A Political Theory of Populism*

Daron Acemoglu
MIT

Georgy Egorov
Northwestern University

Konstantin Sonin
New Economic School

August 2010

Abstract

When voters fear that politicians may have a right-wing bias or that they may be influenced or corrupted by the rich elite, signals of true left-wing conviction are valuable. As a consequence, even a moderate politician seeking reelection choose policies to the left of the median voter as a way of signaling that he is not from the right (while truly right-wing politicians also signal by choosing moderate or even left-of-center policies). This leftist bias of policy is greater when the value of remaining in office is higher for the politician; when there is greater polarization between the policy preferences of the median voter and right-wing politicians; and when politicians are indeed likely to have a hidden right-wing agenda. We show that similar results apply when some politicians can be corrupted or influenced through other non-electoral means by the rich elite.

Keywords: political economy, inequality, populism, voting, signaling.
JEL Classification: D71, D74, C71.

Work in Progress. Comments Welcome.

*We thank Timothy Feddersen, Emir Kamenica, Jean Tirole, Francesco Trebbi, participants of Political Economy in the Chicago Area conference, Non-Democratic regimes conference at Yale, Society for Economic Dynamics conference in Montreal, the Econometric Society World Congress in Shanghai, and seminars at Harvard and the New Economic School for valuable comments.

1 Introduction

There has recently been a resurgence of 'populist' politicians in many developing countries, particularly in Latin America. Hugo Chavez in Venezuela, the Kirchners in Argentina, Evo Morales in Bolivia, Alan Garcia in Peru, and Rafael Correa in Ecuador are some of the examples. The label 'populist' is often used to emphasize that these politicians use the rhetoric of aggressively defending the interests of the 'common man' against the privileged elite. Hawkins (2003), for example, describes the rise of Hugo Chavez in Venezuela in these terms, and writes:

> "If we define populism in strictly political terms—as the presence of what some scholars call a charismatic mode of linkage between voters and politicians, and a democratic discourse that relies on the idea of a popular will and a struggle between 'the people' and 'the elite'—then Chavismo is clearly a populist phenomenon."

Given the high levels of inequality in many of these societies, political platforms built on redistribution are not surprising. But populist rhetoric and policies frequently appear to be to the left of the median voter, and such policies arguably often harm rather than help the majority of the population. In the context of macroeconomic policy, Rudiger Dornbush and Sebastian Edwards (1991) emphasized this 'left of the median' aspect of populism and wrote:

> "Populist regimes have historically tried to deal with income inequality problems through the use of overly expansive macroeconomic policies. These policies, which have relied on deficit financing, generalized controls, and a disregard for basic economic equilibria, have almost unavoidably resulted in major macroeconomic crises that have ended up hurting the poorer segments of society."

In this paper, we offer a simple model of 'populism' defined, following Dornbusch and Edwards (1991), as the implementation of policies receiving support from a significant fraction of the population, but ultimately hurting the economic interests of this majority.[1] More formally, populist policies will be those that are to the left of the political bliss point of the median voter but still receive support from the median voter. The key challenge is therefore to understand why politicians adopt such policies and receive electoral support after doing so. Our starting point is that, as the above examples suggest, the economies in question feature high levels of inequality and sufficiently weak political institutions which enable the rich elites to have a disproportionate influence on politics relative to their numbers. In fact, in many of these societies

[1] Note that this is different from the related but distinct definition of populism originating from the People's Party in the 19th century United States.

political corruption and 'political betrayal,' where politicians using redistributive rhetoric end up choosing policies in line with the interests of the rich elite, are quite common (examples would include the rule of the PRI in Mexico, the policies of traditional parties in Venezuela and Ecuador, Fujimori's reign in Peru and Menem in Argentina). This implies that voters often distrust politicians and believe that they may adopt a rhetoric of redistribution, of leveling the playing field and of defending the interests of the 'common man', but then will pursue policies in the interests of the rich elite. This makes it valuable for politicians to signal to voters that they do not secretly have a 'right-wing agenda' and are not 'in the pockets' of the rich elite.

Formally, an incumbent politician chooses a policy x on the real line, and obtains utility both from remaining in office and also from the distance between the policy and his 'political bliss point' or ideal point (e.g., Wittman, 1973, Calvert, 1985, Besley and Coate, 1997, Osborne and Slivinski, 1996). An incumbent politician can be of two types: (1) left-center (moderate), and (2) right-wing. Alternatively, we also consider a version where the two types are: (1) honest, and (2) dishonest, and dishonest politicians can be bribed by the rich elite.[2] We normalize the political bliss point of the median voter to 0, and assume that this is also the bliss point of the moderate politician.[3] The bliss point of the right-wing politician is some $r > 0$ (i.e., to the right of the median voter's bliss point). Voters observe a noisy signal s of the policy x of the incumbent, and decide whether to reelect him for a second term or replace him by a new politician. Reflecting our discussion in the previous paragraph, the median voter's main concern is that the politician may in truth be a right-winger and will implement a right-wing policy in his second term (or that he is dishonest and will be corrupted by the rich elite).

Our main result is that in order to signal that he is not right-wing, moderate politicians will choose 'populist' policies to the left of the median voter's bliss point, i.e., $x < 0$. Moreover, a truly right-wing politician will also choose a policy to the left of his bliss point, i.e., $x < r$, and may even choose a policy to the left of the median (i.e., $x < 0$) when the value of political office is sufficiently high. It is interesting that what produces the left-wing bias in politics is precisely the strength of right-wing groups (in that either the incumbent politician may be secretly aligned with these groups or they can bribe and influence him). Because of their fear of reelecting a politician who is in truth a right-winger, voters support politicians choosing policies to the left of their preferences, which can loosely be interpreted as policies that are not in their 'best interest' as in our definition of populism.

[2] Although this alternative might be a better representation of the fears of many voters in democracies with weak institutions, we start with our baseline model because it illustrates the main ideas in a more transparent manner.

[3] Our results do not require the preferences of the moderate politician to coincide with those of the median voter; it is sufficient for them to be closer to those of the median voter than are the preferences of the right-wing politician.

In addition to providing a novel explanation for the emergence of populist policies and leaders, our model is tractable and leads to a range of intuitive comparative static results. First, policies are more likely to be populist (or will have greater left-wing bias) when the value of reelection to politicians is greater, since in this case both moderate and right-wing politicians will try to signal to voters by choosing more left-wing policies. Second, populist policies are also more likely when the probability that the politician is indeed a right-winger is higher. Third, they are also more likely when the probability that a politician can be corrupted is greater. Finally, we also show, under an additional but reasonable assumption, that populist policies are also more likely when there is greater 'polarization' in society, meaning a bigger gap between the political bliss points of the median voter and the moderate politician on the one hand and that of the right-wing politician on the other. This is because, with greater polarization, the benefit from reelection to both types of politicians is greater, encouraging more populist policies in the first period. However, counteracting this effect is that greater polarization also makes it more costly for right-wing politicians to adopt populist policies. Our additional assumption ensures that this second effect is dominated.

Interestingly, in the version of the model with corruption, we find that the rich elite can be worse off precisely because of its ability to bribe politicians. In particular, the anticipation of such bribes to dishonest politicians changes the political equilibrium towards more left-wing policies in the initial period, which is costly to the elite. This again highlights that the underlying problem leading to populist politics in this model is the weakness of democratic institutions and the potential non-electoral power of the elite.

Our paper is related to a number of literatures. First, there is now a sizable literature on signaling in elections. Formal models that incorporate 'the cost of betrayal' and signaling concerns into the platform choice by a politician seeking his first election date back to the pioneering work of Banks (1990) and Harrington (1993). In both models, voters learn about candidates' behavior through repeated sampling. Callander and Wilkie (2007) consider signaling equilibria in elections in which participating politicians have different propensity to misinform voters about their true preferences. Kartik and McAfee (2007) study a spatial model of elections in which a candidate might be of the type committed to fulfill his campaign pledge. Thus, a political position is a signal to the voters about the candidate's 'honesty'. As a result, a candidate who is perceived to be more likely to stick to his positition might win on an unpopular platform over an opponent who caters to the median voter's preferences.

Second, our paper is also related to several other works that use models in which politicians or decision-makers have private information and are judged on the basis of performance or messages that they send. Prendergast (1993) shows that workers have an incentive to conform to the opinions and expectations of their superiors. Morris (2001) studies 'political correctness' using a similar approach. Canes-Wrone, Herron, and Shotts (2001) and Maskin and Tirole (2004) use similar ideas to show why leaders or elected officials may 'pander' to the electorate. None of these papers discuss or derive populist bias in politics. In addition, the framework we present is more tractable than many of the models used in past work (because voters observe noisy signals rather than choices), and as a result, it generates a unique equilibrium and a rich set of comparative statics.

Third, our paper is also related to various models of political agency and the careers of politicians. Austen-Smith and Banks (1989), Besley (2005, 2006) and Persson and Tabellini (2000) present several different approaches to political agency and the selection of politicians of different competences (see also Caselli and Morelli, 2004, Messner and Polborn, 2004, Mattozzi and Merlo, 2007, and Acemoglu, Egorov, and Sonin, 2010). Diermeier, Keane, and Merlo (2005) model and estimate the careers of US congressmen.

Fourth, our work is related to the emerging literature on the elite capture of democratic politics. Acemoglu and Robinson (2008) and Acemoglu, Ticchi and Vindigni (2010) emphasize how a rich elite may be able to capture democratic politics and prevent redistributive policies. Bates and La Ferrara (2001), Lizzeri and Persico (2005) and Padro-i-Miquel (2007), among others, construct models in which certain forms of democratic competition may be detrimental to the interests of the majority. Acemoglu, Robinson and Torvik (2010) analyze a model of endogenous checks and balances. They show that in weakly-institutionalized democracies, voters may voluntarily dismantle checks and balances on presidents as a way of increasing their rents and making them more expensive to bribe for a better organized rich elite lobby. None of these papers note or analyze the possibility of populist (left of the median) policies.

Finally, our paper is also related to a few papers investigating other aspects of populist politics. Sachs (1989) discusses 'the populist cycle,' where high inequality leads to policies that make all groups worse off (because voters are shortsighted). Alesina (1989) emphasizes how redistributive policies are captured by special interest groups.

The rest of the paper is organized as follows. Section 2 introduces a basic model with politicians differing in their policy preferences. In Section 3 we analyze the equilibria of the model and study the comparative statics. In Section 4, we supplement our analysis by studying the case where some politicians can accept bribes. Section 5 concludes.

2 Model

There is a population consisting of two groups of citizens, a rich elite and a poor majority, and a pool of politicians. Policy space is represented by $\mathbb{R}$. There are two periods, $t = 1, 2$, and in each period there is a politician in power who chooses policy $x_t \in \mathbb{R}$. Citizens only care about policy outcomes. In particular, we assume that the utility of citizen i is given by

$$u_i(x_1, x_2) = u^r(x_1, x_2) = -\sum_{t=1}^{2} (x_t - \gamma^r)^2 \tag{1}$$

if the citizen is rich, and by

$$u_i(x_1, x_2) = u^p(x_1, x_2) = -\sum_{t=1}^{2} (x_t - \gamma^p)^2 \tag{2}$$

if he is poor. These preferences imply that both types of citizens are averse to deviations from their bliss policy. Without loss of generality, we set $\gamma^p = 0$, and $\gamma^r = r > 0$. We interpret this as the poor preferring more 'left-wing' policies, such as creating a more level playing field, redistribution or investment in public goods favoring the poor, whereas the rich prefer lower redistribution. Policies corresponding to $x < 0$ would be even more left-wing than the preferences of the poor. The assumption that there is no discounting is adopted to save on notation.

Since there are only two groups of voters, and the poor form the majority, the median voter is a poor agent, and $\gamma^p = 0$ corresponds to the political bliss point of the median voter. It is straightforward to see that preferences here are single peaked, and all our results would apply if there are more types of voters differing according to their political bliss points (again with the convention that $\gamma^p = 0$ is the median).

The politicians care about policy, office, and potentially bribes. Their utility is given by

$$v(x_1, x_2) = -\sum_{t=1}^{2} \left\{ \alpha (x_t - \gamma)^2 + W \mathbf{I}_{\{\text{in office at } t\}} + B_t \right\}. \tag{3}$$

Here, x_t is the policy implemented at time t (by this or another politician), γ is the politician's ideal policy, and $\alpha > 0$ is the sensitivity to policy choice; W is the utility from being in office, and B_t is the bribe that he may receive at time t. We consider two possible alternative assumptions about the politicians' type. First, we rule bribes out, and assume that there are two types of politician that differ in their policy preferences. A fraction μ of politicians are 'moderates' and have a political bliss point identical to that of the median voter, which is a poor agent, with political bliss point $\gamma = 0$.[4] Motivated by the discussion of Latin American politics in the

[4] All of our results generalize to an environment where, instead of two groups, there is a distribution of preferences, provided that the moderate politician is closer to the median voter than is the right-wing politician. We chose the model with two groups to clearly emphasize that populist policies are not adopted to cater to the preferences of some subgroup—in our model, they will always be to the left of the preferences of all voters.

Introduction, where voters may be concerned about politicians choosing more right-wing policies than they would like, we assume that the remaining fraction $1-\mu$ are 'right-wingers' and have a political bliss point coinciding with that of the rich, $\gamma = r$. Voters do not directly observe the type of the politician (so this type may be interpreted as the 'secret' leaning of the politician). The assumption that these political bliss points coincide with those of poor and rich voters again serves to reduce notation. The important feature is that poor voters should prefer to have a moderate in office in the second period. In Section 4, we consider a slight variant, which may be easier to map to the context of Latin American politics, where all politicians have the same policy preferences, but differ in their honesty. In particular, a fraction μ will be 'honest,' while the remaining fraction $1-\mu$ will be 'dishonest,' and thus can be bribed by the elite. The results in these two variants are very similar.

At the end of the first period, there will be an election in which the median voter will decide whether to reelect the politician or appoint a new politician from the pool of potential politicians. In particular, we model this by assuming that at the end of the first period there is a challenger of unknown type running against the challenger. Prior to the elections, voters receive a noisy (common) signal $s = x_1 + z$ about the policy x_1 chosen by the incumbent in the first period, where z is noise. Our interpretation for why voters observe a signal rather than the actual policy is that the welfare implications of policies of the take time to be fully realized and understood. All voters use this signal to update their priors about the politician's type and vote on the reelection of the incumbent politician. We assume that they do so in a forward-looking ('rational') manner and vote to retain the incumbent only if their posterior that he is a moderate type (or honest in the version in Section 4) is sufficiently high that they receive at least as high utility retaining him as appointing a new politician.

We will look for a pure-strategy perfect Bayesian equilibrium of the game (in undominated strategies), which ensures that the incumbent politician will be kept in power if the expected utility of poor voters is greater when he remains in power than when a new politician is appointed.[5]

The exact timing of events is as follows:

1. The politician in power at time $t = 1$ chooses policy $x_1 \in \mathbb{R}$.

2. Voters obtain the signal $s = x_1 + z$.

3. Elections take place, and each voter either supports the incumbent or the contender.[6]

[5] The requirement that the perfect Bayesian equilibrium should be in undominated strategies is for the usual reason that in voting games, non-intuitive equilibria can be supported when voters use weakly dominated strategies.

[6] For simplicity, we assume that whenever a voter is indifferent, she supports the incumbent. In equilibrium

4. The politician in power at time $t = 2$ (the incumbent or newly elected politician) chooses policy $x_2 \in \mathbb{R}$.

5. All agents learn the realizations of x_1 and x_2, and payoffs are realized (according to (1)–(3)).

We next impose the following two assumptions, which will be useful in establishing well-defined unique best responses.

Assumption 1 *The noise variable z has a distribution with support on $(-\infty, +\infty)$ with c.d.f. $F(z)$ and p.d.f. $f(z)$. The p.d.f. $f(z)$ is symmetric around 0 and everywhere differentiable, and satisfies $f'(z) < 0$ for all $z > 0$ (and thus $f'(z) > 0$ for all $z < 0$).*

This assumption ensures that the density f is single-peaked. Naturally, any mean zero normal distribution $\mathcal{N}(0, \sigma^2)$ as well as several other standard distributions satisfy this assumption.

Assumption 2 *The p.d.f. of the noise variable z is sufficiently smooth in the sense that*

$$\left|f'(z)\right| < \frac{1}{\frac{r^2}{2} + \frac{W}{2\alpha}} \tag{4}$$

for all z.

Lemma 1 in the Appendix shows that this assumption implies that $f(0) < \frac{2}{r}$. In other words, values close to zero are not too likely, and there is a sufficient probability that there will be large (negative or positive) realizations of the noise, i.e., $\Pr\left(|z| > \frac{r}{4}\right) > \frac{1}{4}$. This feature ensures that there is sufficient uncertainty, which will guarantee that the relevant second-order conditions hold and thus the existence and uniqueness of equilibrium. It is straightforward to see that a normal distribution would satisfy this assumption if its variance is sufficiently large, in particular, if $\sigma^2 > \frac{\frac{r^2}{2} + \frac{W}{2\alpha}}{\sqrt{2\pi e}}$. One can also verify that until Section 4, it is sufficient to impose (4) only for $|z| < r$.

3 Analysis

We proceed by backward induction. At date $t = 2$, the politician in power solves the problem

$$\max_{x_2 \in \mathbb{R}} -\alpha \left(x_2 - \gamma\right)^2,$$

and therefore chooses his bliss point $x_2 = \gamma$ (which equals $\gamma^p = 0$ for a moderate politician and $\gamma^r = r$ for a right-wing politician). This implies that the poor majority strictly prefers to have

this happens with zero probability.

a moderate politician than a right-wing politician in power at date $t = 2$. Since the contender is a moderate with probability μ, the incumbent will win the elections only if the voters' posterior that he is moderate is no less than μ.

Let us denote the equilibrium policy that a moderate politician chooses at time $t = 1$ by $x_1 = a$, and the policy that a right-wing politician chooses by $x_1 = b$. It is intuitive that $a < b$, and this result is formally established in Proposition 1. Clearly, the probability density of signal s when policy x is chosen is given by $f(s - x)$. Given the prior μ, Bayesian updating gives the posterior that the incumbent is a moderate by

$$\hat{\mu} = \frac{\mu f(s-a)}{\mu f(s-a) + (1-\mu) f(s-b)}. \tag{5}$$

Inspection of (5) shows that the posterior $\hat{\mu}$ satisfies $\hat{\mu} \geq \mu$ if and only if

$$f(s-a) \geq f(s-b). \tag{6}$$

Intuitively, the right-hand side of (5) depends on a and b only through the likelihood ratio $\frac{\mu f(s-a)}{(1-\mu)f(s-b)}$, which must be at least $\frac{\mu}{1-\mu}$ (the corresponding ratio for the contender) if the incumbent is to be reelected; hence $\frac{f(s-a)}{f(s-b)} \geq 1$. By Assumption 1, $f(\cdot)$ is symmetric and single-peaked, and thus (6) is equivalent to

$$s \leq \frac{a+b}{2}. \tag{7}$$

The incumbent will be reelected if and only if condition (7) is satisfied. Therefore, the expected probability of reelection for an incumbent as a function of his choice of policy x is

$$\begin{aligned} \pi(x) &= \Pr\left(x + z \leq \frac{a+b}{2}\right) \\ &= F\left(\frac{a+b}{2} - x\right). \end{aligned} \tag{8}$$

Note that this probability does not depend on the type of the incumbent, only on his choice of policy (since his type is private and does not affect the realization of the signal beyond his choice of policy).

We next establish that $a < b$ and summarize the discussion about incumbent reelection.

Proposition 1 *Denote the equilibrium policy of a moderate politician at $t = 1$ by $x_1 = a$ and that of a right-wing politician by $x_1 = b$. Then:*

1. *$a < b$, i.e., moderate politicians always choose a more left-wing policy than right-wing politicians;*

2. *the incumbent politician is reelected if and only if $s \leq \frac{a+b}{2}$.*

Proof. See the Appendix. ■

Let us next investigate choices at date $t = 1$ more closely. A moderate incumbent solves the following problem at date $t = 1$:

$$\max_{x \in \mathbb{R}} -\alpha x^2 + W\pi(x) - (1-\mu)\alpha r^2 (1 - \pi(x)). \tag{9}$$

Here, $-\alpha x^2$ is this politician's first period utility, W is his second period utility if he is reelected (since in this case he will choose $x_2 = 0$, i.e., equal to his political bliss point), and $-(1-\mu)\alpha r^2$ is his expected second period utility if he is not reelected (with probability μ the contender is a moderate and will choose $x_2 = 0$, while with probability $1-\mu$, a right-winger will come to power and will choose $x_2 = r$). The first-order condition for this problem is

$$-2\alpha x - \left(W + (1-\mu)\alpha r^2\right) f\left(\frac{a+b}{2} - x\right) = 0. \tag{10}$$

Similarly, a right-wing incumbent solves the problem

$$\max_{x \in \mathbb{R}} -\alpha (x-r)^2 + W\pi(x) - \mu\alpha r^2 (1 - \pi(x)). \tag{11}$$

The explanation for this expression is identical and relies on the fact that the right-wing incumbent will incur utility cost αr^2 only if he is not reelected and is replaced by a moderate politician. The first-order condition for this problem is

$$-2\alpha (x-r) - \left(W + \mu\alpha r^2\right) f\left(\frac{a+b}{2} - x\right) = 0. \tag{12}$$

In equilibrium, (10) must hold when $x = a$, and (12) must hold when $x = b$. The symmetry of f implies

$$f\left(\frac{a+b}{2} - a\right) = f\left(\frac{b-a}{2}\right) = f\left(\frac{a-b}{2}\right) = f\left(\frac{a+b}{2} - b\right),$$

so the equilibrium can be characterized in terms of the intersection of two reaction curves:[7]

$$-2\alpha a - \left(W + (1-\mu)\alpha r^2\right) f\left(\frac{b-a}{2}\right) = 0, \tag{13}$$

$$-2\alpha (b-r) - \left(W + \mu\alpha r^2\right) f\left(\frac{b-a}{2}\right) = 0. \tag{14}$$

Mathematically, (10) depicts the equilibrium value of the policy of moderate politicians, a, when right-wing politicians are choosing policy b. Conversely, (12) depicts the equilibrium value of the policy of right-wingers when moderates are choosing a. Figure 1 plots these two curves in the relevant region $a \leq b$.

[7] Note, however, that these are not 'best response maps' as we have already substituted for the equilibrium conditions that (10) must hold at $x = a$ and (12) at $x = b$.

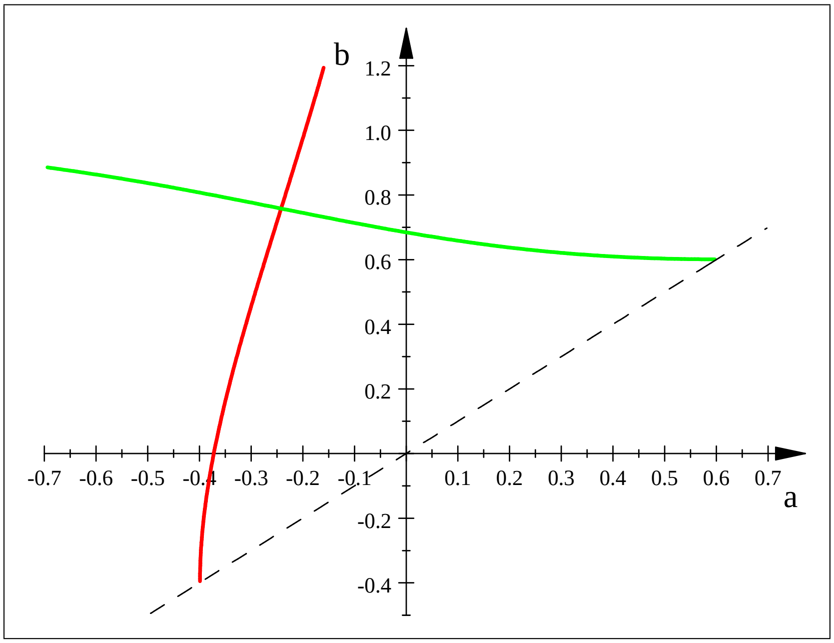

Figure 1: Reaction curves of moderate (red/thick) and right-wing (green/thin) politicians. The distribution f is taken to be normal and the parameter choices are $W = \mu = \sigma = \frac{1}{2}$ and $r = \alpha = 1$.

Conditions (13) and (14) immediately imply that $a < 0$ and $b < r$, i.e., both types of politicians in equilibrium choose policies which lie to the left of their political bliss points. For moderate politician, this implies a populist policy choice—i.e., to the left of the median voter's political bliss point (we will also see that we might even have $b < 0$). This is for an intuitive reason: for both types of politicians, a move to the left starting from their political bliss points creates a second-order loss in the first period, but delivers a first-order increase in the probability of reelection and thus a first-order expected gain.

The result that there will be a left bias in policies does not rely on positive benefits from holding office ($W > 0$), though we will establish later that higher levels of W increase this bias. This is because even when $W = 0$, each politician wants to be reelected because otherwise his preferred policy will be implemented with probability less than one. This effect alone is sufficient for a left bias in policy choice of both types.

Inspection of Figure 1 also provides a more detailed intuition for the results and be uniqueness of equilibrium. The reaction curve of moderate politicians is upward-sloping, while the reaction curve of right-wing politicians is downward-sloping. Formally, these statements follow from differentiating (10) and (12) with respect to a and b. The key observation is that the median voter will decide whether to reelect the incumbent politician depending on whether $\frac{f(s-a)}{f(s-b)}$ exceeds 1. A politician may ensure that he is reelected with an arbitrarily large probability if he chooses an extreme left-wing policy, but this is clearly costly as the policy would be very far away from his bliss point. The relevant trade-off for both types of politicians is therefore between choosing a policy close to their bliss point on the one hand and deviating from their bliss point and increasing their reelection probability on the other. But how much this deviation will increase their reelection probability depends on the expectations of the median voter concerning what types of policies both types of politicians will adopt. Formally, the question is whether a small change in policy will increase $\frac{f(s-a)}{f(s-b)}$ from below 1 to above 1 (which thus requires that $\frac{f(s-a)}{f(s-b)}$ is in some ε-neighborhood of 1). Suppose, for example, that right-wing politicians are expected to choose a more left-wing policy than before. This would make the policies of the two types closer, and it becomes harder for voters to distinguish one type of politicians from another (equivalently, $\frac{f(s-a)}{f(s-b)}$ is more likely to be in any given ε-neighborhood of 1). In response, it would be optimal for moderate politicians to choose an even more left-wing policy in order to distinguish themselves and get reelected with a high probability. This is the reason why (13) is upward-sloping.

Why is (14) downward-sloping? Consider the situation in which moderate politicians are expected to choose more left-wing policies. One might, at first, expect that the same reasoning should push right-wing politicians to also choose more left-wing policies. But since $a < b$, a further shift to the left by moderate politicians will make it more likely that the median voter will be able to distinguish moderate than right-wing politicians (or more formally, $\frac{f(s-a)}{f(s-b)}$ is now less likely to be in any given ε-neighborhood of 1, and thus a small shift to the left by right-wing politicians is less likely to win them the election). This reduces the potential gains from choosing further left-wing policies for right-wing politicians and encourages them to choose policies more in line with their own preferences.

The above discussion ensures the uniqueness of equilibrium. Then characterizing this equilibrium is straightforward. Let us combine (13) and (14) to obtain

$$b = a\frac{W + \mu\alpha r^2}{W + (1-\mu)\,\alpha r^2} + r.$$

Therefore, the equilibrium level of populist bias (of moderate politicians), $p = |a| = -a$, is

$$2\alpha p = \left(W + (1-\mu)\,\alpha r^2\right) f\left(\frac{p}{2}\alpha r^2 \frac{1-2\mu}{W + (1-\mu)\,\alpha r^2} + \frac{r}{2}\right). \tag{15}$$

Proposition 2 *There exists a unique (perfect Bayesian) equilibrium (in pure strategies). In this equilibrium, politicians choose their preferred policy in the second period. In the first period, politicians choose policies which lie to the left of their preferred policy. In particular, moderate politicians necessarily choose a 'populist' policy to the left of the political bliss point of the median voter.*

Proof. See the Appendix. ■

Proposition 2 thus shows that a unique equilibrium always exists (given our assumptions), and that this equilibrium will always feature leftist bias by both types of politicians, and populist policies (to the left of the median voter) by moderate politicians. The intuition for this result was already provided above.

We next provide several comparative static results.

Proposition 3 *The populist bias of moderate politicians, $p = |a|$, is higher when:*

1. *W is higher (i.e., politicians value being in office more);*
2. *μ is lower (i.e., poor politicians are rarer);*
3. *α is lower, provided that $W \neq 0$ (i.e., changing political positions is relatively costless for politicians).*

Proof. See the Appendix. ■

These comparative static results are fairly intuitive. A higher utility from remaining in office increases the incentives for reelection and thus encourages greater signaling by choosing more left-wing policies. Consequently, moderate politicians and up choosing more populist policies. The comparative static result with respect to α has a similar intuition: when α is lower, this utility from choosing a policy different from their bliss point is less important relative to the gain of utility from remaining in office, and this encourages more signaling and thus more left-wing policies by both politicians.

The comparative statics with respect to μ—the share of politicians that are moderate—is a little more subtle. When μ is lower, the probability that a new politician, if the incumbent is not reelected, will be moderate is lower. This makes it more costly for moderate politicians not to be reelected and induces them to choose more left policies. There is a countervailing effect, however, because the same calculus implies that right-wing politicians have less to fear from not being reelected, which makes them choose less left policies; because the reaction curve of modern politicians is upward-sloping, this effect induces them to choose less left biased policies. Nevertheless, when the density f is not high enough (in particular when it satisfies Assumption 2), the latter effect is dominated by the former, and the populist bias of moderate politicians increases.

Our next result shows that under an additional assumption on the form of the noise, an increase in polarization (corresponding to a bigger gap between the bliss points of moderate and right-wing politicians) also increases populist bias. To simplify the statement of the result, we will impose that the distribution of the noise z is normal. The proof of the proposition shows that this assumption can be replaced by a stronger version of Assumption 2.

Proposition 4 *Suppose that noise F is normally distributed. Then:*

1. *If $W = 0$ and $r < 2\sigma$, then an increase in polarization increases the populist bias of moderate politicians.*

2. *As W increases, the effect of polarization on the populist bias diminishes.*

Proof. See the Appendix. ■

This proposition shows that, under the additional assumptions we have imposed, populist policies are also more likely when there is greater 'polarization' in society, meaning a bigger gap between the political bliss points of the median voter and the moderate politician on the one hand and that of the right-wing politician on the other. There is a simple intuition underlying this result: with greater polarization, the benefit from reelection to both types of politicians is greater, encouraging more populist policies in the first period. However, the result is not ambiguous (hence the need for imposing the additional conditions) because there is also a countervailing effect: greater polarization also makes it more costly for right-wing politicians to choose populist policies. Our additional assumptions ensure that this second effect is dominated.

Our final result in this section investigates whether right-wing politicians will also choose populist policies (to the left of the median voter).

Proposition 5 1. *In the absense of direct benefits from holding office (i.e., if $W = 0$), right-wing politicians never choose choose policies to the left of median voter's political bliss point (i.e., we have $b > 0 > a$).*

2. *b is decreasing in W.*

3. *If $W > 0$, the equilibrium may involve $a < b < 0$.*

Proof. See the Appendix. ■

When there are no direct benefits or rents from holding office (i.e., $W = 0$), then right-wing politicians will never bias their policies so much to the left as to end up to the left of the median. This is because the cost of not getting reelected is to see the political bliss point of a moderate politician, which is the same as the median voter's bliss point, implemented.[8] However, higher

[8] This result continues to hold, a fortiori, if the future is discounted. But if the second period is more important in politicians' uility than the first period, then there may be populist bias in right-wing politicians' choices even with $W = 0$.

benefits from holding office increase the left bias of not only moderate politicians, but also of right-wing politicians. Consequently, for sufficiently high values of W, remaining in office is so valuable for right-wing politicians that they may also end up choosing populist policies.

4 Corruption

We have so far assumed that politicians have policy preferences that are private information and some politicians may (secretly) have right-wing preferences. Politicians then choose policies to signal to the voters that they do not have such right-wing preferences. An equally (or more) plausible reason why politicians may wish to signal to voters is to convince them that they are not unduly influenced, or captured, by the rich elite. This is particularly relevant in weak democracies, such as many in Latin America, where the rich elite have several non-electoral means of influencing politics (including lobbying or direct bribery) and may have a disproportionate impact on policies. In this section, we investigate the implications of voters' concern that politicians may be unduly influenced by the rich elite.

To focus on the main issue at hand, we now assume that politicians are identical in terms of their policy preferences (and without sacrificing qualitative conclusions, these coincide with the preferences of the majority of the electorate), but instead, a fraction μ of politicians are 'honest' in the sense that they cannot or will not cannot accept bribes, while the remaining $1-\mu$ share are 'dishonest' and thus corruptible. The bribing process is potentially imperfect (inefficient), and we capture this by assuming that whenever there is such a transaction, the parties must incur cost $C \geq 0$ to avoid being detected. This introduces a parameter that will be useful for comparative statics (and we could set $C = 0$). After the cost C is subtracted, the remainder of the surplus is split between the rich elite and the politician in power, and the politician receives a share χ (on the remaining surplus, i.e., the surplus minus the cost C).

The timing of the new game is therefore as follows.

1. The politician in power at time $t = 1$ and the elite bargain over x_1 (if the politician is dishonest), and the politician chooses policy $x_1 \in \mathbb{R}$.

2. Voters obtain the signal $s = x_1 + z$.

3. Voters vote, and decide whether to replace the current incumbent with a random one drawn from the pool of potential politicians.

4. The politician in power at time $t = 2$ (the incumbent or newly elected politician) and the elite bargain over x_2, and the politician chooses policy $x_2 \in \mathbb{R}$.

5. All agents learn the realizations of x_1 and x_2, and payoffs are realized according to (1)-(3) (with B_t given by the share χ of the surplus minus C).

This timing of the sizes that the structure of the game is very similar to that in Section 2, except for the bargaining between the politician and the elite in stages 1 and 4. Note that parameter α which has so far captured the intensity of politicians' policy preferences may now be given different/complementary interpretations. First, a small α may also correspond to (or result from) a large size of the elite, benefiting from right-wing policies, making bribing can be relatively inexpensive for the elite. Conversely, a large α could be interpreted as the elite constituting a relatively small group of the population. Second, a large α can also be interpreted as the politician standing in for a large bureaucracy, thus increasing the monetary and transactional costs of bribing.

As before, we start our analysis with the second period. Then, an honest politician will choose his bliss point $x_2^h = 0$. A corrupt politician, on the other hand, will bargain with the elite, and the equilibrium policy is determined from the joint maximization of the sum of their surpluses (and the bribing will actually take place if the surplus is high enough). Namely, they maximize

$$\max_{x_2^c \in \mathbb{R}} \left(-\alpha \left(x_2^c\right)^2 - \left(x_2^c - r\right)^2 \right),$$

which has the solution

$$x_2^c = \frac{r}{1+\alpha}. \tag{16}$$

Naturally, a higher α, which corresponds either to a greater weight on politician's preferences or smaller elite that will benefit from biasing the policy, implies a policy closer to the politician's political bliss point. As a consequence, the second period joint utility of a dishonest politician in power and the elite is

$$W - \frac{\alpha r^2}{\alpha+1},$$

and the gain in utility, as compared to the status-quo $W - r^2$ (since without the bribe, the politician would choose $x_2^c = 0$), is $\frac{r^2}{\alpha+1}$. This means that bribing in the second period will occur only if

$$C < \frac{r^2}{\alpha+1}. \tag{17}$$

In this case, we can simply determine the level of the bribe B_2 from the equation

$$-\alpha \left(x_2^c\right)^2 + B_2 - C = \chi \frac{r^2}{\alpha+1}.$$

This implies:

$$B_2 = \left(\chi + \frac{\alpha}{\alpha+1} \right) \frac{r^2}{\alpha+1} + C.$$

Interestingly, the effect of the intensity of politician's preferences on the bribe is nonmonotonic: the bribe reaches its maximum at $\alpha = \frac{1-\chi}{1+\chi}$; for lower α, the bribe is smaller because the politician is very cheap to persuade, and for very large α, the politician is too hard to bribe, hence in the limit the bribe disappears.

The following proposition characterizes several useful features of the equilibrium.

Proposition 6 *Suppose that C is such that* (17) *holds. Then:*

1. *In the second period, honest politicians choose $x_2^h = 0$, while corrupt politicians choose $x_2^c = \frac{r}{1+\alpha}$;*

2. *In the first period, honest and corrupt politicians choose policies $x_1^h = a$ and $x_1^c = b$ such that $a < b$;*

3. *The median voter prefers to have an honest politician in the second period, and reelects the incumbent politician if and only if he receives a signal*
$$s \geq \frac{a+b}{2}.$$

Proof. See the Appendix. ■

Part 1 of Proposition 6 was established earlier, while Parts 2 and 3 are similar to the results established in Proposition 1 in the case without bribery.

Let us next turn to the first-period problem. If the second period involves no bribing, then in the first period, politicians have no reelection concerns—recall that when indifferent voters reelect the incumbent. As a consequence, politicians will solve an identical problem in the second period, and thus the solution is the same and involves no corruption. The interesting case where dishonest politicians accept bribes in the second period (which is the case when (17) holds) is the one studied in Proposition 6. In this case, the probability of reelection of a politician who chooses policy x is again given by (8). An honest politician does not accept bribes and thus solves the problem

$$\max_{x \in \mathbb{R}} -\alpha x^2 + W\pi(x) - (1-\mu)\alpha\left(\frac{r}{1+\alpha}\right)^2 (1-\pi(x)). \tag{18}$$

Note that this problem is the same as (9), except that when the honest incumbent is replaced by a dishonest politician, his utility is $-\alpha\left(\frac{r}{1+\alpha}\right)^2$ instead of $-\alpha r^2$ (since the second period policy in this case will be given by (16)).

This maximization problem gives the first-order condition

$$-2\alpha x - \left(W + (1-\mu)\alpha\left(\frac{r}{1+\alpha}\right)^2\right) f\left(\frac{a+b}{2} - x\right) = 0. \tag{19}$$

A dishonest politician bargains with the elite both in the first and second periods. In the first period, bargaining, anticipating the second-period choices, gives the following joint maximization problem

$$\max_{x\in\mathbb{R}}\left\{\begin{array}{c} -\alpha x^2-(x-r)^2+\left(W-\frac{\alpha r^2}{\alpha+1}-C\right)\pi(x) \\ -(1-\mu)\left(\frac{\alpha r^2}{\alpha+1}+\left(\chi+\frac{\alpha}{\alpha+1}\right)\frac{r^2}{\alpha+1}\right)(1-\pi(x))-\mu r^2(1-\pi(x)) \end{array}\right\}. \tag{20}$$

The first two terms relate to the first period's utilities of the incumbent and the elite, respectively. If this (dishonest) politician is reelected, then together with the elite he jointly obtains second-period utility $W-\frac{\alpha r^2}{\alpha+1}-C$. If he is not reelected but another dishonest politician comes to power, their joint utility is $-\frac{\alpha r^2}{\alpha+1}-B_2$ (the same policy is implemented and the elite has to pay the same second-period bribe, but the current incumbent receives neither the direct benefits from holding office nor the bribes). Finally, if an honest politician is elected, the incumbent and the elite together obtain $-r^2$.

The first-order condition of their maximization problem gives:

$$-2\alpha x-2(x-r)-\left(W-C+\frac{r^2(\alpha+\mu)}{(1+\alpha)^2}+\frac{(1-\mu)\chi r^2}{1+\alpha}\right)f\left(\frac{a+b}{2}-x\right)=0. \tag{21}$$

Since first-order conditions (19) and (21) must be satisfied in equilibrium for $x=a$ and $x=b$, respectively, the equilibrium is now characterized by the following two equations:

$$-2\alpha a-\left(W+\frac{(1-\mu)\alpha r^2}{(1+\alpha)^2}\right)f\left(\frac{b-a}{2}\right)=0 \tag{22}$$

and

$$-2\alpha b-2(b-r)-\left(W-C+\frac{r^2(\alpha+\mu)}{(1+\alpha)^2}+\frac{(1-\mu)\chi r^2}{1+\alpha}\right)f\left(\frac{b-a}{2}\right)=0. \tag{23}$$

From (22) and (23), we obtain

$$b=\frac{1}{\alpha+1}\left(r+\frac{\alpha a\left(W-C+\frac{r^2(\alpha+\mu)}{(1+\alpha)^2}+\frac{(1-\mu)\chi r^2}{1+\alpha}\right)}{W+\frac{(1-\mu)\alpha r^2}{(1+\alpha)^2}}\right).$$

Plugging this into (22) and denoting, as before, the populist bias (of honest politicians) by $p=|a|=-a$, we have the following equilibrium condition for populist bias:

$$2\alpha p=\left(W+\frac{(1-\mu)\alpha r^2}{(1+\alpha)^2}\right)f\left(\frac{r}{2(1+\alpha)}+p\frac{W-\frac{(1-\mu)\chi\alpha r^2}{1+\alpha}+\frac{r^2\alpha(1-2\mu-\alpha\mu)}{(1+\alpha)^2}+C}{2(1+\alpha)\left(W+\frac{(1-\mu)\alpha r^2}{(1+\alpha)^2}\right)}\right). \tag{24}$$

The next proposition follows from our analysis so far and from this expression.

Proposition 7 *Suppose (17) holds (so there is corruption in the second period when the politician is dishonest). Then:*

1. *There exists a unique equilibrium (in pure strategies). In this equilibrium, honest politicians choose populist policies in the first period (i.e., $a < 0$). Dishonest politicians accept a bribe.*

2. *When $W = \chi = 0$, dishonest politicians will not choose populist, i.e., $b > 0$. When either $W > 0$ and/or $\chi > 0$, dishonest politicians may choose populist policies, i.e., $b < 0$.*

Proof. See the Appendix. ■

The intuition for why honest politicians now choose populist policies is similar to before: a small move to the left starting from their political bliss point has a second-order cost in terms of first-period utility, and a first-order gain in terms of the probability of reelection and second-period utility. The last part of the result suggests that if dishonest politicians do not value being in power per se (obtaining neither direct benefits from being in office nor positive rights), then the equilibrium will never involve populist policies in the first period. This is because reelection has limited benefits for dishonest politicians and this makes populist policies jointly too costly for the elite and the politician. The converse case is more interesting. In this case, even though dishonest politicians are effectively representing the wishes of the elite (because of bribery), they will still choose populist policies so as to increase their likelihood of coming to power and obtaining higher utility (for themselves and for the elite) in the second period.

The comparative static results of populist bias are again intuitive, but they also help us to further clarify the nature of the results in this case.

Proposition 8 *Suppose that (17) holds. Then, the populist bias of honest politicians, $p = |a|$, is higher when:*

1. *W is higher (greater direct utility from holding office);*

2. *C is lower (greater gains from the election for this honest politicians because bribing is more efficient);*

3. *χ is higher (dishonest politicians have higher bargaining power vis-à-vis the elite);*

4. *μ is lower (honest politicians are relatively rare).*

Proof. See the Appendix. ■

A higher W makes both politicians value reelection more; they thus engage in more signaling (by choosing more populist policies). A lower C increases the joint utility of the elite and the incumbent in case of reelection, and this makes dishonest politicians choose more left-wing policies, which in turn enables honest politicians to do the same. A higher χ makes holding office more valuable for dishonest politicians. This again makes them seek office more aggressively by choosing policies that are more left-wing, and honest politicians respond by also changing their policies to the left. Since they were already to the left of the median voter, this increases the populist bias of honest politicians. Finally, if honest politicians are rare, it makes them even more willing to signal because the population's prior is that dishonest politicians are the norm not the exception.

Our discussion so far suggests that the potential corruption of politicians empowers the elite to secure policies that are more favorable to their interest. However, as we emphasized in the Introduction, there is a countervailing effect: the fact that the elite will be able to influence politics leads to equilibrium signaling by choosing more left-wing policies. In fact, we have seen that this always makes honest politicians choose populist policies to the left of the preferences of the median voter. This raises the possibility that the elite's ability to bribe politicians may actually harm them (by creating a strong left-wing bias in the first period). The next proposition shows that when W is sufficiently large, the elite may be worse off when they are able to bribe than hypothetical world in which they are never able to influence politics via bribes.

Proposition 9 *There exists $\bar{W}$ such that if $W > \bar{W}$, the elite are better off when (17) does not hold as compared to the case in which it holds (or in the extreme, where $C = 0$).*

Proof. See the Appendix. ■

Intuitively, when politicians receive sufficient rents from holding office, policy choices by both honest and dishonest politicians will be more to the left, and as a result, despite their ability to bribe dishonest politicians, the elite will end up with first-period policies that are very far from there preferences. In this case, they may have lower utility than a situation without the possibility of bribery (which would have led to the implementation of the median voter's bliss point in both periods). This result thus shows that weak institutions, which normally empower the elite, may at the end make them worse off because of the endogenous response of democratic policies, even if democracy works only imperfectly.

5 Conclusion

In this paper, we presented a simple theory of populist politics. Populist politics is interpreted as (some) politicians adopting populist policies that are harmful to the rich elite but are not in the best interest of the poor majority or the median voter. More specifically, such policies, which may at least on the surface involve defending the rights of the poor against the elite, establishing redistributive programs and leveling the playing field, are to the left of the bliss point of the median voter, but still receive support from the median because they signal that the politician does not have a secret right-wing agenda and is not unduly influenced by the rich elite. The driving force of populist politics is the weakness of democratic institutions, which makes voters believe that politicians, despite their rhetoric, might have a right-wing agenda or may be corruptible or unduly influenced by the elite. Populist policies thus emerge as a way for politicians to signal that they will choose future policies in line with the interests of the median voter.

We show that moderate politicians will necessarily choose policies to the left of the median voter's preferences, and even right-wing politicians (or those that are captured and bribed by the elite) may end up choosing policies to the left of the median voter. This leftist (populist) bias of policy is greater when the value of remaining in office is higher for the politician; when there is greater polarization between the policy preferences of the median voter and right-wing politicians; and when politicians are indeed likely to have a hidden right-wing agenda.

When (some) politicians can be bribed, we also find that the efficiency of the process of bribery and the share of the surplus that politicians can capture also encourage left- wing bias (because they make politicians more eager to get reelected). Interestingly, in this case, the elite may be worse off than a situation in which institutions are stronger and bribery is not possible (because the equilibrium left-wing bias of first-period policies is more pronounced).

Our paper and model have been motivated by Latin American politics, where populist policies and rhetoric as well as fears of politicians reneging on their redistributive agenda and being excessively influenced by rich and powerful elites right have been commonplace. Nevertheless, the ideas here can be applied to other contexts. If voters are afraid of a secret left-wing agenda, then the equilibrium will create a force towards right-wing policies. Similarly, if bureaucrats are expected to show a bias in favor of a particular group or a particular type of policy, their actions may be biased in the opposite direction to dispel these notions and guarantee good performance evaluation. One could also study variants in which both types of extremism are possible. Finally, our model has focused on a two-period economy to communicate the basic ideas in the clearest fashion. In a multi-period setup, politicians may choose biased policies for several periods. Despite the tractability of our basic model, the infinite-horizon extension turns out to be challenging and is an open area for future research.

References

Acemoglu, Daron (2003) "Why Not a Political Coase Theorem? Social Conflict, Commitment, and Politics," *Journal of Comparative Economics,* 31, 620-652.

Acemoglu, Daron, Egorov, Georgy, and Konstantin Sonin (2010) "Political Selection and Persistence of Bad Governments," forthcoming *Quarterly Journal of Economics*, 125:4.

Acemoglu, Daron and James Robinson (2008) "Persistence of Elites, Power and Institutions," *American Economic Review,* 98: 267-293.

Acemoglu, Daron, James Robinson and Ragnar Torvik (2010) "Equilibrium Checks and Balances," MIT mimeo.

Acemoglu, Daron, Davide Ticchi, and Andrea Vindigni (2010) "Emergence and Persistence of Inefficient States," forthcoming *Journal of European Economic Association*,.

Alesina, Alberto (1989) "The Political Economy of Macroeconomic Stabilization: Myths and Reality," in *Income distribution and high-quality growth*, Vito Tanzi, Keyoung Chu, eds.

Austen-Smith, David and Jeffrey Banks (1989) "Electoral Accountability and Incumbency." In *Models of Strategic Choice in Politics*, edited by Peter Ordeshook, Ann Arbor MI: University of Michigan Press.

Banks, Jeffrey (1990) "A Model of Electoral Competition with Incomplete Information," *Journal of Economic Theory*, 50(2): 309-325.

Bates, Robert and Eliana La Ferrara (2001) "Political Competition in Weak States," *Economics and Politics,* 13: 159–184.

Besley, Timothy (2005) "Political Selection," *Journal of Economic Perspectives*, 19(3): 43-60.

Besley, Timothy (2006) *Principled Agents?* Oxford: Oxford University Press.

Besley, Timothy and Stephen Coate (1997) "An Economic Model of Representative Democracy," *Quarterly Journal of Economics* 112:85-114.

Callander, Steven and Simon Wilkie (2007) "Lies, Damned Lies, and Political Campaigns," *Games and Economic Behavior* 60: 262-286.

Calvert, Randall (1985) "The Value of Biased Information. A Rational Model of Policy Advice," *Journal of Politics*, 47: 530-555.

Canes-Wrone, Brandice, Michael Herron, and Kenneth Shotts (2001) "Leadership and Pandering: A Theory of Executive Policymaking." *American Journal of Political Science* 45:532-550.

Caselli, Francesco and Massimo Morelli (2004) "Bad Politicians," *Journal of Public Economics*, 88(3-4): 759-782

De la Torre, Carlos (2000) *Populist Seduction in Latin America. The Ecuadorian Experience.* Ohio University Press.

Diermeier, Daniel, Michael Keane, and Antonio Merlo (2005) "A Political Economy Model of Congressional Careers," *American Economic Review.* 95(1): 347-373.

Drake, Paul W. (1978) *Socialism and Populism in Chile, 1932-52.* Urbana: University of Illinois Press.

Dornbush, Rudiger and Sebastian Edwards, eds. (1991) *The Macroeconomics of Populism in Latin America.*

Ferejohn, John (1986) "Incumbent Performance and Electral Control," *Public Choice* 50, 5-26.

Harrington, Joseph (1993) "The Impact of Re-election Pressures on the Fulfillment of Campaign Promises," *Games and Economic Behavior* 5, 71-97.

Hawkins, Kirk (2003) "Populism in Venezuela: The rise of Chavismo," *Third World Quarterly*, 2003.

Kartik, Navin and Preston McAfee (2007) "Signaling Character in Electoral Competition," *American Economic Review* 97, 852-870.

Lizzeri, Alessandro and Nicola Persico (2005) "A Drawback Of Electoral Competition," *Journal of the European Economic Association,* 3(6): 1318-1348

Maskin, Eric and Jean Tirole (2004) "The Politician and the Judge: Accountability in Government," *American Economic Review*, 94(4), 034-1054

Mattozzi, Andrea and Antonio Merlo (2007) "Mediocracy," CEPR Discussion Papers 6163.

Messner, Matthias and Mattias Polborn (2004) "Paying politicians," *Journal of Public Economics*, 88(12), 2423-2445

Morris, Stephen (2001) "Political Correctness," *Journal of Political Economy,* 109 (2): 231–265.

Osborne, Martin J. and Al Slivinski (1996) "A Model of Political Competition with Citizen Candidates," *Quarterly Journal of Economics* 111:65-96.

Padro-i-Miguel, Gerard (2007) "The Control of Politicians in Divided Societies: The Politics of Fear," *Review of Economic Studies* 74(4): 1259-1274.

Persson, Torsten and Guido Tabellini (2000) *Political Economics: Explaining Economic Policy.* Cambridge, MA: MIT Press.

Prendergast, Canice (1993) "A Theory of 'Yes Men'," *American Economic Review,* 83 (4): 757–770.

Rodrik, Dani (1996) "Understanding Economic Policy Reform," *Journal of Economic Literature*, 34 (1): 9-41

Sachs, Jeffrey (1989) "Social Conflict and Populist Policies in Latin America," NBER Working Papers 2897.

Schoultz, Lars (1983). *The Populist Challenge: Argentine Electoral Behavior in the Postwar Era.* Chapel Hill: University of North Carolina Press.

Wittman, Donald (1973) "Parties as Utility Maximizers," *American Political Science Review* 67, 490-498.

Appendix: Proofs

The following Lemma is useful to understand the role and consequences of Assumption 2.

Lemma 1 *Suppose that Assumption 2 holds. Then:*

1. $f(0) < \frac{2}{r}$;

2. $\Pr\left(|z| > \frac{r}{5}\right) > \frac{1}{5}$.

Proof of Lemma 1. **Part 1.** Assumption 2 implies that $|f'(x)| < \frac{2}{r^2}$ if $|x| < r$. Hence, whenever $|x| < r$, we have $|f(x) - f(0)| \leq \frac{4}{r^2}|x|$. This implies $f(x) \geq f(0) - \frac{4}{r^2}|x|$. We then have

$$\begin{aligned} 1 &= \int_{-\infty}^{+\infty} f(x)\,dx > \int_{-r/2}^{+r/2} f(x)\,dx \\ &\geq \int_{-r/2}^{+r/2} \left(f(0) - \frac{2}{r^2}|x|\right) dx = rf(0) - 2\int_{-0}^{+r/2} \frac{2}{r^2} x dx \\ &= rf(0) - \frac{4}{r^2}\frac{1}{2}\left(\frac{r}{2}\right)^2 = rf(0) - \frac{1}{3}. \end{aligned}$$

This immediately implies $rf(0) < 2$, and thus $f(0) < \frac{3}{2r} < \frac{2}{r}$.

Part 2. We have

$$\Pr\left(|z| < \frac{r}{4}\right) = \int_{-r/4}^{+r/4} f(x)\,dx < \int_{-r/4}^{+r/4} f(0)\,dx = \frac{r}{2}\frac{3}{2r} = \frac{3}{4}.$$

Hence, $\Pr\left(|z| > \frac{r}{4}\right) > \frac{1}{4}$. ■

Proof of Proposition 1.

Part 1. Consider three cases: $a < b$, $a > b$, and $a = b$. Median voter's posterior that the politician is poor is (by Bayes rule)

$$\hat{\mu} = \frac{\mu f(s-a)}{\mu f(s-a) + (1-\mu) f(s-b)}.$$

If $a = b$, then $\hat{\mu} = \mu$, and thus, by assumption, the incumbent is necessarily reelected. The poor and rich incumbents then choose their bliss points, so $a = 0 < r = b$, which contradicts $a = b$.

If $a \neq b$, then $\hat{\mu} \geq \mu$ if and only if

$$f(s-a) \geq f(s-b),$$

which simplifies to $s \leq (a+b)/2$ if $a < b$ and to $s \geq (a+b)/2$ if $a > b$. But if $a > b$, the likelihood of reelection is given by

$$\pi(x) = \Pr(x + z \geq (a+b)/2) = 1 - F((a+b)/2 - x) = F(x - (a+b)/2)$$

(the last equality follows from the symmetry of distribution). Hence, poor and rich incumbents solve problems

$$\max_x -\alpha x^2 + W\pi(x) - (1-\mu)\alpha r^2 (1-\pi(x)),$$
$$\max_x -\alpha (x-r)^2 + W\pi(x) - \mu\alpha r^2 (1-\pi(x)),$$

respectively.

We have

$$\begin{aligned} -\alpha a^2 + WF\left(\frac{a-b}{2}\right) - (1-\mu)\alpha r^2 F\left(\frac{b-a}{2}\right) &\geq -\alpha b^2 + WF\left(\frac{b-a}{2}\right) - (1-\mu)\alpha r^2 F\left(\frac{a-b}{2}\right), \\ -\alpha (b-r)^2 + WF\left(\frac{b-a}{2}\right) - \mu\alpha r^2 F\left(\frac{a-b}{2}\right) &\geq -\alpha (a-r)^2 + WF\left(\frac{a-b}{2}\right) - \mu\alpha r^2 F\left(\frac{b-a}{2}\right). \end{aligned}$$

Adding these inequalities, dividing by α and simplifying, we get

$$-a^2 - (b-r)^2 \geq -b^2 - (a-r)^2 + r^2 \left(2F\left(\frac{a-b}{2}\right) - 1\right)(2\mu - 1)$$

Further simplification yields

$$a - b \leq \frac{1}{2} r \left(2F\left(\frac{a-b}{2}\right) - 1\right)(1 - 2\mu).$$

If $a > b$, then this condition implies $1 - 2\mu > 0$. However, then we have

$$\frac{a-b}{2} \leq \frac{r}{2}\left(F\left(\frac{a-b}{2}\right) - \frac{1}{2}\right).$$

Due to concavity of F for positive arguments (see Assumption 2), this is only possible if

$$f(0) \geq \frac{2}{r}.$$

However, in that case Assumption 2 is violated.

Part 2. By part 1, the only remaining possibility is $a < b$. The statement for this case is proved in the text. ■

Proof of Proposition 2. The second derivatives for (9) and (11) are

$$-2\alpha + \left(W + (1-\mu)\alpha r^2\right) f'\left(\frac{a+b}{2} - x\right) < 0$$

and

$$-2\alpha + \left(W + \mu\alpha r^2\right) f'\left(\frac{a+b}{2} - x\right) < 0,$$

respectively. Assumption 2 ensures that both are satisfied. Once this is true, we find that (13) implies an increasing a as a function of b whenever $a < b$, whereas (14) implies a decreasing b as a function of a whenever $a < b$. Consequently, if an equilibrium exists, it is unique, since these "best-response" curves may intersect only once. The existence result trivially follows from (15): clearly, the left-hand side and the right-hand side have an intersection at some positive p. ■

Proof of Proposition 3. Part 1. Let us rewrite (15) as

$$G(p, W, \alpha, \mu) = 0, \tag{A1}$$

where

$$G(p, W, \alpha, \mu) = \left(W + (1-\mu)\alpha r^2\right) f\left(\frac{p}{2}\alpha r^2 \frac{1-2\mu}{W + (1-\mu)\alpha r^2} + \frac{r}{2}\right) - 2\alpha p.$$

We denote

$$y = \frac{p}{2}\alpha r^2 \frac{1-2\mu}{W + (1-\mu)\alpha r^2} + \frac{r}{2} > 0$$

($y > 0$ because it equals $\frac{b-a}{2}$, which is positive by Proposition 2). Notice that, since $f(x) < \frac{2}{r}$ for any x, then

$$p < \left(W + (1-\mu)\alpha r^2\right)\frac{1}{r\alpha},$$

and thus

$$\begin{aligned} y &= \frac{p}{2}\alpha r^2 \frac{1-2\mu}{W + (1-\mu)\alpha r^2} + \frac{r}{2} < \frac{1}{2}\left(W + (1-\mu)\alpha r^2\right)\frac{1}{r\alpha}\alpha r^2 \frac{1}{W + (1-\mu)\alpha r^2} + \frac{r}{2} \\ &= \frac{r}{2} + \frac{r}{2} = r. \end{aligned}$$

Consequently, by Assumption 2,

$$\left|f'(y)\right| < \frac{1}{\frac{r^2}{2} + \frac{W}{2\alpha}}. \tag{A2}$$

In addition, (A1) implies

$$p = \frac{W + (1-\mu)\alpha r^2}{2\alpha} f(y). \tag{A3}$$

We now differentiate G with respect to p and W. We have

$$\begin{aligned} \frac{\partial G}{\partial p} &= \alpha r^2 \frac{1-2\mu}{2} f'(y) - 2\alpha \\ &< \alpha r^2 \frac{1}{2}\frac{1}{\frac{r^2}{2} + \frac{W}{2\alpha}} - 2\alpha = \frac{\alpha}{1 + \frac{W}{r^2}} - 2\alpha < 0. \end{aligned}$$

Differentiating by W yields

$$\begin{aligned} \frac{\partial G}{\partial W} &= f(y) - \frac{p}{2}\alpha r^2 \frac{1-2\mu}{\left(W + (1-\mu)\alpha r^2\right)} f'(y) > f(y) - \frac{p}{2}\alpha r^2 \frac{1}{\left(W + (1-\mu)\alpha r^2\right)}\frac{1}{\frac{r^2}{2} + \frac{W}{2\alpha}} \\ &= f(y) - \frac{1}{4} f(y) r^2 \frac{1}{\frac{r^2}{2}} > f(y) - \frac{1}{2} f(y) > 0; \end{aligned}$$

here, we used (A2) and (A3). This implies that

$$\frac{\partial p}{\partial W} = -\frac{\partial G/\partial W}{\partial G/\partial p} > 0,$$

and thus p is increasing in W.

Part 2. Let us differentiate G with respect to μ. We have

$$\begin{aligned}\frac{\partial G}{\partial \mu} &= -\alpha r^2 f(y) - \frac{p}{2}\alpha r^2 \frac{\alpha r^2 + 2W}{W + (1-\mu)\alpha r^2} f'(y) \\ &< -\alpha r^2 f(y) + \frac{p}{2}\alpha r^2 \frac{\alpha r^2 + 2W}{W + (1-\mu)\alpha r^2} \frac{1}{\frac{r^2}{2} + \frac{W}{2\alpha}} \\ &= \alpha r^2 f(y)\left(-1 + \frac{1}{4\alpha}\frac{\alpha r^2 + 2W}{\frac{r^2}{2} + \frac{W}{2\alpha}}\right) = \alpha r^2 f(y)\left(-1 + \frac{\alpha r^2 + 2W}{2\alpha r^2 + 2W}\right) < 0\end{aligned}$$

(we again used (A2) and (A3)). This implies that

$$\frac{\partial p}{\partial \mu} = -\frac{\partial G/\partial \mu}{\partial G/\partial p} < 0,$$

and hence p is decreasing in μ.

Part 3. Observe that (A1) may be rewritten as

$$\left(\frac{W}{\alpha} + (1-\mu) r^2\right) f\left(\frac{p}{2} r^2 \frac{1-2\mu}{\frac{W}{\alpha} + (1-\mu) r^2} + \frac{r}{2}\right) - 2p = 0. \tag{A4}$$

Consequently, if $W = 0$, then equilibrium populist bias p does not depend on α. If $W > 0$, then, since (A4) only depends on W and α through $\frac{W}{\alpha}$, an increase in α has the same effect as a decrease in W. In other words, p is decreasing in α. This completes the proof. ■

Proof of Proposition 4. Part 1. A normal distribution is characterized by density

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}.$$

One can easily check that $\left|\frac{\partial f}{\partial x}\right|$ reaches its maximum at $x = \pm\sigma$, and this maximum equals $\frac{1}{\sqrt{2\pi e \sigma^2}}$. For Assumption 2 to hold, it would be sufficient to require that

$$\sigma^2 > \frac{\frac{r^2}{2} + \frac{W}{2\alpha}}{\sqrt{2\pi e}}, \tag{A5}$$

which is obviously satisfied if $W = 0$ and $r < 2\sigma$.

Now, differentiate G with respect to r. We have

$$\frac{\partial G}{\partial r} = 2(1-\mu)\alpha r f(y) + \frac{W^2 + 2\alpha r (p + r - 2p\mu - r\mu) W + \alpha^2 r^4 (1-\mu)^2}{2(W + (1-\mu)\alpha r^2)} f'(y).$$

If $W = 0$, we have

$$\begin{aligned}\frac{\partial G}{\partial r} &= 2(1-\mu)\alpha r f(y) + \frac{\alpha r^2 (1-\mu)}{2} f'(y) \\ &= \frac{(1-\mu)\alpha r f(y)}{2}\left(4 + r\frac{f'(y)}{f(y)}\right).\end{aligned}$$

For a normal distribution,

$$\frac{f'(y)}{f(y)} = -\frac{y}{\sigma^2},$$

and thus, using (A5), we get that if $W = 0$,

$$\frac{\partial G}{\partial r} = \frac{(1-\mu)\alpha r f(y)}{2}\left(4 - r\frac{y}{\sigma^2}\right) > 0$$

whenever $r^2 < 4\sigma^2$ (because $0 < y < r$). Therefore, for r not too large,

$$\frac{\partial p}{\partial r} = -\frac{\partial G/\partial r}{\partial G/\partial p} > 0$$

(it is worth noting that for this result, we would not need the assumption about normal distribution).

Part 2. As W varies, so does p, and in such a way that $\frac{p}{W+(1-\mu)\alpha r^2}$ remains constant (this follows from (A1)). Consequently, as W increases, y remains constant. This implies, first, that $\frac{\partial G}{\partial p}$ remains constant, and the full derivative of $\frac{\partial G}{\partial r}$ with respect to W equals (we substitute $p = \frac{y-\frac{r}{2}}{\frac{\alpha r^2}{2}(1-2\mu)}\left(W + (1-\mu)\alpha r^2\right)$ before differentiating):

$$\frac{d\frac{\partial G}{\partial r}}{dW} = -\frac{f'(y)}{2r}(r - 4y).$$

Now, since $f'(y) < 0$, then to show that $\frac{\partial G}{\partial r}$ is decreasing in W it suffices to show that $y > \frac{r}{4}$. To do that, we prove that solution to equation

$$y = f(y) r^2 \frac{1-2\mu}{4} + \frac{r}{2},$$

which we obtain from (A1), satisfies $y > \frac{r}{4}$. We can rewrite the equation as

$$\frac{y}{r} = \frac{1}{\sqrt{2\pi}}\frac{r}{\sigma}e^{-\frac{1}{2}\frac{y^2}{r^2}\frac{r^2}{\sigma^2}}\frac{1-2\mu}{4} + \frac{1}{2}.$$

Denote $\frac{y}{r} = z$, $\frac{r}{\sigma} = k$; then this equation may be rewritten as $g(z) = \frac{1}{2}$, where

$$g(z) = z - \frac{1}{\sqrt{2\pi}} k e^{-\frac{1}{2}k^2 z^2}\frac{1-2\mu}{4};$$

obviously,

$$g(z) < z + \frac{1}{4\sqrt{2\pi}} k e^{-\frac{1}{2}k^2 z^2}.$$

Suppose, to obtain a contradiction, that $z < \frac{1}{4}$. Then on the interval $0 < k < \sqrt[4]{8\pi e}$ (which captures all values of $k = \frac{r}{\sigma}$ that satisfy (A5), function $ke^{-\frac{1}{2}k^2 z^2}$ is monotonically increasing, and thus does not exceed its value at $k = \sqrt[4]{8\pi e}$. Therefore,

$$g(z) < z + \frac{\sqrt[4]{8\pi e}}{4\sqrt{2\pi}} e^{-\sqrt{2\pi e} z^2}.$$

One can easily check that the right-hand side is monotonically increasing in z on $\left(0, \frac{1}{4}\right)$, and therefore does not exceed

$$\frac{1}{4} + \frac{\sqrt[4]{8\pi e}}{4\sqrt{2\pi}} e^{-\frac{\sqrt{2\pi e}}{16}} < \frac{1}{2}.$$

This leads us to a contradiction, meaning that for no $z < \frac{1}{4}$ it is possible that $g(z) = \frac{1}{2}$. This completes the proof. ■

Proof of Proposition 5. **Part 1.** If $W = 0$, an elite politician would never choose $x_1 < 0$. Indeed, in this case he would get at most $-(r - x_1)^2 < -r^2$, even if his ideal policy is implemented in the second period. At the same time, he can always guarantee getting $-r^2$ by choosing $x_1 = r$ (in that case, x_2 will be either 0 or r, depending on the election result and new politician's type). Consequently, $x_1 = r$ would be a profitable deviation, and thus $x_1 < 0$ may not be the case in an equilibrium if $W = 0$.

Part 2. Combining (13) and (14) yields

$$a = (b - r) \frac{W + (1 - \mu) \alpha r^2}{W + \mu \alpha r^2};$$

plugging this into (14) yields

$$\left(W + \mu \alpha r^2\right) f\left(\frac{r - b}{2} \alpha r^2 \frac{1 - 2\mu}{W + \mu \alpha r^2} + \frac{r}{2}\right) - 2\alpha (r - b) = 0.$$

The remainder of the proof is completely analogous with the proof of Part 1 of Proposition 3: like there, we show that $r - b$ is increasing in W, and thus b is decreasing in W.

Part 3. If $W > 0$ is sufficiently large, $x_1 < 0$ is possible for both politicians. Indeed, x_1^r and x_1^p are linked by

$$x_1^r = x_1^p \frac{W + \mu \alpha r^2}{W + (1 - \mu) \alpha r^2} + r.$$

If W is sufficiently large, equilibrium populism p can be made arbitrarily large (otherwise in (15), the argument of f would be bounded from below, and thus the right-hand side could be arbitrarily large, which is incompatible with the left-hand side being bounded). But then x_1^p may be arbitrarily large (in absolute value) negative number, and thus x_1^r will also become negative, since $\frac{W + \mu \alpha r^2}{W + (1 - \mu) \alpha r^2}$ has a limit as $W \to \infty$. This completes the proof. ■

Proof of Proposition 6.

The proof of the Proposition is completely analagous to the proof of Proposition 1 and is omitted. ■

Proof of Proposition 7. Part 1. Notice that under Assumption 2, problems (18) and (20) are convex. Indeed, take (18); the second derivative w.r.t. x equals

$$-2\alpha + \left(W + (1-\mu)\,\alpha\left(\frac{r}{1+\alpha}\right)^2\right) f'\left(\frac{a+b}{2} - x\right) < 0,$$

since $1+\alpha > 1$ and thus $W + (1-\mu)\,\alpha\left(\frac{r}{1+\alpha}\right)^2 < W + (1-\mu)\,\alpha r^2$. For problem (20), the second derivative w.r.t. x equals, after simplifications,

$$-2\alpha - 2 + \left(W + \left(\mu + (1-\mu)\left(\chi + \frac{\alpha}{\alpha+1}\right)\right)\frac{r^2}{\alpha+1}\right) f'\left(\frac{a+b}{2} - x\right).$$

For this to be negative, it suffices that

$$\begin{aligned} f'\left(\frac{a+b}{2} - x\right) &< \frac{2(\alpha+1)}{\left(W + \left(\mu + (1-\mu)\left(\chi + \frac{\alpha}{\alpha+1}\right)\right)\frac{r^2}{\alpha+1}\right)} \\ &= \frac{2\alpha}{W\frac{\alpha}{\alpha+1} + \left(\mu + (1-\mu)\left(\chi + \frac{\alpha}{\alpha+1}\right)\right)\frac{r^2}{\alpha+1}\frac{\alpha}{\alpha+1}}. \end{aligned} \tag{A6}$$

Simple algebra yields

$$W\frac{\alpha}{\alpha+1} + \left(\mu + (1-\mu)\left(\chi + \frac{\alpha}{\alpha+1}\right)\right)\frac{r^2}{\alpha+1}\frac{\alpha}{\alpha+1} < W + \alpha r^2,$$

and thus (A6) is negative, provided that 2 is satisfied. As in the Proposition 2 we get that equilibrium is determined by the intersection of a as an increasing function of b and of b as a decreasing function of a. This ensures that equilibrium is unique if it exists, and existence is proved similarly to Proposition 2. Finally, as the first-period policy of an honest politician is given by (19), we immediately obtain $a < 0$.

Part 2. It is straightforward to see that if there is corruption in the second period, then there is one in the first (without bribes, dishonest politicians would choose $b < 0$, which gives the elite even more incentives to bribe them. Given that, we can apply the reasoning we used in the proof of Proposition 5. If $W = \chi = 0$, then the dishonest incumbent has not reelection motives except for influence the policy choice. Consequently, the incumbent and the elite will jointly choose policy $x_1 > 0$, for otherwise they would be better off choosing any policy between 0 and 1 in the first period, and then playing equilibrium strategies in the second. If, however, W is high enough, we can again show that neither a nor b are bounded from below, and thus $b < 0$ is possible. ■

Proof of Proposition 8. **Part 1.** Let us rewrite (24) as

$$H(p,W,C,\mu,\chi)=0, \tag{A7}$$

where

$$H(p,W,C,\mu,\chi)=\left(W+\frac{(1-\mu)\alpha r^2}{(1+\alpha)^2}\right)f\left(p\frac{W-\frac{(1-\mu)\chi\alpha r^2}{1+\alpha}+\frac{r^2\alpha(1-2\mu-\alpha\mu)}{(1+\alpha)^2}+C}{2(1+\alpha)\left(W+\frac{(1-\mu)\alpha r^2}{(1+\alpha)^2}\right)}+\frac{r}{2(1+\alpha)}\right)-2\alpha p.$$

We denote

$$y=p\frac{W-\frac{(1-\mu)\chi\alpha r^2}{1+\alpha}+\frac{r^2\alpha(1-2\mu-\alpha\mu)}{(\alpha+1)^2}+C}{2(\alpha+1)\left(W+\frac{(1-\mu)\alpha r^2}{(1+\alpha)^2}\right)}+\frac{r}{2(1+\alpha)}>0$$

($y>0$ because it equals $\frac{b-a}{2}$, which is positive by Proposition 7). By Assumption 2,

$$\left|f'(y)\right|<\frac{1}{\frac{r^2}{2}+\frac{W}{2\alpha}}. \tag{A8}$$

In addition, (A7) implies

$$p=\frac{W+\frac{(1-\mu)\alpha r^2}{(1+\alpha)^2}}{2\alpha}f(y). \tag{A9}$$

We now differentiate H with respect to p and W. We have

$$\begin{aligned}\frac{\partial H}{\partial p} &= \frac{W-\frac{(1-\mu)\chi\alpha r^2}{1+\alpha}+\frac{r^2\alpha(1-2\mu-\alpha\mu)}{(1+\alpha)^2}+C}{2(\alpha+1)}f'(y)-2\alpha\\ &< \frac{-\frac{(1-\mu)\alpha r^2}{1+\alpha}+\frac{r^2\alpha(1-2\mu-\alpha\mu)}{(1+\alpha)^2}}{2(1+\alpha)}f'(y)-2\alpha=-\frac{r^2\alpha}{2}\frac{\alpha+\mu}{(1+\alpha)^3}f'(y)\\ &< \frac{r^2\alpha}{2}\frac{1}{(1+\alpha)^2}\frac{1}{\frac{r^2}{2}+\frac{W}{2\alpha}}-2\alpha\le\frac{r^2\alpha}{2}\frac{1}{(1+\alpha)^2}\frac{2}{r^2}-2\alpha\\ &\le \alpha-2\alpha<0\end{aligned}$$

(since $f'(y)<0$, $\frac{\partial H}{\partial p}$ would achieve its maximum at $W=0$, $\chi=1$). Differentiating with respect to W yields

$$\begin{aligned}\frac{\partial H}{\partial W} &= f(y)+\frac{p\alpha r^2(\mu+(1-\mu)\chi)}{2(1+\alpha)^2\left(W+\frac{(1-\mu)\alpha r^2}{(1+\alpha)^2}\right)}f'(y)\\ &> f(y)-\frac{r^2(\mu+(1-\mu)\chi)}{4(1+\alpha)^2}\frac{1}{\frac{r^2}{2}+\frac{W}{2\alpha}}f(y)>f(y)-\frac{r^2}{4}\frac{2}{r^2}f(y)>0.\end{aligned}$$

Hence,

$$\frac{\partial p}{\partial W}=-\frac{\partial G/\partial W}{\partial G/\partial p}>0,$$

and thus p is increasing in W.

Part 2. Differentiate H with respect to C. Clearly,

$$\frac{\partial H}{\partial C} < 0,$$

and therefore

$$\frac{\partial p}{\partial C} < 0,$$

so p is decreasing in C.

Part 3. Differentiate H with respect to χ. We have

$$\frac{\partial H}{\partial \chi} = -p\frac{(1-\mu)\,\alpha r^2}{2\,(1+\alpha)^2} f'(y) > 0.$$

Therefore,

$$\frac{\partial p}{\partial \chi} = -\frac{\partial G/\partial \chi}{\partial G/\partial p} > 0,$$

and thus p is increasing in χ.

Part 4. Differentiate H with respect to μ. We have

$$\begin{aligned}
\frac{\partial H}{\partial \mu} &= -\frac{\alpha r^2}{(1+\alpha)^2} f(y) - \frac{p}{2}\frac{\alpha r^2}{(1+\alpha)^4}\frac{W\,(1+\alpha)^2\,(1-\chi) + r^2\alpha}{W + \frac{(1-\mu)\alpha r^2}{(1+\alpha)^2}} f'(y) \\
&= -\frac{\alpha r^2}{(1+\alpha)^2} f(y) - \frac{f(y)}{4\alpha}\frac{\alpha r^2}{(1+\alpha)^4}\left(W\,(1+\alpha)^2\,(1-\chi) + r^2\alpha\right) f'(y) \\
&= \frac{\alpha r^2}{(1+\alpha)^2} f(y)\left(-1 - \frac{1}{4\alpha}\frac{1}{(1+\alpha)^2}\left(W\,(1+\alpha)^2\,(1-\chi) + r^2\alpha\right) f'(y)\right) \\
&< \frac{\alpha r^2}{(1+\alpha)^2} f(y)\left(-1 + \left(W\frac{1-\chi}{4\alpha} + \frac{r^2}{4\,(1+\alpha)^2}\right)\frac{1}{\frac{r^2}{2} + \frac{W}{2\alpha}}\right) \\
&< \frac{\alpha r^2}{(1+\alpha)^2} f(y)\left(-1 + \left(W\frac{1}{2\alpha} + \frac{r^2}{2}\right)\frac{1}{\frac{r^2}{2} + \frac{W}{2\alpha}}\right) = 0.
\end{aligned}$$

Consequently,

$$\frac{\partial p}{\partial \mu} = -\frac{\partial G/\partial \mu}{\partial G/\partial p} < 0,$$

and thus p is decreasing in μ. This completes the proof. ■

Proof of Proposition 9. If W is high enough, then for $C = 0$, both a and b can be arbitrarily low in the first period if. In that case, the utility of the elite may be arbitrarily low, whereas for C sufficiently high, there is no corruption, and the utility of the elite is given by $-2r^2$, as all politicians will choose $x_1 = x_2 = 0$. Consequently, there exists $\bar{W}$ such that for $W > \bar{W}$, the elite is better off under $C = 0$. This completes the proof. ■