Electoral Competition and Partisan Policy Feedbacks

Carlo Prato†

March 1, 2013

Abstract

We study a dynamic model of electoral competition with endogenous occupational choice and partisan policy feedbacks (a situation where policies systematically affect the electorate’s future induced preferences over politicians). Two parties compete for power over redistribution and public employment/public good provision. While parties only have (diverging) preferences over redistribution, they manipulate public employment because of its effect on the voters’ beliefs about redistribution and, as a consequence, their political preferences. We provide an explicit microfoundation for this dynamic linkage and investigate main determinants of the resulting distortions. More forward looking voters or parties increase distortions. Consensual constitutions (as opposed to majoritarian) are associated with more platform divergence (but only when the horizon is finite), less inefficient public good provision, and more redistribution. A hybrid constitution can improve welfare over both.

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

JEL Classification: H00, D 72, D 78.

*I would like to thank David Austen-Smith, Marina Azzimonti, Marco Battaglini, Briana Chang, Eddie Dekel, Matthias Doepke, Georgy Egorov, Hulya Eraslan, Tim Feddersen, Simone Galperti, Edoardo Grillo, Bard Harstad, Igal Hendel, Daisuke Ikeda, Lorenzo Michelozzi, Salvatore Nunnari, Alvaro Parra, Bruno Strulovici, Martin Szydlowski, Stephane Wolton, Galina Zudenkova, as well as seminar and conference participants at Northwestern, EUI, Harris School, PET11, ESEM13, NYU, Caltech, Georgetown, Pompeu Fabra, IAE, for their help and comments. All mistakes are mine.

†Edmund A. Walsh School of Foreign Service, Georgetown University. Email: cp747@georgetown.edu.
Introduction

Together with their main social and economic effects, policies can often have indirect long-term consequences on the electorate’s political attitudes. As E. E. Schattschneider put it, “new policies create a new politics.” In other words, a policy today can have a systematic effect on the electorate’s political preferences. For example, the decision of privatizing public assets might influence voters’ future attitude towards redistribution (Biais and Perotti, 2002), or the decision to fund religious education is likely to influence the electorate’s future attitudes towards a number of public policy issues. The mechanism through which policies affect voters’ future political attitudes is called policy feedback effect (Pierson, 1996).

In this paper, we formally study a class of policy feedbacks where policies affect voters’ future induced preferences over politicians, to which we refer as partisan policy feedback. 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 how the resulting distortions are affected by the type of constitution and some key non-institutional variables (ex ante inequality, time preference).

The simplest instance of a partisan policy feedback is the Curley effect: a politician with a strong socio-ethnic affiliation has an incentive to manipulate policies in order to generate a large inflow of immigrants of his own socio-ethnic group, and encourage an outflow of residents from other groups (Glaeser and Shleifer, 2005). The Curley effect is named after James M. Curley, four-terms mayor of Boston in the period 1914–1950. Curley, the son of an Irish immigrant, became famous for aggressively pursuing various types of populist policies that (a) favored the inflow into Boston of poor, Catholic, Irish immigrants and (b) drove out of the city some of 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 long1 and successful political career2

Similar explanations have been proposed, mostly informally, for a few instances of socially

---

1Besides his four terms as a major of Boston, Curley also served as Governor and Senator of Massachusetts.
2Glaeser and Shleifer (2005), who brought the Curley effect into the economic literature, 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).
undesirable privatization\(^3\) and subsidies to home ownership\(^4\). 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 favor migration not only across space, but also across information sets and economic interests. Moreover, this is the first paper that systematically studies the main institutional and non-institutional determinants of this source of political failure.

We employ 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 (which generates a deadweight loss). Two parties have diverging preferences over the latter dimension only, and can ex ante compete for power by committing to a platform on the public employment/public good dimension. The partisan policy feedback is generated by the assumption\(^5\) that a citizen’s employment status (public vs private sector) influences his beliefs about redistribution: due to their indirect exposure to the “state” of the private sector, public sector workers systematically underestimate the social cost of redistribution vis a vis private sector workers. As a consequence, the initial size of the public sector systematically affects the two parties’s relative electoral strength.

Parties have then an incentive to manipulate public employment (and thus public good provision) in order to improve their future electoral strength even in a Downsian environment\(^6\) where platforms are “pushed” towards the socially optimal level. Parties optimally trade-off current electoral strength for a better future electoral environment, and commit to socially undesirable levels of public employment. 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

\(^3\)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.

\(^4\)Several columnists (e.g., Becker, Stolberg and Labaton in The New York Times, December 20, 2008) have also suggested that the increase in the subsidy to home ownership implemented in the last decades in the United States was motivated by the goal of shifting the electorate in a conservative way. Regarding home ownership, 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, but their policy implications are mostly normative.

\(^5\)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 (1996) and extensively applied in Persson and Tabellini’s Political Economics (2002).
between a dynamic policy feedback with two key elements. First, political actors are differentiated. (In this model, they are associated to different levels of redistribution\(^7\). Second, political platforms are related to implemented policies in a systematic way. (In this paper, through a commitment assumption and a constitution\(^8\). In this paper, the source of the policy feedback is purely informational. While this choice is mostly to improve the model’s tractability, another source could be due to systematic changes in voters’ economic interests\(^9\). In this paper, since employment status (private vs public sector) affects the precision of one’s information about the inefficiency of redistribution, public employment influences a voter’s propensity to choose one or the other party.

This highlights two fundamental differences between this paper and Glaeser and Shleifer (2005). First, in this paper, politicians spread conflict from a redistributive policy dimension to a common value dimension\(^10\) (public good provision), without exploiting an underlying ethnic conflict: 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 derived from informational asymmetries, 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 and implement undesirable policies. More far sighted politicians or voters and, more generally, more political persistence increase the inefficiency in public good provision. This finding seems at odds with what most of the recent literature in dynamic political economy suggests\(^11\). In

\(^7\)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 (Krassa and Polborn, 2009).

\(^8\)We define a constitution as a mapping from electoral outcomes into policy-making rights

\(^9\)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).

\(^10\)This mechanism is similar to Morelli and Van Weelden (2011), but comes from a completely different source.

\(^11\)Several authors have identified short termism as a primary source of political failures, especially in contexts involving redistributive politics (Acemoglu and Robinson, 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, 2009, on excessive investment taxation; Besley and Persson, 2010, on the development of fiscal capacity; Aitd and Dutta, 2007, on long term 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 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.
particular, the idea that more patient voters can lead to larger inefficiencies in the political process is a peculiar feature of partisan policy feedback.

The third contribution of the paper is to study how the distortions induced by partisan policy feedbacks depend on the type of constitution. Building on Lijphart (1999), we compare a winner-take-all majoritarian constitution to a consensual constitution, where parties’ influence over policies is proportional to their electoral strength. The main results are: (1) majoritarian constitutions are associated to less desirable public good provision (because of the absence of the moderating effect of bargaining), and lower redistribution; (2) when the horizon is finite, majoritarian constitutions display lower platform divergence (that is, lower political polarization), while in an infinite horizon model platform divergence is independent of the constitution; (3) regardless of the welfare criterion adopted, majoritarian constitutions are always welfare dominated by either consensual or “hybrid” constitutions (majoritarian on redistribution, proportional on public good provision).

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 the inefficiency of public good provision is ambiguous (parties are more extreme, but on average there is less underprovision of public goods).

1 Economic and Electoral Environment

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

Policies. The government produces a public good and redistributes income from self employed to workers using redistributive lump sum taxation $\tau_t \in \mathbb{R}^+$. The public sector employs 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 marginal product of labor is equalized across sectors.\footnote{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.}

\footnote{In order to keep the problem well behaved, we assume that $x^* < 1/2$, the lower bound on the number}
Income redistribution generates a transfer of per capita amount \( b_t \) to the citizens who choose to become workers. \( b_t \) is financed through the lump sum tax on those who choose to become entrepreneurs. A citizen chooses to become an entrepreneur if and only if \( a_i - \tau_t \geq b_t \). As a consequence, \( b_t \) is implicitly defined by the zero of

\[
\frac{b_t}{b_t} \Pr(a \leq b + \tau) - \tau \Pr(a \geq b + \tau)
\]

Payoffs and first best. Payoffs are linear income and public good provision. Depending on his occupation, citizen \( i \)'s per period indirect utility is then

\[
v_i(x, \tau) = (a_i - \tau) \mathbb{I}_{(a \geq b + \tau)} + b \mathbb{I}_{(a \leq b + \tau)} + g(x)
\]

A social planner placing relative weight \( \gamma \) on high productivity citizens would choose \((x, \tau)\) 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:

- public good inefficiency, \( \mathbb{E}\{-g(.)\} \) (i.e., 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, captures political polarization).

Political process. In each period two parties, \( R \) and \( L \), compete for office, which is associated with policy-making power over \( x \) and \( \tau \). 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 \( \bar{\tau} \). In every period \( t \), each party \( J \) commits to a policy platform \( x^J_t \). On the other hand, no commitment is possible for \( \tau_t \): \( R \), if alone in power, would then set \( \tau_t = 0 \) and \( L \), if alone in power, would implement \( \bar{\tau} \).

of workers (generated by \( \theta = 0 \)). Details in the Appendix.

\(^{14}\)Under the assumption, it is easy to show that there is a unique positive solution to that problem.

\(^{15}\)See, for example, McCarty and Shor (2011), who measure political polarization using data on candidates’ pre-electoral commitments.

\(^{16}\)Since we assume that parties are unitary actors, \( R \) and \( L \) can also be candidates.
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 \( \tau_t = \tau \) 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\(^{[17]}\) Assuming that \( R \) and \( L \) have preferences over redistribution is, instead, essential: removing it would eliminate party differentiation, and, as a consequence, partisan feedback effects. This assumption seems quite natural, since parties typically differ (Boix, 1998; Bradley et al., 2003) in how they balance the trade-off between inequality and efficiency that virtually every polity faces.

**Voting behavior.** Each voter \( i \) computes \( \mathbb{E}_i\{v(x_t^R, \tau^L)\} \), the expected per period payoff associated with each party’s \( x_t^R \) and \( \tau^L \). Although voters know \( \tau^R = 0 \) and \( \tau^L = \tau \), the payoff associated with \( \tau \) depends on \( \sigma_t \). The latter, as we will see, can only be conjectured by some voters. Voting behavior is probabilistic: \( i \) votes for \( R \) iff

\[
\mathbb{E}_i\{v(x_t^R, 0)\} > \mathbb{E}_i\{v(x_t^L, \tau)\} + \xi_t + \delta_t^i
\]

where \( \xi_t \) is the realization of a stationary zero-mean aggregate preference shock \( \xi \), and \( \delta_t^i \) is the realization of a stationary zero-mean, idiosyncratic preference shock \( \delta \). As in one of the standard formulations of the probabilistic voting model\(^{[18]}\) \( \xi_t, \delta_t^i \) 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\(^{[19]}\) 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. \( \xi_t \) measures how much the median voter prefers the \( L \)-candidate over the \( R \)-candidate abstracting from \( x \) and \( \tau \). \( \delta_t^i \), 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 expects to be beneficial for her\(^{[20]}\)

---

\(^{[17]}\)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. 18See Lindbeck and Weibull (1987) and Persson and Tabellini (2002).

\(^{[19]}\)More precisely, the payoff difference between \( L \)'s and \( R \)'s attributes.

\(^{[20]}\)Moreover, voting behavior in this model is compatible with the assumption, often employed in political
Notice also that, since they only look at the current period’s payoff, voters are myopic. This assumption will be relaxed.

**Information and timing.** Citizens have a uninformative prior $U_{\mathbb{R}^+}$. At the beginning of each period, private sector workers observe $\sigma_t$, 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 (also summarized in Figure 1) is then as follows.

1. Depending on their initial sectorial allocation, citizens observe either $\sigma_t$ or $s_{it}$.
2. Parties observe voters’ information, and announce a platform $x^J_t$, $J \in \{R, L\}$.
3. Citizens compute $\{E_i[v(x^J_t, \tau^J)]\}_{J \in \{R, L\}}$, are hit by preference shocks $\{\xi_t, \delta^J_t\}$, and cast their vote.
4. After the constitution maps votes into policy-making power, policies are implemented and voters, upon observing $a_{it}$ and $b_t$, make optimal sectorial and occupational choices, and enjoy their payoff.

---

21 The uninformative prior greatly improves the model’s tractability. In particular, it prevents uninformed voters from extrapolating information about $\sigma_t$ 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.
Expected payoffs. Denote by $D_i$ i’s expected payoff difference between Right and Left. Under the assumptions, the probability that $i$ votes for the right and $R$’s total realized vote share are, respectively

$$\hat{\pi}_{it} = \Pr[\delta < D_i - \xi_t] = \frac{1}{2} + \varphi[D_i - \xi_t]; \hat{\pi}_t = \frac{1}{2} + \varphi\left[\int D_i di - \xi_t\right]$$

Let $D(a)$ be the payoff difference between Right and Left for a citizen with productivity $a$:

$$D_t(a) = g(x^R_t) - g(x^L_t) + \tau\,I\{a \geq b_t + \tau\} - b_t\,I\{a \leq b_t + \tau\} + a\,I\{a \in [0, b_t + \tau]\}.$$ 

Averaging across all possible productivity levels that $i$ might have, and using (1), yields

$$D_i = g(x^R_t) - g(x^L_t) + \int_0^{b_t + \tau} a\,\sigma_t\,da = g(x^R_t) - g(x^L_t) + \tau - \mathbb{E}_t\{b(\sigma_t)\}.$$ 

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

Aggregating over all citizens and signal realization yields

$$\hat{\pi}_t = \frac{1}{2} + \varphi\left\{\frac{g(x^R_t) - g(x^L_t)}{\text{Platform-related gain}} + \frac{\eta_t}{\text{L’s efficiency loss}} - \frac{\lambda_t x_{t-1}}{\text{Informational wedge}} - \frac{\xi_t}{\text{Shock}}\right\}.$$ 

$\eta_t = (\tau - b(\sigma_t))/4$, is the efficiency loss associated with L’s redistributive taxation: the entrepreneurial income of those who, despite having positive productivity, under $\tau$ choose to be workers. $\lambda_t = \int_{\sigma_t} \mathbb{E}\{b(\sigma_t)|s\}f_s(s)\,ds - b(\sigma_t)$, the informational wedge, is the marginal effect of a change in the initial size of the public sector on $R$’s electoral strength. $\lambda_t$ 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 $\lambda > 0$. Public sector workers are, on average, systematically more favorable to redistribution (as empirically documented, among others, by Guillaud, 2011). Public employment then generates a partisan policy feedback.

---

22More specifically, we use the fact that (1) can be rewritten as

$$(b + \tau)/2 - \sigma(b + \tau)(b + \tau) = 0.$$ 

23We are assuming that both parties have a proper prior, which seems natural, given that their electoral chances depend on $\sigma$. 

9
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 of the electorate. When choosing $x_t^l$ (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 $\sigma$) is small.

Two important observations are in order. First, there might be other types of informational asymmetries that can generate a partisan policy feedback (for example, asymmetric information on the administrative cost of implementing redistribution), but as long as voters’ beliefs on the effect of $\tau$ depend on their sectorial affiliation, the linkage will exist.

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.

In order to ensure continuity and differentiability in the objective functions, we also assume that $\varphi$ and $\psi$ are related in such a way that, for every initial value of $x$ and realization of $\sigma$, both actors have a positive probability of obtaining a majority of the votes.\footnote{This assumption is essentially equivalent to the one in Persson and Tabellini (2002).} Moreover, in order to ensure the consistency of proposed policies with future sectorial choices\footnote{That is, that there will always be at least enough workers to make the prescribed platforms feasible} we assume that $x^*$ is small enough. These assumptions, which are technical in nature, are described in the Appendix.

### 1.1 Constitutions and Equilibrium Concept

So far we have left unspecifed the rule mapping electoral outcomes into policy-making power over $x$ and $\tau$. We call such mapping a \textit{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 \textit{majoritarian} (denoted by $M$) and \textit{consensual} (denoted by $C$) constitutions, formally defined below.

#### Majoritarian Constitution

The majority winner, $W_t$\footnote{Formally, $W_t = R^l_{\{x_t \geq 1/2\}} + L^l_{\{x_t \leq 1/2\}}$}, gets full policy-making power on both policy dimensions: he implements his announced platform $x^l_t$ and his preferred $\tau^l$. Denote by $X^M = x^W_t$ the implemented policy under majoritarian constitution. Parties’ per
period expected payoff is then the probability of winning a majority.

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

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 his 1977 book, * Democracy in plural societies, * Lijphart describes the main features of a consensual democracy:

Consociational Democracy can be defined in terms of four characteristics. The first and most important element is 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.

In an effort to balance adherence to the original definition and analytical tractability, we model consensual democracy as a stylized post-electoral bargaining game between \( R \) and \( L \), where the two negotiate over \( x \) and \( \tau \) 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:

\[ X_t^C = \hat{\pi}_t x_t^R + (1 - \hat{\pi}_t) x_t^L; \quad \tau_t^C = (1 - \hat{\pi}_t) \tau. \]

If there is a nonempty set of Pareto improving pairs \((x^{pr}, \tau^{pr})\) 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.\(^{27}\)

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

\(^{27}\)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.
As a consequence, R’s per-period expected payoffs is
\[ \pi_t = \mathbb{E}_t[\hat{\pi}_t] = 1/2 + \varphi[g(x^R_t) - g(x^L_t) + \eta_t - \lambda_t x_{t-1}] \]
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 \(x^j : [0, \mu^h] \to [0, 1] j \in \{R, L\}\), which depend on the payoff-relevant state \(\mu_t(\lambda_t, \eta_t, \sigma_t, x_{t-1}) = \eta_t - \lambda_t x_{t-1}\). \(\mu_t\) is the average expected inefficiency of \(\pi\), and measures how favorable to R is the electoral environment.

### 2 Majoritarian Constitution

In this section we first 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: \(x^R_2 = x^L_2 = x^*\).

ii) *In *\(t = 1\) parties commit to
\[ x^R_1 = x^* - \frac{\Delta M}{2} - \frac{\psi \Delta M}{1 + \psi \Delta^2 M} \mu_1; \quad x^L_1 = x^* + \frac{\Delta M}{2} - \frac{\psi \Delta M}{1 + \psi \Delta^2 M} \mu_1, \]
and platform divergence (denoted by \(\Delta_M\)), solves
\[ \Delta[1 + \psi \beta \Delta] = \beta \Delta / 2. \tag{3} \]

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, \(X^M_1\), 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 \(X^M_1\) 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 $x^R_1$:

$$\frac{dp_1}{dx^R_1} + \beta \frac{dp_1}{dx^R_1} E[p_2(\mu_2|x^R_1) - p_2(\mu_2|x^L_1)] - \beta \frac{dE}{dx^R_1} p_2(\mu_2|x^R_1) = 0.$$  

On the one hand, setting $x^R_1 = x^*$ 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 impact of $R$’s platform on tomorrow’s electoral environment (Legacy component). On the other hand, a marginal reduction in $x^R_1$, conditional on winning the election, increases $R$’s future electoral chances (Curleyan component). $L$ faces an analogous trade-off, with the only difference that he gains from distorting his platform upward. The point at which this trade-off is balanced generates distortions at both platform and implemented policy level.

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 ($\beta$), in the informational wedge ($\lambda$), and in the volatility of the aggregate shock ($\psi^{-1}$).

ii) Public good inefficiency ($-E\{g(X^M)\}$) is increasing in policy divergence and the initial state ($\mu_1$), ambiguous in $\psi^{-1}$.

The parties’ discount factor increases dynamic distortions: by increasing the relative importance of the future electoral environment, $\beta$ increases the incentive to distort political platforms $x$. The intuition behind this result is analogous to Glaeser and Shleifer (2005). A larger $\lambda$ has the same effect of a larger $\beta$: 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 ($\mu_1$) 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 Curleyan component in his intertemporal trade-off, and pushes implemented policy towards a more severe underprovision of public good.

The effect of aggregate ideological volatility ($\psi^{-1}$) 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 \( \mu_1 \) is small enough, the effect on the former dominates, and ideological volatility, by lowering voters’ responsiveness, weakens accountability. When, instead, \( \mu_1 \) 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.\(^{28}\) 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** Forward looking voters increase public good inefficiency \( -\mathbb{E}\{g(X^M)\} \).

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 \( \mu_1 \)), thereby pushing downwards both platforms.\(^{29}\)

### 2.1 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 solved 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 \( \Delta_\infty = \beta \lambda \), platforms are given by

\(^{28}\)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.

\(^{29}\)Simple inspection of equilibrium platforms also shows that having more far sighted voters does not affect political polarization.
\[ x^R(\mu) = x^* - \frac{\Delta_\infty}{2} - \frac{\psi \Delta_\infty}{1 + \psi \Delta_\infty^2} \mu \quad ; \quad x^L(\mu) = x^* + \frac{\Delta_\infty}{2} - \frac{\psi \Delta_\infty}{1 + \psi \Delta_\infty^2} \mu \]

**Corollary 1**

i) Platform divergence is independent of aggregate volatility \( \psi^{-1} \).

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 increase.

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 platform divergence through the Legacy component, which represents the incentive of to maximize the influence of the own platform on the future electoral environment.

In the last period there is no incentive to maximize the influence of the own platform on future electoral environment, the Legacy component enters only one side of the political actors' dynamic trade-off. More generally, in every finite horizon model, the Legacy component will decrease in \( t \) (reflecting a shorter sequence of future elections), generating a change in the incentives for political divergence across periods. In a stationary equilibrium, instead, the Legacy component is be constant over time. In the baseline setting of model, where payoffs are quadratic, it must be that the Legacy component disappears from the dynamic trade-off. This effect is captured by the term \( \psi \beta \bar{\lambda} \Delta^2 \) in (3); removing it yields precisely \( \Delta = \beta \bar{\lambda} \).

**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 \( (\beta, \bar{\beta} = \sum_{t=1}^{\infty} \beta^t) \), and compare the equilibrium with \( T = 2 \) and \( \bar{\beta} \) to the one with \( T = \infty \) and \( \beta \). 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 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.

## 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: \( x_R^2 = x_L^2 = x^* \).

ii) In \( t = 1 \) parties commit to

\[
    x_R^1 = x^* - \frac{\Delta C}{2} - \frac{\varphi \Delta C}{1 + \varphi \Delta C} \mu_1 \quad ; \quad x_L^1 = x^* + \frac{\Delta C}{2} - \frac{\varphi \Delta C}{1 + \varphi \Delta C} \mu_1,
\]

and platform divergence (denoted by \( \Delta_C \)), solves

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

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 (\( \psi^{-1} \)), platforms depend also on idiosyncratic volatility (\( \varphi^{-1} \)): under a majoritarian constitution \( R \) and \( L \) only care about going above half of the votes, 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. It is straightforward to show that the rest of the comparative static (in \( \lambda, \beta \), and \( \sigma^{-1} \)) is as in the \( M \)-game.

### 3.1 Infinite horizon model

In this subsection we study the infinite horizon version of the \( C \)-game and compare it to its two-period version. As for \( M \), let \( x^R(\mu), x^L(\mu) \) denote the equilibrium platform for a given state \( \mu \).
Proposition 7 In the unique stationary differentiable MPE of the infinite horizon $C$-game, platform divergence is given by $\Delta_\infty = \beta \lambda$, platforms are given by $x^R(\mu) = x^* - \frac{\Delta_\infty}{2} - \frac{\varphi \Delta_\infty}{1 + \varphi \Delta_\infty^2} \mu$; $x^L(\mu) = x^* + \frac{\Delta_\infty}{2} - \frac{\varphi \Delta_\infty}{1 + \varphi \Delta_\infty^2} \mu$.

The key difference between two-period and infinite horizon is then that platform divergence 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 ($\varphi^{-1}$) plays the same role that aggregate volatility plays in the $M$-game. The rest of the comparative static (effect of $\beta, \lambda, \mu$) is as in the $M$-game.

3.2 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 (Gerber and Ortúño Ortún, 1998; Persson, Roland and Tabellini, 2000; Persson and Tabellini, 2003 and 2004) and proportional electoral systems (Austen-Smith and Banks, 1988; Milesi-Ferretti, Perotti and Rostagno, 2002; Persson and Tabellini, 2003 and 2004).

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 compensating 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 such increase is mitigated by the compensating effect to a larger
extent. 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. While this latter fact is a by-product of the simple quadratic structure of the baseline model, the general intuition behind this long term neutrality result also holds in more general payoff structures.

Finally, it is worth noting that part (ii) of Proposition (8) echoes Ticchi and Vindigni (2010) and Iversen and Soskice (2006), who find that consensual constitutions and proportional systems are associated with more redistribution because they tend to give left-wing parties more power.

**Semi-consensual constitution.** $C$ and $M$ differ in the allocation of power over two dimensions: public good provision and redistribution. One’s first conjecture might be that, since parties only care about redistribution, what drives differences across constitutions is the difference in the allocation of power over $\tau$. 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[^30] and the one over public good provision is consensual[^31] As the next proposition formalizes, in the semiconsensual constitution parties commit to the exact same levels of public good provision as in the consensual constitution. As a consequence, 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). It must be stressed that is not a consequence of the lack of commitment assumption on redistribution. It can be shown that removing it yields a very similar comparison between $S$, $C$, and $M$.

Policy distortions in $S$ are as large as in $C$, and smaller than in $M$. Furthermore, implemented redistribution under $S$ is majoritarian, which suggests that semi-consensual has lower average redistribution than $C$. One would then conjecture that, from a utilitarian perspective, $S$ dominates both $C$ and $M$. The following proposition shows that this is true, and that something can be said even when one adopts a non-utilitarian welfare criterion, in which the relative weight of the rich and poor are, respectively, $\gamma$ and $1 - \gamma$, with $\gamma \in (0, 1)$.

**Proposition 9** i) In the $S$-game, $x^R_t$ and $x^L_t$ are exactly as in the $C$-game.

ii) Under a utilitarian welfare criterion, a semi-consensual constitution dominates both consensual and majoritarian.

iii) Under any generic welfare criterion, a majoritarian constitution is always dominated.

[^30]: Formally, $\pi_t^S = \pi_t^I(LW_i = L)$.
[^31]: Formally, $X_t^S = \hat{\pi}_t x_t^R + (1 - \hat{\pi}_t) x_t^L$. 

18
A consensual allocation of power on policy dimensions that generate policy feedbacks, coupled with a majoritarian allocation of on policy dimensions over which political actors intrinsically care about the maximizes utilitarian welfare. Under a more egalitarian welfare criterion, C might be optimal, but the more important result from this proposition is that majoritarian constitution is always dominated.

4 The effect of Inequality

This section discussed three empirical predictions that the model delivers about the relationship between ex ante inequality, political polarization, public good provision, and redistribution. These topics have received a lot of attention from researchers across various disciplines in the social sciences and are part of a recent debate that goes beyond academia. In the context of this paper, an increase in inequality corresponds to a downward first order stochastic shift in the distribution of $\sigma$.

**Proposition 10** Higher inequality increases platform divergence, has an ambiguous effect on public good inefficiency.

When the distribution of $\sigma$ shifts towards the origin both the informational wedge and the value of $b()$ increase: the reason is that (1) for given $\tau$, 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. This reduces $\mathbb{E}\{\eta\}$ and increases $\lambda$, thereby reducing the overall size of the Right’s initial electoral advantage.

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. The second result, which follows from the first, is that higher inequality is associated with lower and more volatile public good provision, but also with a weaker electoral advantage to the Right, which reduces the expected underprovision in public good.

---

32 It is also possible to show that the opposite configuration (majoritarianism on public good provision and consociativism over redistribution) would be, under a utilitarian criterion, dominated by both (M) and (C).

33 Notice that these results do not depend on the time horizon: both propositions can also be obtained for the infinite horizon case.
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 that political agents face, 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), 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. In the first two papers (which develop the idea originally illustrated in Dixit and Londregan, 1995), the relevant cleavages are farmers vs manufacturing workers, in Glaeser and Shleifer (2005), Irish/Catholic vs Protestant/Anglo-Saxons citizens.

In all these contributions, inefficiencies arise because political actors can successfully exploit an existing conflict that feeds back into political preferences. In this paper, instead, no citizen gains from a suboptimal level of public employment: inefficiencies arise because of the effect asymmetric information between public and private sector workers on their induced political preferences. As a consequence, any institutional device that mitigates dynamic commitment problems will not reduce the inefficiencies associated with this channel of political failure.

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) inefficiencies are observed in the composition of the bureaucracy (rather than its size) and can only arise in a non-democratic context (rather than the endogenous outcome of electoral competition). Most important, these inefficiencies due to the presence of an underlying interest in the society (the rich) benefitting from these policies, which is absent in our paper (where inefficiencies arise because political actors exploit policy feedback effects).

Krussell and Rios-Rull (1999), and Hassler et al. (2005) are also related to this paper. The first looks at a dynamic version of the Meltzer-Richard model to study the dynamics of redistribution within the context of a neoclassical model. The second studies the evolution of preferences over redistribution in an OLG economy in which young agents have to undertake an investment that improves their future expected productivity and how this interacts with redistributive politics to generate inefficiencies.

By studying the dynamics in public employment, output, and redistribution, this paper is

---

34In 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.
connected to a dynamic public finance literature focusing on exogenous changes in power and inefficient volatility in output and consumption. Acemoglu et al. (2009) explore the effect of stochastic power fluctuations on the allocation of resources in a dynamic production economy where groups differ in their labor-leisure preferences. From a more general perspective, Bai and Lagunoff (2008) consider an environment in which policy-making exhibits “Faustian” dynamics: in every period the identity of the (weighted) median voter depends on the wealth distribution. As a consequence, the policy that maximizes the immediate payoff of the current median voter also shifts away political power from him. This literature shares with the present paper the idea that uncertainty over future allocation of political power constitutes an independent channel for political failures. While in Acemoglu et al. (2009) more patient actors help mitigating these issues, in Bai and Lagunoff they make these distortions more pronounced.

Our paper features a dynamic linkage in policies, whose interaction with the political process generates distortions, and is, in this respect, related to a series of paper in political economy of investment and taxes (Besley and Coate, 1998; Azzimonti, 2009) and public debt (Aghion and Bolton, 1990; Milesi-Ferretti and Spolaore, 1994). This idea is also key in Battaglini and Coate (2008), who look at a dynamic legislative bargain model of redistributive politics where the presence of public debt create incentives to shift costs towards future periods. While in their model more persistence in power reduces pork and the inefficient accumulation of debt, our paper generates the opposite prediction.

This paper is also related to a small literature on short termism as an equilibrium response to some underlying friction in the policy-making process. Garr`ı (2009) explains the policy bias toward the short term public goods using a reputational argument, leading to the conclusion that political short termism might be welfare improving because it enhances selection of congruent politicians. In this paper, short termism is beneficial absent any independent of reputation and selection motives.

The idea that political actors can manipulate policies to improve their future electoral strength is also featured in Hodler et al. (2010), where different politicians are associated with different maps between policies and outcomes. For this reason, an incumbent has incentives to provide inefficient policies in order to shift the salience towards the dimensions in which he thought to be more productive than his competitor. Policy manipulation, then, creates an endogenous incumbency advantage. The present paper shows that a similar manipulation can arise even at platform level and, more importantly, without technological differences between politicians.

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, Roland and Tabellini, 2005; Lizzeri and Persico, 2001; Milesi-Ferretti, Perotti and Rostagno, 2002). In a recent paper, Battaglini (2010) extends the setting of Battaglini and Coate (2008) to a multiple-district environment with probabilistic voting. His main finding is that the basic prediction of the static literature on electoral outcomes on public finance (PR gives leads to overspending and less transfers with respect to majoritarian systems) does necessarily hold in a dynamic setting with endogenous public debt. The reason is that, while proportional representation generates more incentives to overspending, it also results in faster accumulation of public debt, which in turn generates a tighter endogenous constraint on spending. The dynamic linkage in Battaglini’s paper is purely economic (public debt), while in our paper is essentially political (because of the effect of public employment on perceived redistribution). Moreover, our paper focuses on a different institutional comparison, as in Ticchi and Vindigni (2010). Baron, Diermeier and Fong (2011), instead, focus on parliamentary form of government with proportional electoral rule. With respect to the current paper, their model features a richer description of the institutional structure and a more stylized underlying economic environment. Nevertheless, the idea that parties currently endowed with proposal power have incentive to manipulate policies to improve their future bargaining position in future government coalitions is closely related to our model.

Kalandrakis (2009) builds a reputational theory of two party competition in which voters are uncertain about whether a party is controlled by extremists or moderate agents. In his setting, far sightedness of partisan agent has two set of consequences. On a static level, it encourages the adoption of moderate policies for electoral purposes. On a dynamic level, it pushes towards extreme policies because of their impact on reputation: government parties pursue extreme policies to avoid losing elections almost for sure against an opponent on moderate platforms. Although the setting and the source of the intertemporal trade-off are very distant from our model, this 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.

The empirical implications of our paper relate the model to two important bodies of literature in economics and political science. The first investigates the negative relationship between inequality and redistribution, documented empirically, among other, in Enns and Kelly (2010), De Mello and Tiongson (2003), Perotti (1996). These results seriously challenged the conjecture (Meltzer and Richard, 1981) that redistribution increases with inequality. Therefore, in recent years there have been several attempt to produce theories
yielding the opposite prediction (Moene and Wallerstein, 2001; Bénabou, 2000; Bénabou and Ok, 2001; Iversen and Soskice, 2006). Iversen and Soskice (2006), in particular, focuses on constitutional differences and argues that countries with PR are more likely to have center left governments and, as a consequence, higher redistribution and lower inequality than countries with majoritarian electoral systems. Unlike the present paper, their focus is cross sectional, rather than dynamic. Moreover, differences in outcomes are generated by the different incentives to form coalitions, rather than different incentives to distort platforms. The second relevant body of literature investigates the relationship between inequality and political polarization (see McCarty et al., 2008, for a comprehensive study). This paper contributes to each of these literatures by showing the existence of a novel channel that can help explaining how the channel between inequality, redistribution and political polarization is affected by institutional factors.

6 Robustness

This section illustrates how the results of this paper are robust to alternative specifications of various components of the model. Since the setting is meant to capture, albeit in a stylized way, several aspects of an economic and a political system, it is quite natural to wonder how many of the modeling choices play a crucial role in the derivation of the results. This section addresses some of the main potential sources of concern.

Term limits, general heterogeneity in the actors’ intertemporal trade-off. The outcome under C critically depends on the fact that the agents cannot find any Pareto improvement from bargaining jointly on both dimensions. The reason is that, in the baseline setting, the agents have the same intertemporal trade-off. Without this symmetry, the comparison between M and C can change quite dramatically. There are at least two different perturbations of the model that generate a failure in such symmetry. First, the presence of term limits might create asymmetries in political actors’ time horizon. Second, political actors might have different discount factors: 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.

Under a consensual constitution the two actors can then improve their expected utility by bargaining on both dimensions jointly, and splitting power across them asymmetrically. For example, when L is more far sighted, his induced preferences are such that he cares relatively more about public good provision than R. He finds then profitable to trade policy-making
power over redistribution
for policy-making power over public good provision, and $R$ finds it profitable to do the opposite. As a consequence, the allocation of policy-making power differs from the one that arises in the baseline case.

How does this asymmetry affect the incentives to distort platforms? It is possible to show that the \textit{ex ante} incentive for policy distortions increase. The reason is that the connection between the vote shares and each actor’s payoff is weaker than in the symmetric case: in the latter, vote share is essentially the actors’ payoff, in the former it only affects their reversion points. As a consequence, consensual democracy not only allocates power over redistribution overwhelmingly towards the short sighted actor and over public good provision towards the far sighted actor, but displays an even larger platform divergence than in the symmetric case.

The effect of such asymmetry is minimal under a majoritarian constitution, because actors cannot trade power through bargaining. The only effect of the asymmetry is given by a differential 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 under $\mathbf{M}$. A qualitatively similar effect is observed in the $\mathbf{S}$ constitution, where bargaining over both dimensions is not possible, so the basic intuition of Proposition 9 remains.

In conclusion, consensual constitutions are fragile to heterogeneity in the intertemporal trade-off among actors, while majoritarian and semi-consensual are not only robust to it, but do benefit from such asymmetries. This fact also helps explaining why term limits tend to be more often observed and less controversial in majoritarian systems: they not only reduce the incentive for a successful incumbent to distort policies, but also reduce platform divergence.

**Alternative channels linking current policies to future political preferences.** The mechanism of the model is crucially based on two features: the presence of an informational link between the two policy dimensions and the fact that political actors are associated with different levels of redistribution. This is not the only possible source of dynamic political distortions. For example, working in the private sector might directly affect the distribution of workers’ skills in the society, increasing inequality and, therefore, the social cost of redistribution. Below, we briefly discuss two other potential channels that would generate an endogenous link between implemented policies and future electoral environment.

- Ideology (rather than asymmetric information). Suppose that there are two preference
types: most voters are like in the baseline model (rational type, $r$), but a measure $\varepsilon_t$ of them derives a disutility term from inequality (egalitarians, type $e$). Further assume that the public sector has a larger share of egalitarian types (that is $\varepsilon^G_t > \varepsilon^P_t$), and one’s type can switch after interacting with a new working environment. More specifically, assume that a worker who moves from one sector to the other takes the type of the first co-worker she interacts with, which is randomly drawn. The law of motion of the share of types $e$ in this simple environment is then given by

$$\varepsilon_{t+1} = (x_{t+1} - x_t)(\varepsilon^G_t - \varepsilon^P_t) + \varepsilon_t$$

The effect of expending the public sector is then to increase the chance of ideological switches from $r$ to $e$ and vice versa. As a consequence, expanding the public sector today produces an electorate that is more favorable (or less hostile) to redistribution, thereby increasing $L$’s appeal.

- Specialized candidates (rather than policy motivated). Recent theoretical work in political economy focuses on specialized (Krasa and Polborn, 2009, 2010), who differ in their ability to provide policies across different dimensions. When $R$ and $L$ have different abilities in implementing redistribution\footnote{One can think that $L$ has, for various reasons, more expertise in implementing redistributive programs.}, $L$’s electoral strength increases in the demand for redistribution. Since the latter is related to the perceived inefficiency of redistribution, $L$ and $R$ have a similar \textit{ex ante} incentive to manipulate public good provision as in the baseline model.

\textbf{Alternative specification of Consensual Democracy.} Some authors (for example, Battaglini, 2010, or Herrera and Morelli, 2010) have modeled proportional electoral systems (which are a prominent feature of consensual constitutions) as games in which a party gets to implement his platforms with a probability proportional to their realized vote share. Although the setting presented in the baseline better captures Lijphart’s concept, it is quite natural to ask to what extent the results derived in this paper are sensitive to the modeling assumptions of $C$.

To partially address this question, let’s consider an alternative version of consensual democracy $C_2$, where, as in Battaglini (2010), a political actor gets full policy-making rights with probability $\hat{\pi}_t$. It is easy to show that equilibrium strategies are exactly as in the standard model, and the implemented policy is a lottery 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 that policy distortion should not necessarily be larger under the latter, as in the
baseline model. It is possible to show that, when the initial state \( \mu_1 \) is large enough, platform distortion is larger under \( M \), while the opposite is true for low values of \( \mu_1 \). The intuition is that under \( C2 \) platform divergence is larger, but the probability of having the most extreme platform (that is, \( x^R \)) implemented is lower.

In conclusion, although the constitutional comparison is affected by changes to the modeling assumptions of a consensual constitution, the basic message of the paper remains.

**Commitment on redistribution.** It is possible to show that, as long as parties cares about redistribution, allowing them to choose \( \tau \) will not affect the existence of partisan feedback effects.

**General public good production technology.** When, rather than a quadratic loss function for \( g() \), one assumes a generic inverted U-shaped function, most of the qualitative results of the paper for two period models still hold, although an explicit solution is no longer available. Numerical simulations, using a power function of the form \( g(x) = -(x - x^*)^{1+\gamma} \), suggest that the equilibrium still exists, and that most of the results derived in the paper hold. The long run neutrality of the constitution on platform divergence holds in a weaker sense: rather than eliminated, the effect of the constitution becomes less pronounced (platform divergence becomes more similar) as one increases \( T \), while \( C \) remains associated with higher divergence.

7 Conclusion

This paper studies dynamic electoral competition with partisan policy feedback effects: policies today influence the electorate’s induced preferences over parties tomorrow. We study how far sighted 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 on a specific instance of feedback effect, migration policy and the Curley effect (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 setting, public employment affects the precision of a voters’ information about the distribution of labor productivities in the private sector and, as a result, her beliefs about the social consequences of redistribution. Since the parties competing for power are differentiated along this dimension, public employment systematically affects political preferences. Parties have then an incentive to manipulate public employment to increase their future electoral strength, with undesirable consequences for public
good provision.

We study how these consequences depend on various institutional (constitutions) and non-institutional (time preference, ideological volatility) factors. Somewhat contrary to what the recent literature on dynamic political failures would suggest, inefficiencies increase with politicians’ and voters’ patience and, more generally, with political persistence.

We focus on two measures of distortions: platform divergence, capturing political polarization, and inefficiency in public good provision. Majoritarian constitutions 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 good). In an infinite horizon model, platform divergence no longer depends on the type of constitution. In other words, the political polarization induced by partisan policy feedback does depend, in the long run, on institutional factors.

The differences in outcomes between constitutions critically depend on the allocation of policy-making rights over the public good/public employment dimension, and is less sensitive to the allocation of power over redistribution (which is, quite paradoxically, the only dimension over which political actors have preferences). Moreover, from a utilitarian perspective, 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. The main result is that a flatter distribution of productivities 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 in democratic policy-making.

Appendix

**Lemma 2** The function $b(\sigma) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$, defined as the zero of $F(b, \sigma) = (b-\tau)/2 + \sigma(b+\tau)^2$, is strictly decreasing and strictly convex in $\sigma$ for $\tau \geq 1/2\sigma$ (constant otherwise).
Proof. Using the implicit function theorem,
\[
\frac{\partial b}{\partial \sigma} = \frac{-(b + \tau)^2}{1/2 + \sigma 2(b + \tau)} \mathbb{I}_{[\tau \leq 1/2 \sigma]} \leq 0
\]
\[
\frac{\partial^2}{\partial^2 \sigma} \propto \left[-(b + \tau)b' - 2\sigma(b + \tau)^2 b' + 2(b + \tau)^3\right] \mathbb{I}_{[\tau \leq 1/2 \sigma]} \geq 0
\]
\[\blacksquare\]

Ensuring well behaved objective functions. In the paper, we solve the party’s problem under the assumption that their objective function is continuous and differentiable. To ensure that, we need to guarantee that (1) the proposed level for the public sector is always lower than the total pool of workers; (2) both political actors to be competitive in every election. The first condition is equivalent to assuming that \(x^L\) is always smaller than the number of actual workers, which bounded below by 1/2. As we will see later, \(x^L \leq x^* + \beta/4\). As a consequence, we assume
\[x^* + \beta/4 < 1/2.\] (5)

The second condition is equivalent to making assumptions on the relative size of the state space, \(\varphi\) and \(\psi\), as standard in this type of models.\(^{36}\) In order for the the range of the realized vote share to include, for every realization \(\mu_t\), and any platform profile \((x^R, x^L)\), the value 1/2. More formally, given \(\hat{\pi}(\xi, \mu) = 1/2 + \varphi[g(x^R) - g(x^L) + \mu + \xi]\), we must have \(\max_{\xi} \hat{\pi}(\xi, \mu) \in (1/2, 1), \min_{\xi} \hat{\pi}(\xi, \mu) \in (0, 1/2) \forall \mu \in [0, \mu^h]\), \(\forall g(x^R) - g(x^L) \in [- (1 - x^*), 1 - x^*]\). The two conditions can be repressed as
\[
\min\{1/\varphi - 1/\psi, 1/\psi\} > 2(1 - x^* + \mu^h)
\]
\[\text{(6)}\]

Proof of Proposition 1\(^{1}\) In \(t = 2\) equilibrium policies solve \(x^R_2 \in \arg \max\{p_2(x_1)\}, x^L_2 \in \arg \max\{1 - p_2(x_1)\}\), where \(p_2(x_1) = 1/2 + \psi[d(x^R_2, x^L_2) + I_2 - \lambda_2 x_1]\), where \(d(x^R_2, x^L_2) = g(x^R_2) - g(x^L_2)\). As a consequence, we must have \(x^R_2 = x^L_2 = \arg \max d(x^R_2, x^L_2) = x^*.\) \(x^R_1\) and \(x^L_1\), instead, solve
\[
\begin{align*}
x^R_1 & \in \arg \max_{x \in [0, 1]} \left\{p_1(x_0) + \beta p_1(x_0) p^*_2(x) + \beta (1 - p_1(x_0)) p^*_2(x^L_1)\right\} \\
x^L_1 & \in \arg \max_{x \in [0, 1]} 1 + \beta - \left\{p_1(x_0) + \beta p_1(x_0) p^*_2(x) + \beta (1 - p_1(x_0)) p^*_2(x^R_1)\right\}
\end{align*}
\]
\[\text{(7)}\]

where \(p^*_2(x) = \mathbb{E}[p_2(x)] = 1/2 + \psi \mathbb{E}\{\eta\} - \psi \bar{x} \mu\) follows from the observation that \(d(x^R_2, x^L_2) = 0\). The FONC are of the problem (which are also sufficient under the assumptions) define

\(^{36}\)See Persson and Tabellini (2002), Chapter 3.

\(^{37}\)By (5), \(x^* < 1/2\), which implies that \(x = 1\) gives voters the lowest payoff.
the following system

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

(8)

summing the two equations yields the condition \( \Delta[1 + \psi \beta \Delta] = \beta \lambda / 2 \), subtracting the first from the second yields

\[
\psi[2x^* - (x^R + x^L)](1 + \beta \psi \lambda \Delta_M) = \psi \beta \lambda [2p_1(x_0) - 1]
\]

which, after substituting for \( 2p_1(x_0) - 1 = \psi \Delta(x^R + x^L) - 2\psi x^* \Delta + 2\psi \mu_1, \Delta[1 + \psi \beta \lambda \Delta] = \beta \lambda / 2, \)

and rearranging, becomes

\[
2x^* - (x^R + x^L) = \frac{2\psi \Delta_M}{1 + \psi \Delta^2_M} \mu_1
\]

from which \( x_1^R \) and \( x_1^L \) are derived. \( \blacksquare \)

**Proof of Proposition 2.** For i), start observing that, since the RHS of the equation \( \Delta \) is supermodular in \( \beta \lambda \), the positive solution of that equation shifts to the right as \( \beta \lambda \) increases.

ii) Notice that \( \mathbb{E}\{-g(X^M)\} \propto \mathbb{E}\{(X^M - x^*)^2\} \), substituting platforms into \( p_1(x_0) \), and using the fact that we can rewrite \( x_1^R = x^* - \Delta_M p_1(x_0), x_1^L = x^* + \Delta_M (1 - p_1(x_0)) \), the expression simplifies to

\[
\Delta^2_M \left[ 1/4 + 3(\psi \mu_1)^2(1 + \psi \Delta^2_M)^{-2} \right]
\]

(9)

which is increasing in \( \Delta^2_M \) and \( \mu_1 \). To see that it ambiguous in \( \psi \), notice that \( \frac{d}{d\psi} \mathbb{E}\{-g(X^M)\} = \frac{\partial \mathbb{E}\{-g(X^M)\}}{\partial \psi} \frac{\partial \Delta_M}{\partial \psi} + \frac{\partial \mathbb{E}\{X^M - x^*\}^2}{\partial \psi} \frac{\partial \Delta_M}{\partial \psi} \) is proportional to

\[
3\mu_1^2 \psi - \frac{\Delta_M}{4}(1 + \psi \Delta^2_M)^2 - 3\mu_1 \psi^2 \Delta_M \frac{(1 - \psi \Delta^2_M)}{(1 + \psi \Delta^2_M)}
\]

which, depending on the value of \( \mu_1 \), can be positive or negative. \( \blacksquare \)

**Proof of Proposition 3.** Let \( \beta^v \) 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 the following equilibrium policy functions:

\[
x_1^R = x^* - \frac{\Delta_M}{2} - \frac{\psi \Delta_M}{1 + \psi \Delta^2_M} \mu_1 - \psi \beta^v \Phi_1; x_1^L = x^* + \frac{\Delta_M}{2} - \frac{\psi \Delta_M}{1 + \psi \Delta^2_M} \mu_1 - \psi \beta^v \Phi_1.
\]
Proof of Proposition 4

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

$$V^R(\mu) = \max_{x^R \in [0,1]} \{p(x^R, x^L, \mu)[1 + \beta(E[V^R(\mu^R)] - E[V^R(\mu^L)])] + E[V^R(\mu^L)]\}$$ (10)

$$V^L(\mu) = \max_{x^L \in [0,1]} \{\bar{V} - p(x^R, x^L, \mu)[1 + \beta(E[V^R(\mu^R)] - E[V^R(\mu^L)])] - E[V^R(\mu^L)]\}$$ (11)

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 (10) and, given $x^R = x^R(\mu)$, $V^L(\mu)$ solves (11)
2. $x^R(\mu)$ attains the RHS of (10) and $x^L(\mu)$ attains the RHS of (11)

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 Step 1 we verify that the value functions are affine in $\mu$, in Step 2 we solve for the coefficients, in Step 3 we show it is the unique DSMPE.

**Step 1.** A few lines of algebra allow to verify that $p(h^R(\mu), h^L(\mu), \mu)$ is affine in $\mu$: $p(h^R(\mu), h^L(\mu), \mu) = \overline{p}(\mu) = h_p + h_p \mu$. Then the value functions can be re-expressed in the following way:

$$V^R(\mu) = \overline{p}(\mu)(1 + \beta(\overline{p}(\mu | h^R(\mu_0)) - \overline{p}(\mu | h^R(\mu)))) + \beta \overline{p}(\mu | h^R(\mu)) + ...$$

$$V^L(\mu) = (1 - \beta)^{-1} - V^R(\mu)$$

where $\overline{p}(\mu | h^R(\mu)) = \mathbb{E}\{\overline{p}(\eta - \lambda h^L(\mu))\} = \overline{p}(\eta - \lambda h^L(\mu))$. Moreover, $\overline{p}(\mu | h^R(\mu)) - \overline{p}(\mu | h^R(\mu))$ does not depend on $\mu$. Therefore, $\overline{p}(\mu)(1 + \beta(\overline{p}(\mu | h^R(\mu)) - \overline{p}(\mu | h^R(\mu))))$ is affine in $\mu$ and, for the same reason, all subsequent terms of the summation are also affine in $\mu$. Denote by $V_1$ the slope coefficient of $V^R$.

**Step 2.** The FONCs are (the equilibrium must be interior)

$$\left\{ \begin{array}{l} \frac{dp^R}{dx^R}[1 + \beta \overline{p}V_1(h^L - x^R)] = \beta p^R V_1 \overline{\lambda} \\ \frac{dp^L}{dx^L}[1 + \beta \overline{p}V_1(x^L - h^R)] = \beta (1 - p^L)V_1 \overline{\lambda} \end{array} \right.$$
where \( p^R = p(x^R, h^L(\mu), \mu) \) and \( p^L = p(h^R(\mu), x^L, \mu) \); the envelope conditions yield

\[
V_1 = \frac{dp^R}{d\mu}[1 + \beta \bar{V} h^L - x^R] - \beta (1 - p^R) \bar{V} h^R h_1
\]
\[
V_1 = \frac{dp^L}{d\mu}[1 + \beta \bar{V} (x^L - h^R)] - \beta p^L \bar{V} h_1
\]

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 \( \Delta \) 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 \( \mu \). 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 \( V_1 = \psi(1 - \psi(\beta \bar{\lambda})^2)^{-1} \).

**Step 3a.** 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 \( \sigma' \) and \( s'_i \) future values of \( \sigma \) and \( s_i \), \( f(\mu'|x^R, x^L, \mu) = p(x^R, x^L, \mu) f(\sigma', \{s'_i\}) \) 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.

Using the results from Step 1, we show that the M-game has the minimax property by showing that Theorem 4.2 in Jásikiewicz 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 \( (x^L, x^R, \mu) \mapsto \int_0^{\mu^h} U(\mu') f(\mu'|x^R, x^L, \mu) d\mu' \) is continuous
4. There exists a Borel function \( \delta : [0, 1]^2 \times [0, \mu^h] \mapsto [0, 1] \) 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. \( \phi(U) = \int_0^{\mu^h} U(\mu) \phi(d\mu) < \infty \)

\[38\text{Note that, although the distribution of implemented policies is essentially a Bernoulli, the distribution of realized states has full support, due to the assumptions on } f(\sigma', \{s'_i\}).\]
iv. For some \( \rho \in (0, 1) \) and for every \((x^R, x^L, \mu)\)
\[
\int_0^{\mu^h} U(\mu') f(\mu'|x^R, x^L, \mu) d\mu' \leq \rho U(\mu) + \delta(x^R, x^L, \mu) \phi(U) \tag{12}
\]

Define \( \underline{f} = \inf_{(\mu', x^R, x^L, \mu')} f(\mu'|x^R, x^L, \mu) \)
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
\( U \) 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
\( \mu^h = \max \int_0^{b(\sigma) + \pi} a \sigma d\sigma \) is arbitrarily large, there exists \( \bar{\rho} < 1 \) that satisfies (12).

**Step 3b.** Therefore, the following operator can be defined
\[
Val(P[v]) = \max_{x \in [0, 1]^2} \min_{x \in [0, 1]} \{P[v](x^R, x^L, \mu)\} = \min_{x \in [0, 1]^2} \max_{x \in [0, 1]} \{P[v](x^R, x^L, \mu)\}
\]

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

**Lemma UB:** Fix \( \mu \), then for every pair of bounded, continuous, and differentiable \( P[v^1](x^R, x^L; \mu) \)
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
\[
P[v^1](\bar{h}^R, \bar{h}^L) \leq P[v^1](h^R, h^L) \leq P[v^1](\bar{h}^R, \bar{h}^L)
\]
\[
P[v^2](\bar{h}^R, \bar{h}^L) \leq P[v^2](h^R, h^L) \leq P[v^2](\bar{h}^R, \bar{h}^L)
\]
which implies
\[
P[v^1](h^R, h^L) - P[v^2](\bar{h}^R, \bar{h}^L) \leq P[v^1](h^R, \bar{h}^L) - P[v^2](h^R, \bar{h}^L) \leq \max_{(x,y) \in [0,1]^2} |P[v^1](x,y) - P[v^2](x,y)|
\]
\[
P[v^2](\bar{h}^R, \bar{h}^L) - P[v^1](h^R, h^L) \leq P[v^2](\bar{h}^R, h^L) - P[v^1](\bar{h}^R, h^L) \leq \max_{(x,y) \in [0,1]^2} |P[v^1](x,y) - P[v^2](x,y)|
\]
\[\]
\[39\]
Under the assumptions, we know that, \( \forall K = (x^R, x^L, \mu), f(\mu'|K) \) has full support. As a consequence \( \underline{f} > 0. \)
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

$$\|T[v] - T[v']\|_{\infty} = \max_{\mu \in [0, \mu^h]} |Val(P[v]) - Val(P[v'])|$$

$$\leq \max_{\mu \in [0, \mu^h]} \left\{ \max_{(x^R, x^L) \in [0, 1]^2} |P[v^1](x^R, x^L) - P[v^2](x^R, x^L)| \right\}$$

$$= \max_{\mu \in [0, \mu^h]} \left\{ \max_{(x^R, x^L) \in [0, 1]^2} \beta \left| \int_{\mu^l}^{\mu^h} (v(\mu') - v'(\mu'))f(\mu'|x^R, x^L, \mu)d\mu' \right| \right\}$$

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

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

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

**Step 3c.** 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 Step 1. To see that, combining the FONCs and the Envelope Conditions of the problem with the requirement that the value function is linear yields

$$\psi(x^* - x^R)(1 + \beta \lambda V_1(x^L - x^R)) = \beta p V_1 \lambda$$

$$\psi(x^L - x^*)(1 + \beta \lambda V_1(x^L - x^R)) = \beta (1 - p) V_1 \lambda$$

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

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

$$V_1 = \psi[1 + \beta \lambda V_1(x^L - x^R)].$$

(13)
The equation, in turn, implies that the difference $\Delta = x^L - x^R$ is independent of $\mu$. We can then set $x^L = x^R + \Delta$. The difference between the FONCs then becomes $\psi(2x^* - \Delta - 2x^L(\mu))[1 + \beta \lambda V_1 \Delta] = \beta V_1 \lambda \psi(\Delta(2x^R(\mu) + \Delta) - 2x^*(\Delta + \mu))$.

Assuming that $x^R(\mu)$ it is linear leads to the equilibrium already derived in Step 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: $-2\psi[1 + \beta \lambda V_1 \Delta] = \beta V_1 \lambda \psi \Delta 2$. This equation implies that $V_1 = -(2\beta \lambda \Delta)^{-1}.\text{Combining this with (13) yields } V_1 = -\psi/2. The value function associated with the equilibrium obtained in Step 1, instead, is $V_1 = \psi(1 - \psi \beta \lambda \Delta)$: that is a contradiction. As a consequence, the infinite horizon of the M-game must have only one DSMPE.

Proof of Corollary 1: Part i) and the first part of ii) directly follow from inspecting $\Delta_\infty$, the last part follows from $\Delta^2_M[1/4 + 3(\psi \mu_1)^2(1 + \psi \Delta^2_M)^{-2}] > \Delta^2_\infty[1/4 + 3(\psi \mu_1)^2(1 + \psi \Delta^2_\infty)^{-2}]$.

Proof of Proposition 5: Since $\beta = (1 - \beta)^{-1}, (3)$ can be rewritten as

$$\Delta_M[1 - \beta + \psi \beta \lambda \Delta_M] - \beta \lambda = 0$$

since $\Delta_\infty = \beta \lambda$, if one proves that $\psi \lambda \Delta_M < 1$, then $\Delta_M > \Delta_\infty$ and, by proposition 2, $\mathbb{E}\{-g(X^M)\}$ is larger in the two period model. To see that $\psi \lambda \Delta_M < 1$, notice that, $p < 1$ implies $\psi \mu < 1/2$. Since $\mu = \eta - \lambda x > 0, \mu > \lambda$ implies $\psi \lambda < 1$, 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 $x^R_2 \in \text{arg max} \pi_2(x_1), x^L_2 \in \text{arg max}\{1 - \pi_2(x_1)\}$, where $\pi_2(x_1) = 1/2 + \varphi[g(x^R_2) - g(x^L_2) + \mu_2]$; the FONC of the problem define the solution.

ii) $x^R_1$ and $x^L_1$ solve

$$x^R_1 \in \text{arg max}_{x \in [0,1]} \{\pi_1(x_0) + \beta \mathbb{E}\{\hat{\pi}_2(X^C)\}\}$$

$$x^L_1 \in \text{arg max}_{x \in [0,1]} 1 + \beta - \{\pi_1(x_0) + \beta \mathbb{E}\{\hat{\pi}_2(X^C)\}\}$$
where $\mathbb{E}\{\hat{\pi}_2(X^C)\} = 1/2 + \varphi \mathbb{E}\{\eta_2\} - \varphi \mathbb{E}\{X^C\}$, 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

$$
\begin{align*}
\frac{d}{dx_1} \pi_1(x_0)[1 + \beta \varphi \lambda(x_1^R - x_1^L)] + \beta \varphi \lambda \pi_1(x_0) &= 0 \\
\frac{d}{dx_1} \pi_1(x_0)[1 + \beta \varphi \lambda(x_1^R - x_1^L)] + \beta \varphi \lambda(1 - \pi_1(x_0)) &= 0
\end{align*}
$$

(14)

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, $\xi$. Since $X^C = x_1^L - \hat{\pi}_1 \Delta_C$, $X^C \sim U_{[\xi_C, \xi]}$, where $[\xi_C, \xi] = \left[x^E_C - \frac{\varphi \Delta_C}{2}, x^E_C + \frac{\varphi \Delta_C}{2}\right]$ and $x^E_C = x^* - \frac{\varphi \Delta_C}{1 + \varphi \Delta_C} \mu_1$. $\mathbb{E}\{- (X^C - x^*)^2\}$ is then

$$
\Delta_C^2 \left[\varphi^2/(12 \psi^2) + 4(\varphi \mu_1)^2(1 + \varphi \Delta_C^2)^{-2}\right]
$$

which is increasing in $\psi^{-1}$ and $\Delta_C^2$.

\textbf{Proof of Proposition 7 i).} The recursive formulation of the problem solved by $R$ and $L$ under (C) is given by:

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

(15)

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

(16)

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 Step 1 we verify that the value functions are affine in $\mu$, and in Step 2 we solve for the coefficients, and in Step 3 we show that this is the unique DSMPE.

\textbf{Step 1.} A few lines of algebra allow to verify that $\pi(h^R(\mu), h^L(\mu), \mu)$ is an affine function of $\mu$: $\pi(h^R(\mu), h^L(\mu), \mu) = \pi(\mu) = h_p + h_p \mu$, where the realized value of $\pi(\mu)$ is $\pi(\mu) + \varphi \xi$. Then the value functions can be re-expressed in the following way:

$$
V^R(\mu_0) = \pi(\mu_0) + \beta \mathbb{E}\{\pi(\mu_1)\} + \beta^2 \beta \mathbb{E}\{\pi(\mu_2)\} + ... \\
V^L(\mu_0) = (1 - \beta)^{-1} - V^R(\mu_0)
$$

where

$$
\mathbb{E}\{\pi(\mu_1)\} = \mathbb{E}\{\pi(\eta - \lambda(\pi(\mu_{t-1}) + \varphi \xi)h^R(\mu_{t-1}) + (1 - \pi(\mu_{t-1}) - \varphi \xi)h^L(\mu_{t-1}))\}\}
$$
simplifies to \( \pi(\eta - \lambda(\mu_{t-1})(h^R_0 - h^L_0) + h^L_0 + h_1\mu_{t-1})) \), which is affine in \( \mu_{t-1} \). Therefore, all the terms in the summation are compositions of affine functions, therefore affine. Denote by \( V_1 \) the slope coefficient of \( V \).

**Step 2.** The FONCs are (the equilibrium must be interior)

\[
\begin{align*}
\frac{d\pi^R}{dx} \left[ 1 + \beta \lambda V_1(h^L - x^R) \right] &= \beta \pi^R V_1' \lambda \\
\frac{d\pi^L}{dx} \left[ 1 + \beta \lambda V_1(x^L - h^R) \right] &= \beta (1 - \pi^L) V_1' \lambda
\end{align*}
\]

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

\[
\begin{align*}
V_1 &= \frac{d\pi^R}{d\mu} \left[ 1 + \beta \lambda V_1(h^L - x^R) \right] - \beta (1 - \pi^R) \lambda V_1 h_1 \\
V_1 &= \frac{d\pi^L}{d\mu} \left[ 1 + \beta \lambda V_1(x^L - h^R) \right] - \beta \pi^L \lambda V_1 h_1
\end{align*}
\]

re-expressing these 4 equations as functions of \( h^R_0 - h^L_0 = \Delta_\infty, h^R_0 + h^L_0, \) 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 \( \Delta_\infty \) 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 \( h^R_0 + h^L_0, \) and \( h_1, \) which is affine in \( \mu. \) Setting \( \mu = 0 \) gives \( h^R_0 + h^L_0, \) which then allows us to solve for \( h_1. \) The slope of the value function is then \( V_1 = \varphi(1 - \varphi(\beta \lambda)^2)^{-1}. \)

**Step 3a.** 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 \( g(\mu'|x^R, x^L, \mu) = f(A', s') \frac{\psi}{\varphi(x^R - x^L)}. \) The assumptions in Jaskiewicz and Nowak (2006) are still satisfied (following same steps as in the proof in Proposition 3, Step 3a, choose \( \delta(x^R, x^L, \mu) = f\pi(x^R, x^L, \mu), \) and Lemma UB holds. As a consequence, following the same steps as in Proposition 3, Step 3b, 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.

**Step 3b.** It remains to show that no other pair of policy functions \( (x^R(\mu), x^L(\mu)) \) can generate the value function obtained in Step 1. To see that, combining the FONCs and the Envelope Conditions of the problem with the requirement that the value function is linear
Since \( \Delta = x^L - x^R \) is independent of \( \mu \), and subsequently \( V_1 = -\varphi/2 \). Since the value function associated with the equilibrium obtained in Step 1 is \( V_1 = \varphi(1 - \varphi \beta \Delta \infty) \), a contradiction is obtained. As a consequence, the infinite horizon of the C-game must have only one DSMPE.

**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 \( \mathbb{E}\{-g(X^C_1)\} > \mathbb{E}\{-g(X^M_1)\} \), notice that the difference can be rewritten as

\[
\frac{\Delta^2_C}{4} - \frac{\Delta^2_M}{12} \left( \frac{\varphi}{\psi} \right)^2 - 4\mu_1^2 \left( \frac{\Delta_C \varphi}{1 + \varphi \Delta_C^2} \right)^2 + 3\mu_1^2 \left( \frac{\Delta_M \psi}{1 + \psi \Delta_M^2} \right)^2 > 0.
\]

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 \( \psi \Delta_M = 1 - \Delta_M / \beta \lambda > \varphi \Delta_C = 1 - \Delta_C / \beta \lambda \), that \( \Delta_C \varphi/(1 + \varphi \Delta_C^2) < \Delta_M \psi/(1 + \psi \Delta_M^2) \). As a consequence, a sufficient condition for (18) is

\[
\Delta^2_C - \frac{\Delta^2_M}{3} \left( \frac{\varphi}{\psi} \right)^2 - 4\mu_1^2 \left( \frac{\Delta_C \varphi}{1 + \varphi \Delta_C^2} \right)^2 > 0.
\]

Multiplying each side by \( \Delta^2_M \) and using (6), we can derive the following lower bound for (18)

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

**Case 1:** \( \frac{1}{\psi} < \frac{1}{\varphi} - \frac{1}{\psi} \). (19) 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 \( \psi \) is \( \varphi(1 - 2\varphi \mu^h)^{-1} \) Combining this with the fact that \( \frac{1}{\psi} < \frac{1}{\varphi} - \frac{1}{\psi} \), 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

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

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 \( 1 - (1/2)(4/3) > 0 \), (19) holds.
Case 2: \( \frac{1}{\psi} > \frac{1}{\varphi} - \frac{1}{\psi} \), (19) becomes
\[
1 - \left( \frac{\Delta C \varphi}{\Delta M \psi} \right)^2 \left( \frac{4}{3} + \left( \frac{\psi}{\varphi} \right)^2 + 2 \frac{\psi}{\varphi} \right) > 0
\] (21)

Direct inspection allows us to conclude that the expression is minimized when \( \varphi \) is smallest with respect to \( \psi \). Combining this with the restriction \( \frac{1}{\psi} > \frac{1}{\varphi} - \frac{1}{\psi} \), one obtains \( 2\varphi = \psi \).

As a consequence \( 1 - \left( \frac{\Delta C \varphi}{\Delta M \psi} \right)^2 \left( \frac{4}{3} + \left( \frac{\psi}{\varphi} \right)^2 + 2 \frac{\psi}{\varphi} \right) > 0 \) is a lower bound for the LHS of (21). Using (3), (4) and \( 2\varphi = \psi \), one obtains that \( \Delta M / \Delta C \) can be re-expressed as \( \frac{1 - 2\varphi \Delta^2 C}{1 - \varphi \Delta^2 C} \).

Combining that and \( \Delta M / \Delta C < 1 \) yields \( \left( \frac{\Delta C \varphi}{\Delta M \psi} \right)^2 < \frac{1}{2} \) and, as before, we have that \( 1 - (1/2)(4/3) > 0 \) is a lower bound for (19).

Since public good inefficiency in the two-period model are lower under \( C \), despite a larger platform divergence, the same must also hold in the infinite horizon model, where platform divergence is the same.

To see why redistribution is higher, notice that, regardless of the time horizon,
\[
\mathbb{E}\{\tau^M_t\} = (1 - p_t)\tau, \quad \mathbb{E}\{X^M_t\} = x^* - (2p_t - 1)\Delta_M, \quad \mathbb{E}\{X^C_t\} = x^* - (2p_t - 1)\Delta_C
\]

Since \( \psi > \varphi \),
\[
p_t = \frac{1}{2} + \mu_t \psi (1 + \psi \Delta^2 M)^{-1} > \pi_t = \frac{1}{2} + \mu_t \varphi (1 + \varphi \Delta^2 C)^{-1}
\]
\[
(2p_t - 1)\Delta_M = \mu_t \psi \Delta_M (1 + \psi \Delta^2 M)^{-1} > (2\pi_t - 1)\Delta_M = \mu_t \varphi \Delta_C (1 + \varphi \Delta^2 C)^{-1}
\]

which imply that \( \mathbb{E}\{\tau^M_t\} < \mathbb{E}\{\tau_t^M\}, \mathbb{E}\{X^M_t\} < \mathbb{E}\{X^C_t\} < x^* \).

**Proof of Proposition [9]**

1. The FONC of the problem define the solution. For \( t = 1 \) \( X_t^R \) and \( X_t^L \)

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

where \( \mathbb{E}\{p(x_t^S)\} = 1/2 + \psi \mathbb{E}\{\eta\} + \psi \bar{x} \mathbb{E}\{X_t^S\} = 1/2 + \psi \mathbb{E}\{\eta\} + \psi \bar{x} [\pi_1 x_t^R + (1 - \pi_1) x_t^L] \) follows from the observation that \( g(x_t^R_2) - g(x_t^L_2) = 0 \). The FONC are of the problem (which are also sufficient under the assumptions) define the following system
\[
\begin{cases}
\frac{d}{dx_t} p_1(x_0) + \beta \psi \bar{x} \left\{ \pi_1 + (x_t^R - x_t^L) \frac{d}{dx_t} x_t^S \right\} = 0 \\
\frac{d}{dx_t} p_1(x_0) + \beta \psi \bar{x} \left\{ (1 - \pi_1) + (x_t^R - x_t^L) \frac{d}{dx_t} x_t^R \right\} = 0
\end{cases}
\]
which simplifies to
\[
\begin{aligned}
&\left\{\begin{array}{l}
x^* - x_1^R + \beta \lambda \left\{ \pi_1 + \Delta_S \varphi (x^* - x_1^R) \right\} = 0 \\
x_1^L - x^* + \beta \lambda \left\{ (1 - \pi_1) + (x_1^R - x_1^L) \varphi (x_1^L - x^*) \right\} = 0
\end{array}\right.
\end{aligned}
\]
which is the same system as in (14).

ii) First, observe that, by part i), we must have, \(E\{-g(X_S)\} = E\{-g(X_C)\} < E\{-g(X_M)\}\). Next, we show that \(\tau^C_1 > \tau^M_1 > \tau^S_1\). 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
\[
\pi(x_M^R, x_M^L, \mu_1) < p(x_M^R, x_M^L, \mu_1) = \frac{1}{2} + \frac{\mu_1 \psi}{1 + \psi \Delta_M^2} < \frac{1}{2} + \frac{\mu_1 \psi}{1 + \psi \Delta_C^2} = p(x_S^R, x_S^L, \mu_1).
\]
To complete the proof, notice that, in \(t = 2\), \(E\{-g(X_S)\} = E\{-g(X_C)\} = E\{-g(X_M)\} = 0\) and, since \(E\{X_1^M\} < E\{X_1^C\}\)
\[
E\{p_2\} = 1/2 + \psi E\{\eta\} + \psi \lambda E\{X_1^M\} > E(\pi_2) = 1/2 + \varphi E\{\eta\} + \varphi \lambda E\{X_1^C\}
\]
\(E\{\tau^C_2\} > E\{\tau^M_2\} > E\{\tau^S_2\}\).

iii) The generic welfare function is
\[
W(x, \tau) = g(x) + \gamma \int_{b+\tau}^{1/2\sigma} \sigma(a - \tau) \, da + b(\tau, A) \int_{-1/2\sigma}^{b+\tau} \sigma \, da
\]
\[
= g(x) + \gamma \frac{\tau - b}{4} + (1 - \gamma)(1/2 - \sigma b - \sigma \tau)
\]
For \(\gamma\) large enough, the welfare ranking among constitutions is the same as with the utilitarian criterion. For \(\gamma\) low enough, larger redistribution is welfare improving, in which case C welfare dominates S.

\[\blacksquare\]

**Proof of Proposition 10** The effect of higher inequality (lower \(\sigma\)) is to increase \(b()\) (see Lemma 2) and, as a consequence, to decrease \(\eta_t\). On the other hand, we have that
\[
\frac{\partial^2 b}{\partial^2 \sigma} = \frac{3(b + \tau)^3}{(1/2 + 2\sigma(b + \tau))^2} \mathbb{I}_{(\tau \leq 1/2\sigma)}
\]
The numerator is decreasing in \(\sigma\), 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, \(\lambda\) increases for two reasons: (1) higher \(b()\), (2) stronger effect of informational asymmetry. The overall effect of a first order stochastic shift in the distribution of \(\sigma\) is to decrease \(E\{\mu\}\), thereby making electoral competition more balanced, but also increasing the volatility of implemented policy.

\[\blacksquare\]
References


Azzimonti-Renzo, M. (2009), Barriers to investment in polarized societies, mimeo.


Battaglini, M. (2010), Dynamic electoral competition and constitutional design, mimeo.


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


