Access Econ
Scholarly communication without the middle men

Submission Number: PET11-11-00107

Dynamic political distortions under alternative constitutional settings

Carlo Prato
Northwestern University

Abstract

This paper looks at a dynamic model of electoral competition with Downsian voters in which 2 political actors can commit to public employment, not to redistribution (over which they have diverging preferences). Current employment status affects voters' information on the state of the public sector and, indirectly, their political preferences. This dynamic dependence generates inefficiency in political actors' platforms and implemented policies. A constitution determines how the electoral outcome maps into policy making power. Two constitutions are compared: majoritarian and consensual (Lijphart, 1999). The latter is associated with more political polarization (but only when the time horizon is finite), less inefficient public good provision (provided discount factors are close enough) and more redistribution. An optimal constitution allocates policy making rights over the non-commitment policy dimension in a "majoritarian" way, while forces political actors to compromise over the commitment dimension. The dependence of distortions on income inequality, inefficiency of redistribution, voters' responsiveness, and political actors' patience is also analyzed. Finally, the model also delivers implications on the relationship between inequality, redistribution, and political polarization that are consistent with recent empirical evidence.

Submitted: February 28, 2011.

Dynamic political distortions under alternative constitutional settings.

Carlo Prato*

Northwestern University

carlo.prato@u.northwestern.edu

February 28, 2011

Abstract

This paper looks at a dynamic model of electoral competition with Downsian voters in which 2 political actors can commit to public employment, not to redistribution (over which they have diverging preferences). Current employment status affects voters' information on the state of the public sector and, indirectly, their political preferences. This dynamic dependence generates inefficiency in political actors' platforms and implemented policies. A constitution determines how the electoral outcome maps into policy making power. Two constitutions are compared: majoritarian and consensual (Lijphart, 1999). The latter is associated with more political polarization (but only when the time horizon is finite), less inefficient public good provision (provided discount factors are close enough) and more redistribution. An optimal constitution allocates policy making rights over the non-commitment policy dimension in a "majoritarian" way, while forces political actors to compromise over the commitment dimension. The dependence of distortions on income inequality, inefficiency of redistribution, voters' responsiveness, and political actors' patience is also analyzed. Finally, the model also delivers implications on the relationship between inequality, redistribution, and political polarization that are consistent with recent empirical evidence.

Keywords: Dynamic Political Economy, Public Employment, Constitutions.
JEL Classification: D 72, D 78.

*I would like to thank David Austen-Smith, Marina Azzimonti, Marco Battaglini, Briana Chang, Eddie Dekel, Matthias Doepke, Georgy Egorov, Tim Feddersen, Simone Galperti, Edoardo Grillo, Bard Harstad, Igal Hendel, Daisuke Ikeda, Lorenzo Michelozzi, Salvatore Nunnari, Bruno Strulovici, Martin Szydlowski and seminar participants at Northwestern and European University Institute for their helpful comments. All mistakes are obviously mine.

Introduction

Implemented policies can have long term consequences that, by affecting structural aspects of a society, shape the electoral environment in which future political competition will take place. The term electoral environment captures various factors influencing voters' *political* preferences (that is, preferences over candidates or parties) and that political actors take as given when they choose to which political platforms to commit before the elections. This paper studies how the presence of a dynamic link between current implemented policies and future electoral environment affects political competition in a democracy. More specifically, it shows how this dependence generates deviations from efficiency in platforms (political polarization) and implemented policies (policy distortions) and how these distortions critically depend on the constitutional setting and other key factors. (e.g. income inequality, time horizon, political actors' discount factors).

The cleanest example of the above described mechanism is migration policy: a large inflow of poor individuals into a polity affects the future electoral popularity of various forms of welfare programs. Therefore, political actors with a stronger association to welfare programs than others might have a strategic incentive to manipulate migration policy to improve their future electoral strength. As a result, inefficient policies might arise. The mechanism linking migration policy to populist redistribution is called *Curley effect* (Glaeser and Shleifer, 2005) and owes its name to James M. Curley, four-terms serving major of Boston between 1914 and 1950, who also served as Senator and Governor of Massachusetts. Curley pursued various types of populist redistributive policies that favored the inflow into the city of poor, Catholic, Irish immigrants and caused an outflow of wealthier, Protestant, Anglo-Saxon population. By discouraging business and promoting a Anglo-Saxon emigration, his policies hampered Boston's economy. Nevertheless, by enlarging and consolidating his electoral base, they played a key role in building a successful political career, despite the widely available evidence of Curley's deep corruption. Glaeser and Shleifer, 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 conflicts in racially divided polities.[1]

This paper argues that the same logic of the Curley effect can be applied to other important policy domains, and can constitute an independent channel for political failures. More specifically, it analyzes the distortions that the presence of a dynamic link between current public employment and future electoral environment can generate in the context of a public finance model of dynamic electoral competition with public goods and redistribution. The first contribution of the paper is to show how distortions arise because of the interplay of three key elements: first, political actors are *differentiated* (Krasa and Polborn, 2009): either they have different preferences on a policy dimension (in this model, over redistribution), or they have different abilities in delivering utility to voters. Second, announced political platforms are related to implemented policies in a systematic way: a commitment assumption on platforms (in this paper, on public employment) and a *constitution*, which

[1] Other examples include the city of Detroit under Coleman Young (1973-1993), or the state of Zimbabwe under Robert Mugabe (1987-present).

in this model maps electoral outcomes into policy making rights, provide a link between platforms and implemented policies. Third, implemented policies systematically affect future political preferences: in this paper, public employment affects the precision of voters' information about the state of the public sector. The latter, by determining how inefficient redistribution is, influences the future *ex ante* (that is, before they commit to political platforms) relative strength of the political actors competing for office.

This paper is part of a broader research agenda aimed at identifying other instances of this type of distortions (which I call dynamic political distortions) as independent sources of political failures, a topic so far almost unexplored in political economy and development economics. More important, the relationship between these distortions and important aspects of the political process goes against what has been so far concluded by a large and growing body of literature. Specifically, several authors have identified short termism (that is, the inability of incorporating future consequences of current policies) as one of the main sources of political failures. Recent contributions on this literature have explored such failures in dynamic environments, achieving remarkable results in explaining various empirical phenomena. Most notably, Battaglini and Coate (2008) on the accumulation and fluctuation of public debt, Azzimonti (2009) on inefficiently high investment taxation in polarized polities, Aidt and Dutta (2007) on the excessively low implementation of policies with long term benefits, Besley and Persson (2010) on inefficient investment in fiscal capacity, Acemoglu et al (2009) labor supply distortions induced by redistribution policies. In all these papers, the source of the distortion lies in the inability of each period's policy makers to be dynamically consistent: folding future poltical power is an uncertain event, while current payoffs are fully appropriable. In other words, there is a wedge between today's benefits, coming from policies or office holding, and tomorrow's benefits from policies, which will not, with some probability, be enjoyed by the current decision maker. As a consequence, having more persistence in political power and/or more patient political actors would mitigate these distortions.[2] The reason is intuitive: politicians would then internalize a larger share of the long term benefits of efficient policies they would not otherwise implement and a larger share of the long term costs of inefficient policies they would otherwise implement. This logic seems partially reversed in the Curley effect: as actors get more patient, the importance of the future electoral environment might get stronger, thereby leading to larger distortions. On the other hand, having increased persistence in political power should decrease the extent of these distortions as the electoral environment becomes less critical for the electoral outcome. The second contribution of this paper is to show that only the first intuition is correct: more patience does lead to larger dynamic political distortions. On the other hand, increasing the *a priori* relative strength of a political actor over the other will increase, not decrease, the severity of dynamic political distortions.

This paper describes a model of dynamic electoral competition between differentiated political actors with a 2 dimensional policy space. In every period, office is associated with policy making rights over public employment (which corresponds to public good provision)

[2] Increasing persistence means, in Battaglini and Coate's model, assuming some degree of autocorrelation in the identity of the proposer in their legislative bargaining game; in Aidt and Dutta's, it means ensuring that the identity of the politician in charge is the same in every period; in Azzimonti's and Besley and Persson's, it means ensuring that the incumbent advantage is large enough.

and redistribution from rich to poor in a given polity. Political actors have diverging preferences over the latter dimension and cannot *ex ante* commit to a different level than their preferred.[3] Instead, actors can *ex ante* commit to a specific public employment level, over which they do not have preferences. The latter has a systematic effect on voters' information about the inefficiency of future redistribution. Political actors have then the incentive to *ex ante* manipulate public employment for electoral purposes. The assumptions on voting behavior[4] imply that the electoral process will push platforms towards the efficient level. (like in a standard Downsian model) Therefore, political actors trade off current electoral strength with a better future electoral environment. Platforms and implemented policies will then, in general, differ from the efficient ones (dynamic political distortions).

The third contribution of this paper is to analyze how the type of constitution, (along with time horizon, discount factor and income inequality affect the extent of these distortions. In particular, building on the distinction in Lijphart's classic book *Patterns of Democracy (1999)*, two types of constitution are compared: a majoritarian constitution, **(M),** and a consensual constitution, **(C)**. In the former the majority winner gets to enjoy full policy making rights. In a consensual democracy, instead, political actors negotiate over the future implemented policies with bargaining power that is proportional to their vote share.

Three types of distortions are considered: on the platform level, I consider the divergence between announced public employment levels, which I relate to a measure of political polarization. On the implemented public employment level, I consider the expected quadratic deviation from the efficient level. On the implemented redistribution (which is, by assumption, inefficient) level, I simply consider its *ex ante* expectation. While implemented redistribution depends on political actors' preferences, and is not, in a strict sense, a political failure, looking at it will be relevant for welfare analysis and will generate interesting empirical predictions. In a 2 period model, platforms converge at the efficient level only in the terminal period. In the first period, majoritarian constitutions display lower platform divergence, higher expected public employment inefficiency and lower redistribution inefficiency with respect to consensual constitutions. These inter-constitutional differences weaken as actors become more patient or as the extent of the informational asymmetry increases - that is, as the effect of current policies on future discounted electoral benefits gets stronger. Moreover, when the time horizon is infinite, platform divergence, which is larger than in a 2-period model, no longer depends on the type of constitution, highlighting an interesting long term neutrality property; the expected inefficiencies in implemented policies, instead, are still different. The difference between constitutions is entirely driven by the allocation of policy making rights over the public employment dimension, and is independent of how actors share power over redistribution. The latter is a quite surprising fact, given that political actors only care about redistribution. A consensual allocation of these

[3] This assumption, although extreme, is just meant to produce a cleaner analysis, but is not necessary for the existence of dynamic distortions. As long as political actors have diverging preferences and some degree of limited commitment on redistribution, the logic of Hotelling competition will fail.

[4] Voting follows one of the standard version of the probabilistic voting model, pioneered by Lindbeck and Weibull (1987) and Dixit and Lodregan (1996) and extensively applied in Persson and Tabellini's *Political Economics* book

rights, combined with a majoritarian allocation of power over implemented redistribution is showed to welfare dominate both consensual and majoritarian constitutions. Finally, equilibrium outcome in a consensual democracy is much more sensitive to heterogeneity in the discount factors than the one in a majoritaran democracy. While heterogeneity reduces platform divergence and public employment inefficiencies in majoritarian constitutions, it creates asymmetries in the outcome of the bargaining among political actors. These asymmetries are shown to *ex ante* increase incentives to platform divergence and *ex post* increases the inefficiency of the implemented public employment.

The fourth contribution of this paper is to deliver three important empirical implications connecting inequality, redistribution and political polarization: first, higher inequality increases political polarization; second, higher inequality can increase or decrease implemented redistribution, and an inverse relationship is more likely to arise in a majoritarian democracy. Third, implemented redistribution is higher in consensual democracies. Each of these prediction is line and can provide a way to interpret the findings of various recent empirical contributions (Kelly and Enns, 2010; Mc Carty, Poole and Rosenthal, 2008; Persson and Tabellini, 2003).

The paper is organized as follows: the first section describes the basic economic environment, voting behavior and the political process. Subsequently, the outcome of a majoritarian constitution under different assumptions on the time horizon (2 periods vs infinite horizon) is analyzed, and the main comparative static results are derived. Section 3 describes the model of consensual constitution, the solution of the 2 period and infinite horizon model, and compares the two constitutional settings. In Section 4, the main empirical implications of the model are discussed. Section 5 considers the heterogeneous discount factor case in the 2 period version of each constitutional setting, and shows how the constitutional comparison changes. Section 6 reviews the relevant literature, and Section 7 concludes. All proofs are contained in the Appendix.

1 Economic and electoral environment

Economic environment: A society is composed by a unit-mass continuum of citizens and lasts for multiple periods. In every period, each citizen i is endowed with one unit of labor, the only input in the economy, and one unit of a capital good. The latter produces a per period stochastic payoff which can be, with equal probability, 0 or $\omega > 0$ consumption units. I call "rich" a citizen with positive capital income and "poor" a citizen with 0 capital income. In every period labor can be supplied in either the public sector ($i \in G_t$) or in the private sector ($i \in P_t$). The private sector, whose size is denoted by

$$x_t = \int_{P_t} di$$

turns each unit of labor into $Q \in (0, 1)$ units of consumption good. The labor market in the private sector is competitive and there is full mobility of labor. As a result, in every period the gross wage in both sectors equals Q. Since the labor demand from the private sector is undetermined, the labor demand from the public sector, $1 - x_t$, fully determines sectorial allocation.

The public sector turns labor into a public good, g_t, with a technology exhibiting unitary marginal productivity. The public sector is financed through a proportional labor income tax τ_t. Therefore, the budget constraint for G_t implies $\tau_t = 1 - x_t$. Proportional income taxation is associated with inefficiencies, captured by a simple quadratic cost component $\tau_t^2/2$.[5]

The government can also administer a transfer to poor citizens of amount b_t, financed by a capital income tax θ_t. The public sector turns every unit of revenues into $1 - q_t$ units of transfers, where q_t is drawn from the interval $\Omega_t^q \in (0, 1)$. q_t is the marginal cost of public funds raised through capital taxation and captures various factors affecting the productivity of the public sector. Both Ω_t^q and q_t are assumed to be stochastic. The budget constraint for b_t is then $(1 - q_t)\theta_t\omega/2 = b_t/2$, which implies

$$\theta_t = \frac{1}{\omega}\frac{b_t}{1 - q_t} \tag{1}$$

Both consumption good and public good enter citizens' payoff linearly; as a result, a rich citizen's payoff is $Qx_t + g_t - \tau_t^2/2 + \omega(1 - \theta_t)$ and a poor citizen's one is $Qx_t + g_t - \tau_t^2/2 + b_t$.

Substituting he production function for g_t, (1), and the budget constraints for G_t, the per period indirect utility of a rich and a poor citizen are, respectively,

$$\begin{aligned} v^r(x_t, b_t) &= x_t(Q - x_t/2) + 1/2 + \omega - \frac{b_t}{1 - q_t} \\ v^p(x_t, b_t) &= x_t(Q - x_t/2) + 1/2 + b_t \end{aligned}$$

First Best. A utilitarian social planner will implement, in every period, $x^* = Q$ and $b^* = 0$. To see this, notice that the utilitarian social welfare function is given by

$$W(x_t, b_t) = x_t(Q - x_t/2) + 1/2\left\{1 + \omega - \frac{q_t}{1 - q_t}b_t\right\}$$

and

$$x^*, b^* = \arg \max_{\{x_t, b_t\}_{t=0}^{\infty}} \sum_{t=0}^{T} \beta^t W(x_t, b_t).$$

It is also useful to notice that, because of separability of $W(x_t, b_t)$, a Rawlsian social planner would also choose $x^* = Q$, and so will any other social planner whose preferences are a mixture between utilitarian and rawlsian. This will be the benchmark against which the

[5]These assumptions imply that the x-related payoff for each voter is quadratic and strictly concave in x, which makes the mathematical structure of the model very convenient. Numerical simulations of the model shows that the same basic qualitative results derived int this paper hold with a more general formualtion with constant marginal productivity in the private sector and strictly decreasing marginal productivity in the public sector.

concept of distortion will be defined.

Political process. In each period two political actors ($j \in \{R, L\}$) compete in elections for office. Being in office is associated with policy making rights over b and x. At the beginning of every period t, each political actor commits to a policy platform for x_t. On the other hand, no commitment is possible for b_t, over which the two actors have divergent preferences: R prefers no redistribution while L prefers a fully egalitarian society.[6] The lack of commitment assumption implies that R, if alone in power, will implement $b_t = 0$ and L, if alone in power, would implement a fully egalitarian policy, which solves $\omega - \bar{b}_t/(1-q_t) = \bar{b}_t$, and therefore implies

$$\bar{b}_t = \frac{1-q_t}{2-q_t}\omega$$

Political actors do not intrinsically care about the relative size of each sector and derive a payoff (normalized within the unit interval) that is linear in the distance from their preferred transfer level. As a result, R's per period payoff ranges between 1 (when $b_t = 0$ is implemented) and 0 (when $b_t = \bar{b}_t$ is implemented), while the opposite is true for L.[7]

Together, these assumption are meant to capture 2 stylized facts. The first is that parties and candidates typically differ in how they balance the trade off between inequality and efficiency. The second is that pre-electoral commitment to public employment is easier than pre-electoral commitment to transfers, since redistribution can be implemented through a large variety of means and often delegated to lower level government officials, whose functions are often not determined until after the elections. This assumption can be easily relaxed to allow for partial commitment: given that R and L only care about redistribution, the full commitment assumption on public good provision has the same interpretation as in standard Downsian models. On the other hand, assuming partial commitment on redistribution will make the model more complicated, but not affect its main tradeoff. L, who prefers a policy that generates inefficiencies, will have a dynamic tradeoff (current vs future electoral strenght) and also a static trade off between policy goals and electoral appeal. Given that the paper focuses on the dynamic tradeoff, assuming partial commitment will mostly make the analysis of the fundamental mechanisms of the model less clean.

Timing and voting behavior. Every period begins with a given sectorial distribution x_{t-1}. Ω_t^q is private information of the two political actors. After the announcement of the platforms (x_t^R, x_t^L), all the G_{t-1}-workers observe q_t, while P_{t-1}-workers only observe $s_t = q_t + \varepsilon_t$, where ε_t is a zero-mean random variable with support $[-e, e]$.[8]

Each voter i then computes $v(x_t^R, \tilde{b}_t^R)$ and $v(x_t^L, \tilde{b}_t^L)$, the expected per period payoff associated with each actor's announced x and conjectured b. Voters perfectly anticipate that $b^R = 0$ and $b_L = \bar{b}_t$. The latter depends on q_t, and can only be conjectured by workers

[6] Note that, in this setting, egalitarian and rawlsian preferences have the same representation.

[7] As a result, in every period the sum of the political actors' payoffs equals 1.

[8] Notice that for the equilibrium analysis I do not need to assume that the knowledge about c_t is ex post accurate, but only that current public sector workers have an a priori more precise understanding of what is the state of the public sector.

in G_{t-1}. Voting behavior is probabilistic: i votes for R iff

$$v(x_t^R, \tilde{b}_t^R) > v(x_t^L, \tilde{b}_t^L) + \xi_t + \delta_t^i$$

where $\hat{\xi}$ is the realization of a stationary zero-mean aggregate preference shock ξ_t and δ_t^i is the realization of a stationary zero-mean, idiosyncratic preference shock δ. As in one of the standard formulations of the probabilistic voting model,[9] I assume that both shocks are uniformly distributed; more precisely, the aggregate shock is uniformly distributed over the interval $[-1/2\psi, 1/2\psi]$ and the idiosyncratic shock is uniformly distributed over the interval $[-1/2\varphi, 1/2\varphi]$. Without knowing how electoral outcome maps into policies it is not fully transparent how strong the assumptions on voting behavior are. As it will become fully clear in the rest of the paper, they are arguably quite natural under both constitutional settings: a voter will try to push the implemented policies in the direction that he expects to benefit him the most.

For simplicity, I assume no serial correlation in capital income over time, so each citizen has equal probability of being rich or poor.[10] For the $1 - x_{t-1}$ workers who observe q_t, the expected payoff from R's platform is $v(x_t^R, 0) = x_t^R(Q - x_t^R/2) + 1/2 + \omega_t/2$, the one from L's is $v(x_t^R, \bar{b}_t) = x_t^L(Q - x_t^L/2) + 1/2 + \omega/2 - I(q_t)$, where

$$I(q_t) = \frac{q_t}{2(2 - q_t)}\omega$$

is the *redistribution inefficiency* associated with L's egalitarian redistribution. As a result, R's realized vote share among the G_{t-1}-workers is

$$\Pr[\delta < d(x_t^R, x_t^L) + I(b_t) - \hat{\xi}_t] = 1/2 + \varphi[d(x_t^R, x_t^L) + I(q_t) - \xi_t]$$

where $d(x_t^R, x_t^L)$ is the x-related payoff differential $d(x_t^R, x_t^L) = x_t^R(Q - x_t^R/2) - x_t^L(Q - x_t^L/2)$.

Similar computation lead to conclude that R's realized vote share among the P_{t-1}-workers is

$$1/2 + \varphi[d(x_t^R, x_t^L) + \tilde{I}(s_t) - \xi_t]$$

where

$$\tilde{I}(s_t) = E[I(q_t)|s_t] = \int_{s-e}^{s+e} I(z)dF_\varepsilon(z)$$

is the expectation of the redistribution inefficiency conditional on s_t. R's total realized vote share is then the sum of three, separated components

$$\hat{\pi}_t = 1/2 + \varphi[\underbrace{d(x_t^R, x_t^L)}_{\substack{x\text{-related} \\ \text{payoff}}} + \underbrace{I(q_t)}_{\substack{R\text{'s structural} \\ \text{advantage}}} + \underbrace{\lambda_t x_{t-1}}_{\substack{\text{Informational} \\ \text{wedge}}} - \underbrace{\xi_t}_{\substack{\text{Preference} \\ \text{shock}}}]$$

[9] see Lindbeck and Weibull (1987) and Persson and Tabellini (2002).

[10] Assuming serial correlation would have the only consequence of making the notation heavier without any substantial effect on the structure of the model.

where $\lambda_t = \tilde{I}(s_t) - I(q_t)$, which I call *informational wedge*, is the marginal effect of a change in the initial size of the private sector on R's electoral strength.[11] The following lemma states a very useful fact:

Lemma 1 *The unconditional expectation of the informational wedge, $\bar{\lambda}$, is strictly positive.*

This dependence, which comes from the asymmetric information on the expected redistribution inefficiency associated with L's political platform, implies that, *in expectation*, R will have an electoral advantage over L when the initial size of the private sector is large. The reason is that private sector workers will anticipate a larger inefficiency in L's transfer, due to their imperfect information on the cost of the latter. The presence of this dependence is the source of dynamic political distortions.

In order to keep the political actors' problem well behaved, that is their objective functions to be continuous and differentiable, I need to assume that φ and ψ are related in such a way that ensures that, for every initial value of x, both politicians have a positive probability of obtaining a majority of the votes and a zero probability of obtaining all the votes. A detailed description of this assumption is contained in the Appendix.

2 Policy making under majoritarian constitution

The only piece missing to cpmplete the model is a rule that specifies, for every possible electoral result, an allocation of policy making rights over x and r. As already anticipated, a constitution within the context of this model provides exactly such mapping.

In this paper I will look at Arend Lijphart's fundamental dichotomy between consensual democracy, from now onward **(C),** and majoritarian democracy, from now onward **(M)**. Although both types of constitutions have already been considered in the economic literature,[12] **(M)** is by far the simplest and less controversial in terms of modeling choices. Therefore, in this section I formally describe the electoral game under that type constitution, solve it under alternative assumptions about the time horizon, and finally show how changes in the main parameters of the model affect the equilibrium. In the subsequent section, I define **(C)**, solve the corresponding electoral game, and compare the two constitutions. For the latter task, I also introduce a hybrid constitutional type, denoted by **(S)** and called semi-consensual. By combining elements from the previous two, **(S)** will help understanding the role of the policy making rights over public employment (over which commitment is possible) versus redistribution (over which preferences are defined)

[11] Before solving for the equilibrium, I need to verify the internal consistency of the information partition. That is, I need to make sure that the players who are imperfectly informed cannot learn about c_t from the observed equilibrium platforms. This is true since, lacking information on Ω_{c_t}, voters cannot infer the value of $\bar{\lambda}$, which, as will be clear in the rest of the paper, is necessary to extract information about λ from equilibrium platforms.

[12] Ticchi and Vindigni (2010) and Herrera and Morelli (2010) are recent examples of applications of the concept to study, respectively, the endogneous formation of constitutions and voting behavior.

in determining the difference in outcomes between **(C)** and **(M).**

Majoritarian Democracy. In a majoritarian democracy the majority winner gets full policy making rights over b and x. More formally, under **(M)**, the payoffs in every period are $1_{\{R=W\}}$; $1_{\{R=W\}}$, the implemented policies are

$$x_t^W, b_t^W = \begin{cases} x_t^R, 0 & \text{if} \quad \hat{\pi}_t \geq 1/2 \\ x_t^L, \bar{b}_t & \text{if} \quad \hat{\pi}_t < 1/2 \end{cases}$$

where W is the majority winner in the elections held in t, that is

$$W_t = \begin{cases} R & \text{if} \quad \hat{\pi}_t \geq 1/2 \\ L & \text{if} \quad \hat{\pi}_t < 1/2 \end{cases}$$

In this context, political actors' expected payoff is the probability of winning a majority of the votes; this is given, for R, by

$$p_t = \Pr\left[\hat{\pi}_t > \frac{1}{2}\right] = 1/2 + \psi[d(x_t^R, x_t^L) + I(q_t) + \lambda_t x_{t-1}] \tag{2}$$

and by $(1 - p_t)$ for L. Throughout the paper I will restrict to equilibrium strategies that are affine, stationary Markov Perfect (see Maskin and Tirole, 2001).[13] As a result, players's strategies will depend on the previous realization of the payoff-relevant state $\mu_t(q_t, \varepsilon_t, x_{t-1}) = I(q_t) + \lambda_t x_{t-1}$. Since x_{t-1} is the only endogenous part of the state, I will often use the notation $p_t(x_{t-1})$ to highlight the dynamic link between current policy and future electoral environment.

2.1 2-period model

This section describes the political equilibrium in the $T = 2$ 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 electoral equilibrium will be given by two pairs of platform functions of the form

$$X_t^j : [\mu^l, \mu^h] \to [0, 1] \quad j \in \{R, L\}, \ t \in \{1, 2\}.$$

where $\mu^l = I(q^l)$, $\mu^h = I(q^h)$ defines the range of feasible states.[14] The following proposition describes the unique equilibrium of the game:

Proposition 1 *In the unique equilibrium of the 2-period **(M)**-game*
i) In $t = 2$ both platforms converge to the efficient level: $X_2^R = X_2^L = x^*$.
ii) In $t = 1$ the difference between the 2 platforms, $\Delta_{(M)}$, solves

$$\Delta[1 + \psi\beta\bar{\lambda}\Delta] - \beta\bar{\lambda} = 0 \tag{3}$$

[13]More specifically, the Markov restriction will have substantial bite in the infinite horizon game, while in the $T = 2$ case pure strategy Nash equilibrium will be enough. The restriction to affine policy functions is motivated by them baing the qualitative analog of the unique equilibrium of the 2 period models.

[14]To compute these bounds, notice that $\max_{x,\varepsilon,s} \mu$ can be re-expressed as $\max_x\{\max_s\{(1-x)\max_\varepsilon\{I(s+\varepsilon)\} + x\tilde{I}(s)\}\}$, where $c_t^l = \min \Omega_{c_t}$, $c_t^h = \max \Omega_{c_t}$

and the equilibrium platforms are given by

$$\begin{cases} X_1^R = x^* + \frac{\Delta_{(M)}}{2} + \frac{\psi\Delta_{(M)}}{1+\psi\Delta_{(M)}^2}\mu_1 \\ \\ X_1^L = x^* - \frac{\Delta_{(M)}}{2} + \frac{\psi\Delta_{(M)}}{1+\psi\Delta_{(M)}^2}\mu_1 \end{cases}$$

iii) The implemented policy in $t = 1$, $X_1^{(M)}$, is a simple lottery between X_1^R and X_1^L.

Therefore, in the last period, political competition drives both platforms to the efficient level. This is not true in the first period, when platform divergence is observed: R underprovides the public good to secure himself a better electoral environment in $t = 2$ while L does the opposite to achieve the same goal. As a result in $t = 1$ there will be either underprovision or overprovision of public good (or either inefficiently low or inefficiently high levels of public employment), and redistribution will be either 0 or $\bar{b}_t$. Regardless of how large is the Rawlsian component in the adopted welfare criterion, political competition then delivers a second best outcome, because the relative size of each sector is never set at the optimal level. The reason is that in $t = 1$, political actors face a trade-off. On one hand, setting $X_1^j = x^*$ maximizes the chances of winning the upcoming elections, which is valuable for two reasons: it allows to implement the favorite redistribution level today and, by making X_1^j the implemented policy, it maximizes the impact of the own platform on period 2's electoral environment. On the other hand, appropriately distorting X_1^j will increase the chances of winning the elections in period 2 conditional on winning the current elections. The point at which this trade-off is balanced generates distortions at both platform level and at the implemented policy level. These distortions might be large even when the individual incentive to deviate is relatively low, due to the centripetal force exerted by the interplay between platforms that must be best reply to each other.

The following proposition summarizes how the size of these distortions changes with the main parameters of the model. Throughout the paper, I will use expected policy distortion (that is the expected quadratic deviation from x^*, denoted by $\Sigma_1^{(M)}$) to quantify distortions in the implemented policy and platform divergence (in this context, $\Delta_{(M)}$) to quantify distortions at platform level.

Proposition 2 *In the unique equilibrium of the 2-period* ***(M)****-game*
i) Platform divergence is increasing in political actors' discount factor (β), in $\bar{\lambda}$, in wealth inequality (ω), and in the variance of the aggregate shock (ψ^{-1}).
ii) $\Sigma_1^{(M)}$ is increasing in policy divergence, ω, q_1, λ_1, and the initial size of the private sector (x_0); it is ambiguous in ψ^{-1}.

As already suggested, political actors patience increases dynamic distortions. A larger $\bar{\lambda}$ has the same effect, since a larger expected informational wedge implies a larger marginal effect of current distortions on future elections. The effect of ψ is in principle ambiguous: as the aggregate shock becomes more volatile (lower ψ), the connection between platforms and electoral result weakens (which should increase distortions); at the same time, the connection between current platforms and future electoral environment also weakens (which should decrease distortions). The proposition shows how the first effect always dominates the second. On the other hand, the effect on expected deviation from x^* cannot be resolved: one one hand, a larger ψ makes divergence smaller, which decreases $\Sigma_1^{(M)}$, but on the other hand it tilts the probability of winning towards the more likely winner, thereby decreasing the variance of the implemented policy

From the previous proposition one can also see that, as q_1 increases, R's platform becomes more extreme. This means that the electoral process has a "spiraling effect" on policy inefficiencies: as the inefficiency associated with L's redistribution scheme increases (that is, as R's structural advantage increases), the inefficiencies associated with R's public good provision increase and so does R's equilibrium probability of winning. Therefore, while voters become less likely to have an inefficient redistribution scheme implemented, they become more likely to suffer from a greater underprovision of public good. This result allows to relate this paper with most of the recent literature on political economy distortions. q_1 increases, *ceteris paribus*, R's probability of winning, and can be therefore related to a measure of political persistence. This paper shows that, unlike in most of the previous literature, more political persistence can increase, rather than decrease, the severity of political failures. This result, is, to my best knowledge, novel.

2.2 Infinite horizon model

The infinite horizon version of the **(M)** game confirms some of the insights of the 2 periods model, but has also important quantitative and qualitative differences. This subsection also highlights an important virtue of the model: the ability to explicitly solve for the stationary Markov perfect equilibrium, rather than relying on numerical simulations. This feature allows to obtain the same type of comparative static derived in the previous subsection, which permits a direct comparison between the two models.

Due to its recursive structure, I simply denote by $X^R(\mu)$, $X^L(\mu)$ the equilibrium platform given an initial state $\mu = I(q) + \lambda x$

Proposition 3 *In the unique stationary affine MPE of the infinite horizon* ***(M)****-game*
i) platform divergence, Δ_∞, equals $\beta\bar{\lambda}$ and the equilibrium platforms are given by

$$\begin{cases} X^R(\mu) = x^* + \frac{\Delta_\infty}{2} + \frac{\psi\Delta_\infty}{1+\psi\Delta_\infty^2}\mu \\ \\ X^L(\mu) = x^* - \frac{\Delta_\infty}{2} + \frac{\psi\Delta_\infty}{1+\psi\Delta_\infty^2}\mu \end{cases}$$

ii) platform divergence is increasing in β, $\bar{\lambda}$, ω, and independent of ψ^{-1}.
iii) platform divergence and policy distortion are larger than in the 2-period model.

Therefore, as T increases to infinity, the size of the distortions on both platforms and implemented policies increase. Moreover, it is important to notice that policy divergence is no longer dependent on the variance of the aggregate shock. This implies that the two effects of a change in the volatility of the electoral outcome perfectly offset when $T = \infty$:

the effect on the marginal loss (from platform distortion) of winning probability exactly compensates the effect on expected marginal gain from a more favorable future electoral environment. Intuitively, this is due to the fact that, in steady state, higher volatility (lower ψ) means weaker connection between implemented platforms and an infinite sequence of future implemented policies. As a result, the marginal gain from distorting the platform is higher than when $T = 2$ by a factor proportional to the present-value effect on all future elections. Moreover, this gain must be constant over time, as actors' platforms will diverge by the same amount in every period. The latter is equivalent to multiplying the last term in (3) by $(1 + \psi\beta\bar{\lambda}\Delta)$.

3 Consensual constitution

The idea of consensual (also known as consociational) democracy, introduced by Lijphart, is based on the observation that in several countries, especially in northern Europe, constitutional mechanism are such that, rather than assigned to a majoritarian winner like in United Kingdom,[15] political power is shared between different groups within a society. In his 1977 book, *Democracy in plural societies* Lijphart identifies the main features of this type of 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 adhere as much as possible to this definition (and at the same time maintaining comparability with the majoritarian setting), I model consensual democracy as a post-electoral bargaining game between R and L. More specifically, they can negotiate over the implemented x and b (where b can be anything and x should be between the two announced platforms) according to the following protocol:

1) A randomly determined (R with probability $1/2$, otherwise L) proposer offers to implement a pair (x^{pr}, b^{pr}) to the other actor. If the offer is accepted, $x = x^{pr}$, $b = b^{pr}$ are implemented.

2) If the offer is not accepted, the following default policies are implemented

$$X^{(C)} = \hat{\pi}_t x_t^R + (1 - \hat{\pi}_t) x_t^L;\ b^{(C)} = (1 - \hat{\pi}_t)\bar{b}_t.$$

This assumption captures the idea that, lacking a different agreement between the two political actors, the constitution prescribes that each political actor will have an influence on each policy dimension proportional to his realized vote share: as a result, implemented policies will be linear combinations between the two platforms and the two actors' preferred

[15] The alternative concept to consensus democracy is called by Lijphart "Westminster democracy". This comparison, together with a more detailed and slightly updated definition of consensus democracy, is the topic of Lijphart's classic book, *Patterns of Democracy.*

redistribution levels with weights proportional to their vote share. Finally, in order to break indifferences, I assume that bargaining has an infinitesimally small cost. These assumptions generate the following Lemma:

Lemma 2 *Under the assumptions the outcome of the bargaining game is the default policy.*

As a consequence, under **(C)**, R's per-period realized payoff is $\hat{\pi}_t$ and L's payoff is $(1-\hat{\pi}_t)$. The expected payoffs are then, respectively, π_t and $(1-\pi_t)$, where

$$\pi_t = E_t[\hat{\pi}_t] = 1/2 + \varphi[d(x_t^R, x_t^L) + I(q_t) + \lambda_t x_{t-1}]$$

is the expected vote share.

3.1 2-period consensual

In this section I present the unique equilibrium of the 2 period version of the **(C)**-game and compare it with the results from the 2-period version of the **(M)**-game.

Proposition 4 *In the unique equilibrium of the 2-period **(C)**-game*
i) In $t=2$ both platforms converge to the efficient level: $X_2^R = X_2^L = x^*$.
ii) In $t=1$ platform divergence, $\Delta_{(C)}$, solves

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

and the equilibrium platforms are given by

$$\begin{cases} X_1^R = x^* + \frac{\Delta_{(C)}}{2} + \frac{\varphi\Delta_{(C)}}{1+\varphi\Delta_{(C)}^2}\mu_1 \\ \\ X_1^L = x^* - \frac{\Delta_{(C)}}{2} + \frac{\varphi\Delta_{(C)}}{1+\varphi\Delta_{(C)}^2}\mu_1 \end{cases}$$

iii) The implemented policy in $t=1$, $X_1^{(C)}$, is uniformly distributed in

$$\left[x^* + \frac{\varphi\Delta_{(C)}^2}{1+\varphi\Delta_{(C)}^2}\mu_1 - \frac{\varphi\Delta_{(C)}}{2\psi}, x^* + \frac{\varphi\Delta_{(C)}^2}{1+\varphi\Delta_{(C)}^2}\mu_1 + \frac{\varphi\Delta_{(C)}}{2\psi}\right]$$

iv) $\Delta_{(C)}$ is increasing in β, $\bar{\lambda}$, ω, and in the variance of the idiosyncratic shock (φ^{-1}); $\Sigma_1^{(C)}$ is increasing in policy divergence, ω, q_1, λ_1, and the volatility of the aggregate shock (ψ^{-1}); it is ambiguous in φ^{-1}.

Thus, consensual democracy displays a qualitatively similar solution to the majoritarian case, with the key difference that, rather than the variance of the aggregate shock (ψ^{-1}), the relevant parameter becomes the variance of the idiosyncratic shock (φ^{-1}). The reason is that, while political actors in a majoritarian constitution only care about winning a majority, in a consensual democracy every vote has the same marginal effect on the future implemented policy, as the actors' objective function depends on their expected vote share. The other key difference is that, while φ does not affect the variance of the implemented policy in **(M)**, the variance of the aggregate shock plays a key role in determining the variance of the policy outcome in **(C)**, since the latter depends on the realized vote share. The rest of the comparative static has similar intuition as for the majoritarian case. The following proposition compares the equilibria of the two constitutions with $T = 2$.

Proposition 5 *In a 2-period model*
i) platform divergence is larger under ***(C)****.*
ii) policy distortion is larger under ***(M)****, expected redistribution is larger under* ***(C)****.*
iii) larger $\beta\bar{\lambda}$ increases differences among constitutions at platform level and decreases differences at the implemented policy level (that is $\frac{d}{d\beta\bar{\lambda}}\Delta_{(C)} > \frac{d}{d\beta\bar{\lambda}}\Delta_{(M)}$ and $\frac{d}{d\beta\bar{\lambda}}E[X_1^{(M)}] < \frac{d}{d\beta\bar{\lambda}}E[X_1^{(C)}]$).

Therefore, consensual constitutions are associated with more extremism at the platform level and with more redistribution, results that echoes several theoretical findings on coalition governments and proportional electoral rules, two constitutional features that Lijphart strongly associates with consensual democracy.[16] On the other hand, the proposition shows that the moderating effect of consensual democracy on the implemented policy more than offsets the higher divergence at platform level. Finally, part iii) implies that, as the expected reward from distorting the platform increases, constitutional differences become at platform level, and more pronounced at implemented policy level. As political actors become more patient or the expected informational wedge increases, platform divergence in **(M)**, which is smaller than in **(C)**, and the expected implemented policy in **(M)**, which is further above x^* than in **(C)**, both increase by a smaller amount than in **(C)**.

3.2 Semi-consensual constitution

In order to better understand what drives the difference in the distortions observed under different constitutions, in the section I introduce an hybrid constitution, called semi-consensual.

This exercise is valuable because there are 2 fundamental differences between **(C)** and **(M)**, namely the allocation of policy making rights over the public good and redistribution. One's first conjecture might be that, since actors only care about redistribution, it is the allocation of policy making power over this dimension that drives the observed difference among constitutions. Therefore, the 2-period equilibrium of a game in which the allocation of rights over redistribution (b) is majoritarian and the one over public employment (x) is consensual should be quite similar to the one of the 2 period **(M)**-game. The following proposition shows that the opposite is actually true.

[16] See, for example, Austen-Smith and Banks (1988) or Gerber and Ortuño Ortín (1998), or Milesi-Ferretti, Perotti and Rostagno (2002). Notice, also, that while the result on transfers seems to be driven by a similar mechanic as in Persson and Tabellini (aggregate shock has lower variance than idiosyncratic shock), here the reason why the two systems are related to φ and ψ is completely independent on parties trarget median voters at district levels.

To do so, I formally define the semi-consensual constitution **(S)**, as a situation in which the implemented policies are $X_t^{(S)} = X_t^{(C)} = \hat{\pi}_t x_t^R + (1 - \hat{\pi}_t) x_t^L$ and $\bar{b}_t^{(S)} = \bar{b}_t^{(M)} = \bar{b}_t I_{\{L=W\}}$

Proposition 6 *In $t = 1$ the equilibrium of the **(S)** game the platforms are the same as in the consensual constitution.*

This proposition shows that the equilibrium is completely driven by the way the constitution allocates policy making power over public good provision (i.e. public employment), over which political actors have no preferences, but can credibly commit. Therefore, in this environment commitment plays the key role of transmitting distortions across policy dimensions.

Another implication of the proposition is that policy distortions in **(S)** are as large as in **(C)** and smaller than in **(M)**, but the implemented redistribution will be majoritarian, which suggests that, from the overall welfare perspective, a mixed constitution can dominate both **(C)** and **(M)**.

To answer this question, I define the expected redistribution distortion associated with constitution (J) as $\Sigma_b^{(J)}$. thereby implicitly adopting a utilitarian welfare criterion.[17] It's easy to see that, since the equilibrium strategies are the same, expected redistributions are

$$\Sigma_b^{(S)} = (1 - p(X_1^R, X_1^L, \mu_0))\bar{b}_t \ ; \ \Sigma_b^{(C)}(1 - \pi(X_1^R, X_1^L, \mu_0))\bar{b}_t$$

Since, given the structure of the model (and formally proved in the Appendix) $\pi(X_1^R, X_1^L, \mu_0) < p(X_1^R, X_1^L, \mu_0)$, one can conclude that $\Sigma_b^{(C)} > \Sigma_b^{(S)}$. The following proposition formally compares the three constitutions.

Proposition 7 *Under a strict utilitarian welfare criterion, a Semi-consensual constitution dominates both consensual and majoritarian democracy.*

The normative implication is that a majoritarian allocation of power on policy dimensions over which commitment is not possible, coupled with a consensual allocation of power on policy dimensions over which commitment is possible minimizes the overall distortions associated with policy making.[18]

3.3 Infinite horizon model

In this subsection I consider the solution of the infinite horizon version of the **(C)**-game, compare it to the one from its 2 period version and to the infinite horizon version of the **(M)**-game. The main results are a long term neutrality of constitutional choice on platform divergence, and stronger differences at the implemented policy level with respect to the comparison for 2 period models.

As for the **(M)** case, I simply denote by $X^R(\mu)$, $X^L(\mu)$ the equilibrium platform given an initial state $\mu = I(q) + \lambda x$

[17] For this, notice that I do not need to use a quadratic form because b is defined over the semipositive orthant and unambiguously decreases welfare.

[18] At this point it is tempting to guess that the opposite configuration (majoritarianism on policies over which commitment is possible and consociativism on policies over which no commitment is possible) would deliver the worse possible outcome. It is possible to show that this is, indeed the case.

Proposition 8 *In the unique stationary affine MPE of the infinite horizon **(C)**-game*
i) platform divergence, Δ_∞, equals $\beta\bar{\lambda}$ and the equilibrium platforms are given by

$$\begin{cases} X^R(\mu) = x^* + \frac{\Delta_\infty}{2} + \frac{\varphi\Delta_\infty}{1+\varphi\Delta_\infty^2}\mu \\ \\ X^L(\mu) = x^* - \frac{\Delta_\infty}{2} + \frac{\varphi\Delta_\infty}{1+\varphi\Delta_\infty^2}\mu \end{cases}$$

*ii) platform divergence is increasing in β, $\bar{\lambda}$, ω, independent of φ^{-1} and larger than in 2-period **(C)** game; policy distortion is also larger than in 2-period **(C)** game.*

3.4 Constitutional comparison

As for the **(M)** game, when $T = \infty$, the size of the distortions on platforms and implemented policies increase and policy divergence is no longer dependent on the variance of the relevant shock to voters' political preferences. The mechanism is very similar as the one described before: the change in marginal loss (from platform distortion) on expected vote share exactly offsets the change in the expected marginal gain from a more favorable future electoral environment. In steady state, higher volatility (lower φ) means weaker connection between implemented platforms and an infinite sequence of future implemented policies.

The following corollary compares the infinite horizon equilibria under each constitution

Corollary 1 *In the infinite horizon model*
i) Platform divergence is the same across constitutions.
*ii) Policy distortion on public employment are larger under **(M)**, expected redistribution is larger under **(C)**.*

This corollary leads to the surprising conclusion that dynamic distortions at platform level are, in the long run, independent on the type of constitution. The reason is that, when $T = \infty$, the volatility of the relevant shocks (aggregate shock for **(M)** and idiosyncratic shock for **(C)**) no longer affects the trade-off between current electoral strength and future electoral strength.

The second part of the corollary follows from simple inspection of $\Sigma_\infty^{(M)}$ and $\Sigma_\infty^{(C)}$ and implies that the difference in policy distortions, which in the 2 period models was in favor of **(C)**, is even larger in the infinite horizon model. The reason is that, as the larger platform divergence disappears, only the "averaging" effect of bargaining survives.

4 Empirical implications

Independent of the time horizon, the model delivers three important predictions on the link between wealth inequality, political polarization and redistribution. First, across constitutions, increased inequality increases the inefficiency associated with L's proposed redistribution, which, through the informational wedge, increases platform divergence. The latter can be related to commonly used measures of political polarization.[19] Therefore, the model predicts a positive relationship between inequality and political polarization, as widely documented, among others, by Mc Carty, Poole and Rosenthal (2006). The suggested channel for that relationship is that, as inequality rises, so does the perceived inefficiency of redistribution, which in turns magnifies the effect of asymmetric information, which means that political actors have a stronger incentive to distort their platforms in opposite directions.

To develop the second implication, I make use of the following lemma.

Lemma 3 *R's equilibrium probability of winning and expected vote share are increasing in ω in both the 2-period model and the infinite horizon model. Moreover, when the cost of redistribution is high enough, higher inequality decreases expected redistribution.*

The first part of the lemma states that, in the model, inequality has an ambiguous effect on implemented redistribution. On one hand, more inequality increases L's implemented redistribution; on the other, it decreases its influence on implemented policy. When the cost of redistribution is high and initial size of the private sector is large enough, the indirect effect dominates the direct effect. The inverse relationship between inequality and redistribution has been the object of empirical and theoretical investigation,[20] to which this model relates. Enns and Kelly (2010), in particular, have recently documented an inverse relationship between inequality and political support to redistribution among all income levels, which sharply contrasts with the prediction from classic Meltzer-Richard type of public finance models. This model suggests that one potential explanation for this inverse relationship might be that, as inequality increases, the perceived inefficiencies associated with redistribution also increase, thereby leading to a lower demand for redistribution *across both income groups.*

The third empirical implication is that the above described effect is stronger in a majoritarian than in a consensual democracy, thereby yielding the prediction that, *ceteris paribus*, consensual democracies will display higher levels of redistribution. It must be stressed that this result does not directly depend on the averaging effect of policy making under **(C)**, but on how equilibrium political competition balances the fundamental trade-off of the model (current vs future electoral strength) under each constitutional setting.

Finally, it is important to add that, like in the 2-period model, a hybrid semi-consensual constitution will deliver higher utilitarian welfare than both **(C)** and **(M)**. The reason is that, like in the 2-period model platforms and implemented policies for x are as in **(C)**, while **(S)** still delivers a more favorable distribution of implemented b (that is, lower expected redistribution).

In the following section I will show that the superiority of consensual democracy on the implemented public employment level is fragile to a key assumption: homogeneity of discount factors across political actors. It is also worth anticipating that this perturbation will leave **(S)** mostly unaffected.

[19]See, on this the recent work of Mc Carty and Shor on quantifying the amount political polarization using data from surveyed pre-electoral candidates' commitments.

[20]See, for example, De Mello and Tiongson (2003) for cross country evidence, Moene and Wallerstein (2001) or Benabou (2000) for other proposed theoretical explanations.

5 Heterogeneous discount factors

In this section I look at the situation in which political actors differ in their discount factor. To simplify the comparison with the default (homogenous βs) case, I assume a mean preserving perturbation of amount h (which can be either positive or negative) that additively enters $R's$ discount factor (β^R). The two actors' discount factors are then given by $\beta^R = \beta + h$ and $\beta^L = \beta - h$. For this reason, I denote the 2-period perturbed versions of the **(C)** and **(M)** games by $\mathbf{(M)}_h$ and $\mathbf{(C)}_h$, respectively. This exercise is justified by empirical and theoretical reasons. Empirically, candidates often differ in their age and type of commitment to politics: one might be old, the other quite young;[21] one might be a career politician, the other a professional who is expected to go back to his previous job after serving in office.[22] Moreover, parties are also typically ruled by different waves of top executives, who often change in a quite dramatic way.[23] In the model considered in this paper political actors are treated a single decision maker. Nevertheless, it is reasonable to suspect that, when two competing parties are ruled by top executives mostly belonging to different generations, the homogeneous discount factor assumption might be problematic.

On a theoretical level, heterogeneous discount factors allow to uncover an important robustness property of each setting, given that the discount factor plays such an important role in determining the extent of the dynamic distortions this paper studies.

It also important to notice that there are two independent reasons for discount factor heterogeneity to affect the equilibrium. First, it makes the equilibrium asymmetric; second, it affects the bargaining outcome in **(C)**, since Lemma 2 no longer holds. Given that each actor has essentially a Euclidean function embedded in his objective function, and that in the default case equilibrium both actors essentially equate the slope of that concave, single peaked function to a certain number, the marginal gain from distortions is the same for both. With heterogeneous discount factors these marginal gains are no longer the same. Due to the strict concavity, one would expect that the decrease in platform distortion of the less patient actor will be larger than the corresponding increase for the more patient actor. This is the only change in the majoritarian case, and the following proposition confirms this intuition:

Proposition 9 *In the unique $t = 1$ equilibrium of the $\mathbf{(M)}_h$-game, policy divergence is strictly decreasing in $|h|$.*

For the consensual constitution, instead, any heterogeneity in the discount factors creates an opportunity for Pareto superior agreement between the two actors. Therefore, the equilibrium might look very different from the homogeneous β case. In the following proposition I make the additional assumption that the initial value of μ_0 is not too

[21] For example, Barak Obama and John Mc Cain in the 2008 presidential elections.

[22] For example, Gray Davis and Arnold Schwarzenegger, the contenders in the California special gubernatorial election of 2003.

[23] For example, the takeover of the British Labor party by Tony Blair and Gordon Brown in 1994, or the sudden change in the leadership of the Italian Socialist Party in 1976, with the election of Bettino Craxi as the party head.

extreme.[24]

Proposition 10 *In the* ***(C)***$_h$ *game*
i) the more patient actor in equilibrium gets full policy making rights over x*; power over* b *is split according to realized vote share and the identity of the proposer.*
ii) platform divergence is strictly increasing in $|h|$ *and larger than in* ***(C)****.*

Therefore, when discount factors are heterogenous, the outcome of the multidimensional bargaining characterizing consensual democracy ha no longer an averaging effect over implemented policies, and voters suffer from both underprovision of public good and excessive redistribution. The extent of this asymmetry is proportional to the heterogeneity in discount factors (h), which has the opposite effect on policy divergence with respect to **(M)**$_h$. As a result, when h is large enough, **(M)** might welfare dominate **(C)**. It must be stressed that the discontinuous nature of the equilibrium in **(C)**$_h$ (that is, the sudden jump to the edge of the Edgeworth box as h becomes positive) is not an artifact of the linearity of political actors' payoff function. It's easy to show that any power function of the form $[(\bar{b}-b)/\bar{b}]^{\gamma}$, with $\gamma \in (0,1)$, would yield a qualitatively analogous outcome.

6 Related literature

This paper is related to several literatures on political failures in dynamic settings. The first connects exogenous changes in power and lack of commitment with political failures. Acemoglu et al (2009) explore the effect of stochastic power fluctuations on the allocation of resources in a dynamic production economy. In their model a society is divided into N groups differing in their labor-leisure preferences and endowed with the same production technology. In each period a Markov process determines the identity of the group in charge of allocating, after all groups have chosen their labor supply, the total output. Rather than the set of equilibria of the game, the authors characterize the Pareto efficient allocations. The main result is that, if the common discount factor is high enough, the economy converges to the first best allocation in which there is full consumption smoothing. When the discount factor is low, conversely, distortions are a permanent feature of the stationary distribution to which the economy converges, together with inefficient fluctuations in labor supply and output. These distortions arise from the fact that groups that are not in power find it optimal to give the ruling group a smaller output to be divided. Finally, they show that a higher probability of switches of power leads to a larger set of first-best sustainable allocations. From a more general perspective, Bai and Lagunoff (2008) consider an environment in which policy making exhibits 'Faustian' dynamics: due to the presence of a period-by-period wealth-weighted version of the median voter theorem, the policy that maximizes the immediate payoff of the current median voter also shifts away political power from him, thereby inducing dynamic distortions. The benchmark outcome under the permanent authority of a certain individual is compared with the equilibrium dynamics when

[24] This assumption, formally described in the Appendix, is technical in nature and allows to make a clear prediction of the outcomes. Removing the assumption makes the analysis much more complicated without adding any substantial insight.

policy-induced power shifts are allowed. Depending on whether the policy moves the median voter toward individuals with more extreme or more moderate preferences, different type of convergence and steady state are observed, leading to either excessive inertia or to inefficient extremism. This literature shares with the present paper the idea that uncertainty over future allocation of political power can generate an independent channel for political failures. While in Acemoglu et al. more patient actors help mitigating these issues, in Bai and Lagunoff they make these distortions more pronounced.

My model is also related to a literature on political failures in which the presence of a dynamic linkage in policies interacts with the political process to generate distortions. Battaglini and Coate (2008) look at a situation in which political power varies exogenously and the presence of public debt create incentives to shift costs towards future periods. They look at a dynamic neoclassical model with n legislators (each representing a group), who bargain *à la* Baron and Ferejohn (1989) over spending on a public good, whose marginal utility is stochastic, district-specific transfers, taxation, and debt. The main result is that in equilibrium there will be regime switches between a responsible regime, in which policy making is Pareto efficient and a "business-as-usual" regime in which inefficiently high debt and pork spending are observed.

Besley and Coate (1998) look at a 2-period citizen-candidate model in which the political process might hamper efficient public investment. In of the examples concluding the paper, they analyze productivity enhancing investment with 1 large, but non majoritarian, high productivity group and 2 smaller low-productivity groups. A policy that increases at no cost the productivity of one of the 2 small groups is not implemented because it alters the distribution of political power in the future. The group that would benefit from the investment is in favor of redistribution in period 1 but will switch its political preferences against redistribution in period 2, and therefore cannot find support from the other 2 groups to enact the investment. Aghion and Bolton (1990) and Milesi-Ferretti and Spolaore (1994) look at a similar trade-off in the case, respectively, of public debt and the size of government. Also related to the persistence of inefficiently large governments are Krussell and Rios-Rull (1999), Acemoglu et al (2006), and Hassler et al. (2005). The first looks at a dynamic version of the Meltzer-Richard model to study the dynamics of larger-than-efficient redistribution programs. The second looks at large and inefficient state bureaucracies as a commitment device for the elites to counteract the loss of political power associated with democratization. The third 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.

Azzimonti (2009) looks at a neoclassical economy with 2 districts, each expressing a party competing for political power and local public goods, financed through an investment tax. The political process is based on probabilistic voting, with an exogenous incumbency advantage The distortions associated with political competition result in inefficiently low investment rates and excessive large local public good provision Moreover, political stability (proportional to the size of the incumbent advantage) and low polarization (defined as a lower marginal utility for local public goods) are associated with higher level of investment and lower governments. The second result is more intuitive: a society where the level of conflict is exogenously lower will produce lower distortions. The first result, which contradicts Acemoglu et al. (2009), echoes previous theoretical findings in political science, and is based on the fact that it is easier to provide incentives for long termism to political agents that remain in power for longer. In all these papers, having more far sighted actors or more persistence in political power would allow to compensate the dynamic negative externality to a higher extent, thereby reducing the extent of the distortions. In this paper, I show that asymmetric information, lack of commitment and differentiation in political actors can generate the opposite effect.

In all these papers the basic source of distortions is the interaction between political power and a policy dimension that provides a dynamic linkage: in all these papers, providing more persistence in political power or having more patient policy makers mitigates the extent of these distortions.

This paper is also related to a small literature that explains short termism as an equilibrium response to some underlying friction in the policy making process. Aidt and Dutta (2007) consider the provision of 2 public goods, one with short-term benefits and the other with long term ones. Investment in the latter can only be assessed with one period lag, thereby creating a moral hazard problem, since politicians have the possibility to extract rents. As a consequence, voters might not be able to perfectly discipline politicians over time; therefore, their optimal re-election strategy produces inefficient underinvestment in the long term public good. In a similar setting, Garrì (2009) explains the policy bias toward the short term public good using a reputational argument, which leads to the conclusion that short termism might be welfare improving because of improved selection of congruent politicians. In both papers, imperfect information on policy outcomes generates a "good" type of short termism.

My paper is also related to a large body of literature comparing different constitutional features in terms of provision of public goods, transfers and government size: examples include Persson and Tabellini (2004), Persson, Roland and Tabellini (2005), Lizzeri and Persico (2001), and Milesi-Ferretti, Perotti and Rostagno (2002). In a recent paper Battaglini (2010) extends the setting of Battaglini and Coate (2008) to a case with 3 districts and probabilistic voting and compares the outcome of proportional rule to the one of a majoritarian system. The main finding is that the basic prediction of the static literature (PR gives leads to overspending and less transfers) is not necessarily true because of an additional source of dynamic inefficiency. Under proportional representation parties are responsive to (and therefore willing to 'bribe' through transfers) all 3 groups in the society, while under majority rule the only relevant one is the pivotal group. Finally, the paper that is most closely related to mine in terms of institutional comparison is Ticchi and Vindigni (2010), which explicitly compares consensual and majoritarian constitutions, but is more focused on the conditions facilitating the *ex ante* adoption of either type.

Kalandrakis (2009) builds a reputational theory of 2-party competition in which voters are uncertain about whether a party is controlled by extremists or moderate agents. Depending on the persistence of policy preferences within the party, two types of dynamics are observed: if the persistence is high in equilibrium there is going to be low political turnover and moderate policies, while if the persistence is low parties are going to regularly alternate in office and some extreme policies will observed (especially when elections are close). Although the paper is almost orthogonal to my work in terms of economic environment and political process, the setting is very interesting for two reasons. First, since the median voter prefers moderate policies, close races, that in this setting are associated with extreme policies, seem to deliver an outcome that is more divergent from the median voter's preferences. Second, 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 finding themselves in a situation of losing elections almost for sure against an opponent on moderate platforms. This is another instance in which persistence seems to help, but at the same time far-sighted politicians are potentially detrimental for voters, since incorporating future political distortions that have a reinforcing effect on today's distortions.

The empirical implications of the model relate the paper to two important bodies of literature in economics and political science. The first investigates the negative relationship between inequality and redistribution, documented empirically by Enns and Kelly (2010), De Mello and Tiongson (2003). These results seriously challenged the conclusion, pioneered by Meltzer and Richard (1981), that redistribution increases with inequality, and focused the attention on theories yielding the opposite prediction, which include, among others, Moene and Wallerstein (2001) and Bénabou (2000) and Bénabou and Ok (2001). The second relevant body of literature investigates the relationship between inequality and political polarization. The evidence is presented in a comprehensive fashion in Mc Carty, Poole and Rosenthal's book *Polarized America: The Dance of Political Ideology and Unequal Riches*. This paper contributes to each of these literatures by showing the existence of a novel channel that can help explaining how inequality affects redistribution and political polarization.

7 Conclusion

This paper analyzes a model of dynamic electoral competition in which current policies affect political actors' future electoral environment. Unlike most of the recent literature on political failures, the associated distortions decrease with political actors' patience. Moreover, as persistence in R's electoral victory increases, distortions get stronger.

A second contribution of the paper is to show that the induced political failures are not specific to a peculiar set of policy aspects as so far analyzed in economics (that is, migration policy in the Curley effect), but can be rather pervasive and impact some of the most important policy dimensions traditionally identified in public finance: public employment and redistribution.

The third contribution of the paper is to show how these distortions depend on several variables: political actors' time horizon and discount factor, type of constitution (majoritarian vs consensual), wealth inequality, strength of informational asymmetry.

Three types of dynamic distortions are considered: at platform level (which captures a measure of political polarization), and at implemented policy level (deviation from efficient public employment and amount of redistribution). In a 2 period model, platforms converge at the efficient level only in the terminal period. In the first period, majoritarian constitutions display lower platform divergence, more inefficient public employment but less (inefficient) redistribution than in consensual constitutions. These inter-constitutional differences on implemented policies weaken as actors become more patient or as the extent of the informational asymmetry increases (that is, as the effect of current policies on future discounted electoral benefits gets stronger). Moreover, when the time horizon is infinite, platform divergence is larger than in a 2-period model and no longer depends on the type of constitution, highlighting an interesting long term neutrality property. The expected inefficiencies in policies, instead, are still different across constitutions. The difference between constitutions is entirely driven by the allocation of policy making rights over the public employment dimension and is independent of how actors share power over redistribution (the dimension over which they care!). A consensual allocation of these rights, combined with a majoritarian allocation of power over implemented redistribution is showed to welfare dominate both consensual and majoritarian democracy. Finally, the performance of consensual democracy is not robust to heterogeneity in the discount factors. While heterogeneity reduces platform divergence and public employment inefficiencies in majoritarian constitutions, it makes the outcome of the bargaining among political actors fully asymmetric, which *ex ante* increases incentives to platform divergence and *ex post* increases the inefficiency of the implemented public employment.

The fourth contribution of the paper is to deliver various implications that show how three important empirical regularities, each generating an independent and branch of literature, can be interpreted in light of the presence of dynamic political distortions, First, regardless of the constitution, higher inequality is associated with higher political polarization; second, when the public sector is small enough, higher inequality decreases implemented redistribution. This is effect is more likely to be observed in majoritarian democracies. Third, implemented redistribution is higher, *ceteris paribus*, in consensual democracies.

A battery of serious empirical tests is necessary to add empirical validation to this model, although beyond the scope of this paper. These relationships are clearly derived from the analytics of the model, which is flexible enough not to require numerical solution even in its infinite horizon version.

Finally, more theoretical and empirical work is needed to cast light of the presence of dynamic political distortions on other policy dimensions. Despite being disregarded by the literature so far, these phenomena can help understanding the dynamics of policy making on other important areas, such as education and home ownership subsidy.

Appendix

Proof of Lemma 1

The unconditional expectation of λ_t (taken in $t-1$) can be rewritten as

$$\bar{\lambda} = \int_{\Omega_t^q} \int_{\Omega_{s|\gamma}} [\tilde{I}(\sigma) - I(\gamma)] dF_{s|\gamma}(\sigma) dF_q(\gamma).$$

The proof has two steps. First, I establish that

$$\tilde{I}(\sigma) > I(\sigma); \tag{5}$$

then I show that

$$\int_{\Omega_{s|\gamma}} [I(\sigma) - I(\gamma)] dF_{s|\gamma}(\sigma) > 0 \ \forall \ \gamma \in \Omega_t^q \tag{6}$$

To see why (5) must hold, notice that, since σ is an unbiased predictor of q, $\tilde{I}(\sigma) - I(\sigma)$ can be rewritten as

$$E[I(q)|\sigma] - I(\sigma) = E[I(q)|\sigma] - I(E[q|\sigma])$$

which, by Jensen's inequality, is strictly positive. As a consequence, if one shows that (6) holds, then $\bar{\lambda}$ is integrating over a set whose elements are all strictly positive and. as a consequence, must be strictly positive.

To see why(6) must hold, notice that, for fixed γ, $\sigma = \gamma + \varepsilon$. As a consequence, the LHS of (6) can be rewritten as

$$\int_{-e}^{e} [I(\gamma + z) - I(\gamma)] dF_{\varepsilon}(z) = E[I(\gamma + \varepsilon)] - I(E[\gamma + \varepsilon])$$

which is strictly positive, again using Jensen's inequality.

Bounds on φ and ψ

In order to make players' objective functions continuous and differentiable, I need to make assumptions on the relative size of the state space, φ and ψ. This is a standard in this type of models[25]. Given the structure of the model (and, in particular, the different types of constitution considered), I need both political actors to be competitive in every election. That implies that the range of the realized vote share must include, for every realization of the state μ_t and any platform profile (x^R, x^L), the value 1/2.

More formally, given

$$\hat{\pi}(\xi, d, \mu) = 1/2 + \varphi[d + \mu + \xi]$$

we must have $\max_\xi \hat{\pi}(\xi, d, \mu) \in (1/2, 1)$, $\min_\xi \hat{\pi}(\xi, d, \mu) \in (0, 1/2) \ \forall \ \mu, d$, where $\mu \in [\mu^l, \mu^h]$, $d \in [d^l, d^h]$, $d^l = \min\{0, Q - 1/2\} - Q^2/2$, and $d^h = Q^2/2 - \min\{0, Q - 1/2\}$. The two conditions yield 4 equations that can be simplified into

$$\min\{1/\varphi - 1/\psi, 1/\psi\} > 2 \max\{-[d^l + \mu^l], d^h + \mu^h\} \tag{7}$$

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

Proof of Proposition 1

In $t = 2$ equilibrium policies solve $X_2^R \in \arg\max p_2(x_1)$, $X_2^L \in \arg\max\{1 - p_2(x_1)\}$, where

$$p_2(x_1) = 1/2 + \psi[d(x_2^R, x_2^L) + I(q_2) + \lambda_2 x_1].$$

The FONC of the problem define the solution. For $t = 1$ X_1^R and X_1^L solve

$$\begin{cases} X_1^R \in \arg\max_{x\in[0,1]} \left\{p_1(x_0) + \beta p_1(x_0)p_2^*(x) + \beta(1 - p_1(x_0))p_2^*(X_1^L)\right\} \\ \\ X_1^L \in \arg\max_{x\in[0,1]} 1 + \beta - \left\{p_1(x_0) + \beta p_1(x_0)p_2^*(x) + \beta(1 - p_1(x_0))p_2^*(X_1^L)\right\} \end{cases} \tag{8}$$

where $p_2^*(x) = E_1[p_2(x)] = 1/2+\psi E[I(q)]+\psi\bar{\lambda}x$ follows from the observation that $d(X_2^R, X_2^L) = 0$. The FONC are of the problem (which are also sufficient under the assumptions) define the following system

$$\begin{cases} \frac{d}{dX_1^R}p_1(x_0)[1 + \beta(p_2^*(X_1^R) - p_2^*(X_1^L)] + \beta p_1(x_0)\frac{d}{X_1^R}p_2^*(X_1^R) = 0 \\ \frac{d}{dX_1^L}p_1(x_0)[1 + \beta(p_2^*(X_1^R) - p_2^*(X_1^L)] + \beta(1 - p_1(x_0))\frac{d}{X_1^L}p_2^*(X_1^L) = 0 \end{cases} \tag{9}$$

subtracting the first from the second gives (3), while summing them gives

$$\psi(2x^*-(X^R+X^L))(1+\beta\psi\bar{\lambda}\Delta_{(M)}) = \psi\bar{\lambda}\beta(1-2p_1(x_0)) = -\psi\bar{\lambda}\beta[\Delta_{(M)}(2x^*-(X^R+X^L))+\mu]$$

whose unique solution gives the equilibrium in $t = 1$. Part iii) follows from the binary outcome of the (M)-game.

Proof of Proposition 2

For i), start observing that, since the RHS of (3) is supermodular in $\beta\bar{\lambda}$, the positive solution of that equation shifts to the right as $\beta\bar{\lambda}$ increases. Then notice that $\bar{\lambda}$ can be re-written as $\omega \int_{\Omega_t^q} \int_{\Omega_{s|\gamma}} [\tilde{I}_q(\sigma) - I_q(\gamma)]dF_{s|\gamma}(\sigma)dF_q(\gamma)$ where

$$I_q(q) = \frac{q}{2(2-q)} \; ; \; \tilde{I}_q(s) = E[I_q(q)|s]$$

are independent of ω. For ii), notice that $\Sigma_1^{(M)} = E[X_1^{(M)} - Q]^2$ which, since (9) can be re-written as

$$X_1^R = x^* + \Delta_{(M)}p_1(X_1^R, X_1^L, \mu) \; ; \; X_1^L = x^* - \Delta_{(M)}(1 - p_1(X_1^R, X_1^L, \mu))$$

can be written as $\Sigma_1^{(M)} = p_1^3\Delta_{(M)}^2 + (1 - p_1)^3\Delta_{(M)}^2 = (1 - 3p_1(1 - p_1))\Delta_{(M)}^2$, which after explicitly substituting X_1^R, X_1^L into $p(X_1^R, X_1^L, \mu)$ becomes

$$\left[\frac{1}{4} + \frac{3\psi^2}{[1 + \psi\Delta_{(M)}^2]^2}\mu_1^2\right]\Delta_{(M)}^2 \tag{10}$$

which is increasing in $\Delta_{(M)}^2$ and μ_1. Notice that $\mu_1 = I(q_1) + \lambda_1 x_0$ and that (7) implies $\psi\Delta_{(M)}/[1 + \psi\Delta_{(M)}^2] < 1$, which is sufficient to establish that $\frac{d}{d\Delta_{(M)}}\Sigma_1^{(M)} > 0$. To see that it is ambiguous in ψ, notice that $\frac{d}{d\psi}\Sigma_1^{(M)} = \frac{\partial}{\partial\psi}\Sigma_1^{(M)} + \frac{\partial}{\partial\Delta}\Sigma_1^{(M)}\frac{d}{d\psi}\Delta_{(M)}$, which is the sum of a negative term and a positive term.

Proof of Proposition 3

Part i). Since the game has maximum value $\bar{V} = (1-\beta)^{-1}$ for each player, the recursive formulation of the problem solved by R and L under **(M)** is given by:

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

$$V^L(\mu) = \max_{x^L \in [0,1]} \bar{V} - p(x^R, x^L, \mu)[1 + \beta(E[V^R(\mu^R)] - E[V^R(\mu^L)])] - E[V^R(\mu^L)] \quad (12)$$

where $\mu^R = I(q) + \lambda x^R$ and $\mu^L = I(q) + \lambda x^L$.

A Markov Perfect Equilibrium is a pair of value functions $V^R(\mu)$, $V^L(\mu)$ and policy functions $X^R(\mu)$, $X^R(\mu)$ such that,

1) given $X^L(\mu)$, $V^A(\mu)$ solves (11) and, given $X^R(\mu)$, $V^L(\mu)$ solves (12)
2) $X^R(\mu)$ attains the RHS of (11) and $X^L(\mu)$ attains the RHS of (12)

To see that the two platforms solve the system, 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 I verify that the value functions are affine in μ, and in Step 2 I solve for the coefficients.

Step 1.A few lines of algebra allow to verify that $p(h^R(\mu), h^L(\mu), \mu)$ is an affine function of μ

$$p(h^R(\mu), h^L(\mu), \mu) = \bar{p}(\mu) = h_p + h_p\mu$$

Then the value functions can be re-expressed in the following way:

$$V^R(\mu_0) = \bar{p}(\mu_0)(1 + \beta(\bar{p}(\mu_1|h^R(\mu_0)) - \bar{p}(\mu_1|h^R(\mu_0))) + \beta\bar{p}(\mu_1|h^R(\mu_0)) + ...$$

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

where

$$\bar{p}(\mu_t|h^R(\mu_{t-1})) = E[\bar{p}(I(q) + \lambda h^L(\mu_{t-1}))] = \bar{p}(E[I(q)] + \bar{\lambda} h^L(\mu_{t-1})).$$

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

Step 2. The FONC are

$$\begin{cases} \frac{dp^R}{dx^R}[1 + \beta(V(\mu^R) - V(\mu^L))] + \beta p^R V_1 \bar{\lambda} = 0 \\ \frac{dp^L}{dx^L}[1 + \beta(V(\mu^R) - V(\mu^L)] + \beta(1 - p^L) V_1 \bar{\lambda} = 0 \end{cases} \quad (13)$$

where

$$\begin{aligned} p^R &= p(x^R, h^L(\mu), \mu) = p^R(x^R, \mu) \\ p^L &= p(h^R(\mu), x^L, \mu) = p^L(x^L, \mu) \end{aligned}$$

and the envelope conditions give

$$
\begin{aligned}
V_1 &= \frac{dp^R}{d\mu}[1+\beta(V(x^R)-V(h^L(\mu)))]+\beta(1-p^R)\bar{\lambda}V_1h_1 \\
V_1 &= \frac{dp^L}{d\mu}[1+\beta(V(x^R)-V(h^L(\mu)))]+\beta p^L\bar{\lambda}V_1h_1
\end{aligned}
$$

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

To see why this is the unique affine MPE, notice that, given a generic guess $h^L(\mu) = h_0^L + h_1^L\mu$, the FONC for R's problem is given by

$$
[-dp/dx^R][1+\beta(E[V(\mu^R)]-E[V(\mu^L)])]-\beta pE[V_1(\mu^R)]=0
$$

where $\mu^R = I(c)+\lambda x^R$, $\mu^L = I(c)+\lambda h^L(\mu)$ and $E[V_1(\mu^R)] = E\left[\frac{d}{dx^R}V(I(c)+\lambda x^R)\right]$. The FONC of L's problem, given a linear policy function $h^R(\mu) = h_0^R + h_1^R\mu$ is

$$
[-d\tilde{p}/dx^L][1+\beta(E[V(\mu^R)]-E[V(I(c)+\lambda x^L)])]-\beta(1-p)E\left[\frac{d}{dx^L}V(I(c)+\lambda x^L)\right]=0
$$

where $\mu^R = I(c)+\lambda h^R(\mu)$, $\mu^L = I(c)+\lambda h^L(\mu)$ and $E[V_1(\mu^L)] = E\left[\frac{d}{dx^L}V(I(c)+\lambda x^L)\right]$.

After substituting the envelope conditions for each player's problem, imposing equilibrium, and re-arranging, one obtains

$$
\begin{aligned}
V_1(\mu)-\beta(1-p)E[V_1(\mu^L)]h_1^L-\beta pE[V_1(\mu^R)]\frac{dp/d\mu}{[-dp/dx^R]} &= 0 \\
V_1(\mu)-\beta(1-p)E[V_1(\mu^L)]\frac{d\tilde{p}/d\mu}{[-d\tilde{p}/dx^L]}-\beta pE[V_1(\mu^R)]h_1^R &= 0
\end{aligned}
$$

As a consequence, for every $\mu \in \Omega_\mu$, in equilibrium one must have

$$
(1-p)E[V_1(\mu^L)]h_1^L+pE[V_1(\mu^R)]\frac{dp/d\mu}{[-dp/dx^R]}=(1-p)E[V_1(\mu^L)]\frac{d\tilde{p}/d\mu}{[-d\tilde{p}/dx^L]}+pE[V_1(\mu^R)]h_1^R
$$

When $h_1^L \neq h_1^R$, $p(h^R(\mu), h^L(\mu), \mu)$ is quadratic. As a result, the associated value functions cannot be linear. This implies that $V_1^E(\mu^L) \neq V_1^E(\mu^R)$ and that both are proper functions of μ.

Therefore, in order for the equality to hold $\forall\ \mu$, it must be that

$$
\frac{dp/d\mu}{[-dp/dx^R]} = h_1^R\ \forall\ \mu\ \ ;\ \ h_1^L = \frac{d\tilde{p}/d\mu}{[-d\tilde{p}/dx^L]}\ \forall\ \mu
$$

the first equality is equivalent to

$$\frac{1 - Qh_1^L + (h_0^L + h_1^L\mu)h_1^L}{[h_0^R + h_1^R\mu - Q]} = h_1^R \;\forall\; \mu$$

which can only hold if $h_1^L = h_1^R$, which contradicts the initial assumption that $h_1^L \neq h_1^R$.

Part ii) Standard arguments apply to conclude that the process induced by the equilibrium strategies is ergodic.

Part iii) Simple inspection.

Part iv) Substitute $\beta\bar{\lambda}$ into (3) and notice that the result is strictly positive; that implies $\Delta_{(M)} < \Delta_\infty$. To show that expected deviation from x^* must increase, consider that, since (13) can be rewritten as follows (I use the fact that from the envelope conditions, $V_1^{-1} = \beta\bar{\lambda}/\psi\Delta_\infty - \beta\bar{\lambda}\Delta$)

$$X^R = x^* + \Delta_\infty p(X^R, X^L, \mu) \;;\; X^L = x^* - \Delta_\infty(1 - p(X^R, X^L, \mu))$$

the same steps that led to (10) also lead to

$$\Sigma_\infty^{(M)} = \left[\frac{1}{4} + \frac{3\psi^2}{[1 + \psi\Delta_\infty^2]^2}\mu^2\right]\Delta_\infty^2$$

which is larger than $\Sigma_1^{(M)}$, given that $\Delta_\infty > \Delta$.

Proof of Lemma 2

Follows from the observation that, being it a zero sum game and having actors the same discount factor and the same information, every proposed deviation from the status quo (α_x, α_b), to which are associated implemented policies[26]

$$X^{(C)} = \alpha_x x_t^R + (1 - \alpha_x)x_t^L;\; b^{(C)} = (1 - \alpha_b)\bar{b}_t.$$

would weakly benefit at most one player, and would thereby not proposed in equilibrium.

Proof of Proposition 4

i) In $t = 2$ equilibrium policies solve $X_2^R \in \arg\max \pi_2(x_1)$, $X_2^L \in \arg\max\{1 - \pi_2(x_1)\}$, where

$$\pi_2(x_1) = 1/2 + \varphi[d(x_2^R, x_2^L) + I(q_2) + \lambda_2 x_1].$$

The FONC of the problem define the solution. Notice that, since $d(X_2^R, X_2^L) = 0$, then $\hat{\pi}_2(x_1) = 1/2 + \varphi[I(q_2) + \lambda_2 x_1 - \hat{\xi}_2]$

ii) For $t = 1$ X_1^R and X_1^L solve

$$\begin{cases} X_1^R \in \arg\max_{x\in[0,1]} \left\{\pi_1(x_0) + \beta E[\hat{\pi}_2(X^{(C)})]\right\} \\ X_1^L \in \arg\max_{x\in[0,1]} 1 + \beta - \left\{\pi_1(x_0) + \beta E[\hat{\pi}_2(X^{(C)})]\right\} \end{cases}$$

[26] It's easy to see that a pair (α_x, α_b) would completely characterize any possible alloaction of policy-making rights.

where $E[\hat{\pi}_2(X^{(C)})] = 1/2 + \varphi E[I(q)] + \varphi\bar{\lambda}E[X^{(C)}]$ which becomes

$$1/2 + \varphi E[I(q)] + \varphi\bar{\lambda}\{\pi_1(x_0)x_1^R + (1-\pi_1(x_0))x_1^L\}$$

The FONC are of the problem (which are also sufficient under the assumptions) define the following system

$$\left\{ \begin{array}{c} \frac{d}{dX_1^R}\pi_1(x_0)[1+\beta\varphi\bar{\lambda}(X_1^R - X_1^L)] + \beta\varphi\bar{\lambda}\pi_1(x_0) = 0 \\ \frac{d}{dX_1^L}\pi_1(x_0)[1+\beta\varphi\bar{\lambda}(X_1^R - X_1^L)] + \beta\varphi\bar{\lambda}(1-\pi_1(x_0)) = 0 \end{array} \right. \tag{14}$$

whose unique solution gives the equilibrium in $t = 1$. Part iii) follows from the observation that, once platforms are fixed, the only randomness in the implemented policy is given by the realization of the aggregate shock, $\hat{\xi}_1$, and the observation that $X^{(C)} = \hat{\pi}_t\Delta_{(C)} + X_t^L$.

iv) The first part follows from is supermodularity of the RHS of (4) in $\beta\bar{\lambda}$ and the same decomposition of $\bar{\lambda}$ performed above. For the rest, notice that $\Sigma_1^{(C)} = E[X^{(C)} - x^*]^2$ becomes

$$E\left\{\frac{2\hat{\pi}-1}{2}\Delta_{(C)} + \frac{\varphi\Delta_{(C)}}{1+\varphi\Delta_{(C)}^2}\mu_1\right\}^2 = \left\{\frac{2\varphi\Delta_{(C)}}{1+\varphi\Delta_{(C)}^2}\mu_1\right\}^2 + \frac{\varphi^2\Delta_{(C)}^2}{\psi^2 12}$$

which is increasing in $\Delta_{(C)}$ (use (4) and ψ, and ambiguous in φ due to the increasing direct effect and the decreasing indirect effect through $\Delta_{(C)}$.

Proof of Proposition 5

i) Notice that (7) implies that $\varphi^{-1} > \psi^{-1}$; these two parameters are the only difference between (3) and (4); therefore, the result follows.

ii) If one proves that

(a) Expected policy is closer under **(C)** than under **(M)**

(b) the range of $X^{(C)}$ is within the two platforms under (M)

Then it must be that $\Sigma_1^{(C)} < \Sigma_1^{(M)}$, since expected distortion can be re-expressed as $Var(X) + (E(X) - x^*)$.

For part (a), notice that

$$E(X^{(C)}) = x^* + \Delta_{(C)}(2\pi_1 - 1)\ ;\ E(X^{(M)}) = x^* + \Delta_{(M)}(2p_1 - 1) \tag{15}$$

therefore, if

$$\frac{\Delta_{(C)}}{\Delta_{(M)}}\frac{[2\pi_1 - 1]}{[2p_1 - 1]} < 1 \tag{16}$$

then the result is proved. After a few steps of algebra one can obtain

$$2p_1 - 1 = 2\mu_1\psi/[1+\psi\Delta_{(M)}^2]^{-1} \quad ; \quad 2\pi_1 - 1 = 2\mu_1\varphi/[1+\varphi\Delta_{(C)}^2]^{-1}$$

plugging back into (16) yields $\Delta^{(C)}\varphi(1+\psi\Delta_{(M)}^2) < \Delta^{(M)}\psi(1+\varphi\Delta_{(C)}^2)$.Assume wlog that $p_1 > .5$ and $\pi_1 > .5$. To see that (16) holds, rearrange it in the following way

$$\frac{1}{\psi\Delta^{(M)}} + \Delta^{(M)} < \frac{1}{\varphi\Delta^{(C)}} + \Delta^{(C)}$$

and verify using (3) and (4) that $\psi\Delta^{(M)} > \varphi\Delta^{(C)}$.

For part (b), I have to show that

$$\Delta_{(C)}\left[(2\pi_1 - 1) + \frac{\varphi}{2\psi}\right] < p_1\Delta_{(M)} \ ; \ \Delta_{(C)}\left[(2\pi_1 - 1) - \frac{\varphi}{2\psi}\right] > -\Delta_{(M)}(1 - p_1)$$

which become

$$\begin{aligned} \Delta_{(C)}\left[\frac{\varphi}{2\psi} + \frac{\mu_1\varphi}{1 + \varphi\Delta^2_{(C)}}\right] &< \Delta_{(M)}\left[\frac{1}{2} + \frac{\mu_1\psi}{1 + \psi\Delta^2_{(M)}}\right] \\ \Delta_{(C)}\left[\frac{\varphi}{2\psi} - \frac{\mu_1\varphi}{1 + \varphi\Delta^2_{(C)}}\right] &< \Delta_{(M)}\left[\frac{1}{2} - \frac{\mu_1\psi}{1 + \psi\Delta^2_{(M)}}\right] \end{aligned}$$

which become

$$\frac{\Delta_{(M)}}{2} - \frac{\varphi\Delta_{(C)}}{2\psi} > \frac{\mu_1\psi\Delta_{(M)}}{1 + \psi\Delta^2_{(M)}} - \frac{\mu_1\varphi\Delta_{(C)}}{1 + \varphi\Delta^2_{(C)}} > 0 > \frac{\mu_1\varphi\Delta_{(C)}}{1 + \varphi\Delta^2_{(C)}} - \frac{\mu_1\psi\Delta_{(M)}}{1 + \psi\Delta^2_{(M)}}$$

the first inequality can be rearranged, using (3) and (4) into

$$\Delta_{(C)}\varphi\left[\frac{1}{2} - \frac{\mu_1\psi\beta\bar{\lambda}}{2\beta\bar{\lambda} - \Delta_{(C)}}\right] < \Delta_{(M)}\psi\left[\frac{1}{2} - \frac{\mu_1\psi\beta\bar{\lambda}}{2\beta\bar{\lambda} - \Delta_{(M)}}\right]$$

which, given that $\Delta_{(C)}\varphi < \Delta_{(M)}\psi$ and $\Delta_{(C)} > \Delta_{(M)}$, holds. To see why $E(b^{(C)}) = (1 - \pi_t)\bar{b}_t$ is larger than $E(b^{(M)}) = (1 - p_t)\bar{b}_t$, notice that in equilibrium

$$\pi_t = 1/2 + \frac{\varphi}{1 + \varphi\Delta^2_{(C)}}\mu_1 < p_t = 1/2 + \frac{\psi}{1 + \psi\Delta^2_{(M)}}\mu_1$$

which holds, since, by (7) $\varphi < \psi$ and $\Delta^2_{(C)} > \Delta^2_{(M)}$.

iii) Notice that, by applying the Implicit function theorem to (3) and (4), re-arranging, one obtains

$$\frac{d}{d\beta\bar{\lambda}}\Delta_{(C)} = \frac{\Delta_{(C)}}{\beta\bar{\lambda}}\frac{1 - \varphi\Delta^2_{(C)}}{1 + \varphi\Delta^2_{(C)}} > \frac{d}{d\beta\bar{\lambda}}\Delta_{(M)} = \frac{\Delta_{(M)}}{\beta\bar{\lambda}}\frac{1 - \psi\Delta^2_{(M)}}{1 + \psi\Delta^2_{(M)}}$$

Changes in $\beta\bar{\lambda}$ affect $E[X^{(C)}]$ and $E[X^{(M)}]$,which by (15) are above x^* only through $\Delta_{(C)}$ and $\Delta_{(M)}$. The former is more responsive than the latter. Therefore, if one proves that

$$\frac{d}{d\Delta_{(C)}}E[X^{(C)}] > \frac{d}{d\Delta_{(M)}}E[X^{(M)}]$$

the proof is complete. To see this, simple computation lead to

$$\frac{d}{d\Delta_{(C)}}E[X^{(C)}] = \frac{1 - \varphi\Delta^2_{(C)}}{[1 + \varphi\Delta^2_{(C)}]^2} > \frac{d}{d\Delta_{(M)}}E[X^{(M)}] = \frac{1 - \psi\Delta^2_{(M)}}{[1 + \psi\Delta^2_{(M)}]^2}$$

Proof of Proposition 6

The FONC of the problem define the solution. For $t = 1$ X_1^R and X_1^L solve

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

where $E[p(X_1^S)] = 1/2 + \psi E[I(q)] + \psi\bar{\lambda}E[X_1^{(C)}] = 1/2 + \psi E[I(q)] + \psi\bar{\lambda}\{\pi_1 X_1^R + (1-\pi_1)X_1^L\}$ follows from the observation that $d(X_2^R, X_2^L) = 0$. The FONC are of the problem (which are also sufficient under the assumptions) define the following system

$$\begin{cases} \frac{d}{dX_1^R}p_1(x_0) + \beta\psi\bar{\lambda}\left\{\pi_1 + (X_1^R - X_2^L)\frac{d\pi^R}{dX_1^R}\right\} = 0 \\ \frac{d}{dX_1^L}p_1(x_0) + \beta\psi\bar{\lambda}\left\{(1-\pi_1) + (X_1^R - X_2^L)\frac{d\pi^R}{dX_1^R}\right\} = 0 \end{cases}$$

which simplifies to

$$\begin{cases} x^* - X_1^R + \beta\bar{\lambda}\left\{\pi_1 + \Delta_{(S)}\varphi(x^* - X_1^R)\right\} = 0 \\ X_1^L - x^* + \beta\bar{\lambda}\left\{(1-\pi_1) + (X_1^R - X_2^L)\varphi(X_1^L - x^*)\right\} = 0 \end{cases}$$

which is the same system as in (14).

Proof of Proposition 7

Proof

$$\Sigma_1^{(C)} = \Sigma_1^{(S)} \ ; \ \Sigma_b^{(C)} > \Sigma_b^{(S)}$$
$$\Sigma_1^{(M)} > \Sigma_1^{(S)} \ ; \ \Sigma_b^{(M)} > \Sigma_b^{(S)}$$

where the last inequality is equivalent to

$$\bar{b}(1 - p(X_{(M)}^R, X_{(M)}^L, \mu_0) > \bar{b}(1 - p(X_{(S)}^R, X_{(S)}^L, \mu_0)$$

which becomes

$$\bar{b}(1/2 - \psi(d(X_{(M)}^R, X_{(M)}^L, \mu_0) - \mu_0) > \bar{b}(1/2 - \psi(d(X_{(S)}^R, X_{(S)}^L, \mu_0) - \mu_0)$$

which follows from

$$d(X_{(M)}^R, X_{(M)}^L) = -\frac{\psi\Delta_{(M)}^2}{1 + \psi\Delta_{(M)}^2}\mu_0 < -\frac{\varphi\Delta_{(M)}^2}{1 + \varphi\Delta_{(M)}^2}\mu_0 = d(X_{(S)}^R, X_{(S)}^L)$$

Proof of Proposition 8

Part i). The recursive formulation of the problem solved by R and L under **(M)** is given by:

$$V^R(\mu) = \max_{x^R \in [0,1]} \pi(x^R, x^L, \mu) + \beta E[V^R(\mu)|X^{(C)}] \tag{17}$$

$$V^L(\mu) = \max_{x^L \in [0,1]} \bar{V} - \{\pi(x^R, x^L, \mu) + \beta E[V^R(\mu)|X^{(C)}]\} \tag{18}$$

where

$$\beta E[V^R(\mu)|X^{(C)}] = \beta E[V^R(I(q) + \lambda(\hat{\pi}x^R + (1-\hat{\pi})x^L))]$$

A Markov Perfect Equilibrium is a pair of value functions $V^R(\mu)$, $V^L(\mu)$ and policy functions $X^R(\mu)$, $X^R(\mu)$ such that,

1) given $X^L(\mu)$, $V^A(\mu)$ solves (17) and, given $X^R(\mu)$, $V^L(\mu)$ solves (18)
2) $X^R(\mu)$ attains the RHS of (17) and $X^L(\mu)$ attains the RHS of (18)

To see that the two platforms solve the system, 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 I verify that the value functions are affine in μ, and in Step 2 I solve for the coefficients.

Step 1.A few lines of algebra allow to verify that $\pi(h^R(\mu), h^L(\mu), \mu)$ is an affine function of μ

$$\pi(h^R(\mu), h^L(\mu), \mu) = \bar{\pi}(\mu) = h_p + h_p\mu$$

where the realized value of $\bar{\pi}(\mu)$ is $\bar{\pi}(\mu) + \varphi\hat{\xi}$. Then the value functions can be re-expressed in the following way:

$$V^R(\mu_0) = \bar{\pi}(\mu_0) + \beta E[\bar{\pi}(\mu_1)] + \beta^2 E[\bar{\pi}(\mu_2)] + ...$$

$$V^L(\mu_0) = \frac{1}{1-\beta} - V^R(\mu_0)$$

where

$$E[\bar{\pi}(\mu_t)] = E[\bar{\pi}\{I(q) + \lambda[(\bar{\pi}((\mu_{t-1})) + \varphi\hat{\xi})h^R(\mu_{t-1}) + (1 - \bar{\pi}((\mu_{t-1})) - \varphi\hat{\xi})h^L(\mu_{t-1})]\}\}$$

simplifies to $\bar{\pi}\{E[I(q)] + E[\lambda][(\bar{\pi}(h_0^R - h_0^L) + h_0^L - h_1\mu_{t-1})]\}$, which is affine in μ_{t-1} Therefore, all the terms in the summation are compositions of affine functions, therefore affine. Denote by V_1 the slope coefficient of V.

Step 2. The FONC of the problem are

$$\begin{cases} \frac{d\pi^R}{dx^R} + \beta V_1 \bar{\lambda} \left\{ \pi^R + (x^R - h^L)\frac{d\pi^R}{dx^R} \right\} = 0 \\ \frac{d\pi^L}{dx^L} + \beta V_1 \bar{\lambda} \left\{ \pi^L + (h^R - x^L)\frac{d\pi^L}{dx^L} \right\} = 0 \end{cases} \tag{19}$$

where

$$\begin{aligned}\pi^R &= \pi(x^R, h^L(\mu), \mu) = \pi^R(x^R, \mu) \\ \pi^L &= \pi(h^R(\mu), x^L, \mu) = \pi^L(x^L, \mu)\end{aligned}$$

and the envelope conditions are

$$\begin{aligned}V_1 &= \frac{d\pi^R}{d\mu} + \beta\bar{\lambda}V_1\left\{(1-\pi^R)h_1 + (x^R - h^L)\frac{d\pi^R}{d\mu}\right\} \\ V_1 &= \frac{d\pi^L}{d\mu} + \beta\bar{\lambda}V_1\left\{\pi^L h_1 + (h^R - x^L)\frac{d\pi^R}{d\mu}\right\}\end{aligned}$$

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

To see why this is the unique affine MPE, notice that, given a generic guess $h^L(\mu) = h_0^L + h_1^L\mu$, the FONC for R's problem is given by

$$d\pi/dx^R + \beta E[V_1(\mu^C)(\hat{\pi} + (x^R - h^L(\mu))d\pi/dx^R)] = 0 \tag{20}$$

where $\mu^C = I(c) + \lambda[\hat{\pi}x^R + (1-\hat{\pi})h^L(\mu)]$. The FONC of L's problem, given a linear policy function $h^R(\mu) = h_0^R + h_1^R\mu$ is

$$d\tilde{\pi}/dx^L + \beta E[V_1(\mu^C)(1 - \hat{\pi} + (h^R(\mu) - x^L)d\tilde{\pi}/dx^L)] = 0$$

where where $\mu^C = I(c) + \lambda[\hat{\pi}h^R(\mu) + (1-\hat{\pi})x^L]$.

After substituting the envelope conditions for each player's problem, imposing equilibrium, and re-arranging, one obtains

$$\begin{aligned}V_1(\mu) - \beta E[V_1(\mu^C)]h_1^L &= [1 + \beta E[V_1(\mu^C)](h^R(\mu) - h^L(\mu))]\left[h_1^L\frac{d\pi}{dx^R} + \frac{d\pi}{d\mu}\right] = 0 \\ V_1(\mu) - \beta E[V_1(\mu^C)]h_1^R &= [1 + \beta E[V_1(\mu^C)](h^R(\mu) - h^L(\mu))]\left[h_1^R\frac{d\tilde{\pi}}{dx^L} + \frac{d\tilde{\pi}}{d\mu}\right] = 0\end{aligned}$$

summing up these two equations and substituting for $\frac{d\pi}{dx^R}, \frac{d\pi}{d\mu}, \frac{d\tilde{\pi}}{dx^L}, \frac{d\tilde{\pi}}{d\mu}$ yields

$$[1 + \beta E[V_1(\mu^C)](h^R(\mu) - h^L(\mu))]\psi(h^R(\mu) - h^L(\mu)) = \beta E[V_1(\mu^C)]$$

substituting this back into (20) gives

$$\frac{-d\pi/dx^R}{\varphi} = h^R(\mu) - Q = \hat{\pi}[h^R(\mu) - h^L(\mu)]$$

to see why this is not compatible with $h_1^R \neq h_1^L$, notice that $\hat{\pi} = 1/2 + \varphi Q(h^R(\mu) - h^L(\mu)) - \varphi\frac{1}{2}(h^R(\mu) + h^L(\mu))(h^R(\mu) - h^L(\mu)) + \varphi\mu$.Taking derivatives of both sides one gets

$$\begin{aligned} h_1^R &= [h^R(\mu) - h^L(\mu)]\phi[1 + \varphi(h_1^R - h_1^L)(Q - \frac{1}{2}(h^R(\mu) + h^L(\mu)) + \\ &\quad -\varphi\frac{1}{2}(h_1^R + h_1^R)(h^R(\mu) - h^L(\mu))] + (h_1^R - h_1^L)\hat{\pi}\ \forall\ \mu \end{aligned}$$

which equates a constant to a quadratic function of μ.

Part ii) Same steps as in Proposition 3, from which one concludes that

$$\Sigma_\infty^{(C)} = \left\{\frac{2\varphi\Delta_\infty}{1 + \varphi\Delta_\infty^2}\mu_1\right\}^2 + \frac{\varphi^2\Delta_\infty^2}{\psi^2 12}$$

Proof of Lemma 3

I prove the dependence of equilibrium π on ω; the proof is completely analogous for p. First, notice that, in equilibrium, the expected vote share is

$$\pi = \frac{1}{2} + \varphi\frac{\mu}{1 + \varphi\Delta^2}$$

denote by μ_ω, λ_ω the ratios μ/ω and λ/ω, which are independent of ω. The total derivative of π wrt ω is then $\frac{\partial}{\partial\omega}\pi + \frac{\partial}{\partial\Delta}\pi\frac{d}{d\omega}\Delta$. Using the implicit function theorem, $\frac{d}{d\omega}\Delta = \beta\lambda[1 + \varphi\Delta^2][1 + 2\varphi\Delta\beta\lambda_\omega\omega]^{-1}$. After a few steps of algebra, the total derivative simplifies to

$$\varphi\frac{I_\omega + \bar{\lambda}_\omega x}{1 + \varphi\Delta^2}\left[1 - \frac{2\varphi\Delta\beta\lambda_\omega\omega}{1 + 2\varphi\Delta\beta\lambda_\omega\omega}\right] > 0.$$

The expected redistribution is then given by $\bar{b}(1 - \pi)$. Denote by b_ω the ratio $\bar{b}/\omega$ which is independent of ω. Computing the total derivative and rearranging yields

$$b_\omega\left\{\frac{1}{2} - \varphi\frac{\mu_\omega\omega}{1 + \varphi\Delta^2}\right\} - b_\omega\omega\varphi\frac{\mu_\omega}{1 + \varphi\Delta^2}\frac{1 - \varphi\Delta^2}{1 + \varphi\Delta^2}$$

which is negative iff

$$\varphi^{-1} < \frac{4}{1 + \varphi\Delta^2}\mu$$

which, for μ high enough, holds without violating (7).

Proof of Proposition 9

I solve under the assumption that $h > 0$. The proof for the opposite case is completely analogous. The equilibrium strategies in $t = 1$ solve a problem exactly like in (8), only with actor-specific discount factors. The associated FONC are given by

$$\begin{cases} \frac{d}{dX_1^R}p_1(x_0)[1 + \beta^R(p_2^*(X_1^R) - p_2^*(X_1^L)] + \beta^R p_1(x_0)\frac{d}{X_1^R}p_2^*(X_1^R) = 0 \\ \frac{d}{dX_1^L}p_1(x_0)[1 + \beta^L(p_2^*(X_1^R) - p_2^*(X_1^L)] + \beta^L(1 - p_1(x_0))\frac{d}{X_1^L}p_2^*(X_1^L) = 0 \end{cases}$$

where, like in **(M),** $p_2^*(x) = 1/2 + \psi E[I(q)] + \psi\bar{\lambda}x$. Denote by $\Delta^h_{(M)}$ the platform divergence. FONCs become

$$(x^* - X^R)[1 + (\beta + h)\psi\bar{\lambda}\Delta^h_{(M)}] + (\beta + h)p_1(x_0)\bar{\lambda} = 0$$
$$(X^L - x^*)[1 + (\beta - h)\psi\bar{\lambda}\Delta^h_{(M)}] + (\beta - h)(1 - p_1(x_0))\bar{\lambda} = 0$$

from which one obtains

$$\Delta^h_{(M)}[1 + \beta\psi\bar{\lambda}\Delta^h_{(M)}] - \beta\bar{\lambda} = h\bar{\lambda}[(2p_1(x_0) - 1) - (X^R + X^L - 2Q)\psi\Delta^h_{(M)}]$$

the RHS is the same equation that gives $\Delta_{(M)}$. Since $2p_1(x_0) - 1 = 2\psi(dv + \mu_0) = 2\psi(\Delta^h_{(M)}(Q - \frac{X^R + X^L}{2}) + \mu_0)$, the RHS simplifies to

$$h\bar{\lambda}2\psi\mu_0 > 0$$

which implies that $\Delta^h_{(M)}$ is strictly decreasing in h. Simple inspection of the FONCs allows to conclude that, although the problem is no longer linear, the best reply correspondences $X^R(X^L)$ and $X^L(X^R)$ only cross once.

Proof of Proposition 10

A convenient way to represent the space of possible policy choices of R and L, given their platforms (X_t^R, X_t^L), is by using (α_x, α_b), which give implemented policies $b_t = (1 - \alpha_b)\bar{b}_t$ and $x_t = \alpha_x X_t^R + (1 - \alpha_x)X_t^L$. I then define R's acceptance region as $[\alpha_b^R(\alpha_x), 1]$, where the lower bound is the share of power over redistribution that, given an allocation of power of amount α_x over x, would make R indifferent between (α_x, α_b) and the default policies. L's acceptance region is given by $[0, \alpha_b^L(\alpha_x)]$, where the upper bound is defined analogously.

$$\begin{aligned}
\alpha_b^R(\alpha_x) &= \hat{\pi}_1 - \beta^R\varphi\bar{\lambda}\Delta^h_{(C)}(\alpha_x - \hat{\pi}_1) \\
\alpha_b^L(\alpha_x) &= \hat{\pi}_1 - \beta^L\varphi\bar{\lambda}\Delta^h_{(C)}(\alpha_x - \hat{\pi}_1)
\end{aligned}$$

Due to the difference in discount factors, the two actors' acceptance regions have a proper intersection, moreover, since preferences over (α_x, α_b) are linear, the outcome of bargaining will be on the edge of the unit square. For intermediate values of the equilibrium realized vote share, that is

$$\hat{\pi}_1 \in [1 - \pi^h, \pi^h], \tag{21}$$

where $\pi^h = [1 + (\beta + h)\varphi\bar{\lambda}\Delta^h_{(C)}]^{-1}$, then the outcome of the bargaining game has $\alpha_x \in \{0, 1\}$ and α_b fractional. More specifically, for $h > 0$ the outcome is the following:

$$\alpha_x = 1 \; ; \; \alpha_b = \begin{array}{ll} \alpha_b^L(1) & \text{wp } 1/2 \\ \alpha_b^R(1) & \text{wp } 1/2 \end{array}$$

Therefore, I will solve for the equilibrium and verify *ex post* that (21) holds. Equilibrium platforms solve the following problem

$$\begin{cases}
X_1^R \in \arg\max_{x \in [0,1]} E[\alpha_b^R(1) + \alpha_b^L(1)]/2 + \beta^R E[\hat{\pi}_2(X_1^R)] \\
\\
X_1^L \in \arg\max_{x \in [0,1]} 1 + \beta^L - \left\{E[\alpha_b^R(1) + \alpha_b^L(1)]/2 + \beta^R E[\hat{\pi}_2(X_1^R)]\right\}
\end{cases}$$

where $E[\hat{\pi}_2(X^R)] = 1/2+\varphi E[I(q)]+\varphi\bar{\lambda}X_1^R$ and $E[\alpha_b^R(1)+\alpha_b^L(1)]/2 = \hat{\pi}_1-\beta\varphi\bar{\lambda}\Delta_{(C)}^h(1-\hat{\pi}_1)$. The FONC characterize the solution, and are given by

$$\begin{cases} \frac{d}{dX_1^R}\pi_1[1+\beta\varphi\bar{\lambda}\Delta_{(C)}^h] - (1-\pi_1)\beta\varphi\bar{\lambda} + (\beta+h)\varphi\bar{\lambda} = 0 \\ \frac{d}{dX_1^L}\pi_1[1+\beta\varphi\bar{\lambda}\Delta_{(C)}^h] + (1-\pi_1)\beta\varphi\bar{\lambda} = 0 \end{cases}$$

summing them up gives

$$\Delta_{(C)}^h[1+\beta\varphi\bar{\lambda}\Delta_{(C)}^h] - (\beta+h)\bar{\lambda} = 0 \tag{22}$$

, whose solution is strictly increasing in h. Solving for equilibrium strategies and plugging them back into π_1, one obtains

$$\pi_1 = \frac{1}{2} + \varphi\left[\frac{(1+\varphi\bar{\lambda}\beta\Delta_{(C)}^h)\mu_0 - h\bar{\lambda}\Delta_{(C)}^h}{1+2\varphi\bar{\lambda}\beta\Delta_{(C)}^h}\right]$$

therefore, the realized vote share must lie within $\pi_1 \pm \varphi/2\psi$. For $h>0$, one needs $\hat{\pi} > 1-\pi^h$ (while for the opposite case one would need to check that $\hat{\pi} < \pi^h$, which simplifies into a similar condition to the one that follows) Therefore, In order for the conjectured equilibrium to hold, we then need

$$\pi_1 - \varphi/2\psi > 1-\pi^h \tag{23}$$

From (7), one obtains $\frac{\psi-\varphi}{2\psi} > \varphi[\mu^h+d^h]$, and substituting (22) into the expression of $1-\pi^h$, one can see that $1-\pi^h < \varphi[\Delta_{(C)}^h]^2$. As a result, the following inequality would imply (23):

$$\varphi[\mu^h+d^h] + \varphi\left[\frac{(1+\varphi\bar{\lambda}\beta\Delta_{(C)}^h)\mu_0 - h\bar{\lambda}\Delta_{(C)}^h}{1+2\varphi\bar{\lambda}\beta\Delta_{(C)}^h}\right] > \varphi[\Delta_{(C)}^h]^2$$

Notice that, since $\mu^h = \max\{I+\lambda x\} > \bar{\lambda} > \Delta_{(C)}^h$, one just needs

$$d^h[1+2\varphi\bar{\lambda}\beta\Delta_{(C)}^h] + (1+\varphi\bar{\lambda}\beta\Delta_{(C)}^h)\mu_0 - h\bar{\lambda}\Delta_{(C)}^h > 0$$

which, using again (22), $2d^h \geq [x^*]^2$ and the fact that $\beta > h$, is implied by

$$[x^*]^2(1+\mu_0) > [\Delta_{(C)}^h]^2(1-\varphi\beta).$$

Solving the equilibrium for the $h<0$ yields a similar condition that puts an equally mild upper bound on μ_0.

References

[1] Acemoglu, D., Golosov, M., and Tsyvinski, A. Power fluctuations and political economy. Nber working papers, National Bureau of Economic Research, Inc, 2009.

[2] Acemoglu, D., Ticchi, D., and Vindigni, A. Emergence and persistence of inefficient states. NBER Working Papers 12748, National Bureau of Economic Research, Inc, 2006.

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

[4] Aidt, T. S., and Dutta, J. Policy myopia and economic growth. *European Journal of Political Economy 23*, 3 (2007), 734–753.

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

[6] Azzimonti-Renzo, M. Barriers to investment in polarized societies. mimeo, 2009.

[7] Bai, J., and Lagunoff, R. On the faustian dynamics of policy and political power. Working paper, Georgetown University, Department of Economics, 2008.

[8] Baron, D. P., and Ferejohn, J. A. Bargaining in legislatures. *American Political Science Review 83*, 4 (1989).

[9] Battaglini, M. Dynamic electoral competition and constitutional design. mimeo, 2010.

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

[11] Bénabou, R. Unequal societies: Income distribution and the social contract. *American Economic Review 90*, 1 (2000).

[12] Bénabou, R., and Ok, E. A. Social mobility and the demand for redistribution: The poum hypothesis. *The Quarterly Journal of Economics 116*, 2 (2001).

[13] Besley, T., and Coate, S. Sources of inefficiency in a representative democracy: A dynamic analysis. *American Economic Review 88*, 1 (1998), 139–56.

[14] Besley, T., and Persson, T. State capacity, conflict and development. STICERD - Economic Organisation and Public Policy Discussion Papers Series 010, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE, 2009.

[15] De Mello, L. R., and Tiongson, E. R. Income inequality and redistributive government spending. IMF Working Paper Series 03/14, 2003.

[16] Dixit, A., and Londregan, J. The determinants of success of special interests in redistributive politics. *The Journal of Politics 58*, 4 (1996).

[17] Garrí, I. Political short-termism: a possible explanation. *Public Choice 145*, 1/2 (2009).

[18] Gerber, A., and Ortuno-Ortiдn, I. Political compromise and endogenous formation of coalitions. *Social Choice and Welfare 15*, 3 (1998).

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

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

[21] Herrera, H., and Morelli, M. Turnout and power sharing. mimeo, 2010.

[22] Kalandrakis, T. A reputational theory of two party competition. Wallis Working Papers WP57, University of Rochester - Wallis Institute of Political Economy, 2008.

[23] Kelly, N. J., and Enns, P. K. Inequality and the dynamics of public opinion: The self-reinforcing link between economic inequality and mass preferences. *American Journal of Political Science 54*, 4 (2010).

[24] Krasa, S., and Polborn, M. Political competition between differentiated candidates. CESifo Working Paper Series 2560, 2009.

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

[26] Lijphart, A. *Democracy in Plural Societies: A Comparative Exploration.* Yale University Press, New Haven, 1980.

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

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

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

[30] Maskin, E., and Tirole, J. Markov perfect equilibrium: I. observable actions. *Journal of Economic Theory 100*, 2 (2001).

[31] McCarty, N., Poole, K. T., and Rosenthal, H. *Polarized America: The Dance of Ideology and Unequal Riches.* The MIT Press, 2008.

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

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

[34] Moene, K. O., and Wallerstein, M. Inequality, Social Insurance, and Redistribution. *The American Political Science Review 95*, 4 (2001).

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

[36] Persson, T., and Tabellini, G. *Political Economics: Explaining Economic Policy*. The MIT Press, 2002.

[37] Persson, T., and Tabellini, G. *The Economic Effects of Constitutions (Munich Lectures)*. The MIT Press, 2003.

[38] Persson, T., and Tabellini, G. Constitutional rules and fiscal policy outcomes. *American Economic Review 94*, 1 (2004), 25–45.

[39] Ticchi, D., and Vindigni, A. Endogenous constitutions. *Economic Journal 120*, 543 (2010).