a \geq b + \tau) \tag{1}$$

Payoffs are linear income and public good provision. Depending on his occupation, citizen i's per period indirect utility is then $v(x, \tau) = \max\{a_i - \tau, b\} + g(x)$. A social planner placing relative weight γ on high productivity citizens would choose (x, τ) to maximize

$$W(x, \tau) = g(x) + \gamma \int_{b+\tau}^{1/2\sigma} \sigma(a - \tau) da + b \int_{-1/2\sigma}^{b+\tau} \sigma da$$

Due to the additive separability of W, any social planner–including utilitarian ($\gamma = 1$) and rawlsian ($\gamma = 0$)–would choose $x_t = x^* \; \forall t$. x^* is then a natural benchmark for the political process. In this paper, we focus on two measures to quantify distortions:

[13] Since the wage is the same across sectors, the demand for labor from the public sector fully characterizes the labor market equilibrium. Moreover, under the assumption the size of the public sector will never exceed the number of available workers.

[14] In order to keep the problem well behaved, we assume that $x^* < 1/2$, the lower bound on the number of workers (generated by $\theta = 0$). Details in the Appendix.

[15] Under the assumption, it is easy to show that there is a unique positive solution to that problem.

- *public good inefficiency* $\mathbb{E}\{-g(.)\}$, the expected quadratic deviation from x^*;
- *platform divergence* $\Delta = x^R - x^L$, the difference between the parties' platforms (which, coherently with some proposed empirical measures,[16] captures political polarization).

Political process. In each period two parties[17] R and L compete for office, which is associated with policy-making power over x and τ. Political actors do not intrinsically care about x and differ in their preferred level of redistribution: R is utilitarian, and therefore opposes redistributive taxation, while L prefers some exogenously given level of redistributive taxation $\overline{\tau}$. In every period t, each party J commits to a policy platform x_t^J. On the other hand, no commitment is possible for τ_t: R, if alone in power, would then set $\tau_t = 0$ and L, if alone in power, would implement $\overline{\tau}$.

Political actors's payoff (normalized, in every period, within the unit interval) is linear in the distance from their preferred transfer level: R's per period payoff ranges between 1 (when $\tau_t = 0$ is implemented) and 0 (when $\tau_t = \overline{\tau}$ is implemented), while the opposite is true for L.

The main role of the no commitment assumption is to improve the model's tractability: in the Robustness Section we show that weakening it generates the same qualitative insights of the baseline model, but distortions are even larger.[18] Assuming that R and L have preferences over redistribution is, instead, essential:

[16] See, for example, McCarty and Shor (2011), who measure political polarization using data on candidates' pre-electoral commitments.

[17] Since we assume that parties are unitary actors, R and L can also be candidates.

[18] The assumption is also motivated by the idea that pre-electoral commitment to public good provision is easier than pre-electoral commitment to transfers: redistribution can be implemented through a large variety of means and is often delegated to lower level government officials, with larger agency problems.

removing it would eliminate party differentiation, and, as a consequence, partisan feedback effects. This assumption seems quite natural, given the large observed partisan effects on redistributive outcomes (Boix, 1998; Bradley et al., 2003). We simply assume that R and L differ in how they balance the trade-off between inequality and efficiency that virtually every polity faces.

Each voter i computes $\mathbb{E}_i\{v(x_t^J, \tau^J)\}$, the expected per period payoff associated with each party's x_t^J and τ^J. Although voters know $\tau^R = 0$ and $\tau^L = \overline{\tau}$, the payoff associated with $\overline{\tau}$ depends on σ_t. The latter, as we will see, can only be conjectured by some voters. Voting behavior is probabilistic: i votes for R iff $\mathbb{E}_i\{v(x_t^R, 0)\} > \mathbb{E}_i\{v(x_t^L, \overline{\tau})\} + \xi_t + \delta_t^i$, where ξ_t is the realization of a stationary zero-mean aggregate preference shock ξ, and δ_t^i is the realization of a stationary zero-mean idiosyncratic preference shock δ. As in one of the standard formulations of the probabilistic voting model,[19] ξ_t, δ_t^i are *iid* over time and drawn from uniform distributions with support, respectively $[-1/2\psi, 1/2\psi]$ and $[-1/2\varphi, 1/2\varphi]$.

Both shocks capture preferences on a vector of attributes[20] not explicitly modeled and assumed to be orthogonal to public good provision and redistribution. Examples of such attributes are positions on abortion or foreign policy, or personal charisma. ξ_t measures how much the median voter prefers the L-candidate over the R-candidate *abstracting from x and τ*. δ_t^i, instead, measures i's individual-specific deviation from the median bias.

Without knowing how electoral outcome maps into policies, it is hard to evaluate the assumptions on voting behavior. As it will become fully clear in the rest of the paper, they are arguably quite natural under each constitutional setting considered: a voter tries to "pull" the implemented policies in the direction that she

[19] See Lindbeck and Weibull (1987) and Persson and Tabellini (2002).

[20] More precisely, the payoff difference between L's and R's attributes.

expects to be beneficial for her.[21] Notice also that, since they only look at the current period's payoff, voters are myopic. This assumption will be relaxed.

Information, timing, and voting behavior. Citizens have an uninformative prior $U_{\mathbb{R}^+}$.[22] At the beginning of each period, private sector workers observe σ_t, while public sector workers, not being exposed to the large array of factors affecting the distribution of returns of entrepreneurial activity in the private sector (changes in technology, demand shocks, etc.), can only rely on a noisy signal $s_{it} \sim F_{s|\sigma}()$.

In every period t, the timing of the game is then as follows. (1) Depending on their initial sectorial allocation, citizens observe either σ_t or s_{it}. (2) Parties observe voters' information, and announce a platform x_t^J, $J \in \{R, L\}$. (3) Citizens compute $\{\mathbb{E}_i\{v(x_t^J, \tau^J)\}\}_{J\in\{R,L\}}$, are hit by preference shocks $\{\xi_t, \delta_t^i\}$, and cast their vote. (4) After the *constitution* maps votes into policy-making power, policies are implemented, voters make optimal sectorial and occupational choices upon observing a_{it} and b_t, and receive their payoff.

Denote by D_i i's expected payoff difference between Right and Left. Under the assumptions, R's total realized vote share are is given by $\hat{\pi}_t = 1/2 + \varphi\left[\int D_i di - \xi_t\right]$. Let $D(a)$ be the payoff difference between Right and Left for

[21]Moreover, voting behavior in this model is compatible with the assumption, often employed in political economy, that a voter acts as if she was pivotal.

[22]The uninformative prior greatly improves the model's tractability. In particular, it prevents uninformed voters from extrapolating information about σ_t from the parties' equilibrium platforms, which would make the model extremely complicated. An alternative approach would be assuming that (1) voters are constrained in their ability to process information; (2) the signal s_{it} reflects an optimal allocation of attention, and cannot be improved upon by looking at platforms; (3) the choice of s_{it} is not highly responsive to platforms. In such setting, parties will ignore the direct informational content of their platforms (which are electorally costly to distort) and the equilibrium will be analogous to the one analyzed in the paper.

a citizen with productivity a:

$$D_t(a) = g(x_t^R) - g(x_t^L) + \overline{\tau}\mathbb{I}_{\{a \geq b_t + \overline{\tau}\}} - b_t\mathbb{I}_{\{a \leq b_t + \overline{\tau}\}} + a\mathbb{I}_{\{a \in [0, b_t + \overline{\tau}]\}},$$

Averaging across all possible productivity levels, and using (1),[23] yields

$$D_i = g(x_t^R) - g(x_t^L) + \int_0^{b(\sigma_t)+\overline{\tau}} a\sigma_t da = g(x_t^R) - g(x_t^L) + \frac{\overline{\tau} - \mathbb{E}_i\{b(\sigma_t)\}}{4}$$

For a citizen i who was among the x_{t-1} citizens formerly employed in the public sector and observed a signal s, we have $\mathbb{E}_i\{b(\sigma_t)\} = \mathbb{E}\{b(\sigma_t)|s_{it}\}$. For a citizen formerly employed in the private sector, $\mathbb{E}_i\{b(\sigma_t)\} = b(\sigma_t)$. Aggregating over all citizens and signal realization yields

$$\hat{\pi}_t = \frac{1}{2} + \varphi\left[g(x_t^R) - g(x_t^L) + \overbrace{\underbrace{\eta_t}_{\text{Efficiency loss}} - \underbrace{\lambda_t x_{t-1}}_{\text{Informational wedge}}}^{\text{Electoral Environment}} - \underbrace{\xi_t}_{\text{Shock}}\right].$$

$\eta_t = (\overline{\tau} - b(\sigma_t))/4$, is the *efficiency loss* associated with L's redistributive taxation: the entrepreneurial income of those who, despite having positive productivity, under $\overline{\tau}$ choose to be workers. $\lambda_t = \int_{\Omega_{s|\sigma_t}} \mathbb{E}\{b(\sigma_t)|s\} f_{s|\sigma}(s)ds - b(\sigma_t)$, the *informational wedge*, is the marginal effect of a change in the initial size of the public sector on R's electoral strength. λ_t can be rewritten as $\mathbb{E}_{s|\sigma_t}\{b(\sigma_t)\} - b(\mathbb{E}_{s|\sigma_t}\{\sigma_t\})$. Using the fact that $b(.)$ is a strictly convex (proved it in the Appendix), Jensen's inequality implies $\lambda_t > 0$ and, as a consequence, the same is true for the parties' ex ante expectation of it, denoted by $\overline{\lambda} > 0$.[24] Public sector workers are, on average, systematically more favorable to redistribution, in line with the available evidence (Blais et al., 1990, Cusak et al., 2006; Guillaud, 2011). Public employment then generates a *partisan policy feedback.*

[23]More specifically, we use the fact that (1) can be rewritten as $(b-\tau)/2 - \sigma(b+\tau)(b+\tau) = 0$.

[24]We are assuming that both parties have a proper prior, which seems natural, given that their electoral chances depend on σ.

Since redistribution entails inefficiencies, the initial electoral environment (i.e., the $\tau-$related term in the aggregate vote share) is always favorable to R. But the extent of this advantage depends on the initial sectorial distribution. When choosing x_t^J (and facing at least another election in the future), political actors take into account the effect of x_t on future electoral outcomes. R has a larger electoral advantage over L when the initial size of the public sector (hence, the number of voters with noisy information on σ) is small.

Two important observations are in order. First, there might be other types of informational asymmetries that can generate a partisan policy feedbacks. For example, asymmetric information on the administrative cost of implementing redistribution, or "rational inattention" by public servants. By assuming noisy information on the distribution of productivities, we are capturing in a stylized and tractable way the idea that exposure to market forces influences the beliefs about redistribution in a systematic way (Jensen et al., 2009; Tepe, 2012).

Second, there might be other, non-informational sources of policy feedbacks. Citizens' actual economic interest might depend on their sector of employment (e.g., due to investment or depreciation of sector-specific human capital). Although the model can be modified to accommodate for that, focusing on an informational source improves the model's tractability and leads to a cleaner analysis: it allows us to abstract from policies' long term *economic* consequences, which we would have to disentangle from their *political* consequences (the focus of this paper).

In order to ensure continuity and differentiability in the objective functions, we also assume that φ and ψ are related in such a way that, for every initial value of x and realization of σ, both actors have a positive probability of obtaining a majority of the votes,[25]. Moreover, in order to ensure the consistency of proposed

[25]This assumption is essentially equivalent to the one in Persson and Tabellini (2002).

policies with future sectorial choices,[26] we assume that x^* is small enough. These assumptions, which are technical in nature, are described in the Appendix.

So far we have left unspecified the rule mapping electoral outcomes into policy-making power over x and τ. We call such mapping a *constitution*. Following the theoretical distinction introduced by Lijphart (1999) and already employed in formal political theory (see Ticchi and Vindigni, 2010, or Herrera and Morelli, 2010), in this paper we study *majoritarian* (denoted by **M)** and *consensual* (denoted by **C)** constitutions, formally defined below.

Majoritarian Constitution. The majority winner, W_t,[27] gets full policy-making power on both policy dimensions: he implements his announced platform x_t^J and his preferred τ^J. Denote by $X^M = x_t^W$ the implemented policy under majoritarian constitution. Parties' per period expected payoff is then the probability of winning a majority.

$$p_t = \Pr(\hat{\pi}_t \geq 1/2) = 1/2 + \psi[g(x_t^R) - g(x_t^L) + \eta_t - \lambda_t x_{t-1}]; \tag{2}$$

for R , $1 - p_t$ for L.

Consensual Constitution. Consensual (also known as consociational) democracy is based on the observation that in several countries (especially in northern and central Europe) constitutional rules effectively impose power sharing among different political actors. In an effort to balance adherence to the original definition[28] and

[26]That is, that there will always be at least enough workers to make the prescribed platforms feasible

[27]Formally, $W_t = R\mathbb{I}_{\{\hat{\pi}_t \geq 1/2\}} + L\mathbb{I}_{\{\hat{\pi}_t \leq 1/2\}}$.

[28]Lijphart (1980) describes the main features of a consensual democracy:

> [...] government by a grand coalition of the political leaders of all significant segments of the plural society. (...) The other three basic elements are (1) the mutual veto (...) (2) proportionality (...), and (3) a high degree of autonomy for each segment.

analytical tractability, we model consensual democracy as a stylized post-electoral bargaining game between R and L, where the two negotiate over x and τ with bargaining power proportional to their vote share. The default option is to bargain separately over each dimension, which implies that the following policies would be implemented: $X_t^C = \hat{\pi}_t x_t^R + (1-\hat{\pi}_t)x_t^L$, and $\tau_t^C = (1-\hat{\pi}_t)\overline{\tau}$.

If there is a nonempty set of Pareto improving pairs (x^{pr}, τ^{pr}) that would allow a randomly determined proposer to strictly increase his expected payoff with respect to the default option, he will choose his preferred pair within that set and the other will accept it. As a consequence, lacking a different agreement between R and L, the constitution prescribes that each party will have an influence on each policy dimension proportional to his electoral strength.[29]

LEMMA 1 *Under the assumptions, bargaining separately over each dimension has no Pareto improvement.*

As a consequence, R's per-period expected payoffs is

$$\pi_t = \mathbb{E}_t[\hat{\pi}_t] = 1/2 + \varphi[g(x_t^R) - g(x_t^L) + \eta_t - \lambda_t x_{t-1}]$$

while L's per period payoff is $(1-\pi_t)$.

Equilibrium concept. For finite T we use subgame perfect Nash, and for infinite T we restrict to equilibrium strategies that are differentiable, stationary, and Markov (Maskin and Tirole, 2001). As a result, players's strategies will be a pair of platform functions of the form $x^J : [0, \mu^h] \to [0,1] \quad j \in \{R, L\}$, which depend on the payoff-relevant state $\mu_t(\lambda_t, \eta_t, \sigma_t, x_{t-1}) = \eta_t - \lambda_t x_{t-1}$. μ_t is the average expected inefficiency of $\overline{\tau}$, and measures how favorable to R is the electoral environment.

[29]In Western democracies, it is possible to find several formal and informal mechanisms explicitly tying the number and type of cabinet positions to a party's vote share. For example, the so called *Cencelli manual*, used to distribute cabinet positions in pre-1994 Italy.

2 Majoritarian Constitution

In this section we characterize the equilibrium of the two-period model, and describe the intertemporal trade-off that parties always face, with minimal modifications depending on timing and constitutions, in this model.

In the two-period version of the **M**-game, the economy starts with an initial sectorial distribution, x_0, and the first elections take place at the end of $t = 0$. The following proposition describes the unique equilibrium of the game.

PROPOSITION 1 *In the unique equilibrium of the two-period **M**-game*
i) In $t = 2$ both platforms converge to the efficient level: $x_2^R = x_2^L = x^*$.
ii) In $t = 1$ parties commit to

$$x_1^R = x^* - \frac{\Delta_M}{2} - \frac{\psi\Delta_M}{1+\psi\Delta_M^2}\mu_1 \quad ; \quad x_1^L = x^* + \frac{\Delta_M}{2} - \frac{\psi\Delta_M}{1+\psi\Delta_M^2}\mu_1,$$

and platform divergence (denoted by Δ_M), solves

$$\Delta[1 + \psi\beta\overline{\lambda}\Delta] = \beta\overline{\lambda}/2. \tag{3}$$

In the last period we have the standard Downsian result that political competition drives both platforms to the efficient level. This is not true in the first period: in order to improve her future electoral environment, R commits to an inefficiently low x, L to an inefficiently high one. As a result, X_1^M, the implemented policy, is either inefficiently large or inefficiently small. Furthermore, since in equilibrium R is more likely to win the election has, in expectation X_1^M is below x^*.

The dynamic trade-off. In this model, political competition delivers a second best outcome because parties face a dynamic trade-off, which can be analyzed by decomposing the FONC for the optimal choice of x_1^R:

$$\underbrace{\frac{dp_1}{dx_1^R}}_{\text{Downsian}} + \underbrace{\beta\frac{dp_1}{dx_1^R}\mathbb{E}[p_2(\mu_2|x_1^R) - p_2(\mu_2|x_1^L)]}_{\text{Legacy}} - \underbrace{\beta p_1\frac{d\mathbb{E}[p_2(\mu_2|x_1^R)]}{dx_1^R}}_{\text{Curleyian}} = 0.$$

On the one hand, setting $x_1^R = x^*$ maximizes the chances of winning the upcoming elections, which is valuable for two reasons. First, it allows R to implement his favorite redistribution level today (*Downsian* component). Second, it maximizes the influence of R's platform on tomorrow's electoral environment (*Legacy* component). On the other hand, a marginal reduction in x_1^R, conditional on winning the election, increases R's future electoral chances (*Curleyian* component). L faces an analogous trade-off when choosing the upward distortion for his platform.

The point at which this trade-off is balanced generates distortions at both platform and implemented policy level. It is also worth pointing out that this description allows to distinguish the *benefit* of distorting a platform on future electoral environment (Curleyian) from the *influence* of the own platform on future electoral environment (Legacy): these two forces, both related to dynamic considerations, push political platform in opposite directions.

Comparative statics. The following proposition summarizes how the size of the distortions changes with the main parameters of the model.

PROPOSITION 2 *In the two-period **M**-game*
i) Platform divergence is increasing in political actors' discount factor (β), in the informational wedge ($\overline{\lambda}$), and in the volatility of the aggregate shock (ψ^{-1}).
ii) Public good inefficiency ($-\mathbb{E}\{g(X^M)\}$) is increasing in policy divergence and the initial state (μ_1), ambiguous in ψ^{-1}.

The parties' discount factor increases distortions: by increasing the relative importance of the future electoral environment, β increases the incentive to distort political platforms x. The intuition behind this result is analogous to Glaeser and Shleifer (2005). A larger $\overline{\lambda}$ has the same effect of a larger β: increasing the size of the informational wedge increases the electoral return of distorting the size of the public sector.

The current level of perceived inefficiency of redistribution (μ_1) does not affect platform divergence, but increases the expected policy distortion: by making R *ceteris paribus* more likely to win, it increases the relative weight of the Curleyian component in his intertemporal trade-off, and pushes implemented policy towards a more severe underprovision of public good.

The effect of aggregate ideological volatility (ψ^{-1}) on policy distortions is ambiguous: aggregate volatility makes voters more responsive to platforms, but also to information. In other words, it increases both the legacy and the Curleyian component of the dynamic trade-off: when μ_1 is small enough, the effect on the former dominates, and ideological volatility, by lowering voters' responsiveness, weakens accountability. When, instead, μ_1 is large enough, ideology can have an *insulating effect* and reduce policy distortion.

Forward looking voters. As already mentioned, an important assumption in the baseline model is that voters are myopic. One might then suspect that the interaction between far sighted politicians and fully myopic voters might play a key role in generating the distortions analyzed in this paper. It is possible to show that distortions not only persist, but are stronger than in the baseline specification when voters are far sighted.[30] In that case, at $t = 1$ voters also take into account the effect of implemented policies on their expected payoff at $t = 2$. In the second period platforms converge to x^*, so the only relevant effect is on implemented redistribution. Since implemented redistribution depends on $p_2(x_1)$, forward looking voters have an additional incentive to support R (the promise of lower future redistribution inefficiency):

PROPOSITION 3 *Public good inefficiency ($-\mathbb{E}\{g(X^M)\}$) increases in voters' dis-*

[30]Given our assumption on voters' prior, in order for them to be far sighted we need to assume that they receive some subjective signal about the distribution of A. See Appendix for details.

count factor.

Having more forward looking voters creates an additional channel through which initial sectorial allocation affects policies: since voters care also about the second period's inefficiency, the *ex ante* advantage of R increases (as if he was enjoying a larger μ_1), thereby decreasing both platforms.[31]

Infinite horizon model. The analysis of infinite horizon version of the **M** game confirms most of the insights of the two-period model, but also uncovers other important aspects. It also highlights the model's tractability: there is a unique MPE and can be characterized analytically.

Given its recursive structure, we denote by $x^R(\mu)$, $x^L(\mu)$ the equilibrium platform given an initial state $\mu = \eta - \lambda x$

PROPOSITION 4 *In the unique stationary differentiable MPE of the infinite horizon **M**-game, platform divergence is given by $\Delta_\infty = \beta\overline{\lambda}$, platforms are given by*

$$x^R(\mu) = x^* - \frac{\Delta_\infty}{2} - \frac{\psi\Delta_\infty}{1+\psi\Delta_\infty^2}\mu \quad ; \quad x^L(\mu) = x^* + \frac{\Delta_\infty}{2} - \frac{\psi\Delta_\infty}{1+\psi\Delta_\infty^2}\mu$$

COROLLARY 1 *i) Platform divergence is independent of aggregate volatility ψ^{-1}. ii) Platform divergence and expected inefficiency in X^M are larger than in the two-period model.*

Both parts of Corollary (1) are quite intuitive: as the time horizon increases to infinity, the marginal value of platform distortion increases. As a consequence, distortions on both platforms and implemented policy are higher than in the two-period model.

Far more surprising is that policy divergence is no longer dependent on the variance of the aggregate shock. As argued before, aggregate volatility only affects

[31] Simple inspection of equilibrium platforms also shows that having more far sighted voters does not affect political polarization.

platform divergence through the Legacy component (the incentive of to maximize the influence of the own platform on the future electoral environment). In the last period there is no Legacy component, so in the first period the Legacy component enters only one side of the political actors' dynamic trade-off. More generally, in every finite horizon model, the Legacy component has a stronger weight on the "current" part of the dynamic trade-off (reflecting a shorter sequence of future elections), rather than the "future" one, generating a non stationary sequence of incentives for political divergence across periods. In a stationary equilibrium, instead, the Legacy component has to affect current and future be constant over time. In the baseline setting of model, where payoffs are quadratic, the Legacy component effectively disappears from the dynamic trade-off.[32]

Keeping the intertemporal trade-off constant. To properly compare the distortions in the two-period model and in the infinite horizon model, one should control for the strength of the agents' intertemporal trade-off, rather than keeping constant the discount factor. To achieve that goal, we consider a pair $\left(\beta, \overline{\beta} = \sum_{t=1}^{\infty} \beta^t\right)$, and compare the equilibrium with $T = 2$ and $\overline{\beta}$ to the one with $T = \infty$ and β. The following proposition describes the comparison:

PROPOSITION 5 *If one keeps the intertemporal trade-off constant, platform divergence and expected policy distortion are lower in the infinite horizon model.*

When the intertemporal trade-off is the same across the two models, the only difference between two-period and infinite horizon is the expectation of future platform divergence. The location of these platforms with respect to x^* depends on the future electoral environment: the more the latter is favorable to R, the more inefficient his platform, and the closer L's one is to x^*. As a result, in the infinite horizon

[32]Analytically, this effect is captured by the term $\psi\beta\overline{\lambda}\Delta^2$ in (3); removing it yields precisely $\Delta = \beta\overline{\lambda}$.

model, future platform divergence exerts a *mitigating effect* on the impact of future electoral environment on electoral outcome, thereby decreasing the marginal gain from platform distortion with respect to a finite horizon model.

3 Consensual constitution

The following proposition describes the unique equilibrium of the two-period **C**-game.

PROPOSITION 6 *In the unique equilibrium of the two-period **C**-game*
i) In $t = 2$ both platforms converge to the efficient level: $x_2^R = x_2^L = x^*$.
ii) In $t = 1$ parties commit to

$$x_1^R = x^* - \frac{\Delta_C}{2} - \frac{\varphi\Delta_C}{1+\varphi\Delta_C^2}\mu_1 \quad ; \quad x_1^L = x^* + \frac{\Delta_C}{2} - \frac{\varphi\Delta_C}{1+\varphi\Delta_C^2}\mu_1,$$

and platform divergence (denoted by Δ_C), solves

$$\Delta[1+\varphi\beta\overline{\lambda}\Delta] - \beta\overline{\lambda} = 0 \tag{4}$$

iii) $\mathbb{E}\{-g(X^M)\}$ is increasing in aggregate volatility.

The key difference between consensual and majoritarian case is that, rather than only on aggregate volatility (ψ^{-1}), platforms depend also on idiosyncratic volatility (φ^{-1}): under a majoritarian constitution R and L only care about obtaining a majority, under a consensual democracy every vote has the same marginal effect on the future implemented policy, because it has the same effect on a party's bargaining power. As clearly illustrated by equilibrium platforms, idiosyncratic volatility in the **C**-game plays the has same role that aggregate volatility plays in the **M**-game. On the other hand, while the former has no effect on the equilibrium of the **M**-game, the latter also affects outcomes in the **C**-game. As the last part of Proposition (6), aggregate volatility increases public good inefficiency. The rest of the comparative static (in $\overline{\lambda}$, β, and σ^{-1}) is as in the **M**-game.

Infinite horizon model. As for **M**, let $x^R(\mu)$, $x^L(\mu)$ denote the equilibrium platform for a given state μ.

PROPOSITION 7 *In the unique stationary differentiable MPE of the infinite horizon **C**-game, platform divergence is given by $\Delta_\infty = \beta\overline{\lambda}$, platforms are given by*

$$x^R(\mu) = x^* - \frac{\Delta_\infty}{2} - \frac{\varphi\Delta_\infty}{1+\varphi\Delta_\infty^2}\mu \quad ; \quad x^L(\mu) = x^* + \frac{\Delta_\infty}{2} - \frac{\varphi\Delta_\infty}{1+\varphi\Delta_\infty^2}\mu$$

The key difference between two-period and infinite horizon is then that platform diverge no longer depends on the idiosyncratic volatility. The intuition is similar to the majoritarian case: in a stationary equilibrium, platform divergence is constant over time. As a consequence, the current and future Legacy components in the actors' dynamic trade-off offset each other. Like in the 2-period game, the idiosyncratic volatility (φ^{-1}) plays the same role that aggregate volatility plays in the **M**-game. The rest of the comparative static (effect of β, $\overline{\lambda}$, μ) is as in the **M**-game.

Comparing consensual and majoritarian. The following proposition compares the equilibria of the **C** and **M** games.

PROPOSITION 8 *With respect to a majoritarian constitution, a consensual constitution is associated with*

i) Larger platform divergence, but only when T is finite

ii) Smaller expected public good inefficiency, more public good provision, and more redistribution

These results echo several theoretical and empirical findings on two constitutional features that Lijphart explicitly associates with consensual democracy: parliamentarism and proportional electoral systems (see discussion in Section 4).

It is important to highlight that, when $T = \infty$, political polarization is independent of the constitution. This result is generated by the joint presence of two effects: (1) the volatility of the relevant preference shocks (aggregate shock for **M**, idiosyncratic shock for **C**) no longer affects the parties' dynamic trade-off; and (2) the mitigating effect of future polarization behind Proposition (5). An increase in T, when the latter is finite, has two effects: first, the incentive to manipulate platforms is higher, so polarization increases under each constitution; second, the compensating effect must be stronger under a consensual constitution, because of the higher platform divergence. In other words, while polarization increases under both constitutions, in the **C**-game the mitigating effect is *relatively stronger*. Polarization in a majoritarian democracy is then more reactive to changes in T. As T goes to infinity, constitutional difference disappear because the compensating effect has to be both stationary and independent on the relevant shock.

Semi-consensual constitution. **C** and **M** differ in the allocation of power over two dimensions: public good provision and redistribution. One might then conjecture that, since parties only care about redistribution, what drives differences across constitutions is the difference in the allocation of power over τ. In order to explore this conjecture, we consider a hybrid type of constitution, called semi-consensual (**S**), where the allocation of policy-making power over redistribution is majoritarian[33] and the one over public good provision is consensual.[34] In the Appendix, we show that in the semiconsensual constitution parties commit to the exact same levels of public good provision as in the consensual constitution. As a consequence,

[33]Formally, $\tau_t^S = \overline{\tau}\mathbb{I}_{\{LW_t=L\}}$.

[34]Formally, $X_t^S = \hat{\pi}_t x_t^R + (1-\hat{\pi}_t)x_t^L$.

the equilibrium is mostly driven by the constitutional allocation of policy-making power over public good provision, over which political actors have no preferences (but can credibly commit). Moreover, from a utilitarian perspective, **S** welfare dominates both **C** and **M** (also proved in the Appendix).

4 Empirical Implications

Constitutions and fiscal policy. Part (ii) of Proposition (8) is consistent with several empirical contributions on fiscal policy and constitutions. First, a large body of literature has documented how parliamentarism (Gerber and Ortuño Ortín, 1998; Persson et al., 2000; Persson and Tabellini, 2003 and 2004) and proportional representation (Austen-Smith and Banks, 198; Milesi-Ferretti et al., 2002; Persson and Tabellini, 2002 and 2004) seem to be associated with higher levels of redistribution and public good provision. Second, Ticchi and Vindigni (2010) and Iversen and Soskice (2006), find that consensual constitutions and proportional systems are associated with more redistribution because they tend to give left-wing parties more power.

The effect of Inequality. The model yields a few interesting implications on the relationship between *ex ante* inequality, political polarization, public good provision, and redistribution. In the context of this paper, an increase in inequality corresponds to a downward first order stochastic shift in $F(\sigma)$.

PROPOSITION 9 *Higher inequality increases platform divergence, has an ambiguous effect on public good inefficiency.*

When the distribution of σ shifts towards the origin both the informational wedge and the value of $b()$ increase: the reason is that (1) for given $\overline{\tau}$, less people with positive probability choose to become workers, and (2) the second derivative of $b()$ increases, and public sector workers' systematic overestimation of $b()$ gets larger. As a consequence, the overall size of the Right's initial electoral advantage shrinks.

The model suggests a positive relationship between inequality and political polarization, as documented by the work of McCarty, Poole, and Rosenthal (McCarty et al., 2006), among others. The suggested channel for that relationship is that, as inequality rises, so does the size of partisan feedback effect. Political actors have, then, a stronger incentive to distort their platforms and become more extreme, but political competition is also more balanced. Higher inequality is then associated with more volatile public good provision, but also with a weaker electoral advantage to the Right, which reduces the expected underprovision in public good.

5 Related literature

This paper is related to a large literature on the emergence and persistence of inefficient policies. Virtually all papers focus on dynamic commitment problems, and most of them are based on the presence of some underlying conflict in the society that directly generates these inefficiencies. In Acemoglu and Robinson (2001), Kundu (2007),[35] Glaeser and Shleifer (2005), a polity is exogenously divided into cleavages and political actors have incentive to manipulate their relative size to improve their electoral success. Inefficiencies arise because political actors can successfully exploit an existing conflict that feeds back into political preferences.[36] In

[35]Both paper develop the idea originally illustrated in Dixit and Londregan (1995).

[36]In Glaeser and Shleifer (2005) Irish workers are more likely to vote, for exogenous reasons, for an Irish candidate. In the other two papers, inefficient redistribution to agricultural workers arise because it is politically unfeasible (due to lack of dynamic commitment to future redistribution) for parties to induce a transition to a more efficient equilibrium, which requires existing workers to suffer a short term loss.

this paper, instead, inefficiencies arise because of the effect asymmetric information between public and private sector workers on their induced political preferences.

Like Acemoglu et al. (2011), this paper features a redistributive conflict feeds back into inefficiencies in some aspect of the public sector. Unlike this paper, in Acemoglu et al. (2011) (1) inefficiencies are observed in the composition of the bureaucracy (rather than its size), can only arise in a non-democratic context, and are generated by an underlying interest in the society (the rich) benefitting from these policies.

By studying the dynamics in public employment, output, and redistribution, this paper is connected to a dynamic public finance literature focusing on endogenous (Bai and Lagunoff, 2011; Krussell and Rios-Rull, 1999; Hassler et al. 2005) or exogenous (Acemoglu et al., 2009; Besley and Persson, 2010) changes in power and inefficient volatility in output and consumption. This literature shares with the present paper the idea that uncertainty over future allocation of political power constitutes an independent channel for political failures.

By studying how a dynamic linkage (the policy feedback) interacts with the political process generating distortions, our paper is also related to a literature on the political economy of investment and taxes (Besley and Coate, 1998; Azzimonti, 2011) and public debt (Aghion and Bolton, 1990; Milesi-Ferretti and Spolaore, 1994, Battaglini and Coate, 2008). Contrary to all these paper, political persistence does not mitigate distortions.

The idea that political actors can manipulate policies to improve their future electoral strength is also featured in Baron, Diermeier and Fong (2011), and Hodler et al. (2010). In the latter, in particular, different politicians are associated with different maps between policies and outcomes and, as a consequence, use policies to shifts salience across issues. Policy manipulation, then, creates an endogenous incumbency advantage that is similar to the one generates by partisan policy feedbacks in this paper (where technological asymmetries among politicians absent).

This paper also contributes to a large body of literature investigating the relationship between constitutional features and public finance outcomes, such as public good provision, transfers, government size (Persson and Tabellini, 2004; Persson et a., 2005; Lizzeri and Persico, 2001; Milesi-Ferretti, Perotti and Rostagno, 2002; Battaglini, 2013). With the exception of Battaglini (2013), where the dynamic linkage is purely economic (public debt), all the other papers feature static models. Moreover, our paper focuses on a different institutional comparison, as in Ticchi and Vindigni (2010).

Although the setting and the source of the intertemporal trade-off are very distant from our model, Kalandrakis (2009) is one of the few papers that shares with ours the idea that far-sighted politicians can be potentially detrimental for voters, due to a complementarity in current and future distortions.

6 Robustness

Term limits, general heterogeneity in the actors' intertemporal trade-off. The outcome under **C** is fragile to the symmetry in parties' intertemporal incentives, which can fail, for example, due to term limits or differences in discount factors.[37] R and L can improve their expected utility by splitting power asymmetrically across dimensions:[38] power over redistribution will go overwhelmingly towards the short sighted actor, and platform divergence will be larger. Since under a majoritarian (or

[37]Candidates typically differ in age and, more broadly, in the expected length of their political career. Analogously, parties are also typically ruled by different waves of top executives, who often belong to different generations.

[38]For example, when L is more far sighted than R, they find profitable to trade policy-making power over redistribution for policy-making power over public good provision.

hybrid) constitution the effect of such asymmetry is instead minimal,[39] Proposition 8 is sensitive to large heterogeneity in β.

Alternative specification of Consensual Democracy. Some authors (Battaglini, 2013; Herrera and Morelli, 2010) have modeled proportional representation (a prominent feature of consensual constitutions) as a setting where a party implements his platforms with a probability equal to his realized vote share ($\hat{\pi}_t$). Under that assumption, equilibrium strategies are exactly as in the baseline **C**-game, but X^C becomes a lottery between the two platforms (with probabilities $\hat{\pi}_t$ and $1-\hat{\pi}_t$). Since in a two-period model platform divergence is larger under consensual, one would expect Proposition 8.ii to be reversed. It is possible to show that[40] when the initial state μ_1 is large enough platform distortions are still larger under **M**.

Full commitment on redistribution. Assuming that political actors can fully commit to a certain redistribution level does not eliminate dynamic distortions. If one assumes that $\overline{\tau}$ is not too large, it is possible to show[41] that (1) in $t=2$, L will commit to $\tau^L=\overline{\tau}$; (2) if, in $t=1$, $\tau^L<\overline{\tau}$, the marginal incentive to distort platform for the left is even larger.[42]

[39] The reason is that under **M** (or **S**) parties cannot trade power across dimensions through bargaining. The only effect of asymmetric β is given by a different incentive to distort policies (larger for the far sighted side, lower for the short sighted side). Given the strict concavity of the actor's payoffs over x, the asymmetry actually reduces equilibrium platform divergence, thereby *improving* the outcome.

[40] The intuition is that, while platform divergence is larger, the lottery associated with X^C is less skewed towards x^R (π_t is smaller than p_t). Formally, $Var(X^C) > Var(X^M)$ and $E\{X^M\} < E\{X^C\} < x^*$. The negative effect of higher volatility is then counterbalanced by a lower underprovision. Details are available upon request.

[41] Details are available upon request.

[42] To see that, suppose that $\tau_1^L \sim 0$. Regardless of x_1^R and x_1^L, in $t=1$ R will then capture the whole surplus, and platforms for x will only be relevant for their effect on period 2's electoral environment. This will result in a larger incentive to distort public employment with respect to the

7 Conclusion

This paper studies dynamic electoral competition with partisan policy feedbacks: policies today influence the electorate's induced preferences over parties tomorrow. We study how politicians strategically choose, on a common value dimension, socially undesirable political platforms with the goals of exploiting these effects in their favor.

While previous literature has focused migration policy (Glaeser and Shleifer, 2005), we show how the same logic can be much more general, and arise in a relatively standard dynamic public finance setting with endogenous occupational choice. In this environment, employment status (public vs private sector) affects the precision of voters' information about productivity in the private sector and, as a result, their beliefs about the social consequences of redistribution. Since parties are assumed to associated with different redistributive taxation, the size of the public sector systematically affects the electorate's political preferences. Parties then choose the size of the public sector with the goal of increasing their future electoral strength, with undesirable consequences for public good provision.

We study the main determinants of the associated distortions (constitution, time preference, ideological volatility), and focus on two measures of distortions: platform divergence (capturing political polarization) and inefficiency in public good provision. Contrary to what recent literature on dynamic political failures would suggest, inefficiencies increase with politicians' *and voters'* patience and, more generally, with political persistence.

Majoritarian constitutions, with respect to consensual, display weakly lower platform divergence (strictly lower if the time horizon is finite) but more inefficient public good provision and, on average, a more severe underprovision of public

baseline model.

good. Interestingly, in an infinite horizon model, platform divergence does not depend on the type of constitution. In other words, political polarization induced by partisan policy feedback does not depend, in the long run, on institutional factors. Under a utilitarian welfare criterion, a hybrid constitution with an allocation of power that is consensual on public good provision a majoritarian on redistribution welfare-dominates both consensual and majoritarian democracy.

Finally, we study how income inequality affects the results from the model. Higher inequality (i.e., more dispersion in productivity) increases polarization (McCarty et al., 2008), but does not necessarily result in a more undesirable public good provision.

More work is needed to cast light on the presence of policy feedback effects on other important policy domains, such as subsidies to home ownership and agriculture, or the public funding of religious education. More important, more empirical work is necessary to exactly quantify the importance of these effects on policy-making.

Appendix A. Proofs

LEMMA 2 *The function $b(\sigma) : \mathbb{R}^+ \to \mathbb{R}^+$, defined as the zero of $F(b, \sigma) = (b - \tau)/2 + \sigma(b + \tau)^2$, is strictly decreasing and strictly convex in σ for $\tau \geq 1/2\sigma$ (constant otherwise).*

Proof. Using the implicit function theorem, $\partial b/\partial\sigma = -(b + \tau)^2 \mathbb{I}_{\{\overline{\tau} \leq 1/2\sigma\}}(1/2 + 2\sigma(b + \tau))^{-1} \leq 0$, and $\frac{\partial^2}{\partial^2 \sigma} \propto (-b'(b + \tau)(1 + 2\sigma) + 2(b + \tau)^3)\mathbb{I}_{\{\overline{\tau} \leq 1/2\sigma\}} \geq 0$. ■.

Ensuring well behaved objective functions. To ensure that the parties' objective function is continuous and differentiable, we need to guarantee that (1) the proposed x^J is always lower than the total pool of workers, and (2) both parties are competitive in every election. (1) is guaranteed by assuming that x^L is always smaller than the number of actual workers (bounded below by $1/2$). From Propositions 4 and 8, $x^L \leq x^* + \beta/4$. As a consequence, we assume

$$x^* + \beta/4 < 1/2. \tag{5}$$

(1) is guaranteed by assuming that the support of μ, φ and ψ, as standard in this type of models,[43] are such that the support of $\hat{\pi}$ always includes $1/2$. More formally, given $\hat{\pi}(\xi, \mu) = 1/2 + \varphi[g(x^R) - g(x^L) + \mu + \xi]$, we must have $\max_\xi \hat{\pi}(\xi, \mu) \in (1/2, 1)$, $\min_\xi \hat{\pi}(\xi, \mu) \in (0, 1/2)\ \forall\ \mu \in [0, \mu^h]$, $\forall\ d(x^R) - g(x^L) \in [-(1 - x^*), 1 - x^*]$.[44] The two conditions can be reexpressed as

$$\min\{1/\varphi - 1/\psi, 1/\psi\} > 2(1 - x^* + \mu^h) \tag{6}$$

Proof of Proposition 1. In $t = 2$ equilibrium policies solve $x_2^R \in \arg\max\{p_2(x_1)\}$, $x_2^L \in \arg\max\{1 - p_2(x_1)\}$, where $p_2(x_1) = 1/2 + \psi[d(x_2^R, x_2^L) + I_2 - \lambda_2 x_1]$, where $d(x_2^R, x_2^L) = g(x_2^R) - g(x_2^L)$. As a consequence, we must have $x_2^R = x_2^L = \arg\max d(x_2^R, x_2^L) = x^*$. x_1^R and x_1^L, instead, solve

$$\begin{cases} x_1^R \in \arg\max_{x\in[0,1]} \left\{p_1(x_0) + \beta p_1(x_0) p_2^*(x) + \beta(1 - p_1(x_0)) p_2^*(x_1^L)\right\} \\ x_1^L \in \arg\max_{x\in[0,1]} 1 + \beta - \left\{p_1(x_0) + \beta p_1(x_0) p_2^*(x) + \beta(1 - p_1(x_0)) p_2^*(x_1^R)\right\} \end{cases}$$

where $p_2^*(x) = \mathbb{E}_1[p_2(x)] = 1/2 + \psi\mathbb{E}\{\eta\} - \psi\overline{\lambda}x$ follows from the observation that $d(x_2^R, x_2^L) = 0$. The FONC are of the problem (which are also sufficient under the assumptions) define the following system

$$\begin{cases} \frac{d}{dx_1^R} p_1(x_0)[1 + \beta(p_2^*(x_1^R) - p_2^*(x_1^L)] + \beta p_1(x_0) \frac{d}{x_1^R} p_2^*(x_1^R) = 0 \\ \frac{d}{dx_1^L} p_1(x_0)[1 + \beta(p_2^*(x_1^R) - p_2^*(x_1^L)] + \beta(1 - p_1(x_0)) \frac{d}{x_1^L} p_2^*(x_1^L) = 0 \end{cases}$$

[43] See Persson and Tabellini (2002), Chapter 3.

[44] By (5), $x^* < 1/2$, which implies that $x = 1$ gives voters the lowest payoff.

summing the two equations yields the condition $\Delta[1+\psi\beta\overline{\lambda}\Delta] = \beta\overline{\lambda}/2$, subtracting the first from the second yields (after substituting for $2p_1(x_0)-1 = \psi\Delta(x^R+x^L) - 2\psi x^*\Delta + 2\psi\mu_1$ and (3) $2x^* - (x_1^R + x_1^L) = 2\psi\Delta_M\mu_1(1+\psi\Delta_M^2)^{-1}$, from which x_1^R and x_1^L are derived. ■

Proof of Proposition 2. For i), start observing that, since the RHS of the equation defining Δ is supermodular in $\beta\lambda$, the positive solution of that equation shifts to the right as $\beta\overline{\lambda}$ increases. ii) Substituting platforms into $p_1(x_0)$, and using the fact that we can rewrite $x_1^R = x^* - \Delta_M p_1(x_0)$, $x_1^R = x^* + \Delta_M(1 - p_1(x_0))$, $\mathbb{E}\{-g(X^M)\}$ simplifies to $\Delta_M^2\,[1/4 + 3(\psi\mu_1)^2(1+\psi\Delta_M^2)^{-2}]$, which is increasing in Δ_M^2 and μ_1. To see that it ambiguous in ψ, notice that $\frac{d}{d\psi}\mathbb{E}\{-g(X^M)\} = \frac{\partial\mathbb{E}\{-g(X^M)\}}{\partial\psi} + \frac{\partial\mathbb{E}\{g(X^M)\}}{\partial\Delta}\frac{d\Delta_M}{d\psi}$ is proportional to

$$3\mu_1^2\psi - \Delta_M(1+\psi\Delta_M^2)^2/4 - 3\mu_1\psi^2\Delta_M(1-\psi\Delta_M^2)(1+\psi\Delta_M^2)^{-1}$$

which, depending on the value of μ_1, can be positive or negative. ■

Proof of Proposition 3. Let β^v denotes the voters' discount factor and $\Phi_1 = \mathbb{E}_1\{\eta_2\lambda_2\} > 0$. R's realized vote share is then $1/2 + \varphi[g(x_1^R) - g(x_1^L) + \eta_1 - \lambda_1 x_0 + \beta^v\psi(x_1^L - x_1^R)\Phi_1 - \xi_1]$, where $\psi(x_1^L - x_1^R)\Phi_1$ is the expected change in the expected payoff in $t = 2$ when R, as opposed to L, wins the election in $t = 1$. Solving the model yields $x_1^R = x_1^{RM} - \psi\beta^v\Phi_1$ and $x_1^L = x_1^{LM} - \psi\beta^v\Phi_1$, where x_1^{RM} and x_1^{LM} are the equilibrium platform in Proposition (1). ■

Proof of Proposition 4. i) In Appendix B, Part 1, we define and compute (by guess and verify) a DSMPE. In part 2, we show the uniqueness of the associated value function by using a contraction argument, based on Theorem 4.2 in Jáskiewicz and Nowak (2006), and the game has the minimax property. In Part 3 we show by contradiction that the policy functions in Part 1 are unique.

Proof of Corollary 1. Part i) and the first part of ii) directly follow from inspecting

Δ_∞, the last part follows from

$$\Delta_M^2\left[1/4+3(\psi\mu_1)^2(1+\psi\Delta_M^2)^{-2}\right]>\Delta_\infty^2\left[1/4+3(\psi\mu_1)^2(1+\psi\Delta_\infty^2)^{-2}\right]$$

■

Proof of Proposition 5. Since $\overline{\beta}=\beta(1-\beta)^{-1}$, (3) can be rewritten as $\Delta_M[1-\beta+\psi\beta\overline{\lambda}\Delta_M]-\beta\overline{\lambda}=0$. Since $\Delta_\infty=\beta\overline{\lambda}$, if one proves that $\psi\overline{\lambda}\Delta_M<1$, then $\Delta_M>\Delta_\infty$ and, by proposition 2, $\mathbb{E}\{-g(X^M)\}$ is larger in the two period model. To see that $\psi\overline{\lambda}\Delta_M<1$, notice that, $p<1$ implies $\psi\mu<1/2$. Since $\mu=\eta-\overline{\lambda}x>0$, $\mu>\overline{\lambda}$ implies $\psi\overline{\lambda}<1$, which, combined with $\Delta_M<1$, yields the result. Since $\mathbb{E}\{-g(X^M)\}$ only depends on the discount factor through platform divergence. the rest of the proposition directly follows. ■

Proof of Proposition 6. i) In $t=2$ equilibrium policies solve $x_2^R\in\arg\max\pi_2(x_1)$, $x_2^L\in\arg\max\{1-\pi_2(x_1)\}$, where $\pi_2(x_1)=1/2+\varphi[g(x_2^R)-g(x_2^L)+\mu_2]$; the FONC of the problem define the solution. ii) x_1^R and x_1^L solve

$$\begin{cases} x_1^R\in\arg\max_{x\in[0,1]}\left\{\pi_1(x_0)+\beta\mathbb{E}\{\hat{\pi}_2(X^C)\}\right\}\\ x_1^L\in\arg\max_{x\in[0,1]}1+\beta-\left\{\pi_1(x_0)+\beta\mathbb{E}\{\hat{\pi}_2(X^C)\}\right\}\end{cases}$$

where $\mathbb{E}\{\hat{\pi}_2(X^C)\}=1/2+\varphi\mathbb{E}\{\eta_2\}-\varphi\overline{\lambda}\mathbb{E}\{X^C\}$, and $\mathbb{E}\{X^C\}=\pi_1(x_0)x_1^R+(1-\pi_1(x_0))x_1^L$. The FONC of the problem (which are also sufficient under the assumptions) define the following system

$$\begin{cases} \frac{d}{dx_1^R}\pi_1(x_0)[1+\beta\varphi\overline{\lambda}(x_1^R-x_1^L)]+\beta\varphi\overline{\lambda}\pi_1(x_0)=0\\ \frac{d}{dx_1^L}\pi_1(x_0)[1+\beta\varphi\overline{\lambda}(x_1^R-x_1^L)]+\beta\varphi\overline{\lambda}(1-\pi_1(x_0))=0\end{cases}\tag{7}$$

whose unique solution gives the equilibrium at $t=1$, using the same steps as in **(M)**.

iii) follows from the observation that, once platforms are fixed, the only randomness in the implemented policy is given by the realization of the aggregate shock, $\hat{\xi}_1$.

Since $X^C = x_t^L - \hat{\pi}_t \Delta_C$, $X^C \sim U_{[\underline{x}_C, \overline{x}_C]}$, where$[\underline{x}_C, \overline{x}_C] = \left[x_C^E - \frac{\varphi}{\psi}\frac{\Delta_C}{2}, x_C^E + \frac{\varphi}{\psi}\frac{\Delta_C}{2}\right]$ and $x_C^E = x^* - \frac{\varphi \Delta_C}{1+\varphi\Delta_C^2}\mu_1$. $\mathbb{E}\{-g(X^C)\}$ is then $\Delta_C^2\left[\varphi^2/(12\psi^2) + (2\varphi\mu_1)^2(1+\varphi\Delta_C^2)^{-2}\right]$, which is increasing in ψ^{-1} and Δ_C^2. ■

Proof of Proposition 7. In Appendix C, we show that the same steps used in the proof of Proposition **4** can be used to compute and verify uniqueness of the equilibrium described in the proposition.

Proof of Proposition 8 i) Notice that (6) implies that $\varphi^{-1} > \psi^{-1}$; these two parameters are the only difference between (3) and (4); therefore, the result follows from inspection of Propositions (**2**), (**4**), (**6**), and (**7**). ii) To see that $\mathbb{E}\{-g(X_1^M)\} > \mathbb{E}\{-g(X_1^C)\}$, notice that the difference can be rewritten as

$$\frac{\Delta_M^2}{4} - \frac{\Delta_C^2}{12}\left(\frac{\varphi}{\psi}\right)^2 - 4\mu_1^2\left(\frac{\Delta_C\varphi}{1+\varphi\Delta_C^2}\right)^2 + 3\mu_1^2\left(\frac{\Delta_M\psi}{1+\psi\Delta_M^2}\right)^2 \tag{8}$$

Since $\Delta_C\varphi/(1+\varphi\Delta_C^2)$ and $\Delta_M\psi/(1+\psi\Delta_M^2)$.can be rewritten as $((\Delta_C\varphi)^{-1} + \Delta_C)^{-1}$ and $((\Delta_M\psi)^{-1} + \Delta_M)^{-1}$, it is possible to conclude, using $\psi\Delta_M = 1 - \Delta_M/\beta\overline{\lambda} > \varphi\Delta_C = 1 - \Delta_C/\beta\overline{\lambda}$, that $\Delta_C\varphi/(1+\varphi\Delta_C^2) < \Delta_M\psi/(1+\psi\Delta_M^2)$. As a consequence, a sufficient condition for (8) is

$$\Delta_M^2 - \Delta_C^2\left(\varphi/\psi\right)^2/3 - 4(\Delta_C\varphi\mu_1)^2\left(1+\varphi\Delta_C^2\right)^{-2} > 0.$$

Multiplying each side by Δ_M^{-2} and using (6), we can derive the following lower bound for (8)

$$1 - \left(\Delta_C\varphi/\Delta_M\right)^2\left(1/3\psi^2 + \left[\min\left\{1/\psi, 1/\varphi - 1/\psi\right\}\right]^2\right) > 0. \tag{9}$$

Case 1: $1/\psi < 1/\varphi - 1/\psi$. (9) simplifies to $1 - \left(\frac{\Delta_C\varphi}{\Delta_M\psi}\right)^2\frac{4}{3} > 0$, which, by the implicit function theorem, is strictly increasing. Using (6), an upper bound for ψ is $\varphi(1-2\varphi\mu^h)^{-1}$ Combining this with $1/\psi < 1/\varphi - 1/\psi$, one obtains $(1-2\varphi\mu^h) < 1/2$. Moreover, $\Delta_{M/}\Delta_C$ can be re-expressed, using (3), (4), and

$\psi = \varphi(1-2\varphi\mu^h)^{-1}$, as

$$(1-\varphi\Delta_C^2)^{-1} - \varphi\Delta_M^2[(1-2\varphi\mu^h)(1-\varphi\Delta_C^2)]^{-1} \tag{10}$$

since $\Delta_M < \Delta_C$, the ratio must be below one, which implies $\Delta_C^2 < \Delta_M^2/(1-2\varphi\mu^h)$. Combining the latter inequality with $(1-2\varphi\mu^h) < 1/2$ yields $(\Delta_C\varphi/\Delta_M\psi)^2 < (1-2\varphi\mu^h) < 1/2$. Since $1-(1/2)(4/3) > 0$, (9) holds.

Case 2: $\frac{1}{\psi} > \frac{1}{\varphi} - \frac{1}{\psi}$.(9) becomes

$$1 - (\Delta_C\varphi/\Delta_M\psi)^2\left(4/3 + (\psi/\varphi)^2 + 2\psi/\varphi\right) > 0 \tag{11}$$

Direct inspection allows us to conclude that the expression is minimized when φ is smallest with respect to ψ. Combining this with the restriction $\frac{1}{\psi} > \frac{1}{\varphi} - \frac{1}{\psi}$, one obtains $2\varphi = \psi$. As a consequence $1-\left(\frac{\Delta_C\varphi}{\Delta_M\psi}\right)^2\frac{4}{3} > 0$ is a lower bound for the LHS of (11). Using (3), (4) and $2\varphi = \psi$, one obtains that Δ_M/Δ_C can be re-expressed as $\frac{1-2\varphi\Delta_M^2}{1-\varphi\Delta_C^2}$. Combining that and $\Delta_M/\Delta_C < 1$ yields $\left(\frac{\Delta_C}{\Delta_M}\right)^2 < 2$. As a consequence $\left(\frac{\Delta_C\varphi}{\Delta_M\psi}\right)^2 < \frac{1}{2}$ and, as before, we have that $1-(1/2)(4/3) > 0$ is a lower bound for (9). Since $\mathbb{E}\ -g(x)$ in the two-period model is lower under **C**, the same must also hold in the infinite horizon model, where platform divergence no longer depends on the constitution.

To see why redistribution is higher, notice that, regardless of the time horizon, $\mathbb{E}\{\tau_t^M\} = (1-p_t)\overline{\tau}$, $\mathbb{E}\{\tau_t^M\} = (1-\pi_t)\overline{\tau}$ and $\mathbb{E}\{X_t^M\} = x^* - (2p_t-1)\Delta_M$, $\mathbb{E}\{X_t^C\} = x^* - (2\pi_t-1)\Delta_C$. Since $\psi > \varphi$, $p_t = 1/2 + \mu_t\psi(1+\psi\Delta_M^2)^{-1} > \pi_t = 1/2 + \mu_t\varphi(1+\varphi\Delta_C^2)^{-1}$, and $(2p_t-1)\Delta_M > (2\pi_t-1)\Delta_C \Leftrightarrow \mu_t\psi\Delta_M(1+\psi\Delta_M^2)^{-1} > \mu_t\varphi\Delta_C(1+\varphi\Delta_C^2)^{-1}$, which implies $\mathbb{E}\{\tau_t^M\} < \mathbb{E}\{\tau_t^M\}$, $\mathbb{E}\{X_t^M\} < \mathbb{E}\{X_t^C\} < x^*$.
■

Proof of Proposition 9. The effect of higher inequality (lower σ) is to increase $b()$ (see Lemma 2) and, as a consequence, to decrease η_t. On the other hand, we have that $\partial^2 b/\partial^2\sigma = 3(b+\tau)^3\mathbb{I}_{\{\overline{\tau}\leq 1/2\sigma\}}(1/2+2\sigma(b+\tau))^{-2}$. The numerator is decreasing in σ, and the denominator is increasing, since $2(b+\tau)+2\sigma(\partial b/d\sigma) \propto 1/2+2\sigma(b+\tau)-\sigma(b+\tau)>0$. As a consequence, $\overline{\lambda}$ increases for two reasons: (1) higher $b()$, (2) higher b''(). The overall effect of a first order stochastic shift in the distribution of σ is to decrease $\mathbb{E}\{\mu\}$, thereby making electoral competition more balanced and platforms more extreme. ■

References

Acemoglu, D., Golosov, M. & Tsyvinski, A. (2011), 'Power fluctuations and political economy', *Journal of Economic Theory* **146**(3), 1009–1041.

Acemoglu, D. & Robinson, J. A. (2000), 'Political losers as a barrier to economic development', *The American Economic Review* **90**(2), pp. 126–130.

Acemoglu, D. & Robinson, J. A. (2001), 'Inefficient redistribution', *American Political Science Review* **95**(03), 649–661.

Acemoglu, D., Ticchi, D. & Vindigni, A. (2011), 'Emergence and persistence of inefficient states', *Journal of the European Economic Association* **9**(2), 177–208.

Aghion, P. & Bolton, P. (1990), Government domestic debt and the risk of default: A political-economic model of the strategic role of debt, DELTA Working Papers 90-11, DELTA (Ecole normale supérieure).

Aidt, T. S. & Dutta, J. (2007), 'Policy myopia and economic growth', *European Journal of Political Economy* **23**(3), 734–753.

Austen-Smith, D. & Banks, J. (1988), 'Elections, coalitions, and legislative outcomes', *American Political Science Review* **82**(2).

Azzimonti, M. (2011), 'Barriers to investment in polarized societies', *American Economic Review* **101**(5), 2182–2204.

Bai, J. H. & Lagunoff, R. (2011), 'On the faustian dynamics of policy and political power', *Review of Economic Studies* **78**(1), 17–48.

Baron, D. P., Diermeier, D. & Fong, P. (2012), 'A dynamic theory of parliamentary democracy', *Economic Theory* **49**, 703–738.

Battaglini, M. (2013), 'A dynamic theory of electoral competition', *Theoretical Economics* p. forthcoming.

Battaglini, M. & Coate, S. (2008), 'A dynamic theory of public spending, taxation, and debt', *American Economic Review* **98**(1), 201–36.

Becker, J., Stolberg, S. G. & Labaton, S. (2008), 'White house philosophy stoked mortgage bonfire', *The New York Times* .

Besley, T. & Coate, S. (1998), 'Sources of inefficiency in a representative democracy: A dynamic analysis', *American Economic Review* **88**(1), 139–56.

Besley, T. & Persson, T. (2010), 'State capacity, conflict, and development', *Econometrica* **78**(1).

Biais, B. & Perotti, E. (2002), 'Machiavellian privatization', *The American Economic Review* **92**, 240–258.

Blais, A., Blake, D. & Dion, S. (1990), 'The public/private sector cleavage in north america: The political behavior and attitudes of public sector employees', *Comparative Political Studies* **23**(3), 381–403.

Boix, C. (1998), *Political Parties, Growth and Equality: Conservative and Social Democratic Economic Strategies in the World Economy*, Cambridge University Press.

Bradley, D., Huber, E., Moller, S., Nielsen, F. & Stephens, J. D. (2003), 'Distribution and Redistribution in Postindustrial Democracies', *World Politics* **55**(2), 193–228.

Cusack, T., Iversen, T. & Rehm, P. (2006), 'Risks at work: The demand and supply sides of government redistribution', *Oxford Review of Economic Policy* **22**(3), 365–389.

Dixit, A. & Londregan, J. (1995), 'Redistributive politics and economic efficiency', *The American Political Science Review* **89**(4), 856–866.

Gerber, A. & Ortuño Ortìn, I. (1998), 'Political compromise and endogenous formation of coalitions', *Social Choice and Welfare* **15**(3).

Glaeser, E. L. & Shleifer, A. (2005), 'The curley effect: The economics of shaping the electorate', *Journal of Law, Economics and Organization* **21**(1), 1–19.

Hassler, J., Rodriguez-Mora, J. V., Storesletten, K. & Zilibotti, F. (2003), 'The survival of the welfare state', *American Economic Review* **93**(1), 87–112.

Herrera, H. & Morelli, M. (2010), Turnout and power sharing, mimeo.

Hodler, R., Loertscher, S. & Rohner, D. (2010), 'Inefficient policies and incumbency advantage', *Journal of Public Economics* **94**(9-10), 761–767.

Iversen, T. & Soskice, D. (2006), 'Electoral institutions and the politics of coalitions: Why some democracies redistribute more than others', *American Political Science Review* **100**, 165–181.

Jáskiewicz, A. & Nowak, A. S. (2006), 'Zero-sum ergodic stochastic games with feller transition probabilities', *Siam Journal on Control and Optimization* **45**, 773–789.

Jensen, J. L., Sum, P. E. & Flynn, D. T. (2009), 'Political orientations and behavior of public employees: A cross-national comparison', *Journal of Public Administration Research and Theory* **19**(4), 709–730.

Kalandrakis, T. (2009), 'A reputational theory of two party competition', *Quarterly Journal of Political Science* (4), 343378.

Krasa, S. & Polborn, M. (2009), Political competition between differentiated candidates, CESifo Working Paper Series 2560.

Krusell, P. & Rios-Rull, J.-V. (1999), 'On the size of u.s. government: Political economy in the neoclassical growth model', *American Economic Review* **89**(5), 1156–1181.

Kundu, T. (2007), Can democracy always lead to efficient economic transitions?, Esop working paper, Univeristy of Oslo.

Lijphart, A. (1980), *Democracy in Plural Societies: A Comparative Exploration*, Yale University Press, New Haven.

Lijphart, A. (1999), *Patterns of Democracy. Government Forms and Performance in Thirty-Six Countries*, Yale University Press, New Haven.

Lindbeck, A. & Weibull, J. W. (1987), 'Balanced-budget redistribution as the outcome of political competition', *Public Choice* **52**(3), 273–297.

Lizzeri, A. & Persico, N. (2001), 'The provision of public goods under alternative electoral incentives', *The American Economic Review* **91**(1), 225–239.

Maskin, E. & Tirole, J. (2001), 'Markov perfect equilibrium: I. observable actions', *Journal of Economic Theory* **100**(2).

McCarty, N., Poole, K. T. & Rosenthal, H. (2008), *Polarized America: The Dance of Ideology and Unequal Riches*, The MIT Press.

Milesi-Ferretti, G. M., Perotti, R. & Rostagno, M. (2002), 'Electoral systems and public spending', *Quarterly Journal of Economics* **117**(2), 609–657.

Milesi-Ferretti, G. M. & Spolaore, E. (1994), 'How cynical can an incumbent be? strategic policy in a model of government spending', *Journal of Public Economics* **55**(1), 121–140.

Ortalo-Magné, F. & Prat, A. (2011), On the political economy of urban growth: Homeownership versus affordability, CEPR Discussion Papers 8243, C.E.P.R. Discussion Papers.

Persson, T., Roland, G. & Tabellini, G. (2000), 'Comparative politics and public finance', *Journal of Political Economy* **108**(6), 1121–1161.

Persson, T., Roland, G. & Tabellini, G. (2005), Electoral rules and government spending in parliamentary democracies, Levine's working paper archive, David K. Levine.

Persson, T. & Tabellini, G. (2002), *Political Economics: Explaining Economic Policy*, The MIT Press.

Persson, T. & Tabellini, G. (2003), *The Economic Effects of Constitutions (Munich Lectures)*, The MIT Press.

Persson, T. & Tabellini, G. (2004), 'Constitutional rules and fiscal policy outcomes', *American Economic Review* **94**(1), 25–45.

Pierson, P. (1996), ‘The new politics of the welfare state’, *World Politics* **48**(2), pp. 143–179.

Popkin, S. (1991), *The Reasoning Voter*, University of Chicago Press, Chicago.

Pratt, G. (1986), ‘Why canadians homeowners are more conservative than tenants’, *Urban Geography* **7**(3), pp. 187–209.

Schattschneider, E. E. (1935), *Politics, pressures, and the tariff*, Prentice-Hall.

Shor, B. & McCarty, N. (n.d.), ‘The ideological mapping of american legislatures’, *American Political Science Review* **105**.

Tepe, M. (2012), ‘The public/private sector cleavage revisited: the impact of government employment on political attitudes and behaviour in 11 west european countries’, *Public Administration* **90**(1), 230–261.

Ticchi, D. & Vindigni, A. (2010), ‘Endogenous constitutions’, *Economic Journal* **120**(543).

Appendix B. (For Online Publication)

Proof of Proposition 4

Since the game has a fixed total value $\bar{V} = (1-\beta)^{-1}$ for each player, the recursive problem solved by R and L can be written as:

$$\begin{aligned} V^R(\mu) &= \max_{x^R \in [0,1]} p(x^R, x^L, \mu)[1 + \beta(\mathbb{E}[V^R(\mu^R)] - \mathbb{E}[V^R(\mu^L)])] + \mathbb{E}[V^R(\mu^L)] \\ V^L(\mu) &= \max_{x^L \in [0,1]} \overline{V} - p(x^R, x^L, \mu)[1 + \beta(\mathbb{E}[V^R(\mu^R)] - \mathbb{E}[V^R(\mu^L)])] - \mathbb{E}[V^R(\mu^L)] \end{aligned}$$

where $\mu^R = \eta - \lambda x^R$ and $\mu^L = \eta - \lambda x^L$.

A differentiable stationary Markov perfect equilibrium (DSMPE) is a pair of differentiable value functions $V^R(\mu)$, $V^L(\mu)$ and differentiable policy functions $x^R(\mu)$, $x^R(\mu)$ such that (1) given $x^L = x^L(\mu)$, $V^R(\mu)$ solves the firs equation and, given $x^R = x^R(\mu)$, $V^L(\mu)$ solves the second equation; (2) $x^R(\mu)$ attains the RHS of the first equation and $x^L(\mu)$ attains the RHS of the second equation. To see that the two platforms constitute a DSMPE, start with two affine guesses of the form $h^R(\mu) = h_0^R + h_1\mu$, $h^L(\mu) = h_0^L + h_1\mu$ and plug them into the problem. In Part 1 we verify that the value functions are affine in μ and solve for the coefficients, in Parts 2 and 3 we show it is the unique DSMPE.

Part 1. A few lines of algebra allow to verify that $p(h^R(\mu), h^L(\mu), \mu)$ is affine in μ: $p(h^R(\mu), h^L(\mu), \mu) = \overline{p}(\mu) = h_p + h_p\mu$. Then the value functions can be re-expressed in the following way: $V^L(\mu) = (1-\beta)^{-1} - V^R(\mu)$, $V^R(\mu) = \overline{p}(\mu)(1 + \beta(\overline{p}(\mu'|h^R(\mu_0)) - \overline{p}(\mu'|h^R(\mu))) + \beta\overline{p}(\mu'|h^R(\mu)) + ...$, where $\overline{p}(\mu'|h^R(\mu)) = \mathbb{E}\{\overline{p}(\eta - \lambda' h^L(\mu))\} = \overline{p}(\eta - \overline{\lambda} h^L(\mu))$. Moreover, $\overline{p}(\mu'|h^R(\mu)) - \overline{p}(\mu'|h^R(\mu))$ does not depend on μ. Therefore, $\overline{p}(\mu)(1 + \beta(\overline{p}(\mu'|h^R(\mu)) - \overline{p}(\mu'|h^R(\mu)))$ is affine in μ and, for the same reason, all subsequent terms of the summation are also affine in μ. Denote by

V_1 the slope coefficient of V^R. The FONCs are (the equilibrium must be interior)

$$\begin{cases} \frac{dp^R}{dx^R}[1+\beta\overline{\lambda}V_1(h^L-x^R)] = \beta p^R V_1\overline{\lambda} \\ \frac{dp^L}{dx^L}[1+\beta\overline{\lambda}V_1(x^L-h^R)] = \beta(1-p^L)V_1\overline{\lambda} \end{cases}$$

where $p^R = p(x^R, h^L(\mu), \mu)$ and $p^L = p(h^R(\mu), x^L, \mu)$; the envelope conditions yield

$$V_1 = \frac{dp^R}{d\mu}[1+\beta\overline{\lambda}V_1(h^L-x^R)]-\beta(1-p^R)\overline{\lambda}V_1h_1 = \frac{dp^L}{d\mu}[1+\beta\overline{\lambda}V_1(x^L-h^R)]-\beta p^L\overline{\lambda}V_1h_1$$

re-expressing these 4 equations as functions of $h_0^L - h_0^R = \Delta$, $h_0^R + h_0^L$, and h_1 yields a unique solution, given in the proposition. To obtain the solution, impose equilibrium, which gives $p^R = p^L$, $h^R = x^R$, $h^L = x^L$, then solve for Δ by summing the two FONCs to get V_1^{-1} and equating the resulting expression to the V_1^{-1} obtained from each envelope condition. After that, sum the two first order conditions and obtain an equation in $h_0^R + h_0^L$, and h_1, which is affine in μ. Setting $\mu = 0$ gives $h_0^R + h_0^L = 2x^*$, which then allows to solve for h_1. The slope of the value function is then $V_1 = \psi(1-\psi(\beta\overline{\lambda})^2)^{-1}$.

Part 2. To show that this is the unique differentiable MPE, we first show that the value function must be unique by showing that the operator defining it is a contraction. Subsequently, we show that no pair of policy functions other than the one previously derived can generate the same value function. First, for every bounded, continuous, and increasing value function of the game $v(\mu)$ define

$$P[v](x^R, x^L, \mu) = p(x^R, x^L, \mu) + \beta\int_0^{\mu^h} v(\mu')f(\mu'|x^R, x^L, \mu)d\mu'$$

where, denoting by σ' and s_i' future values, $f(\mu'|x^R, x^L, \mu) = p(x^R, x^L, \mu)f(\sigma', \{s_i'\})$ if $\mu' = \eta' - \lambda' x^R$ $f(\mu'|x^R, x^L, \mu) = (1-p(x^R, x^L, \mu))f(\sigma', \{s_i'\})$ if $\mu' = \eta' - \lambda' x^L$, and zero otherwise.[45] Using the results from Part 1, we show that the **M**-game

[45] Note that, although the distribution of *implemented policies* is essentially a Bernoulli, the distribution of *realized states* has full support, due to the assumptions on $f(\sigma', \{s_i'\})$.

has the minimax property by showing that Theorem 4.2 in Jáskiewicz and Nowak (2006) holds: for their theorem to apply, we need that:

1. $p(x^L, x^R, \mu)$ and $f(\mu'|x^R, x^L, \mu)$ are continuous

2. There exists $U(\mu) : |p(x^L, x^R, \mu)| < U(\mu)\ \forall\ (x^R, x^L, \mu)$

3. The mapping $(x^L, x^R, \mu) \longmapsto \int_0^{\mu^h} U(\mu')f(\mu'|x^R, x^L, \mu)d\mu'$ is continuous

4. There exists a Borel function $\delta : [0,1]^2 \times [0, \mu^h] \longmapsto [0,1]$ and a probability measure $\phi(\mu)$ such that

i. $f(M|x^R, x^L, \mu) \geq \delta(x^R, x^L, \mu)\phi(M)\ \forall\ (x^R, x^L, \mu)$ and every Borel set $M \subset [0, \mu^h]$

ii. $\int_0^{\mu^h} \inf_{x^R\in[0,1]} \inf_{x^L\in[0,1]} \delta(x^R, x^L, \mu)\phi(\mu)d\mu > 0$

iii. $\phi(U) = \int_0^{\mu^h} U(\mu)\phi(d\mu) < \infty$

iv. For some $\rho \in (0,1)$ and for every (x^R, x^L, μ)

$$\int_0^{\mu^h} U(\mu')f(\mu'|x^R, x^L, \mu)d\mu' \leq \rho U(\mu) + \delta(x^R, x^L, \mu)\phi(U) \tag{12}$$

Define $\underline{f} = \inf_{(\mu', x^R, x^L, \mu)} f(\mu'|x^R, x^L, \mu)$[46] and choose $U(\mu) = 1$, $\delta(x^R, x^L, \mu) = \underline{f}p(x^R, x^L, \mu)$ and $\phi(\mu)$ uniform, so that $\phi = 1/\mu^h$. Then conditions 1-4.iii are trivially satisfied. To see why 4.iv must also hold, notice that (12) becomes $U \leq \rho + \underline{f}p(x^R, x^L, \mu)/\mu^h$. Since, under the assumptions, $p(x^R, x^L, \mu) > 0$ and $\mu^h = \max \int_0^{b(\sigma)+\overline{\tau}} a\sigma da$ is arbitrarily large, there exists $\bar{\rho} < 1$ that satisfies (12). Therefore, the following operator can be defined

$$Val(P[v]) = \max_{x^R\in[0,1]} \min_{x^L\in[0,1]} \{P[v](x^R, x^L, \mu)\} = \min_{x^L\in[0,1]} \max_{x^R\in[0,1]} \{P[v](x^R, x^L, \mu)\}$$

To show that the value function is unique, we make use of the following Lemma.

Lemma UB: Fix μ, then for every pair of bounded, continuous, and differentiable

[46]Under the assumptions, we know that, $\forall\ K = (x^R, x^L, \mu)$, $f(\mu'|K)$ has full support. As a consequence $\underline{f} > 0$.

$P[v^1](x^R, x^L; \mu)$ and $P[v^2](x^R, x^L; \mu)$ with support in $[0,1]^2$, we have:

$$|Val(P[v^1]) - Val(P[v^2])| \leq \max_{(x,y)\in[0,1]^2} |P[v^1](x,y) - P[v^2](x,y)|$$

Proof. Let (h^R, h^L) be a pair of policy functions generating an MPE of $P[v^1]$ and let $(\bar{h}^R, \bar{h}^L)$ be its analog for $P[v^2]$. Then it must be that $P[v^1](\bar{h}^R, h^L) \leq P[v^1](h^R, h^L) \leq P[v^1](h^R, \bar{h}^L)$ and $P[v^2](h^R, \bar{h}^L) \leq P[v^2](\bar{h}^R, \bar{h}^L) \leq P[v^2](\bar{h}^R, h^L)$, which implies

$$\begin{aligned}
P[v^1](h^R, h^L) - P[v^2](\bar{h}^R, \bar{h}^L) &\leq \max_{(x,y)\in[0,1]^2} |P[v^1](x,y) - P[v^2](x,y)| \\
P[v^2](\bar{h}^R, \bar{h}^L) - P[v^1](h^R, h^L) &\leq \max_{(x,y)\in[0,1]^2} |P[v^1](x,y) - P[v^2](x,y)|
\end{aligned}$$

■.

Now, let's define the operator T, mapping the space of bounded, continuous, and differentiable functions (with domain in $[0, \mu^h]$) into itself:

$$T[v](\mu) = Val\left(p(x^R, x^L, \mu) + \beta \int_{\mu^l}^{\mu^h} v(\mu') f(\mu' | x^R, x^L, \mu) d\mu'\right).$$

For every bounded, continuous $v, v' : [0,1] \to \mathbb{R}$, we have

$$\begin{aligned}
||T[v] - T[v']||_\infty &= \max_{\mu\in[0,\mu^h]} |Val(P[v]) - Val(P[v'])| \\
&\leq \max_{\mu\in[0,\mu^h]} \left\{ \max_{(x^R,x^L)\in[0,1]^2} |P[v^1](x^R, x^L) - P[v^2](x^R, x^L)| \right\} \\
&= \max_{\mu\in[0,\mu^h]} \left\{ \max_{(x^R,x^L)\in[0,1]^2} \beta \left| \int_0^{\mu^h} (v(\mu') - v'(\mu')) f(\mu' | x^R, x^L, \mu) d\mu' \right| \right\}
\end{aligned}$$

where the inequality follows from Lemma UB. Now define $\overline{D} = \max_{\mu'\in[0,\mu^h]} |v(\mu') - v'(\mu')|$. It must then be that

$$||T[v] - T[v']||_\infty \leq \max_{\mu\in[0,\mu^h]} \left\{ \max_{(x^R,x^L)\in[0,1]^2} \beta \overline{D} \left| \int_{\mu^l}^{\mu^h} f(\mu' | x^R, x^L, \mu) d\mu' \right| \right\}$$

since $\beta\overline{D}$ does not depend on x^R, x^L, and μ, we can move them to the left of the maximum operators. Since $\int_0^{\mu^h} f(\mu'|x^R, x^L, \mu)d\mu' = 1 \,\forall\, (x^R, x^L, \mu)$, one obtains

$$||T[v] - T[v']||_\infty \leq \beta\overline{D} = \beta \sup_{\mu' \in [0,\mu^h]} |v(\mu') - v'(\mu')|$$

This implies that T is a contraction, and the value function associated with the infinite horizon **M**-game is unique.

Part 3. To complete the proof, we need to show that no other pair of policy functions $(x^R(\mu), x^L(\mu))$ can generate the value function obtained in Part 1. To see that, combining the FONCs and the Envelope Conditions of the problem with the requirement that the value function is linear yields

$$\begin{cases} \psi(x^* - x^R)[1 + \beta\overline{\lambda}V_1(x^L - x^R)] = \beta p V_1\overline{\lambda} \\ \psi(x^L - x^*)[1 + \beta\overline{\lambda}V_1(x^L - x^R)] = \beta(1-p)V_1\overline{\lambda} \\ V_1 = \psi(1 + \frac{dx^L}{d\mu}(x^L - x^*))[1 + \beta\overline{\lambda}V_1(x^L - x^R)] - \beta\overline{\lambda}V_1(1-p)\frac{dx^L}{d\mu} \\ V_1 = \psi(1 + \frac{dx^R}{d\mu}(x^* - x^R))[1 + \beta\overline{\lambda}V_1(x^L - x^R)] - \beta\overline{\lambda}V_1 p\frac{dx^R}{d\mu} \end{cases}$$

substituting the first equation into the fourth and the second into the third gives, in both cases,

$$V_1 = \psi[1 + \beta\overline{\lambda}V_1(x^L - x^R)]. \tag{13}$$

The equation, in turn, implies that the difference $\Delta = x^L - x^R$ is independent of μ. We can then set $x^L = x^R + \Delta$. The difference between the FONCs then becomes $\psi(2x^* - \Delta - 2x^L(\mu))[1 + \beta\overline{\lambda}V_1\Delta] = \beta V_1\overline{\lambda}(2p-1)$. Substituting for the expression of p yields

$$\psi(2x^* - \Delta - 2x^L(\mu))[1 + \beta\overline{\lambda}V_1\Delta] = \beta V_1\overline{\lambda}\psi(\Delta(2x^R(\mu) + \Delta) - 2x^*\Delta + \mu).$$

Assuming that $x^R(\mu)$ it is linear leads to the equilibrium already derived in Part 1. As a consequence, one must rule out the existence of a positive non linear component in $x^R(\mu)$. Suppose, wlog, that $x^R(\mu) = x_0 + x_1\mu + x(\mu)$, where $x(\mu)$ is a continuous, differentiable and bounded non-linear function. For that equation to hold, it must be that the non-linear coefficients on each side must be equal. That means that the following equation must hold: $-2\psi[1+\beta\overline{\lambda}V_1\Delta] = \beta V_1\overline{\lambda}\psi\Delta 2$. This equation implies that $V_1 = -(2\beta\overline{\lambda}\Delta)^{-1}$.Combining this with (13) yields $V_1 = -\psi/2$. The value function associated with the equilibrium obtained in Part 1, instead, is $V_1 = \psi(1-\psi\beta\overline{\lambda}\Delta)$: that is a contradiction. As a consequence, the infinite horizon of the **M**-game must have only one DSMPE.

Proof of Proposition 7

The recursive formulation of the problem solved by R and L under **(C)** is given by:

$$\begin{aligned} V^R(\mu) &= \max_{x^R\in[0,1]} \pi(x^R, x^L, \mu) + \beta\mathbb{E}\{V^R(\mu)|X^C\} \\ V^L(\mu) &= \max_{x^L\in[0,1]} (1-\beta)^{-1} - \{\pi(x^R, x^L, \mu) + \beta\mathbb{E}\{V^R(\mu)|X^C\}\} \end{aligned} \tag{14}$$

where $\beta\mathbb{E}\{V^R(\mu)|X^C\} = \beta\mathbb{E}\{V^R(\eta - \lambda(\hat{\pi}x^R + (1-\hat{\pi})x^L))\}$. To see that the two platforms are a DSMPE, start with two affine guesses of the form $h^R(\mu) = h_0^R + h_1\mu$, $h^L(\mu) = h_0^L + h_1\mu$ and plug them into the problem. In Part 1 we verify that the value functions are affine in μ and solve for the coefficients, in Parts 2 and 3 we show it is the unique DSMPE.

Part 1. A few lines of algebra allow to verify that $\pi(h^R(\mu), h^L(\mu), \mu)$ is an affine function of μ: $\pi(h^R(\mu), h^L(\mu), \mu) = \overline{\pi}(\mu) = h_p + h_p\mu$, where the realized value of $\overline{\pi}(\mu)$ is $\pi(\mu) + \varphi\hat{\xi}$. Then the value functions can be re-expressed in the following way: $V^L(\mu_0) = (1-\beta)^{-1} - V^R(\mu_0)$, $V^R(\mu_0) = \overline{\pi}(\mu_0) + \beta\mathbb{E}\{\overline{\pi}(\mu_1)\} + \beta^2\beta\mathbb{E}\{\overline{\pi}(\mu_2)\} + ...$, where

$$\mathbb{E}\{\overline{\pi}(\mu_1)\} = \mathbb{E}\{\overline{\pi}(\eta - \lambda[(\overline{\pi}(\mu_{t-1}) + \varphi\hat{\xi})h^R(\mu_{t-1}) + (1-\overline{\pi}(\mu_{t-1}) - \varphi\hat{\xi})h^L(\mu_{t-1})])\}$$

simplifies to $\overline{\pi}(\{\eta - \overline{\lambda}[\overline{\pi}(\mu_{t-1})(h_0^R - h_0^L) + h_0^L + h_1\mu_{t-1}]\})$, which is affine in μ_{t-1} Therefore, all the terms in the summation are compositions of affine functions, therefore affine. Denote by V_1 the slope coefficient of V. The FONCs are (the equilibrium must be interior)

$$\begin{cases} \frac{d\pi^R}{dx^R}[1+\beta\overline{\lambda}V_1(h^L-x^R)] = \beta\pi^R V_1\overline{\lambda} \\ \frac{d\pi^L}{dx^L}[1+\beta\overline{\lambda}V_1(x^L-h^R)] = \beta(1-\pi^L)V_1\overline{\lambda} \end{cases} \tag{15}$$

where $\pi^R = \pi(x^R, h^L(\mu), \mu)$ and $\pi^L = \pi(h^R(\mu), x^L, \mu)$; the envelope conditions yield

$$V_1 = \frac{d\pi^R}{d\mu}[1+\beta\overline{\lambda}V_1(h^L-x^R)] - \beta(1-\pi^R)\overline{\lambda}V_1 h_1 = \frac{d\pi^L}{d\mu}[1+\beta\overline{\lambda}V_1(x^L-h^R)] - \beta\pi^L\overline{\lambda}V_1 h_1$$

re-expressing these 4 equations as functions of $h_0^R - h_0^L = \Delta_\infty$, $h_0^R + h_0^L$, and h_1 yields a unique solution, given in the proposition. To obtain the solution, impose equilibrium, which gives $\pi^R = \pi^L$, $h^R = x^R$, $h^L = x^L$, then solve for Δ_∞ summing the two FONCs to get V_1^{-1} and equating the resulting expression to the V_1^{-1} obtained from each envelope condition. After that, sum the two first order conditions and obtain an equation in $h_0^R + h_0^L$, and h_1, which is affine in μ. Setting $\mu = 0$ gives $h_0^R + h_0^L$, which then allows us to solve for h_1. The slope of the value function is then $V_1 = \varphi(1-\varphi(\beta\overline{\lambda})^2)^{-1}$.

Part 2. The proof for the uniqueness has the same structure as the one for the **M** game: for every bounded, continuous, and differentiable function $v(\mu)$ define

$$\Pi[v](x^R, x^L, \mu) = \pi(x^R, x^L, \mu) + \beta\int_0^{\mu^h} v(\mu')g(\mu'|x^R, x^L, \mu)d\mu'$$

where $g(\mu'|x^R, x^L, \mu) = f(A', s')\frac{\psi}{\varphi(x^L-x^R)}$. The assumptions in Jáskiewicz and Nowak (2006) are still satisfied (following same steps as in the proof in Proposition 4 (Part 2), choose $\delta(x^R, x^L, \mu) = \underline{f}\pi(x^R, x^L, \mu))$, and Lemma UB holds. As a consequence, following the same steps as in Proposition 4 (Part 2), the operator

$$T[v](\mu) = Val\left(\pi(x^R, x^L, \mu) + \beta\int_0^{\mu^h} v(\mu')g(\mu'|x^R, x^L, \mu)d\mu'\right) = Val(\Pi[v])$$

is a contraction: the value function in the infinite horizon **C**-game is also unique.

Part 3. It remains to show that no other pair of policy functions $(x^R(\mu), x^L(\mu))$ can generate the value function obtained in Part 1. To see that, combining the FONCs and the Envelope Conditions of the problem with the requirement that the value function is linear yields

$$\begin{cases} \varphi(x^* - x^R)[1 + \beta\overline{\lambda}V_1(x^L - x^R)] = \beta\pi V_1\overline{\lambda} \\ \varphi(x^L - x^*)[1 + \beta\overline{\lambda}V_1(x^L - x^R)] = \beta(1-\pi)V_1\overline{\lambda} \\ V_1 = \varphi(1 + \frac{dx^L}{d\mu}(x^L - x^*))[1 + \beta\overline{\lambda}V_1(x^L - x^R)] - \beta\overline{\lambda}V_1(1-\pi)\frac{dx^L}{d\mu} \\ V_1 = \varphi(1 + \frac{dx^R}{d\mu}(x^* - x^R))[1 + \beta\overline{\lambda}V_1(x^L - x^R)] - \beta\overline{\lambda}V_1\pi\frac{dx^R}{d\mu} \end{cases}$$

following the same method as in Proposition 3 allows us to establish that the difference $\Delta = x^L - x^R$ is independent of μ, and subsequently $V_1 = -\varphi/2$. Since the value function associated with the equilibrium obtained in Part 1 is $V_1 = \varphi(1 - \varphi\beta\overline{\lambda}\Delta_\infty)$, a contradiction is obtained. As a consequence, the infinite horizon of the **C**-game must have only one DSMPE.

Semi-consensual constitution

Under a semi-consensual constitution, the FONC of the problem define the solution. For $t = 1$ X_1^R and X_1^L solve

$$\begin{cases} X_1^R \in \arg\max_{x\in[0,1]} \left\{p_1(x_0) + \beta\mathbb{E}\{p(X_1^S)\}\right\} \\ \\ X_1^L \in \arg\max_{x\in[0,1]} 1 + \beta - \left\{p_1(x_0) + \beta\mathbb{E}\{p(X_1^S)\}\right\} \end{cases}$$

where$\mathbb{E}\{p(X_1^S)\} = 1/2 + \psi\mathbb{E}\{\eta\} + \psi\overline{\lambda}\mathbb{E}\{X_1^S\} = 1/2 + \psi\mathbb{E}\{\eta\} + \psi\overline{\lambda}[\pi_1 x_1^R + (1-\pi_1)x_1^L]$ follows from the observation that $g(x_2^R) - g(x_2^L) = 0$. The FONC of the problem (which are also sufficient under the assumptions) define the following system

$$\begin{cases} \frac{d}{dx_1^R}p_1(x_0) + \beta\psi\overline{\lambda}\left\{\pi_1 + (x_1^R - x_2^L)\frac{d\pi^R}{dx_1^R}\right\} = 0 \\ \frac{d}{dx_1^L}p_1(x_0) + \beta\psi\overline{\lambda}\left\{(1-\pi_1) + (x_1^R - x_2^L)\frac{d\pi^R}{dx_1^L}\right\} = 0 \end{cases}$$

which simplifies to

$$\begin{cases} x^* - x_1^R + \beta\overline{\lambda}\left\{\pi_1 + \Delta_S\varphi(x^* - x_1^R)\right\} = 0 \\ x_1^L - x^* + \beta\overline{\lambda}\left\{(1-\pi_1) + (x_1^R - x_2^L)\varphi(x_1^L - x^*)\right\} = 0 \end{cases}$$

which is the same system as in (7). To show that **S** welfare dominates both **C** and **M**, it suffices to show that redistribution is lower than under both (the reason is that we already know public good provision is as in **C**). First, observe that we must have, $\mathbb{E}\{-g(X^S)\} = \mathbb{E}\{-g(X^C)\} < \mathbb{E}\{-g(X^M)\}$. Next, we show that $\tau_1^C > \tau_1^M > \tau_1^S$. To see that the inequality must hold, notice that it is equivalent to $\tau_1(1 - p(x_M^R, x_M^L, \mu_1)) > \tau_1(1 - p(x_S^R, x_S^L, \mu_1))$, which follows from

$$\pi(x_C^R, x_C^L, \mu_1) < p(x_M^R, x_M^L, \mu_1) = \frac{1}{2} + \frac{\mu_1\psi}{1+\psi\Delta_M^2} < \frac{1}{2} + \frac{\mu_1\psi}{1+\varphi\Delta_C^2} = p(x_S^R, x_S^L, \mu_1).$$

To complete the proof, notice that, in $t = 2$, $\mathbb{E}\{-g(X^S)\} = \mathbb{E}\{-g(X^C)\} = \mathbb{E}\{-g(X^M)\} = 0$ and, since $\mathbb{E}\{X_1^M\} < \mathbb{E}\{X_1^C\}$

$$\mathbb{E}\{p_2\} = 1/2 + \psi\mathbb{E}\{\eta\} + \psi\overline{\lambda}\mathbb{E}\{X_1^M\} > E(\pi_2) = 1/2 + \varphi\mathbb{E}\{\eta\} + \varphi\overline{\lambda}\mathbb{E}\{X_1^C\}$$

$\mathbb{E}\{\tau_2^C\} > \mathbb{E}\{\tau_2^M\} > \mathbb{E}\{\tau_2^S\}$.