A Theory of Political Entrenchment[*]

Gilles Saint-Paul
TSE

Davide Ticchi
University of Urbino

Andrea Vindigni
Princeton

April 2011

Abstract

We develop a theory of endogenous political entrenchment in a simple two-parties dynamic model of income redistribution with probabilistic voting. A partially self-interested left-wing political party may deliberately choose to implement policies that reduce the income of *its own* constituency, the lower class (which includes the voter with median income), in order to consolidate its future power. This is because such policies increase the net gain that low-skill agents obtain from the fiscal redistribution of income, which only the Left (but not the Right) can credibly commit to provide. Hence, they can help offsetting a potential future aggregate ideological shock averse to the left-wing party. We show that political entrenchment by the Left occurs only if incumbency rents are sufficiently high (for example due to weak checks and balances) and that voters may vote for it even though they rationally expect entrenchment. Constitutionally legislated term limits also reduce the scope for entrenchment, and potentially increase social welfare, but only if party discipline is relatively weak. The theory sheds light on some puzzles such as why left-wing parties or politicians often support immigration of unskilled workers, are sometime in favor of free trade with less developed economies and of globalization more generally, or fail to reform plainly "dysfunctional" public educational systems damaging the lower classes.

Preliminary and incomplete. Please do not distribute.

1 Introduction

There is now a large literature in political economy explaining why governments in a variety of institutional environments relatively often pursue policies reducing the welfare of the society.[1] A common feature of some of these models is that socially efficient policies involve a potential reallocation of political power from one social group, in power in the *status quo*, to another. Therefore, in absence of commitment, partisan politicians, i.e. representing a certain social constituency, may prefer to implement policies harmful to the society at large in order to prevent an adversary group from gaining power. This type of explanation of why socially harmful policies exist and are relatively pervasive is quite compelling and has a wide scope. Nevertheless, it does not explain (the more intriguing puzzle of) why occasionally incumbent politicians appear to pursue policies that are harmful not only for the society as a whole, but also—and especially—for their *own* constituency. Examples of policies with this apparently paradoxical feature (discussed in greater detail in the following Section) include the dysfunctional educational policies implemented by a number of Latin American "populist" governments, and the liberal immigration policies supported by left-wing parties in Europe. Also potentially puzzling are the cases of policies such as the support of NAFTA by the Clinton Administration in the early 1990's, or the reluctance of many left-wing Latin American governments in the late 1990's and in the 2000's to abandon the pro-globalization policies implemented by their right-wing predecessors. These policies are not inefficient but still damage the economic interests of a relatively significant part of the constituency of the governments implementing them.

Why, and when, once in office politicians implement policies that are not beneficial for the very same people who brought them in power? In this paper, we propose a simple game-theoretical model aimed at addressing this puzzling question. Ours is a relatively standard dynamic (two-periods) model of redistribution, where citizens have both preferences over an economic issue, which is affected by policy, and over an exogenous noneconomic or ideological issue, as in the spirit of probabilistic voting models à la Lindbeck and Weibull (1987) or Dixit and Londregan (1995 and 1998). Citizens have a specific taste for which party is in power, which is heterogeneously distributed in the population. Furthermore, this distribution is subject to aggregate shocks, which generates electoral uncertainty.

---

[1] Example of such policies include the blocking of technological progress (e.g. Krusell and Rios-Rull, 1996), failing to make cost-effective investments in human capital (e.g. Besley and Coate, 1998), subsidizing declining industries (Dixit and Londregan, 1995 and 1998; Brainard and Verdier, 1997; Coate and Morris, 1999; Acemoglu and Robinson, 2001). Other examples include investing in inefficient state institutions with limited capacity to tax or coerce citizens (Acemoglu, Ticchi and Vindigni, 2010a and 2011), protecting unproductive jobs with high firing costs (Saint-Paul, 1993 and 2003), creating inefficient social infrastructures (Robinson and Torvik, 2005), or underpricing the shares of privatized companies (Biais and Perotti, 2002).

For simplicity, we assume that there are only two income classes, and two political parties. Each party cares about the welfare of one of the two social groups as in the spirit of partisan models of political economy (e.g. Alesina, 1988; Besley and Coate, 1997; Osborne and Slivinsky, 1997), but also derives rents from being in power, as in the spirit of the classic Downsian model of political competition. Finally, as in models of partisan politics, parties cannot *ex ante* commit to implement a policy different from their preferred one once in office.

The main innovation of our model is the assumption that the second period pre-tax income of the lower class (or "the poor" or unskilled) can be *reduced* by the government in office in the first period. This policy has two main features: it is straightforwardly Pareto-inefficient and it *increases* income inequality by reducing the income of the poor in absolute terms and relative to the income of the upper class (or "rich" or skilled). Nevertheless, we prove that under some conditions a *leftist* party, rationally elected in office by its natural constituency, may find optimal implementing precisely such a policy. This policy is rational for the party since it allows it to *entrench* itself more, that is to tie its own constituency more to itself, and therefore to remain in power in the future in spite of a potential ideological shock in the favor of the right. The source of such political entrenchment comes from the fact that the policy described makes the poor worse-off in both absolute and relative terms relative to the rich (and reduces the median income relative to the mean income). Therefore, it increases the value for the poor of the redistributive fiscal policy which only the Left will actually implement once in office.

In other words, by pursuing such policies, the Left makes sure that future income inequality will be larger, which enhances its reelection prospects. An essential feature is that pre-tax income inequality is a state variable while redistributive policies are not. As a result, parties cannot commit on which redistributive policies will be undertaken once they are in office, and only the Left, given its preferences, will be inclined to redistribute. At the same time, it can make sure that its voters will be more "dependent" on redistribution by pursuing policies that will deliver more inequality at the time of the election.

A few features of our model are worth emphasizing. First, voting for the Left in the first period can well be rational under some conditions for the lower class, even though the policy of entrenchment (which becomes effective in the second period) is correctly anticipated. This is because in the first period the Left, unlike the Right, provides some redistribution of income which might compensate the poor for their *future* income loss generated by the entrenchment policy. Since the left-wing party is the only party that can provide income redistribution, the lower class has no other choice but voting for the Left, despite that it is harmed by the

entrenchment policies. Therefore, the existence of a equilibrium with potential entrenchment in our model does not rely on any form of myopia or irrationality on either side of the political market.

Second, our argument generalizes naturally to any policy that increases income inequality, regardless of the sign of its impact on each group. We focus on the case where the poor's pre-tax income is reduced and the rich's is unchanged because we want our model to be able to shed light on why governments may pursue inefficient policies that harm their own constituencies. But in some examples the policy may be efficient but inflict losses on the poor and generate gains to the rich, or it could harm the poor more than the rich, or benefit the rich more than the poor. Furthermore, while in our model entrenchment is associated with pro-active policies that reduce the relative position of the Left's constituency, it may also result in failure to implement policies that improve this relative position. In our redistribution model, the reduction in the relative status of the left's constituency is the distinctive feature of political entrenchment; often, but not systematically, this may put the economy off its efficiency frontier and/or harm the poor in absolute terms.

Third, even though we present a model where only the Left may decide to entrench itself, we are by no means arguing that entrenchment is a policy exclusive of Left. It is easy to imagine a different framework, where the purpose of the government is not to provide pure income redistribution, but some public good giving a relatively high direct utility to the natural constituency of the right-wing party (e.g., national defence), where some form of entrenchment is pursued by the Right (e.g., an unnecessarily aggressive foreign policy).

Our model delivers some interesting comparative statics predictions and has also some implications for the design of political institutions. For example, we show that political entrenchment is more likely to occur in equilibrium when the rents potentially appropriated by the party in power are higher. This suggests that we should expect more entrenchment in political systems with relatively limited checks and balances. This is the case in many Latin American countries whose institutions feature a form of presidential government with limited separation of powers with the legislature and the judiciary (see for example Stokes, 2009). Moreover, we show that constitutionally prescribed term limits may reduce the political entrenchment because they reduce the value of capturing power in the future for the incumbent leader. This result is important not only in its own right, but also because it sheds some light on the puzzle of why many real world constitutions do actually prescribe some form of term limits, whereas a key prediction of the standard model of political replacement à la Barro (1973) and Ferejohn (1986) and of its extension by Persson, Roland and Tabellini (1997) is

that term limits are always welfare reducing.[2]

Other related literatures includes the strategic theory of budget deficits (Persson and Svensson, 1989; Alesina and Tabellini, 1990; Bolton and Roland, 1990; Tabellini and Alesina, 1990). In these models, public debt is used strategically by an incumbent partisan leader in order to manipulate the future spending policy or the future allocation of political power. Also related are dynamic models of income redistribution such as Saint-Paul and Verdier (1997), Bénabou (2000, 2002) and Hassler, Rodrìguez Mora, Storesletten and Zilibotti (2003) where, unlike in the classic static models of Romer (1975) and Meltzer and Richards (1981) income inequality is an endogenous state variable and voting over fiscal policy is forward-looking. Finally, the paper is closely related to the recent work on the persistence of power and institutions (e.g. Acemoglu and Robinson, 2008; Acemoglu, Ticchi and Vindigni, 2010b and 2011); yet, the source of political persistence in our paper is rather different from any of these works. In particular, our notion of political entrenchment has some similarity with the concept of *managerial* entrenchment in the theory of corporate finance (e.g. Weisbach, 1988; Shleifer and Vishny, 1989). Managers may entrench themselves by making manager-specific investments, which reduce the probability of being replaced and allow them to earn higher wages. Similarly, political entrenchment makes it more costly for the lower class to vote for the Right, and allows incumbent leftists politicians or parties to appropriate of higher intertemporal rents.

Finally, to clarify better the originality of our notion of political entrenchment it is useful to compare our paper with the paper of Besley and Coate (1998) and Padró-i-Miguel (2007). Within the large literature on inefficient policies (briefly cited in Footnote 1), our paper is most closely related to the paper of Besley and Coate (1998). In that paper, a leader representing the perfect agent of a specific social constituency, the low-skill workers with low-ability, may decide not to implement a cost-efficient policy like investing in education, in order to preserve the power of its own constituency. This would become a minority in the future if the low-skill workers with high-ability, whose skills and income are potentially increased by investment in human capital, were educated. Hence, the political cost of public education in Besley and Coate (1998) is the future emergence of an anti-redistribution constituency. In our paper instead, the Left makes an inefficient decision (that can be assimilated to not upgrade the skills of the lower class) out of its pure self-interest, and *against* the preferences of the lower class. This result highlights the importance of our assumption that political agency is not perfect (for

---

[2]This is because in such models, repeated elections are sufficient to provide incentives to politicians in power to refrain from appropriating too many rents. Term limits represent a harmful self-imposed constraint since they force voters to replace politicians even if the latter have behaved well in office. As a result, elections prove to be less valuable as a discipline device.

example due to the existence of a system of checks and balances providing weak constraints on the executive).

Padró-i-Miguel (2007) shows that the leaders of competing ethnic groups, once in power, exploit not only the adversary groups but also in part their *own* group, which is willing to comply with some degree of group-specific taxation financing the consumption of its leader. This is because of the fear that the replacement of the incumbent leader may bring to power the leader of the competing group, who would exploit them even more. However, in Padró-i-Miguel's paper the leader in office does not deliberately reduce the productivity of its own constituency to consolidate its own future power by manipulating a state variable of the dynamic game, which is the key novel idea of our paper.

The rest of the paper is organized as follows. Section 2 presents some evidence for our story. Section 3 describes the framework of the model. Section 4 characterizes the equilibrium. Section 5 presents the comparative statics analysis and Section 6 discusses the role of term limits in our model. Section 7 concludes while the Appendix contains some more details on the comparative statics analysis.

2   Some Evidence

This Section illustrates briefly some examples of policies that are harmful to the constituency of the government implementing them.

1. *Educational policies in Latin America.* In a recent book, Edwards (2010) argues that dysfunctional educational policies characteristic of many Latin American countries have a premier role in explaining both the persistence of underdevelopment and of income inequality in that continent. Edwards (p. 179) writes that, "But without any doubt the most important cause of Latin America's social ills—including poverty and income inequality—is the historical dreadful state of the region's educational system. By neglecting education the vast majority of the Latin American countries have failed to upgrade their labor force skills and have lagged behind other nations in the key areas of human capital formation and productivity growth. It is not an exaggeration to say that workers in many Latin American countries are among the least prepared to meet the high skill requirements of the twenty-first century."

While the quality of the educational systems in most Latin American countries has been historically very low, it is interesting to remark that, according to Edwards (p. 181), that the efforts occasionally made to reform and modernize the educational system have been strongly, and successfully, opposed by teachers' unions and left-of-center political parties, the natural

political references of the main potential beneficiary of such reforms.

Examples of failed educational reforms include the effort undertook by the respected Argentinian minister Juan Llach to improve the quality of the school system and the ability of Argentina to compete more effectively in a technologically global economy. Llach's program, while ingenious and ambitious, was not ultimately implemented due to the opposition of the powerful teacher's union, of the Peronist opposition party and especially of President Fernando de la Rúa's own party, the Union Civica Radical, an old and traditional center-left political party. Not surprisingly the quality of education in Argentina has gradually deteriorated. While there is a considerable variation across provinces, with the rich areas doing relatively well, the educational system in the poorer regions of the country lays in disarray, performing "no better than the most destitute countries of the world."

Many other educational reforms promoted by left-wing or "populist" Latin American leaders and aimed at improving the skills and human capital of the lower classes had a similar dismal outcome.[3] For example, the government of Hugo Chávez in Perú launched a variety of ambitious projects aimed, in principle, at improving education at various levels. These reforms included a literacy campaign (the Misión Robinson), a program aimed at improving quality and coverage of public education (the Misión Robinson II), a program dealing with high-school students and dropouts (the Misión Ribas), and finally a program aimed at reforming higher education (the Misión Sucre). Many observers, though, have criticized all such programs as ineffective and fraught with corruption. Indeed, despite massive investments, illiteracy barely declined during the Chávez presidency.

Other types of policies implemented by Chávez's government with the goal of reorienting state priorities to benefit the poor, have also appeared to harm rather than to improve the social conditions of segments of the constituency of the leader. For example, economists Chang-Tai Hsieh, Edward Miguel, Daniel Ortega and Francisco Rodrigúez have demonstrated using a large dataset with information on political opinion of voters that many of the original *supporters* of Chávez experienced a 4% decline in personal income after the referendum. Overall, despite the original promises, after twelve year of Chávez's presidency the economic performance of the lower classes has stagnated at best, but more likely has even *declined*.[4]

---

[3]Examples include the educational reforms attempted by Evo Morales in Bolivia, or by Daniel Ortega in Nicaragua, aimed at improving the coverage of education at promoting literacy campaigns (Edwards, 2000, p. 181).

[4]Using official statistics Francisco Rodrigúez has argued in an article in *Foreign Affairs* (March/April 2008) that, "Most health and human development indicators have shown no significant improvement beyond that which is normal in the midst of an oil boom. Indeed, some have deteriorated worryingly, and official estimates indicate that income inequality has increased. The 'Chávez is good for the poor' hypothesis is inconsistent with

All of these examples naturally raise the question of why the masses in Latin America have been so much inclined to support populist parties and leaders. Our answer is that, quite simply, they had probably little alternative to the choice of populist redistribution.

2. *Immigration policies in Western Europe.* Migration and immigrant integration policies in Europe are increasingly determined at the European Union (EU) level. EU rules now cover the full gamut of "migration policies," from entry, residence, and economic rights of immigrants to societal integration of immigrants and their descendents; in addition, the European Parliament has significant amendment and veto powers in the adoption of these policies.

On the basis of economic factors, it is reasonable to expect that, at the individual level, unskilled workers and unemployed workers are more likely to be opposed to immigration than more highly skilled workers or owners of capital, since they more likely to be in competition for jobs with immigrants than higher-skilled workers.[5] Given this premise, one may also expect left-wing parties to support *conservative* migration policies, in line with the economic interests of many of their voters, who often compete with immigrants for unskilled jobs.

In a recent empirical study, Hix and Noury (2007) address the question of which interests EU politicians promote when making migration policies, by looking at the passage of six pieces of migration related legislation in the fifth directly elected European Parliament (1999–2004). In reporting the results of their empirical analysis Hix and Noury (2007, p. 184) write that, "We find that the strongest determinants of behavior in the European Parliament on migration and related issues are the general left-right ideological positions of the European parliamentarians and the transnational European parties. These are stronger predictors of political behavior in the European Parliament on these issues than the economic preferences of the European parliamentarians' constituents, or the economic interests or political preferences of the EU member states. In other words, left-wing politicians support liberal migration policies, despite the economic interests of many of their voters, who often compete with immigrants for unskilled jobs. Meanwhile, right-wing politicians support restrictive migration policies, despite the economic interests of many of their supporters, who benefit from increasing returns on capital investment which results from greater immigration."

Given that European parliamentary elections are based on a proportional system, we expect Euro-MPs to face little personal accountability for deviating from their constituencies' preferences; electoral logic and party discipline will prevail, and in such a context rents from

the fact."

[5] Scheve and Slaughter (2001) find robust evidence that in the U.S. less-skilled workers prefer anti-immigration policies. Mayda (2006) find analogous results in a cross-section of countries.

being in office are likely to be more influential than the constituents' well-being. Our model predicts that this is when entrenchment is most likely to arise, and this is in line with the above evidence on the MEP's voting pattern on migration.[6]

3. Another example where political entrenchment may have played a role is the passing of the North American Free Trade Agreement (NAFTA) by the democratic-controlled Congress in 1992 with the support of President Clinton. While free trade has positive aggregate gains, the Stolper-Samuelson theorem implies that the low skilled workers in the US will lose, while the high-skilled workers will gain. Therefore it is surprising that the agreement was signed by a democratic administration, which supposedly represents the unskilled. Indeed, the union movement, which we may interpret as representing the welfare of the incumbent, median, low-skilled workers, and which clearly derives less rents from Democratic politicians being in office than those politicians themselves, strongly opposed NAFTA. Our model suggests that one of the possible reasons why the democratic administration went ahead with NAFTA is that, by widening the earnings gap between skilled and unskilled, it would increase future political support for the redistributive programs that are traditionally implemented by the democratic party. While this is not the only effect (a substantial fraction of democratic voters are rich and benefit from NAFTA), this effect may have contributed to the overall outcome.[7]

Some interesting features are worth noting. First, the Republicans supported the agreement more than the democrats, and they are the ones who initiated it. Indeed, if only democratic congressmen had voted, it would not have passed.[8] Does that contradict our model? Not necessarily: to the extent that the policy generates aggregate gains, it is possible for the Republicans to benefit from it too, since the direct economic gains accrue to the upper class, even though their reelection probability suffers. Also, it may well be that the President is more

---

[6] In our model, parties internalize part of the welfare of their constituents. This is because their candidates are (partly) "citizen-candidates" who belong to the same social class as their constituents. In practice, though, one observes that most politicians on the Left in fact come from the upper-middle class. Hence it is not as persons that they internalize their constituents' welfare. Rather, this component of their utility is best viewed as a metaphor for their individual strategies in a context where they have to maintain a reputation with their voters. Again, given the proportional system and the aloofness of the European Parliament to most voters, we expect such an internalization to be weaker here. While this will mean that office rents play a bigger role, it may also imply that the politicians' own personal tastes may also affect their voting behavior. This latter mechanism, though, is absent from our model.

[7] Another possibility is that people expected redistribution to transfer enough of the gains to the losers so as to make them better-off *ex ante*. Here we would have an example of entrenchment working to the benefit of the constituency, provided it "does the right thing" and votes for the left. However, this interpretation is not supported by the unions' opposition and by the congressional voting patterns.

[8] Accounts on the determinants of congressional voting on NAFTA differ. But both Kang and Greene (1999) and Kahane (1996) find that these determinants obey a conventional logic, in that congressional districts with a greater proportion of potential losers were more likely to oppose the bill.

prone to entrenchment than representatives. In our model, entrenchment is more likely to arise, the greater is "state capacity," i.e. the greater the amount of money that the government can redistribute. Since the US (as opposed to Europe) is characterized by low party discipline, each individual congressman is accountable to his or her constituency, and has little ability to implement a redistributive programme at the congressional district level. Hence "state capacity" is low for congressmen, while it is larger for the president, implying the latter may be more prone to entrenchment than the former.[9]

4. Finally, it is interesting to remark (see for example Stokes, 2009) that the many left-wing Presidents elected in office in Latin American countries during the late 1990's and the 2000's have generally refrained to reverse the pro-globalization policies implemented by their right-wing predecessors during the previous two decades. This may seem puzzling in the light of the views (e.g. Rodrik, 1997) that globalization potentially damages the natural constituency of the Left by exposing the low-skilled workers to greater volatility in wages and employment.

Yet, many Latin American Presidents supported by leftists or center-left coalitions have pursued strategies of "re-equilibration," in the attempt to compensate the globalization losers by expanding the public economy. Although, there can be other reasonable explanations, the strategy of supporting globalization and at the same time increasing redistribution is quite in line with our theory of entrenchment. The globalization of the Latin American economies was originally pursued by the Right, which implemented policies leading to a massive expansion of economic openness coupled with a contraction of social spending and a raise in inequality, the remarkable electoral success of the Left in Latin America since the end of the 1990's can be arguably regarded as an "unintended" consequence of such policies (Stokes, 2009). Therefore, from a purely electoral point of view, it can be hardly denied that globalization has not favored the Left in Latin America, and that the Left may well have, therefore, a potential vested interested in preserving it.

3   The Model

We consider an economy populated by a continuum of measure one of citizens, and by two partially selfish political parties. There are two periods of time, $t = 0$ and $t = 1$, and citizen i

---

[9]It is interesting to observe that Obama himself initially took a negative stance over NAFTA, linking it to lost jobs in the United States more than once during the electoral campaign of 2007-2008. Yet once in office, Obama appeared to change his mind rather quickly. In his first foreign trip as President, Obama announced, in the presence of Canadian Prime Minister Stephen Harper, "I want to grow trade and not contract it" (quoted in Stokes, 2009, p. 26).

has preferences represented by the following expected utility function

$$U_0^i = \mathbb{E}_0 \sum_{t=0}^{1} \beta^t \left(u_t^i \right). \tag{1}$$

In this expression u_t^i is the per period utility function of group i and $\mathbb{E}_0$ is the expected value operator conditional on the information available at date $t = 0$. The function u_t^i is assumed to be linear in private good consumption, equal to post-tax income since no savings are allowed, in public good provision G_t, and in a partisan component ξ_t^i reflecting specific preferences for which party is in power, and described more precisely below. It follows from this set of assumptions, and from the additional assumption that the only fiscal instrument available to the government to collect revenues is proportional taxation at rate τ, that the utility of agent $i \in \{p, r\}$ at time t can be represented as

$$u_t^i = (1 - \tau_t) a_t^i + G_t + \xi_t^i, \tag{2}$$

where a_t^i denotes the productivity of agent i at time t. Besides for their preferences on the non-economic policies, citizens differ in terms of their productivity and pre-tax income, equal to a^r for a mass $\lambda \in (0, 1/2)$ of high-skill or rich agents, and equal to $a^p(\theta_t) = \theta_t a^p$ for the remaining low-skill or poor agents, where $0 < a^p < a^r$. Hence we assume the poor are more numerous than the rich, which in this simple two-group model reproduces the skewness of actual income distributions.

The productivity of high skill-agents is constant over time, whereas the productivity of the low-skilled is potentially time-varying since it depends on an endogenous state variable θ_t described below. The aggregate (and average) income produced at time t conditionally on θ_t is thus equal to

$$\bar{a}(\theta_t) \equiv \lambda a^r + (1 - \lambda) \theta_t a^p, \tag{3}$$

The state variable θ_t is equal to a number in the interval $\left[\theta^L, 1\right]$, where $0 < \theta^L < 1$,[10] and it is determined as follows: the value of θ_0 is given by history, with $\theta_0 > \theta^L$; the value of θ_1 is instead chosen by the government in office at period $t = 0$ at no cost. Therefore, potential output is always maximized by setting $\theta_1 = 1$, and any $\theta_1 < 1$ represents an inefficient policy choice; moreover, lower levels of θ_1 imply *a poorer and more unequal society* (and vice versa) since only the unskilled workers experience a productivity loss. The choice of θ_1 can be interpreted in a

---

[10] The lower bound θ^L will not play any special role in our analysis. It is introduced for realism (it does not seem plausible that the Left can completely starve-off the lower class for strategic reasons). In addition, increasing inequality beyond a certain point may be suboptimal for the Left, to the extent that the turnout rate of the poor declines as the society becomes more unequal (e.g. Bénabou, 2000).

variety of ways. For example, setting θ_1 at a level lower than 1 can represent an investment in educational or training programs that are made obsolete by the arrival of new technologies (not explicitly modelled in the paper), even though better alternatives are available at no cost (or, more generally, there exist more cost-effective educational investments). An alternative possible interpretation is a policy increasing unemployment among unskilled workers.[11]

The political process is based on a simple dynamic version of the standard probabilistic voting model of Lindbeck and Weibull (1987) and Dixit and Londregan (1995 and 1998), where the outcome of elections is potentially effected by exogenous ideological shocks.

There are two parties, $J \in \{P, R\}$, where party P denotes the Left and R the party of the Right. The preferences of party J are represented by

$$W_0^J = \mathbb{E}_0 \sum_{t=0}^{1} \beta^t \left(w_t^J \right) \tag{4}$$

with

$$w_t^P = (1 - \tau_t) \, a_t^p + G_t + \delta_t, \tag{5}$$

and

$$w_t^R = (1 - \tau_t) \, a_t^r + G_t + \delta_t, \tag{6}$$

where $\delta_t = \delta$ represents the exogenous potential benefit of each party from being in power (if the party is not in power, then $\delta_t = 0$).[12] Expressions (5) and (6) reflect the assumption that political parties are partially benevolent, i.e. they care about the economic utility of one specific social class, and partially self-interested, as they care about the rents they obtain if in power.[13] As in the spirit of the partisan models of politics pioneered by Alesina (1988), we assume that parties cannot commit to implement a policy different from their own ideal policy.

The government raises revenues with proportional taxation, and its budget constraint reads

$$G_t \leq \tau_t \bar{a} \left(\theta_t \right)) \equiv \tau_t \left[\lambda a^r + (1 - \lambda) \theta_t a^p \right]. \tag{7}$$

---

[11] As discussed in the Introduction, we could also allow government policy to affect the income of the rich and get similar results regarding policies that increase inequality and may or may not be efficient in the aggregate. For simplicity, we assume in the baseline model that θ only affects the income of the poor and focus on the stark situation where the entrenchment policy is both inefficient and harmful to the poor.

[12] Incumbency rents may depend on political institutions, featuring different degrees of checks and balances on the executive power. For example according to a large literature in comparative politics, the extent of separation of powers and checks and balances featured by the U.S. constitution and by many Latin American constitutions are significantly different, and the executive power is much more independent in the latter than in the former system. Other possible interpretations of δ include psychological benefits from being in office (the standard "ego rents") or the income generated by natural resources.

[13] A possible micro-foundation of this assumption is that politicians are citizen-candidates as in the models of Besley and Coate (1997) and Osborne and Slivinsky (1997), and therefore care about the policy implemented. However, politicians also care to some extent about being in power *per se* because they are partially selfish, and because incumbent leaders appropriate of an office rent. See Grossman and Helpman (2001) for a discussion of generalized citizen-candidate models.

As in the literature, we assume that taxes create distortions so that full redistribution is precluded. Instead of introducing elastic labor supply or some convex output loss from taxation, we take a simpler route and assume that taxes higher than $\hat{\tau} \in (0, 1)$ are not feasible given the available fiscal technology. The threshold $\hat{\tau}$ is exogenous and represents the "fiscal capacity" of the state. The identity of the government in office at time t will be denoted by $\rho_t \in \{P, R\}$.

The partisan component ξ_t^i in the per period utility of the citizens (2) can be represented as

$$\xi_t^i = \begin{cases} \varepsilon_t^i + \eta_t, \text{ if } \rho_t = P \\ 0, \text{ if } \rho_t = R. \end{cases} \tag{8}$$

The first term on the right hand side of (8) decomposes the overall ideological bias of citizen i in favor of party P at time t, ξ_t^i, in two components, ε_t^i and η_t; both shocks reflect the preferences of the agent for the non-economic policies potentially implemented by the party in office at time t. In particular, as standard in probabilistic voting models (see for example Lindbeck and Weibull, 1987; Dixit and Londregan, 1995, 1998), we assume that ε represents an idiosyncratic ideological shock that varies both across agents and across time, and whose realizations are $i.i.d.$ over time for each agent and drawn from a continuous distribution function $F(\varepsilon)$ with density $f(\varepsilon) \equiv F'(\varepsilon)$ and zero mean. A positive value of ε^i reflects an idiosyncratic bias of agent i in favor of party P, whereas a negative value of ε^i reflects a bias of agent i against it.[14] We assume that the density function $f(\varepsilon)$ has the following properties.

Assumption 1 $f(x) = f(-x)$;

Assumption 2 $xf'(x) \leq 0$.

Assumptions 1 and 2 imply that the density function of the idiosyncratic ideological shock ε is symmetric around its mean and hump-shaped.

The random variable η_t represents an aggregate ideological shock that is equal for all agents at each point in time and it is $i.i.d.$ over time. The realizations of η_t are drawn from a continuous distribution function $\Phi(\eta)$ with density $\phi(\eta) \equiv \Phi'(\eta)$ and zero mean. A positive value of η_t reflects that the existence of an aggregate bias in favor of party P at time t, whereas a negative value of η_t reflects an aggregate bias against it. While the analysis do not require any distributional restriction on $\Phi(\eta)$, in the comparative static analysis we will assume that $\phi(\eta)$ is hump-shaped and maximum at $\eta = 0$, but not necessarily symmetric.

Assumption 3 $x\phi'(x) \leq 0$.

---

[14]Note, however, that our model differs from a standard probabilistic voting model since we are not allowing parties to commit to any policy (other than their own preferred policy).

We also assume that individuals vote sincerely, which is a weakly dominant strategy in a two-parties system (e.g., Grossman and Helpman, 2001).

In the dynamic political game considered, events take place according to the following timing.

- At the beginning of time $t = 0$, the realization of η_0 and of ε_0^i for each i is revealed.

- Citizens vote for either party P or party R conditionally on θ_0, on η_0 and on ε_0^i; a government is elected depending on the outcome of the voting process.

- The elected government chooses and implements the policy vector $\langle \tau_0, G_0, \theta_1 \rangle$.

- At the beginning of time $t = 1$, the realization of η_1 and of ε_1^i for each i is revealed.

- Citizens vote for either party P or party R conditionally on θ_1, on η_1 and on ε_1^i; a government is elected depending on the outcome of the voting process.

- The elected government chooses and implements the policy vector $\langle \tau_1, G_1 \rangle$, and the game ends.

We will now proceed to characterize the Subgame Perfect Equilibria (SPE) in pure strategies of this game.

4   Characterization of the Equilibrium

4.1   Equilibrium in the Last Period

We solve the model by backward induction, beginning from period $t = 1$. At that point, the realizations of the exogenous ideological shocks are known, and θ_1 has been chosen by the government in office in period $t = 0$. To compute the political equilibrium at $t = 1$, we characterize the equilibrium of the subgame played after elections have been held and a government appointed, for every possible electoral outcome. If in office at period 1, party $J \in \{P, R\}$ implements the fiscal policy solving

$$V^J (\theta_1 | \rho_1 = J) = \max_{\tau_1, G_1} \left\{ (1 - \tau_1) a^J (\theta_1) + G_1 + \delta \right\} \tag{9}$$
$$s.t. \ (7),$$

where $V^J (\theta_1 | \rho_1 = J)$ denotes the indirect utility function of being in power of party J conditionally on θ_1. The preferred fiscal policies of the two political parties at time $t = 1$ are described by the following propositions.

Proposition 1 *A right-wing government always sets taxes at the level $\tau_1^R \equiv 0$ and provides no public good, so that $G_1^R \equiv 0$.*

Proof. Substituting the government budget constraint (that in equilibrium holds as equality) into the objective function of party R leads to

$$V^R\left(\theta_1|\rho_1 = R\right) = \max_{\tau_1}\left\{(1 - \tau_1)\,a^r + \tau_1\left[\lambda a^r + (1 - \lambda)\,\theta_t a^p\right] + \delta\right\}.$$

This function decreasing in τ_1 (as $a^r < \theta_t a^p$) and, therefore, it is maximized at $\tau_1 = 0$. ■

In words, a right-wing government provides no public good since it is not optimal for the rich to spend any amount of fiscal revenues in G. Using Proposition 1, the per period utility of a low-skill producer and of a high-skill producer from voting for party R are respectively

$$u^{i,p}\left(\eta_1, \varepsilon_1^i, \theta_1|v_1^i = R\right) = \theta_1 a^p, \tag{10}$$

and

$$u^{i,r}\left(\eta_1, \varepsilon_1^i, \theta_1|v_1^i = R\right) = a^r. \tag{11}$$

These expressions incorporate the fact the ideological bias in favor of party R is normalized to zero according to (8).

The fiscal policy of the left-wing party is characterized by the following proposition.

Proposition 2 *Conditionally on $\theta_1 \in \left[\theta^L, 1\right]$, a left-wing government implements the following fiscal policy: $\tau_1^P = \hat{\tau}$ and $G_1^P = \hat{\tau}\bar{a}\left(\theta_1\right)$.*

Proof. Taking into account the government budget constraint, the maximization problems of the left-wing party is

$$V^P\left(\theta_1|\rho_1 = P\right) = \max_{\tau_1}\left\{(1 - \tau_1)\,\theta_1 a^p + \tau_1\left[\lambda a^r + (1 - \lambda)\,\theta_1 a^p\right] + \delta\right\}.$$

The objective function of party P is increasing in τ_1 for any θ_1, and is therefore maximized at $\tau_1 = \hat{\tau}$. ■

Using Proposition 2, the per period utility of a low-skill producer from voting for party P is

$$u^{i,p}\left(\eta_1, \varepsilon_1^i, \theta_1|v_1^i = P\right) = \theta_1 a^p + \Delta^P\left(\theta_1\right) + \varepsilon_1^i + \eta_1, \tag{12}$$

where

$$\Delta^P\left(\theta_1\right) = \hat{\tau}\lambda\left(a^r - \theta_1 a^p\right) > 0, \tag{13}$$

14

Similarly, the per period utility of a high-skill producer from voting for party P reads

$$u^{i,r}\left(\eta_1, \varepsilon_1^i, \theta_1 | v_1^i = P\right) = a^r + \Delta^r\left(\theta_1\right) + \varepsilon_1^i + \eta_1, \tag{14}$$

where

$$\Delta^r\left(\theta_1\right) = -\hat{\tau}\left(1 - \lambda\right)\left(a^r - \theta_1 a^p\right) < 0. \tag{15}$$

denotes the net value of fiscal redistribution for the upper class.

Expression (12) decomposes the per period utility of a low-skill worker into the (endogenous) pre-tax income, $\theta_1 a^p$, the net static economic gain from voting for the Left represented by the fiscal redistribution $\Delta^p\left(\theta_1\right)$, i.e. the fiscal transfer net of taxes, for the lower class, and the overall (idiosyncratic plus aggregate) ideological bias in favor of the Left. Since the Right provides no fiscal redistribution ($\Delta^p\left(\theta_1 | R\right) = 0$), $\Delta^p\left(\theta_1\right)$ also represents the relative net economic gain for the low-skill agents from voting for party P rather than for party R. Expression (14) has a similar interpretation for a high-skill worker, with the difference that the net get from fiscal redistribution for the rich, $\Delta^r\left(\theta_1\right)$, is now negative.

A central aspect is that the gain from fiscal redistribution obtained by the lower class is decreasing in θ_1 as

$$\partial\Delta^p\left(\theta_1\right)/\partial\theta_1 \equiv -\hat{\tau}\lambda a^p < 0. \tag{16}$$

This is because as θ_1 increases both the income of each low-skill person and average (and aggregate) income increase; however, aggregate income increases proportionally less since the income of the high-skill persons does not depend on θ_1. It follows that the difference between average and personal income for the lower class (and therefore the net value of fiscal redistribution) decreases with θ_1. This result is important because it implies that the Left can artificially increase its comparative politico-economic value (relative to the Right) for the lower class, $\Delta^p\left(\theta_1\right)$, by reducing the pre-tax income of the unskilled.

In addition, we have that the net gain from fiscal redistribution obtained by the upper class in increasing in θ_1. This is because the difference between average income and personal income, which does not depend on θ_1, for this group is increasing in θ_1; in particular, it is the case that

$$\partial\Delta^r\left(\theta_1\right)/\partial\theta_1 \equiv \hat{\tau}\left(1 - \lambda\right) a^p > 0. \tag{17}$$

Citizen $i \in \{p, r\}$ votes for party P if

$$V^i\left(\eta_1, \varepsilon_1^i, \theta_1 | v_1^i = P\right) \geq V^i\left(\eta_1, \varepsilon_1^i, \theta_1 | v_1^i = R\right), \tag{18}$$

namely if the value that the citizen expects to gain from having party P in office is greater than the value that it expects to obtain if party R is in office. Using (12) and (10), condition (18) can be written for the low-skill citizens $i = p$ as

$$\varepsilon_1^{i,p} \geq -\Delta^p(\theta_1) - \eta_1. \tag{19}$$

For high-skill citizens, the substitution of (11) into (18) leads to

$$\varepsilon_1^{i,r} \geq -\Delta^r(\theta_1) - \eta_1. \tag{20}$$

Since $\Delta^p(\theta_1) > 0$ and $\Delta^r(\theta_1) < 0$, and the distribution of ε is the same across income groups, the proportion of low-skill citizens voting for the Left is higher than the proportion of high-skill citizens, for any value of η_1. This result comes from the fact that low-skill agents benefit from fiscal redistribution while the high-skills are harmed by it; hence, the former group is more inclined to support party P than the latter. Using (19) and (20), and the distributional assumptions on the ideological shocks, the total number of votes that party P obtains in state (η_1, θ_1) is

$$S^P(\eta_1, \theta_1) = (1-\lambda)[1 - F(-\Delta^p(\theta_1) - \eta_1)] + \lambda[1 - F(-\Delta^r(\theta_1) - \eta_1)]. \tag{21}$$

The first term on the right hand side of (21) represents the number of votes that the Left obtains from the low-skill citizens, which is given by the mass $(1-\lambda)$ of low-skill agents multiplied by the fraction of them $[1 - F(-\Delta^p(\theta_1|P) - \eta_1)]$ voting for party P. The second term on the right hand side of (21) represents instead the number of votes that party P obtains from the high-skill citizens and, again, this is given by the number of high-skill agents, λ, multiplied by the fraction voting for party R, $[1 - F(-\Delta^r(\theta_1|P) - \eta_1)]$.

Conditionally on the realization of η_1, the outcome of the election in the final period is straightforward: the Left wins the election with certainty if $S^P(\eta_1, \theta_1) > 1/2$, and it looses the with certainty if $S^P(\eta_1, \theta_1) < 1/2$.[15] Since $S^P(\eta_1, \theta_1)$ is strictly increasing in η_1 for any value of θ_1, the condition ensuring the victory of the Left with certainty is equivalent to $\eta > \eta^*(\theta_1)$, where the cutoff $\eta^*(\theta_1)$ is implicitly defined by the equation

$$S^P(\eta_1, \theta_1) = 1/2. \tag{22}$$

Using expression (21), this condition can be rewritten as

$$(1-\lambda)F(-\Delta^p(\theta_1) - \eta^*(\theta_1)) + \lambda F(-\Delta^r(\theta_1) - \eta^*(\theta_1)) = 1/2. \tag{23}$$

The following proposition summarizes the results obtained up to this point.

---

[15] In the case where $S^P(\eta_1, \theta_1) = 1/2$, both parties win with probablity $1/2$.

Proposition 3 *Conditionally on the realization of the aggregate ideological shock η_1, the equilibrium of the subgame beginning at period $t = 1$ is the following.*

1. *If $\eta_1 < \eta^*(\theta_1)$, party R wins the elections with certainty, i.e. $\rho(\eta_1, \theta_1) = R$, and implements the fiscal policy described by Proposition 1.*

2. *If $\eta_1 > \eta^*(\theta_1)$, party P wins the elections with certainty, i.e. $\rho(\eta_1, \theta_1) = P$, and implements the fiscal policy described by Proposition 2.*

3. *If $\eta_1 = \eta^*(\theta_1)$, party R and party P win the elections with probability $1/2$, and the winner implements its preferred fiscal policy.*

Proposition 3 can be used to compute the *ex ante* probability that the Right or the Left are in power at time $t = 1$, namely the probability that the realization of the random variable $S^P(\eta_1, \theta_1)$ is respectively lower or higher than $1/2$. These results are contained in the following corollary.

Corollary 1 *The* ex ante *probability that party R wins the elections in the final period is*

$$\sigma^R(\theta_1) \equiv \Pr\left\{S^P(\eta_1, \theta_1) < 1/2\right\} = \Pr\left\{\eta < \eta^*(\theta_1)\right\} = \Phi(\eta^*(\theta_1)), \tag{24}$$

and ex ante *the probability that party P wins the elections is*

$$\sigma^P(\theta_1) \equiv 1 - \Phi(\eta^*(\theta_1)). \tag{25}$$

We conclude the description of the equilibrium at period $t = 1$ by stating some additional results which will be used to characterize the equilibrium at period $t = 0$.

Lemma 1 *In the equilibrium of the subgame beginning in period $t = 1$, $\eta^*(\theta_1) < 0$ for all θ_1.*

Proof. Let $\psi(h, \eta) \equiv (1 - \lambda)F(-\lambda h - \eta) + \lambda F((1 - \lambda)h - \eta)$. Clearly, $\partial\psi/\partial\eta < 0$. Furthermore, $\psi(0, 0) = F(0) = 1/2$ and $\partial\psi(h, 0)/\partial h = \lambda(1 - \lambda)[f((1 - \lambda)h) - f(-\lambda h)]$. Since $\lambda < 1/2$, $|(1 - \lambda)h| > |-\lambda h|$. Therefore, $f((1 - \lambda)h) < f(-\lambda h)$ and $\partial\psi(h, 0)/\partial h < 0$. Consequently, $\psi(h, 0) < 1/2$ for all $h > 0$. Next, note that $\eta^*(\theta_1)$ is solution to $\psi(\hat{\tau}(a^r - \theta_1 a^p), \eta^*(\theta_1)) = 1/2$. Since $\psi(h, .)$ is decreasing in η and $\psi(\hat{\tau}(a^r - \theta_1 a^p), 0) < 1/2$, it must be that $\eta^*(\theta_1) < 0$. ∎

Intuitively, because there are more low-skill than high-skill agents, and because low-skills are on average more inclined to vote for the Left than for the Right due to the redistribution bias of the former party, the two parties obtain the same number of votes in equilibrium if

17

there is a sufficiently large aggregate ideological shock in favor of the Right. This is the case if the realization of η falls below some negative threshold.

The following lemma clarifies how the threshold $\eta^*(\theta_1)$ depends on θ_1.

Lemma 2 *In the equilibrium of the subgame beginning in period $t = 1$, $\eta^{*\prime}(\theta_1) \equiv d\eta^*(\theta_1)/d\theta_1 > 0$ for all θ_1.*

Proof. Let us define

$$z_P \equiv -\Delta^p(\theta_1) - \eta^*(\theta_1), \tag{26}$$

and

$$z_R \equiv -\Delta^r(\theta_1) - \eta^*(\theta_1). \tag{27}$$

From $\eta^*(\theta_1) < 0$, (13) and (15) we have $z_R > 0 > z_P > -z_R$. Differentiating (23) with respect to θ_1, we get

$$\eta^{*\prime}(\theta_1) = \lambda(1-\lambda)a^P\hat{\tau}\frac{f(z_P) - f(z_R)}{(1-\lambda)f(z_P) + \lambda f(z_R)}. \tag{28}$$

Since $|z_P| < |z_R|$, from Assumptions 1 and 2 it follows that $f(z_P) > f(z_R)$ and, therefore, $\eta^{*\prime}(\theta_1) > 0$. ∎

Intuitively, the threshold $\eta^*(\theta_1)$ is increasing in the last period productivity of low-skill agents because as these become richer, they gain less from fiscal redistribution, and therefore they are less inclined to voting for the Left. It is also true that a greater θ_1 increases the fraction of rich who will vote for the left. But, as the rich are less numerous than the poor, the net effect is that the overall probability of the Left winning the election falls with θ_1. This is a key result of the model, which will explain why, under some conditions, the Left may choose to reduce the income of its own constituency in order to consolidate its own future political power.

Obviously, if the rich are more numerous than the poor the incentives of the two parties are just reversed: it is now the Right that has electoral gains from a reduction in θ_1. This is because more inequality increases the overall the political support for the Right, as the rich get more averse to redistribution (if θ_1 is lower they get back a lower transfer for any level of taxation) and there are more rich than poor around. This result confirms a general result already discussed in the introduction, namely that entrenchment is by no means a policy exclusively pursued by the Left.

4.2 Equilibrium in the Initial Period

We now complete the backward induction solution of the political game by computing the equilibrium of the subgame of the stage game played at time $t = 0$ after elections have been

18

held and a government has been appointed. Conditionally on the exogenous level of θ_0, the optimal policy of party J at that point in time solves the following problem

$$V^J(\theta_0) = \max_{\tau_0, G_0, \theta_1} \left\{ \left[(1 - \tau_0) a^J(\theta_0) + G_0 \right] + \delta + \beta \mathbb{E}_\eta \left[V^J(\eta, \theta_1) \right] \right\}, \tag{29}$$
$$\text{s.t. (7),}$$

where $\mathbb{E}_\eta \left[V^J(\eta, \theta_1) \right]$ is the expected continuation value of party J conditionally on θ_1, defined as

$$\mathbb{E}_\eta \left[V^J(\eta, \theta_1) \right] = \int V^J(\theta_1 | \rho(\eta, \theta_1)) \phi(\eta) \, d\eta. \tag{30}$$

Using Proposition 3, this expression can be written as

$$\mathbb{E}_\eta \left[V^J(\eta, \theta_1) \right] = \sigma^P(\theta_1) V^J(\theta_1 | \rho_1 = P) + \sigma^R(\theta_1) V^J(\theta_1 | \rho_1 = R). \tag{31}$$

Expression (31) reflects the expectation that party P will be in office in period $t = 1$ with probability $\sigma^P(\theta_1)$, in which case party J obtains utility $V^J(\theta_1 | \rho_1 = P)$, and that party R will be in office in period $t = 1$ with the complementary probability $\sigma^R(\theta_1)$, in which case party J obtains utility $V^J(\theta_1 | \rho_1 = R)$. The expected continuation value of party J is affected by the uncertainty relative to the exogenous future realization of η, i.e. η_1, the endogenous choice of θ_1, and the future allocation of political power ρ_1 that also depends on these two aggregate state variables of the model.[16]

Since the choice of θ_1 is made at no resource cost, it can be separated from the fiscal policy decisions, which are still described by Propositions 1 and 2. Party J chooses the optimal value of θ_1 solving the following maximization problem

$$\theta_1^J = \arg \max_{\theta_1 \in [\theta^L, 1]} \mathbb{E}_\eta \left[V^J(\eta, \theta_1) \right]. \tag{32}$$

Using Propositions 1 and 2, Corollary 1, the expected value $\mathbb{E}_\eta \left[V^J(\eta, \theta_1) \right]$ of party $J = P$ in period $t = 1$ as a function of θ_1 and conditionally on δ can be written as

$$V^P(\theta_1, \delta) = \theta_1 a^p + [1 - \Phi(\eta^*(\theta_1))] [\delta + \Delta^p(\theta_1)]. \tag{33}$$

To gain a better understanding of the logic of political entrenchment, it is useful to decompose the derivative of this objective function with respect to θ_1 into its different components

$$\frac{\partial V^P(\theta_1, \delta)}{\partial \theta_1} = a^p - [1 - \Phi(\eta^*(\theta_1))] \hat{\tau} \lambda a^p - \phi(\eta^*(\theta_1)) \eta^{*\prime}(\theta_1) [\delta + \Delta^p(\theta_1)]. \tag{34}$$

---

[16] The notation used reflects that η affects the welfare of party J only indirectly, i.e. by influencing the probability of winning elections and therefore the future allocation of political power.

19

The sign of this derivative is in general ambiguous since it involves the sum of terms of opposite sign. The first term is positive since it reflects the standard welfare gain that the low-skill agents obtain when they become more productive at the margin. The second term is negative since it reflects the loss in fiscal redistribution that the low-skilled experience when their productivity becomes higher, whenever the Left is in power.[17] The third term is also negative, since $\eta^{*\prime}(\theta_1) > 0$ (see Lemma 2) as it represents the loss that party P obtains at period $t = 1$ due to the reduced chances of winning office. This utility loss includes both the office rent (which party P does not get if the Right is in power) and the fiscal transfer potentially benefiting the lower class (which party P internalizes do to its partial altruism).

Similarly, the expected value $\mathbb{E}_\eta \left[V^J(\eta, \theta_1) \right]$ of party $J = R$ in period $t = 1$ is given by

$$V^R(\theta_1, \delta) = a^r + \Phi(\eta^*(\theta_1)) \delta + [1 - \Phi(\eta^*(\theta_1))] \Delta^r(\theta_1), \tag{35}$$

and the derivative of the expression with respect to the policy variable reads

$$\frac{\partial V^R(\theta_1, \delta)}{\partial \theta_1} = [1 - \Phi(\eta^*(\theta_1))] \hat{\tau}(1 - \lambda) a^p + \phi(\eta^*(\theta_1)) \eta^{*\prime}(\theta_1) [\delta - \Delta^r(\theta_1)] > 0. \tag{36}$$

Equation (36), unlike (34), contains only positive terms; the objective function of party R is thus strictly increasing in θ_1 for any value of δ. This is because a higher value θ_1 benefits the Right both directly (first term of (36)) and indirectly (second term). In particular if θ_1 increases, both the income of the poor and aggregate income increase, and therefore the rich benefit of a greater fiscal transfer when the Left is in power (remind that $\Delta^r(\theta_1) < 0$); moreover, if θ_1 increases the probability of electoral victory of the Right increases. Both effects increase the utility of party R.

The following proposition summarizes these results.

Lemma 3 *The optimal level of productivity of the low-skill producers θ_1 of party R and of party P at time $t = 1$ is the following.*

1. *Party R always sets $\theta_1^R = 1$.*

2. *The optimal value of θ_1 for party P is the solution of the following maximization problem*

$$\theta_1^* \equiv \theta_1^P = \arg \max_{\theta_1 \in [\theta^L, 1]} V^P(\theta_1, \delta) = \theta_1 a^p + [1 - \Phi(\eta^*(\theta_1))][\delta + \Delta^p(\theta_1)]. \tag{37}$$

Proof. In the text. ∎

---

[17] If the Right is in power in period 1 there is no fiscal redistribution, and therefore this effect is absent.

Remark 1 Since the program of party P involves the maximization of a continuous function over a compact interval, it has a solution by Weierstrass theorem. If there is more than one solution, we invoke a standard Paretian efficiency argument to select the largest value of θ as the equilibrium value of θ_1 when the Left is in power.

Remark 2 Since the function (34) is continuous and bounded in θ_1, there exists a finite value of δ independent of θ_1, δ^*, such that $V^P(\theta_1, \delta)$ is everywhere decreasing in θ_1 for any $\delta > \delta^*$.[18] In this case, political rents are sufficiently large to insure that the program of party P has the corner solution $\theta_1^* = \theta^L$.

4.3 Election Results at $t = 0$

We now discuss how election results are determined at $t = 0$. Anticipating the policy vector potentially chosen by both parties, citizen i vote at the beginning of time $t = 0$ for party P if

$$V^{i,p}\left(\eta_0, \varepsilon_0^{i,p}, \theta_0 | v_0^{i,p} = P\right) \geq V^{i,p}\left(\eta_0, \varepsilon_0^{i,p}, \theta_0 | v_0^{i,p} = R\right). \tag{38}$$

This condition can be rewritten as

$$(1 - \hat{\tau})\theta_0 a^p + \hat{\tau}\left[\lambda a^r + (1 - \lambda)\theta_0 a^p\right] + \varepsilon_0^{i,p} + \eta_0 + \beta V^p\left(\theta_1^P\right) \geq \theta_0 a^p + \beta V^p(1),$$

where $V^p\left(\theta_1^P\right)$ and $V^p(1) \equiv V^p\left(\theta_1^R\right)$ denote the expected value of the low-skill producers at time $t = 1$ when respectively party P and party R win the elections at time $t = 0$. In the first case θ_1 is set to $\theta_1^P \equiv \theta_1^*$ given by (37), while $\theta_1^R = 1$ when party R is in power in the first period. Without repeating in detail the analysis of the voting decision, which is similar to the one presented above for period $t = 1$, we obtain that a low-skill citizen votes for party P if

$$\varepsilon_0^{i,p} \geq -\hat{\Delta}^p(\theta_0 | P) - \eta_0, \tag{39}$$

where

$$\hat{\Delta}^p(\theta_0 | P) = \hat{\tau}\lambda(a^r - \theta_0 a^p) + \beta\left[V^p(\theta_1^*) - V^p(1)\right] > 0. \tag{40}$$

Similarly, a high-skill citizen votes for the party P at time $t = 0$ if

$$V^{i,r}\left(\eta_0, \varepsilon_0^{i,r}, \theta_0 | v_0^{i,r} = P\right) \geq V^{i,r}\left(\eta_0, \varepsilon_0^{i,r}, \theta_0 | v_0^{i,r} = R\right),$$

which is equivalent to

$$\varepsilon_0^{i,r} \geq -\hat{\Delta}^r(\theta_0 | P) - \eta_0, \tag{41}$$

---

[18] This follows upon inspection of equation (34), which is continuous and bounded in θ_1 since η^* is so by the implicit function theorem and $\phi(\cdot)$ is a smooth density function. Finally $\eta^{*\prime}(\theta_1)$ is also continuous and bounded, as can be immediately verified by inspecting equation (28).

with

$$\hat{\Delta}^r \left(\theta_0 | P \right) \equiv -\hat{\tau} \left(1 - \lambda \right) \left(a^r - \theta_0 a^p \right) + \beta \left[V^r \left(\theta_1^* \right) - V^r \left(1 \right) \right] < 0. \tag{42}$$

Again, $V^r \left(\theta_1^* \right)$ and $V^r \left(1 \right)$ are the expected value of the high-skill producers at time $t = 1$ when party P and party R are in power at $t = 0$ respectively.

It is worth noting that these conditions differ from those derived for the second period of the game since they are *dynamic* voting rules, which reflect the anticipation of the utility that the citizens obtain in period 1 depending on which political party will be in office. This political outcome potentially depends on the voting decision of the citizens at period $t = 0$ because the government initially elected chooses the value of θ_1. Thereby, it can "manipulate" the electoral outcome in the final period through the strategic choice of the endogenous state variable. In particular, when the Left is also expected to choose $\theta_1^P = 1$ if in power, the continuation value implicitly promised by both parties to the rich and to the poor is the same. In this case, the structure of the voting rules in the initial period is the same as the structure of the voting rules in the second period (with the only exception that the state variable may be different: θ_0 versus θ_1) and agents discriminate between the two parties depending only on the present welfare that they implicitly promise to the voters.[19] This is the equilibrium *with no* political entrenchment. This equilibrium has two important features. First, it is Pareto efficient since potential output is maximized. Second, and more subtly, the equilibrium features no persistence of political power, in the sense that the probability that party P is in power at time $t = 1$, $\sigma^P(\theta_1 = 1)$, does not depend on which party is in power at time $t = 0$.

When instead the Left is expected to choose $\theta_1^P < 1$ if in power, the continuation value implicitly promised by the two parties to the rich and to the poor differs. This is the equilibrium *with* political entrenchment: a party deliberately reduce the income of its constituency in order to consolidate its power. This equilibrium has two important features, specular to the features of the equilibrium without entrenchment. First, it is Pareto inefficient since potential output is not maximized. Second, the equilibrium features persistence of political power, in the sense that the probability that party P is power at time $t = 1$, $\sigma^P(\theta_1)$, depends on which party is in power at time $t = 0$. In particular,

$$\Pr \left\{ \rho_1 = P | \rho_0 = P \right\} = \sigma^P \left(\theta_1 \right) > \Pr \left\{ \rho_1 = P | \rho_0 = R \right\} = \sigma^P \left(1 \right) \tag{43}$$

for all $\theta_1 < 1$, as $\sigma^P \left(\theta_1 \right) = 1 - \Phi \left(\eta^* \left(\theta_1 \right) \right)$ and $\eta^{*\prime} \left(\theta_1 \right) > 0$.

---

[19] The same result obtains in the special case where $\beta = 0$ since in this case future agents do not value future welfare, and therefore dynamic issues do not play any role in voting at the initial period.

22

The number of votes that party P obtains at time $t = 0$ is

$$S^P (\eta_0, \theta_0) = (1 - \lambda) \left[1 - F(-\hat{\Delta}^p (\theta_0|P) - \eta_0)\right] + \lambda \left[1 - F(-\hat{\Delta}^r (\theta_0|P) - \eta_0)\right]. \qquad (44)$$

As before, party P wins election with certainty if $S^P (\eta_0, \theta_0) > 1/2$, which is the case if $\eta > \eta^* (\theta_0)$ where $\eta^* (\theta_0)$ is defined as the value of η_0 that satisfies the condition

$$S^P (\eta^* (\theta_0), \theta_0) = 1/2. \qquad (45)$$

The following proposition provides a complete characterization of the equilibrium in the first period of the game.

Proposition 4 *The equilibrium of the subgame beginning at period $t = 0$ is as follows:*

1. *If $\eta_0 < \eta^* (\theta_0)$, party R wins elections, implements the fiscal policy ($\tau_0^R = 0, G_0^R = 0$), and sets $\theta_1^R = 1$.*

2. *If $\eta_0 > \eta^* (\theta_0)$, party P wins elections and implements the fiscal policy ($\tau_0^P = \hat{\tau}, G_0^P = \hat{\tau}\bar{a}(\theta_0)$). Moreover, party P chooses $\theta_1^P \equiv \theta_1^* \in [\theta^L, 1]$ as defined in Lemma 3.*

3. *If $\eta_0 = \eta^* (\theta_0)$, party R wins the elections with probability $1/2$ and implements the policy described in case 1; party P wins the elections with probability $1/2$ and implements the policy described in case 2.*

Proof. In the text. ∎

The equilibrium at time 0 defined in Proposition 4 is similar to that of period 1 contained in Proposition 3. It is worth noting that the threshold $\eta^* (\theta_0)$ is increasing in θ_0, since the potential gain from fiscal redistribution and therefore the bias of the low-skill agents in favor of the Left decreases as their pre-tax income increases relative to the average income. An important implication of this result and of Proposition 4 is that the equilibrium of the model potentially exhibits a certain form of *path dependence*. This is because the threshold $\eta^* (\theta_0)$, and therefore the probability that the Right is in power in period $t = 0$, increases with θ_0. In turn, if the Right is in power in period $t = 0$ it sets $\theta_1^R = 1$, which maximizes the probability that it will remain in power in the following period. Similarly, the lower is θ_0, the higher is the probability that the Left is in power in both periods.

5 Comparative Statics

We now move to analyze how some parameters of the model affect the level of entrenchment of the Left, namely the level of productivity of the low-skilled θ_1^* chosen by the Left when in power in the initial period. In particular, we determine how θ_1^* changes when there is a variation in the benefit δ of the party from being in power, a variation in the balance of power and a change in state capacity $\hat{\tau}$.

As the objective function of the Left (33) is not generally concave in θ_1, we cannot apply the standard methods (based on the implicit differentiation of the first order condition) to characterize the comparative statics properties of the equilibrium. For this reason the analysis will be conducted using the techniques of monotone comparative statics (e.g., Milgrom and Shannon, 1994; Topkis, 1998).

5.1 An Increase in Office Rents

An increase in the rents δ from power always increase the incentive of the Left for entrenchment, namely the choice of a (weakly) lower θ_1. This result comes from the fact that $V^P(\theta_1, \delta)$ is submodular in (θ_1, δ), i.e. $\partial^2 V^P(\theta_1, \delta)/\partial\theta_1\partial\delta < 0$, so that $\partial\theta_1^*/\partial\delta \leq 0$ by Topkis' theorem (Topkis, 1998). In fact, from (34) we obtain that

$$\frac{\partial^2 V^P(\theta_1, \delta)}{\partial\theta_1\partial\delta} = -\phi(\eta^*(\theta_1))\eta^{*\prime}(\theta_1),$$

which is always negative. By Topkis' theorem, this implies that θ_1^* is non-increasing in δ (i.e., $\partial\theta_1^*/\partial\delta \leq 0$). The intuition for this result is straightforward. As the gains from being in office increase, party P will want to increase the probability of winning the elections at time 1, which is possible by decreasing the pre-tax income of its constituency (namely by choosing a lower value of θ_1).

5.2 Effects of the Balance of Power

We now consider the effect of an exogenous variation in the balance of power in favor of the Left. A way to model this bias is by considering a shift to the right in the distribution function of the ideological bias ε of the citizens for party P. Formally, the cumulative distribution function of ε can be rewritten as $F(\varepsilon - \nu)$, with the density function equal to $f(\varepsilon - \nu) \equiv F'(\varepsilon - \nu)$, where ν is the mean of the distribution ($\nu = 0$ corresponds to our baseline model)[20]. An

---

[20]Note that Assumptions 1 and 2 are no longer valid. However, if ν is not too large, the critical property that the marginal density of swing voters is lower among the rich than among the poor, which is behind Lemmas 1 and 2, still holds.

increase in ν implies a bias in favor of the Left. In order to determine the sign of $\partial \theta_1^* / \partial \nu$, we explore whether $V^P(\theta_1, \nu)$ is supermodular or submodular in (θ_1, ν). From (34), we have

$$\frac{\partial^2 V^P(\theta_1, \nu)}{\partial \theta_1 \partial \nu} = \phi(\eta^*(\theta_1)) \frac{\partial \eta^*(\theta_1)}{\partial \nu} \hat{\tau} \lambda a^p - \phi'(\eta^*(\theta_1)) \frac{\partial \eta^*(\theta_1)}{\partial \nu} \eta^{*\prime}(\theta_1) [\delta + \Delta^p(\theta_1)]$$
$$- \phi(\eta^*(\theta_1)) \frac{\partial \eta^{*\prime}(\theta_1)}{\partial \nu} [\delta + \Delta^p(\theta_1)].$$

By deriving (23), it can be easily shown that $\partial \eta^*(\theta_1) / \partial \nu = -1$ and $\partial \eta^{*\prime}(\theta_1) / \partial \nu = 0$, so that

$$\frac{\partial^2 V^P(\theta_1, \nu)}{\partial \theta_1 \partial \nu} = -\phi(\eta^*(\theta_1)) \hat{\tau} \lambda a^p + \phi'(\eta^*(\theta_1)) \eta^{*\prime}(\theta_1) [\delta + \Delta^p(\theta_1)]. \tag{46}$$

The first term in (46) is negative and represents the *appropriation effect*. It implies that θ_1^* is non-increasing in ν (i.e., $\partial \theta_1^* / \partial \nu \leq 0$). The intuition is the following. A higher ν means that citizens have a positive bias in favor of the Left party P, that will then be in power more often. This implies that the poor will get income redistribution more often and, therefore, the cost of entrenchment (i.e., of a lower pre-tax income) is smaller.[21] This leads to more entrenchment, namely to a lower level of θ_1^*. The second term in (46) is positive and represents the *decreasing marginal political return effect*, which leads to θ_1^* being non-decreasing in ν (i.e., $\partial \theta_1^* / \partial \nu \geq 0$). A positive bias in favor of party P (higher ν) reduces the mass of citizens at the margin. This is the result of the fact that $\partial \eta^*(\theta_1) / \partial \nu < 0$ and Assumption 3 which implies that $\phi'(\eta^*(\theta_1)) > 0$ as $\eta^*(\theta_1) < 0$. Given the existence of fewer citizens at the margin (i.e., $\phi'(\eta^*(\theta_1)) \eta^{*\prime}(\theta_1) > 0$), a reduction of θ_1 would lead to a lower increase in the votes gained by the Left, which means that the marginal political return from entrenchment (a lower θ_1) decreases with ν. This leads to a higher θ_1^* and to less entrenchment. Therefore, the effect of a positive bias in favor of the Left on entrenchment will generally be ambiguous. However, one may observe that the second term in (46) depends on δ while the first term does not. This means that when the rents δ from office are sufficiently high, the decreasing marginal political return effect is likely to dominate over the appropriation effect, which implies that a positive bias of the citizens for the Left leads to less entrenchment ($\partial^2 V^P(\theta_1, \nu) / \partial \theta_1 \partial \nu > 0$ and $\partial \theta_1^* / \partial \nu \geq 0$). More fundamentally, the appropriation effect comes from the "citizen-candidate" part of the politicians' objective function: as individuals, they consider they also lose from entrenchment, and any reduction in this loss will increase their incentives to entrench. On the other hand, the decreasing marginal political returns effect is driven by the cross-effect of the change in the parameter of interest (here, ν) and of θ on reelection probability. Here it tells us that entrenchment yields less

---

[21] We remind that income redistribution partially offset the income loss of the poor generated by a lower level of their productivity θ_1.

political leverage, the higher ν, because people like the left more and their voting behavior is therefore less elastic to the choice of θ.[22]

5.3 The Effect of State Capacity

We now analyze the effect of a higher state capacity $\hat{\tau}$ on the entrenchment of the Left. Again, from (34), we obtain

$$
\begin{aligned}
\frac{\partial^2 V^P(\theta_1, \hat{\tau})}{\partial \theta_1 \partial \hat{\tau}} &= -\phi(\eta^*(\theta_1))\eta^{*\prime}(\theta_1)\,\lambda\,(a^r - \theta_1 a^p) - [1 - \Phi(\eta^*(\theta_1))]\,\lambda a^p \\
&\quad + \hat{\tau}\lambda a^p \phi(\eta^*(\theta_1))\frac{\partial \eta^*(\theta_1)}{\partial \hat{\tau}} - \phi'(\eta^*(\theta_1))\eta^{*\prime}(\theta_1)\frac{\partial \eta^*(\theta_1)}{\partial \hat{\tau}}\,[\delta + \Delta^p(\theta_1)] \\
&\quad - \phi(\eta^*(\theta_1))\frac{\partial^2 \eta^*(\theta_1)}{\partial \theta_1 \partial \hat{\tau}}\,[\delta + \Delta^p(\theta_1)].
\end{aligned}
\tag{47}
$$

By deriving the implicit function (23) defining $\eta^*(\theta_1)$ with respect to $\hat{\tau}$ and using (26) and (27), we get

$$
\frac{\partial \eta^*(\theta_1)}{\partial \hat{\tau}} = -\lambda(1-\lambda)\,(a^r - \theta_1 a^p)\,\frac{f(z_P) - f(z_R)}{(1-\lambda)f(z_P) + \lambda f(z_R)} < 0
$$

because $f(z_P) > f(z_R)$. Moreover,

$$
\begin{aligned}
\frac{\partial^2 \eta^*(\theta_1)}{\partial \theta_1 \partial \hat{\tau}} &= \lambda(1-\lambda)a^p \frac{f(z_P) - f(z_R)}{(1-\lambda)f(z_P) + \lambda f(z_R)} \\
&\quad - \lambda(1-\lambda)\hat{\tau}a^p\,(a^r - \theta_1 a^p)\,\frac{[(1-\lambda)\,f(z_P)f'(z_R) + \lambda f(z_R)f'(z_P)]}{[(1-\lambda)f(z_P) + \lambda f(z_R)]^2} \\
&\quad - \lambda(1-\lambda)\hat{\tau}a^p \frac{\partial \eta^*(\theta_1)}{\partial \hat{\tau}}\frac{f(z_R)f'(z_P) - f(z_P)f'(z_R)}{[(1-\lambda)f(z_P) + \lambda f(z_R)]^2},
\end{aligned}
$$

which is generally ambiguous as the first term is positive, the second is ambiguous and the last term is negative since $f'(z_P) > 0$ and $f'(z_R) < 0$ (from $z_R > 0 > z_P$ and Assumption 2).

The first three terms in (47) are negative and they represent the *appropriation effect* generated by a higher feasible tax rate. These components push towards a reduction of θ_1^* as $\hat{\tau}$ increases $(\partial \theta_1^*/\partial \hat{\tau} \leq 0)$. The interpretation of these components is the following. The first term comes from the fact that the poor gain more from the Left in power when $\hat{\tau}$ is higher. Hence, party P will want to reduce θ_1 more in order to increase the probability of being elected. The second term captures the increase in redistribution generated by a higher $\hat{\tau}$ when the Left is in power. This reduces the cost of a lower pre-tax income of the poor (i.e., of a lower θ_1^*) so favoring entrenchment. The third term accounts for the fact that a higher $\hat{\tau}$ reduces $\eta^*(\theta_1)$

---

[22] This is true because in the original situation, the left already has more power due to the fact that the poor are more numerous. This makes the marginal swing voters atypically averse to the left. An increase in ν further increases the Left's political power, thus making the swing voters even more atypical, i.e. less numerous. Consequently, the electoral benefits from entrenchment are reduced.

making the Left more popular and more often in power. Again, this reduces the cost of a lower pre-tax income of the poor and favors entrenchment. The last two terms represent the *marginal political return effect* generated by the increase in state capacity. The fourth term is generated by the fact that the increase in $\hat{\tau}$ reduces $\eta^*(\theta_1)$ (the Left is more popular), and this reduces the mass of citizens at the margin. Hence, a reduction of θ_1 would lead to a lower increase in the votes gained by the Left, which reduces the scope for entrenchment, so leading to a higher θ_1^*. The final term captures the direct effect of $\hat{\tau}$ on the responsiveness of election outcomes to entrenchment, as captured by $\eta^{*\prime}(\theta_1)$, and it can be positive or negative. As the *appropriation effect* and the *marginal political return effect* may have opposite sign, the total effect of higher state capacity on the choice of θ_1^* will generally be ambiguous. Nevertheless, in some special cases, which are described next, the effect of state capacity of entrenchment is not ambiguous.

Lemma 4 *If $\hat{\tau}$ is closed enough to zero, higher state capacity generates more political entrenchments.*

Proof. Straightforward computations using (34) and the fact that $\eta^{*\prime}(\theta_1) \downarrow 0$ as $\hat{\tau} \downarrow 0$ show that

$$\lim_{\hat{\tau} \to 0} \frac{\partial^2 V^P(\theta_1, \hat{\tau})}{\partial \theta_1 \partial \hat{\tau}} = -\left[1 - \Phi(\eta^*(\theta_1))\right] \lambda a^p < 0.$$

Therefore, if state capacity is limited, by submodularity follows that higher state capacity weakly increases entrenchment by the Left. ∎

Lemma 4 is of interests since it makes clear that more state capacity is not *always* beneficial for the society, but may actually turn out to be a double-edge sword. In particular, higher state capacity may be harmful for society by giving to the party in power more incentives to consolidate its own future power. Note also that low state capacity is more characteristic of developing than of developed economies. Therefore, Lemma 4 suggests that the potentially harmful consequences of an increase in state capacity, through the political economy interactions identified in this paper, may be more serious for developing than for developed countries.

A higher degree of income inequality has an effect on the choice of θ_1 very similar to that of state capacity. The analysis is reported in Appendix.

The following lemma summarizes the above results.

Lemma 5 *The optimal level of productivity of the low-skill producers θ_1^* of party P is affected by the following.*

1. *An increase in the rents δ from office (weakly) reduces θ_1^* (i.e., $\partial \theta_1^*/\partial \delta \leq 0$).*

2. *A positive bias ν in favor of the poor has an ambiguous effect on θ_1^*, and reduces entrenchment when the rents from office are sufficiently high (i.e., $\partial\theta_1^*/\partial\nu \geq 0$).*

3. *A higher state capacity $\hat{\tau}$ has in general an ambiguous effect on θ_1^*. See Lemma 4 for some special results.*

6   One Application: Term Limits

According to the standard models of Barro (1973) and Ferejohn (1986) elections allow the citizens to solve (partially) the potential moral hazard problem of incumbent politicians, who can be induced not to appropriate for themselves of too many resources from the public purse,[23] by the implicit threat of non re-election in the future. Persson, Roland and Tabellini (1997) reach essentially the same conclusion in a model featuring a richer description of political institutions.

In these frameworks, a term limit is unambiguously welfare reducing since it forces voters not to reappoint a politician in office *regardless* on its performance. Therefore, under a term limit elections lose part of their efficacy as discipline device, which implies that the implicit contract offered by the voters to incumbent politicians must allow the latter to appropriate greater rents. This result raises the natural (but surprisingly under-researched) question of why many real world constitutions prescribe term limits for elections held either at the national or at the local level.

In this Section, we take a first step toward addressing this puzzle, by extending the baseline model assuming the existence of a constitutionally legislated term limit, which prevents an individual politician, *but not its own party*, in office in the first period to compete for office in the second period. The main result will be that term limits may be beneficial for society (i.e. the sum of rich and poor citizens, not including the political parties) by reducing the scope for political entrenchment.[24]

We assume that individual politicians initially in power are partially altruistic toward their own party, in the sense that their utility in the second period of the game (where they are

---

[23] In a type of political replacement game à la Barro-Ferejohn, "corruption" is not zero in equilibrium since the politicians in office can always appropriate of some incumbency rents. In particular, the strategy of "going for broke," that is of steeling as much as possible from the public purse at the cost of not being re-elected is credible for incumbent politicians if voters condition the re-election of the incumbent on a too high level of performance (i.e. allow too little rent appropriation).

[24] Here we do not provide a complete positive theory of term limits (which is beyond the purpose of this paper) since in our framework term limits will have only benefits for the society (i.e. reducing the scope for entrenchment), but no costs. A general positive theory of term limits should incorporate both the benefits (such as those we emphasize) and the costs (such as those emphasized by the models of Barro, 1973, and of Ferejohn, 1986), and explain how the corresponding trade-off is resolved.

never in office) from the rent potentially appropriated by their party at that point is equal to a fraction $\epsilon \in [0,1]$ of δ. The parameter ϵ can be interpreted in a variety of ways. For example, it can reflect a genuine concern of the politicians for its own party. Alternatively, ϵ can capture the degree of "party discipline," defined as the capacity of a party to align the preferences (and decisions) of its members in office to its own preferences (see Grossman and Helpman, 2008). More generally, this assumption is in the spirit of the observation (e.g., Alesina and Spear, 1988) that political parties are more durable, and have longer horizons, than politicians.

Formally, the lifetime preferences of a politician from party $J \in \{P, R\}$ under term limits can be represented as

$$W_0^{J,TL} = \mathbb{E}_0 \sum_{t=0}^{1} (\beta)^t \left(w_t^{J,TL} \right), \tag{48}$$

where

$$w_t^{J,TL} = (1 - \tau_t) a_t^j + G_t + (t(\epsilon - 1) + 1) \delta_t. \tag{49}$$

The analysis of the equilibrium in presence of terms limits is straightforward since the objective function of party J at time is the same as before at time $t = 0$, and has $\epsilon\delta$ rather than δ at time $t = 1$. It follows that the structure of the equilibrium and its comparative statics properties are essentially the same as before. In particular, the objective function of party P at time $t = 1$ is submodular in ϵ; by standard monotone comparative statics results this implies that, conditionally on the presence of term limits, greater party discipline leads to a (weakly) higher entrenchment as it induces the politicians to act more in line with the preferences of their own party.

In the limit case of $\epsilon = 1$, party discipline is so strong that the politician in office at period $t = 0$ behaves as perfect agent of its own party, maximizing the party's continuation value, and choosing the same value of θ_1 that it would choose in absence of term limits. In this case, term limits are of course irrelevant. In the polar case of $\epsilon = 0$, conversely, a term limit is maximally effective since it induces a left-wing politician to act as perfect agent of its own constituency, just as if it would if there were no political rents ($\delta = 0$), and therefore to set $\theta_1^P = 1$.

We conclude this section by observing that term limits themselves are not necessary a panacea to political moral hazard problems, as the extent to which term limits actually reduce the incentives of incumbent politicians to implement policies of entrenchment depends critically on the degree of party discipline.

7   Conclusions

This paper has posed a simple but, to best of our knowledge, novel question: why do politicians occasionally implement policies damaging the economic interests of the same people that brought them in power?

To address this question, we have studied a simple dynamic model of voting over a purely redistributive fiscal policy in a two-parties system. The Left party alone can credibly commit to provide some redistribution in favor of the lower classes, but the economic gain of voting for it may be outweighed by the emergence of a strong ideological bias in favor of the Right. We have found that policies that reduce the income of the poor relative to the average income, such as failing to upgrade the skills of the workers and preventing their erosion by new, skill-biased, technologies, paradoxically consolidates the political power of the Left. This is because the policies in question make the natural constituency of a left-wing party endogenously more dependent on it, and therefore have some political appeal for the party itself. An equilibrium with entrenchment features relative economic stagnation and in particular falling unskilled wages, higher inequality, *and* persistence of the power of the Left. Importantly, in our model such an equilibrium is not based on any form of myopia or irrationality on either politicians or voters. The equilibrium with entrenchment is more likely to emerge, *inter alia*, when the political rents appropriated by the incumbent leaders are relatively high, which is the case for example in weakly-institutionalized polities.

It is important to remark the scope for political entrenchment is much wider then suggested by our simple model focusing only on the politics of pure income redistribution. It is relatively easy to think to other scenarios (involving for example the provision of different public goods) where the Right, as opposed to the Left, has a strategic incentive to entrench itself in power.

An interesting application of the theory is in the area of the positive analysis of institutional design, by providing one rationale for the existence of term limits in a number of democratic constitutions around the world. We have demonstrated that term limits are potentially beneficial for society to the extent that they reduce the incentives of incumbent politicians in engaging in socially inefficient entrenchment by reducing their political time-horizon. Nevertheless, term limits work only in those political systems where party discipline is relatively weak and become irrelevant (if not harmful) where parties exert as sufficiently strong control on the choices made by their own members elected in office.

8 Appendix

We here analyze the effect of income inequality on entrenchment. We rewrite

$$a^r = \frac{\mu}{\lambda}a \quad \text{and} \quad a^p = \frac{1-\mu}{1-\lambda}a, \tag{50}$$

with $\mu \in (\lambda, 1)$ representing an index of income inequality and a the average income.

From (34), we obtain

$$
\begin{aligned}
\frac{\partial^2 V^P (\theta_1, \mu)}{\partial \theta_1 \partial \mu} &= -\frac{a}{1-\lambda}[1 - (1 - \Phi(\eta^*(\theta_1)))\hat{\tau}\lambda] \\
&\quad -\phi(\eta^*(\theta_1))\eta^{*\prime}(\theta_1)\left[1 + \hat{\tau}\lambda\left(\frac{a}{\lambda} + \frac{\theta_1 a}{1-\lambda}\right)\right] + \phi(\eta^*(\theta_1))\frac{\partial \eta^*(\theta_1)}{\partial \mu}\hat{\tau}\lambda a^p \\
&\quad -\phi'(\eta^*(\theta_1))\frac{\partial \eta^*(\theta_1)}{\partial \mu}\eta^{*\prime}(\theta_1)[\delta + \Delta^p(\theta_1)] - \phi(\eta^*(\theta_1))\frac{\partial \eta^{*\prime}(\theta_1)}{\partial \mu}[\delta + \Delta^p(\theta_1)].
\end{aligned}
\tag{51}
$$

Deriving (23) with respect to μ and using (50), we have

$$\frac{\partial \eta^*(\theta_1)}{\partial \mu} = -\hat{\tau}\lambda(1-\lambda)\left(\frac{a}{\lambda} + \frac{\theta_1 a}{1-\lambda}\right)\frac{f(z_P) - f(z_R)}{(1-\lambda)f(z_P) + \lambda f(z_R)} < 0.$$

The derivative of (28) with respect to μ leads to

$$
\begin{aligned}
\frac{\partial \eta^{*\prime}(\theta_1)}{\partial \mu} &= -\lambda\hat{\tau}a\frac{f(z_P) - f(z_R)}{(1-\lambda)f(z_P) + \lambda f(z_R)} \\
&\quad -\lambda(1-\lambda)\hat{\tau}^2 a^p\left(\frac{a}{\lambda} + \frac{\theta_1 a}{1-\lambda}\right)\frac{[\lambda f(z_R)f'(z_P) - (1-\lambda)f(z_P)f'(z_R)]}{[(1-\lambda)f(z_P) + \lambda f(z_R)]^2} \\
&\quad -\lambda(1-\lambda)\hat{\tau}a^p\frac{\partial \eta^*(\theta_1)}{\partial \mu}\frac{f(z_R)f'(z_P) - f(z_P)f'(z_R)}{[(1-\lambda)f(z_P) + \lambda f(z_R)]^2},
\end{aligned}
$$

whose sign is generally ambiguous as the first term is negative and the other two are positive.

The effect of higher income inequality on entrenchment is ambiguous and very similar to the effect of state capacity. In fact, the first three terms in (51) are negative and they represent the *appropriation effect* generated by higher inequality. These components push towards a reduction of θ_1^* as inequality represented by μ increases $(\partial \theta_1^*/\partial \mu \leq 0)$. The last two terms are the *marginal political return effect* from higher inequality. The fourth term is positive and come from the fact that an increase in μ makes the Left more popular so reducing the mass of citizens at the margin. This reduces the scope for entrenchment and leads to a higher θ_1^* $(\partial \theta_1^*/\partial \mu \geq 0)$. The final term has a similar interpretation and will generally be ambiguous. As the *appropriation effect* and the *marginal political return effect* may have opposite sign, the total effect of income inequality on θ_1^* will generally be ambiguous.

9 References

Acemoglu, Daron and James A. Robinson (2001) "Inefficient Redistribution," *American Political Science Review*, 95, pp. 649-662.

Acemoglu, Daron and James A. Robinson (2008) "Persistence of Power, Elites and Institutions," *American Economic Review*, 98(1), pp. 267-93.

Acemoglu, Daron, Davide Ticchi and Andrea Vindigni (2010a) "A Theory of Military Dictatorships," *American Economic Journal: Macroeconomics*, January, 2(1), pp. 1–42.

Acemoglu, Daron, Davide Ticchi and Andrea Vindigni (2010b) "Persistence of Civil Wars," *Journal of the European Economic Association*, April-May, 8(2–3), pp. 664–676.

Acemoglu, Daron, Davide Ticchi and Andrea Vindigni (2011) "Emergence and Persistence of Inefficient States," Forthcoming at the *Journal of the European Economic Association.*

Aghion, Philippe and Patrick Bolton (1990) "Government Domestic Debt and the Risk of Default: A Political-Economic Model of the Strategic Role of Debt," in *Capital Markets and Debt Management*, R. Dornbusch and M. Draghi, eds., Cambridge, MA: MIT Press.

Alesina, Alberto (1988) "Credibility and Policy Convergence in a Two-Party System with Rational Voters," *American Economic Review*, 78, pp. 796-806.

Alesina, Alberto, and Stephen Spear (1988) "An Overlapping Generations Model of Electoral Competition," *Journal of Public Economics*, 37, pp. 359-379.

Alberto Alesina, and Guido Tabellini (1990) "A Positive Theory of Fiscal Deficits and Government Debt," *The Review of Economic Studies*, Vol. 57(3), pp. 403-414.

Barro, Robert (1973) "The Control of Politicians: An Economic Model," *Public Choice*, 14(1), pp. 19-42.

Bénabou, Roland (2000) "Unequal Societies: Income Distribution and the Social Contract," *American Economic Review*, 90(1), pp. 96-129.

Bénabou, Roland (2002) "Tax and Education Policy in a Heterogeneous Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?," *Econometrica*, 70(2), March, pp. 481-517.

Besley, Timothy and Stephen Coate (1997) "An Economic Model of Representative Democracy," *Quarterly Journal of Economics*, 108, pp. 85-114.

Besley, Timothy and Stephen Coate (1998) "Sources of Inefficiency in a Representative Democracy: A Dynamic Analysis," *American Economic Review* 88(1), pp. 139-156.

Biais, Bruno and Enrico Perotti (2002) "Machiavellian Privatization," *American Economic Review*, 92(1), pp. 240-258.

Brainard, S. Lael and Thierry Verdier (1997) "The Political Economy of Declining Industries: Senescent Industry Collapse Revisited," *Journal of International Economics*, 42(1), pp. 221-238.

Kristin F. Butcher and David Card (1991) "Immigration and Wages: Evidence from the 1980's," *American Economic Review*, 81(2), Papers and Proceedings, pp. 292-296.

LaLonde, Robert and Robert Topel (1991) "Labor Market Adjustments to Increased Immigration." In *Immigration, Trade, and the Labor Market*, ed. J. Abowd and R. Freeman. Chicago: University of Chicago Press.

Coate, Stephen and Stephen Morris (1995) "On the Form of Transfers to Special Interests," *Journal of Political Economy*, 103(6), pp. 1210-1235.

Coate, Stephen and Stephen Morris (1999) "Policy Persistence," *American Economic Review*, 89(5), pp. 1327-1336.

Dixit, Avinash, and John Londregan (1995) "Redistributive Politics and Economic Efficiency." *American Political Science Review* 89 (4), pp. 856-66.

Dixit, Avinash, and John Londregan (1998) "Ideology, Tactics, and Efficiency in Redistributive Politics," *Quarterly Journal of Economics* 113, pp. 497-529.

Dornbusch, Rudiger and Sebastian Edwards (1992) *The Macroeconomics of Populism in Latin America.* Chicago IL: University Of Chicago Press.

Edwards, Sebastian (2010) *Left Behind: Latin America and the False Promise of Populism.* Chicago, IL: University Of Chicago Press.

Ferejohn, John (1986) "Incumbent Performance and Electoral Control," *Public Choice*, 50, pp. 5-25.

Grossman, Gene and Elhanan Helpman (2001) *Special Interest Politics.* Cambridge, MA: The MIT Press.

Grossman, Gene and Elhanan Helpman (2008) "Party Discipline and Pork-Barrel Politics," in *Institution and Economic Performance*, edited by Elhanan Helpman. Cambridge, MA: Harvard University Press.

Hassler, John, José V. Rodrìguez Mora, Kjetil Storesletten and Fabrizio Zilibotti (2003) "The Survival of the Welfare State," *American Economic Review*, 93(1), pp. 87-112.

Hix, Simon and Abdul Noury (2007) "Politics not Economic Interests: Determinants of Migration Policies in the European Union," *International Migration Review* 41(1) pp. 182-205.

Kahane, Leo H. (1996) "Congressional Voting Patterns on NAFTA: An Empirical Analysis," *American Journal of Economics and Sociology,* 55(4), 395-409.

Kang, In-Bong and Kenneth Greene (1999) "A Political Economic Analysis of Congressional Voting Patterns on NAFTA," *Public Choice,* 98(3/4), 385-397.

Lindbeck, Assar and Jörgen W. Weibull (1987) "Balanced-budget Redistribution as the Outcome of Political Competition," *Public Choice* 52(3): pp. 272–97.

Mayda, Anna Maria (2006) "Who is against immigration? A Cross-Country Investigation of Individual Attitudes Toward Immigrants," *Review of Economics and Statistics*, 88(3), pp. 510-530.

Milgrom, Paul and Shannon, Chris (1994) "Monotone Comparative Statics," *Econometrica*, 62, pp. 157–180.

Meltzer, Allan and Scott Richards (1981) "A Rational Theory of the Size of Government," *Journal of Political Economy*, 89 (5): pp. 914-927.

Osborne, Martin J. and Al. Slivinsky (1996) "A Model of Political Competition with Citizen-Candidates," *Quarterly Journal of Economics*, 111(1), pp. 65-96.

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

Persson, Torsten, Gérard Roland and Guido Tabellini (1997) "Separation of Powers and Political Accountability," *Quarterly Journal of Economics*, 112(4), 1163-1202.

Persson, Torsten and Lars E. O. Svensson (1989) "Why a Stubborn Conservative would Run a Deficit: Policy with Time-Inconsistent Preferences," *Quarterly Journal of Economics*, 104(2), pp. 325-345.

Robinson, James A. and Ragnar Torvik (2005) "White Elephants," *Journal of Public Economics*, 89, pp. 197-210.

Rodrik, Dani (1997) "Has Globalization Gone Too Far?," Washington DC: Institute for International Economics.

Romer, Thomas (1975) "Individual Welfare, Majority Voting, and the Properties of a Linear Income Tax," *Journal of Public Economics*, Vol. 4(2), pp. 163-185.

Saint-Paul, Gilles (1993) "On the Political Economy of Labor Market Flexibility," *NBER Macroeconomics Annual*, pp. 151-95.

Saint-Paul, Gilles and Thierry Verdier (1997) "Power, Distributive Conflicts, and Multiple Growth Paths," *Journal of Economic Growth*, vol. 2, pp. 155-168.

Saint-Paul, Gilles (2002) "The Political Economy of Employment Protection," *Journal of Political Economy*, 110(3), pp. 672-704.

Scheve, Kenneth F. and Matthew Slaughter (2001) "Labor-Market Competition and Individual Preferences Over Immigration Policy," *Review of Economics and Statistics*, 83, pp.

133-145.

Shleifer, Andrei and Robert W. Vishny (1989) "Management Entrenchment. The Case of Manager-Specific Investments," *Journal of Financial Economics*, 25, pp. 123-139.

Stokes, Susan C. (2009) "Globalization and the Left in Latin America," Yale mimeo.

Tabellini, Guido and Alberto Alesina (1990) "Voting on the Budget Deficit," *American Economic Review*, 80(1), pp. 37-49.

Topkis, Donald M. (1998) *Supermodularity and Complementarity*, Princeton, NJ: University Press.

Weisbach, Michael S. (1988) "Outside Directors and CEO Turnover," *Journal of Financial Economics*, 20, pp. 431-460.