A Theory of Political Entrenchment

GILLES SAINT PAUL, DAVID TICCHI AND

ANDREA VINDIGNI

A Theory of Political Entrenchment[*]

Gilles Saint-Paul
TSE

Davide Ticchi
IMT Lucca

Andrea Vindigni
IMT Lucca

March 2012

Abstract

We develop a theory of endogenous political entrenchment in a simple two-party dynamic model of income redistribution with probabilistic voting. A partially self-interested left-wing party may implement (entrenchment) policies reducing the income of its own constituency, the lower class, in order to consolidate its future political power. Such policies increase the net gain that low-skill agents obtain from income redistribution, which only the Left (but not the Right) can credibly commit to provide, and therefore may help offsetting a potential future aggregate ideological shock averse to the left-wing party. We demonstrate that political entrenchment by the Left occurs only if incumbency rents are sufficiently high and that low-skill citizens may vote for this party even though they rationally expect the adoption of these policies. We also discuss the case where the left-wing party may have the incentive to ex-ante commit to not pursue entrenchment policies once in power. Finally, we show that, in a more general framework, the entrenchment policies can be implemented also by the right-wing party. The comparative statics analyzes the effects of state capacity, a positive bias of voters for one party and income inequality on the incentives of the incumbent party to pursue entrenchment policies. The importance of our theory for constitutionally legislated term limits is also discussed. The theory sheds light on why left-wing parties or politicians often support liberal immigration policies 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.

Keywords: Political entrenchment; constituencies; inequality; inefficient redistribution; checks and balances; political rents; state fiscal capacity.
JEL Classification: D72, P16.

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 particular 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 liberal immigration policies supported by left-wing parties in Europe and in the US, and the dysfunctional educational policies implemented by a number of Latin American "populist" governments. 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 are also potentially puzzling. These policies are not necessarily inefficient but still damage the economic interests of a relatively significant part of the constituency of the governments implementing them.

Why, and under what conditions, is it the case that 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 dynamic (two-periods) model addressing this question.

In our framework, individuals have both preferences over an economic issue, which is affected by policy (i.e., income redistribution), 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, 1998). Therefore, each citizen has a specific taste for which party is in power and the distribution of this taste is subject to aggregate shocks, which generates

---

[1] Examples of such policies include the blocking of technological progress (e.g., Krusell and Rios-Rull, 1996) and failing to make cost-effective investments in human capital (e.g., Besley and Coate, 1998), subsidizing declining industries (e.g., Dixit and Londregan, 1995, 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, 2011; Besley and Persson, 2009), protecting unproductive jobs with high firing costs (Saint-Paul, 1993, 2002), creating inefficient social infrastructures (Coate and Morris, 1995; Robinson and Torvik, 2005), or underpricing the shares of privatized companies (Biais and Perotti, 2002).

uncertainty on the outcome of the electoral competition.

We assume that there are only two income classes, the low-skill (or the poor) and the high-skill (or the rich), and two political parties which compete in majoritarian (i.e., winner-takes-all) elections. 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; Osborne and Slivinsky, 1996; Besley and Coate, 1997), but also derives rents from being in power, as in the spirit of the classic Downsian model of political competition. As in models of partisan politics, parties cannot ex-ante commit to implement a policy different from their preferred one once in office. We will call the *Left* and the *Right* the party that cares about the welfare of the poor and the rich respectively.

There are two periods and, in each period, the party in office chooses the degree of income redistribution. Moreover, and this is the main innovation of the model, we assume that the second period pre-tax income of the low-skill agents can be set at a level lower than the potential one 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 as well as relative to the mean income. Nevertheless, we show that, under some conditions, the left-wing party may find it optimal implementing such a policy that damages the members of its own natural constituency.

The rationale for the adoption of this policy is that by reducing the income of the low-skill individuals, income inequality increases and income redistribution becomes more valuable for the poor, so that they will have a higher incentive to vote for the left-wing party. In other words, partisan politicians cannot commit on which redistributive policies will be undertaken once they are in office, and only the Left, given its partisan preferences, promotes income redistribution policies. In other words, if the Right could commit to implement enough redistribution in advance, the damaging policy would not be implemented by the Left.

Hence, the Left, if in power in the first period, may find it optimal implementing policies that will make the poor more "dependent" on income redistribution in the future, so that the economic incentives of the poor to vote for this party increase. In our model, there are aggregate preference shocks for the identity of the party in power. Given the distribution of those shocks, the Left is more likely to win the election, the more the poor have a genuine economic benefit from its policies. This benefit is in turn stronger, the poorer the poor relative to the rich. We will refer to a policy with that effect as *entrenchment* because its implementation allows the party to tie its own natural constituency more to itself.

A number of features of our model are worth emphasizing. First, the existence of an

equilibrium with potential entrenchment does not rely on any form of myopia or irrationality on either side of the political arena. In fact, it may be rational for the poor to vote for the Left in the first period 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 more than compensate the poor for their future income losses generated by the entrenchment policy.[2] Second, our argument generalizes naturally to any policy that increases income inequality regardless of the sign of its impact on the income of 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.[3] Third, even though we present a model where only the left-wing party implements entrenchment policies, we are by no means arguing that entrenchment is a policy exclusive of Left. In fact, we also propose an extension of our baseline model where the entrenchment strategy may be optimal for the Right. Moreover, the concept of political entrenchment is more general and can be applied to different frameworks, such as the provision of other public goods. For example, if we consider the case of national defence and assume that this gives a relatively higher utility to the natural constituency of the right-wing party and that such a party has an advantage in providing it, then it is possible that the Right pursues entrenchment strategies through an unnecessarily aggressive foreign policy.

The comparative static analysis shows that political entrenchment is more likely to occur when the rents appropriated by the party in power are higher, which suggests that we should expect more entrenchment in political systems with relatively limited checks and balances.[4] The effect of a positive bias in favor of the Left, a higher income inequality and a higher level of state capacity have in general an ambiguous effect on entrenchment. However, under some conditions on the distribution of the ideological taste shock or when state capacity is low, we show that a higher level of state capacity increases the incentive of the Left to implement entrenchment strategies. Since a low level of state capacity is characteristic of developing economies, this result also suggests that an autonomous increase in state capacity is particularly likely to have harmful consequences for developing countries.

---

[2] Furthermore, in the second period, the poor's post-redistribution income may still be higher than if the right were in power, despite a lower pre-redistribution income.

[3] 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.

[4] This could be 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, Mainwaring and Shugart, 1997, and Stokes, 2009).

3

We also show that, under some conditions such as the presence of a limited state fiscal capacity, the left-wing party may have an incentive to ex-ante commit not to pursue entrenchment policies once in power. In fact, entrenchment policies increase the probability of the Left's winning future elections, but they reduce the welfare of the citizens from having this party in power and, therefore, its probability of winning current elections.

Our theory also provides a rationale for constitutionally prescribed term limits. In fact, term limits may reduce the adoption of entrenchment policies since they lower 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 helps explaining why many real world constitutions prescribe some form of term limits in spite of the fact that a key prediction of the standard model of political replacement is that term limits are always welfare reducing (e.g., Barro, 1973; Ferejohn, 1986; Persson, Roland and Tabellini, 1997).[5]

This paper is related to the literature on the strategic theory of budget deficits where an incumbent partisan leader may strategically use public debt in order to manipulate the future spending policy or the future allocation of political power (e.g., Persson and Svensson, 1989; Aghion and Bolton, 1990; Alesina and Tabellini, 1990; Tabellini and Alesina, 1990). In a similar spirit, Milesi-Ferretti and Spolaore (1994) and Milesi-Ferretti (1995) demonstrate that incumbent leaders may choose to strategically manipulate some macroeconomic state variables, such as exchange rates or the composition of government spending, in order to increase their future electoral comparative advantage. None of these papers, however, addresses the issue of why partially self-interested governments may deliberately pursue policies harmful for their own natural constituency.

A literature in economics (e.g., Harrington, 1993) and in political science (e.g., Stokes, 2001) has addressed the question of why occasionally elected politicians deviate from their electoral promises, creating a discrepancy between mandates and actual policies.[6] According to these authors, once in office (benevolent) politicians may deviate from their original electoral promises if they believe that such policies do not serve best the interest of their own constituency. Therefore, violations of mandates are actually made in the interest of some political constituencies, and are not necessarily punished ex-post by rational voters. It is worth

---

[5] This is because in such models, repeated elections are sufficient to provide incentives to politicians in power to refrain from appropriating too much 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 in presence of term limits.

[6] A prominent example is the introduction and consolidation of a vast array of neoliberal policies in many Latin American countries in the last few decades of the previous century, which stood in contrast to the electoral promises made by successful politicians to their constituency (see Stokes, 2001, for a broad discussion on this point).

emphasizing that our paper asks a rather different question, namely why self-interested politicians hurt their natural constituencies and yet the latter vote (and reappoint) them in office despite anticipating correctly the equilibrium strategy of politicians.

Our notion of political entrenchment, which is the source of persistence of power in our model, has some similarity with the concept of managerial entrenchment in the theory of corporate finance (e.g., Weisbach, 1988; Shleifer and Vishny, 1989). In this literature, managers may entrench themselves by making manager-specific investments because this reduces the probability of being replaced by raising the cost of such action for the firm and allows them to earn higher wages. Similarly, in our model, political entrenchment makes it more costly for the lower class to vote for the Right, and therefore allows incumbent leftists politicians or parties to appropriate of higher intertemporal office rents by consolidating their power.

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) and Padró-i-Miguel (2007). In the former 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. In fact, this would lead to the emergence of an anti-redistribution majority 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. In our paper instead, the left-wing party implements an inefficient policy (which is very similar to not upgrading the skills of the poor) for self-interest motivation and against the preferences of its constituency. This result highlights the importance of our assumption that political agency is not perfect contrary to what is assumed in Besley and Coate's paper. In Padró-i-Miguel's paper the leaders of competing ethnic groups, once in power, exploit not only the adversary groups but also in part their own group. This is possible because the members of this group 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 political game, which is the key novel idea of our paper.

The paper also relates to the dynamic models of income redistribution, such as Saint-Paul and Verdier (1997), Bénabou (2000, 2002) and Hassler et al. (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, our paper is closely related to the recent work on the persistence of political power and political

institutions (e.g., Acemoglu and Robinson, 2008; Acemoglu et al., 2010b, 2011); yet, the origin of political persistence in our paper is rather different from any of these works.

The rest of the paper is organized as follows. Section 2 presents some evidence for our theory. Section 3 describes the framework. Section 4 characterizes the equilibrium. Section 5 presents the comparative statics analysis. Section 6 discusses under what conditions the Left may want to limit the possibility of pursuing entrenchment policies in the future, while Section 7 proposes an extension of our framework where also the Right may find it optimal implementing policies of entrenchment. Section 8 discusses the importance of our theory for term limits. Section 9 concludes and the Appendix contains some proofs omitted from the main text.

2 Some Evidence

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

1. Immigration policies in Western Europe and in the US

"After World War I, laws were passed severely limiting immigration. Only a trickle of immigrants has been admitted since then... By keeping labor supply down, immigration policy tends to keep wages high." Paul Samuelson (quoted in Borjas, 2003, p. 1335).

We next present some evidence on the pattern of immigration laws passed in recent years by the European Union (EU) and in the US through the 20th century consistent with the main prediction of our theory.

Migration and immigrant integration policies in Europe are increasingly determined at the 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 standard labor demand theory (e.g., Hamermesh, 1993), it is reasonable to expect that, to the extent that immigrants are disproportionately unskilled, unskilled and unemployed workers are more likely to be opposed to immigration than more highly skilled workers or capital owners, since they are more likely to be in competition for jobs with immigrants

than with higher-skilled workers.[7] Consistently with this expectation, Scheve and Slaughter (2001) find robust evidence that in the US less skilled workers prefer anti-immigration policies; Mayda (2006) finds similar results in a cross-section of countries.[8] Therefore, one should also expect that left-wing parties support *conservative* migration policies, in line with the economic interests of a large share 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 relatively little personal accountability for deviating from their constituencies' preferences;[9] in such a context rents from 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.[10]

---

[7] The empirical evidence on the effects of immigration of unskilled workers on the labor market outcomes of their native peers is mixed. Some papers (e.g., LaLonde and Topel, 1991; Card, 2001), based on the computation of the spatial correlation between native wages and the extent of the penetration of immigrants in local labor markets, indicate that immigrants have little or no effect on the employment opportunities of competing native workers. Borjas (2003) argues instead that such spatial correlation does not necessarily uncover a causal effect for several reasons; immigrants may not be randomly assigned to labor markets and natives may respond to immigration by moving elsewhere. Borjas provides alternative evidence which is supportive of the standard textbook view summarized by the quotation of Paul Samuelson reported at the beginning of this Section (see also Borjas and Aydemir, 2007, 2011).

[8] It is worth emphasizing that, for our theory, it is more important what the unskilled workers *believe* about the effects of immigration on their wages and welfare rather than its *real* effects on them.

[9] See for example Persson and Tabellini (2000).

[10] In our model, parties partly internalize the welfare of their constituents. This is because the politicians are (partly) "citizen-candidates" who belong to the same social class as their constituents. In practice, one observes

Some evidence broadly consistent with our main proposition that left-leaning political parties are inclined to support liberal immigration policies potentially harmful to (at least part of) their natural constituency, is also provided by the history of immigration laws in the United States through the 20th century. In particular, several examples of legislative reforms clearly document that the Republican Party has been generally in favor of restricting immigration, whereas the Democratic Party has often assumed the opposite stance on this issue.

The immigration policy in the US became more restrictive in the second decades of the 20th century with Immigration Act of 1924, or Johnson–Reed Act, including the National Origins Act, and Asian Exclusion Act, passed by the 68th Congress.[11] This new legislation limited the annual number of immigrants who could be admitted from any country to 2% of the number of people from that country who were already living in the US in 1890, down from the 3% cap set by the Immigration Restriction Act of 1921, according to the Census of 1890. Congressman Albert Johnson and Senator David Reed, both Republicans, were the two main architects of the reform, and both the House of Representatives and the Senate had a Republican majority.[12]

The Immigration Act of 1965, also known as the Hart-Celler Act, fundamentally reshaped the American immigration policy for the remainder of the twentieth century and beyond. It abolished the national origins system, set up in the Immigration Act of 1924 and modified by the Immigration Act of 1952. While seeming to maintain the principle of numerical restriction, it so increased the categories of persons who could enter "without numerical limitation" as to make its putative numerical caps—170,000 annually for the Eastern Hemisphere with a maximum of 20,000 per nation plus 120,000 annually for the Western Hemisphere with no national limitations—virtually meaningless within a few years.

By changing long-held immigration policies, the act resulted in new immigration from non-European nations which changed the ethnic make-up of the US. Immigration doubled between 1965 and 1970, and doubled again between 1970 and 1990. The most dramatic effect was to shift immigration from Europe to Asia and Central and South America.

Although the percentage of high school dropouts among immigrants has fallen somewhat,

---

that some politicians on the Left comes from the upper-middle class. In this case, 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 there. While this means that office rents play a bigger role, it may also imply that the politicians' own personal tastes affect their voting behavior. This latter mechanism is absent from our model however.

[11] A detailed historical analysis of the US immigration politicies can be found, for example, in Zolberg (2008).

[12] There were only nine dissenting votes in the Senate and a handful of opponents in the House, the most vigorous of whom was the Brooklyn Democrat Representative Emanuel Celler.

the gap between natives and the foreign born has grown significantly, with immigrants more than twice as likely as native-born Americans not to have completed high school. This contributes to a growing pool of blue-collar workers competing for a shrinking number of well-paying jobs.

Finally, the One Hundred First United States Congress, in which both chambers had a Democratic majority, passed in 1990 the Immigration Act of 1990 increased the number of legal immigrants allowed into the United States each year. The law also provided for exceptions to the English testing process required for naturalization set forth by the Naturalization Act of 1906. After it became law, the United States would admit 700,000 new immigrants annually, up from 500,000 before the bill's passage.

2. The North American Free Trade Agreement

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 may be expected to give a relatively high political voice to the lower classes. In fact, 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 the 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 arguably not the only reason (a substantial fraction of Democratic voters are rich and benefit from NAFTA), it may have nonetheless contributed to the overall outcome.

Some interesting features are worth noting. First, the Republicans supported the agreement more than the Democrats, and they are the ones who initiated it; in fact, if only democratic congressmen had voted it, it would not have passed.[13] 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,

---

[13] 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.

even though their reelection probability suffers. Second, it may well be that the President is more 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 that the latter may be more prone to entrenchment than the former.[14]

3. 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.[15] 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), 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 Argentinian minister Juan Llach to improve the quality of the school system and the ability of Argentina to compete more effectively in a 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

---

[14]It is interesting to observe that Barack 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. However, 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).

[15]See also on this topic the classic contribution of Dornbusch and Edwards (1992), which focuses on the harmful consequences of the "populist" macroeconomic policies widely implemented by Latin American governments.

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" (Edwards, 2010, p. 182).

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.[16] For example, the government of Hugo Chávez in Venezuela launched a variety of ambitious projects aimed, in principle, at improving education at various levels.[17] Many observers have criticized all such programs as ineffective and fraught with corruption however. 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 some segments of the constituency of the leader. For example, Hsieh, Ortega, Miguel, and Rodrigúez (2011) 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.[18]

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 theory suggests that they probably had little alternative to the policy choice of populist redistribution appealing for the Left.

---

[16]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).

[17]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).

[18]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 the facts."

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, $t \in \{0, 1\}$, and citizen i has preferences represented by the following expected utility function

$$\mathbf{u}_0^i = \mathbb{E}_0 \sum_{t=0}^{1} \beta^t u_t^i,$$

where u_t^i is the per period utility function of agent i, $\mathbb{E}_0$ is the expected value operator conditional on the information available at date $t = 0$, and $\beta \in (0, 1)$ is the discount factor. The function u_t^i is assumed to be linear in private good consumption, equal to post-tax income, in the level of the public good provided G_t, financed through proportional taxation at rate τ, and in an ideological component ξ_t^i reflecting specific preferences for which political party is in power (that will be described more precisely below). The instantaneous utility can therefore be written as

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

where a_t^i denotes the income of agent i at time t.

We assume that citizens also differ in terms of their productivity and pre-tax income, and that a mass $\lambda \in (0, 1/2)$ are high-skill or rich with pre-tax income equal to a^r, while the remaining fraction of agents $1 - \lambda$ are low-skill or poor with pre-tax income equal to $a^p(\theta_t) = \theta_t a^p$, where $0 < a^p < a^r$.[19] The productivity of high-skill agents is assumed to be constant over time, whereas the productivity of the low-skilled is potentially time-varying and depends on an endogenous state variable $\theta_t \in [\theta^L, 1]$, where $0 \leq \theta^L \leq 1$.[20] The initial value of θ, $\theta_0 > \theta^L$, is exogenously given, while the value of θ_1 is chosen by the government in office at period $t = 0$ at zero cost, so that potential output is always maximized by setting $\theta_1 = 1$, while any $\theta_1 < 1$ represents an inefficient policy choice. Also note that lower levels of θ_1 imply a poorer and a more unequal society since only the unskilled workers experience a productivity loss. The aggregate and average income at time t is equal to

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

---

[19] We are assuming that the poor are more numerous than the rich, which in this simple two-groups model reproduces the skewness of the empirical income distributions usually observed.

[20] As it will be clear, the lower bound θ^L will not play any special role in our analysis and results are unchanged when $\theta^L = 0$. The existence of a lower bound to the income of the poor will allow us to analyze under what conditions the possibility of entrenchment is ex-ante desirable for the left-wing party and when it is not (see on this point the analysis in Section 6).

and the government budget constraint reads

$$G_t \leq \tau_t \bar{a}_t(\theta_t) \equiv \tau_t \left[\lambda a^r + (1 - \lambda) \theta_t a^p \right]. \tag{2}$$

To simplify the analysis, we assume that taxes create no distortions for all $\tau \leq \hat{\tau}$, where $\hat{\tau} \in (0, 1)$ is an exogenous level of "fiscal capacity" of the state, while distortions are prohibitively high for $\tau > \hat{\tau}$.

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, 1998), where the outcome of elections is potentially effected by exogenous ideological shocks.

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

$$\mathbf{v}_0^J = \mathbb{E}_0 \sum_{t=0}^{1} \beta^t v_t^J,$$

with

$$v_t^P = (1 - \tau_t) a_t^p + G_t + \delta_t, \tag{3}$$

and

$$v_t^R = (1 - \tau_t) a_t^r + G_t + \delta_t, \tag{4}$$

where δ_t represents the private benefit from being in power. We assume that $\delta_t = \delta > 0$ if the party is in power, and $\delta_t = 0$ if it is not. Expressions (3) and (4) reflect the assumption that political parties are partially benevolent, i.e., they care about the economic utility of one specific social class (the Left party cares of the poor and the Right of the rich), and partially self-interested, as they care about the rents from being in office.[21] As in the spirit of the partisan models of politics (e.g., Alesina, 1988), we assume that parties cannot commit to implement a policy different from their own ideal one and that individuals vote sincerely, which is a weakly dominant strategy in a two-parties system (e.g., Grossman and Helpman, 2001). We denote with $\varphi_t^i \in \{P, R\}$ the voting decision of citizen i at time t, and the identity of the government in office at time t will be denoted by $\rho_t \in \{P, R\}$.

The ideological component ξ_t^i in the per period utility of the citizens (1) can be represented as follows

$$\xi_t^i = \begin{cases} \varepsilon_t^i + \eta_t, & \text{if } \rho_t = P \\ 0, & \text{if } \rho_t = R. \end{cases} \tag{5}$$

---

[21] A possible microfoundation of this assumption is that politicians are citizen-candidates as in the models of Osborne and Slivinsky (1996) and Besley and Coate (1997), and therefore care about the policy implemented. However, politicians also care to some extent about being in power *per se* because this allows them to appropriate some rents from office (either due to political institutions or psychological factors, such as "ego rents"). See Grossman and Helpman (2001) for a discussion of citizen-candidate models.

The first term on the right hand side of (5) 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, while the bias is normalized to zero when the Right is in power.[22]

In particular, 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 smooth (i.e., differentiable with continuity) 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. 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 assumed to be *i.i.d.* over time. The realizations of η_t are drawn from a continuous distribution function $\Phi(\eta)$, with smooth density $\phi(\eta) \equiv \Phi'(\eta)$ and zero mean. A positive value of η_t reflects the existence of an aggregate bias in favor of party P at time t, whereas a negative value of η_t represents an aggregate bias against such party. While the computation of the political equilibrium does not require any distributional restriction on $\Phi(\eta)$, in the comparative static analysis we will assume that $\phi(\eta)$ is hump-shaped and reaches a maximum at $\eta = 0$, but it is not necessarily symmetric.

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

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.

- (First election). 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.

---

[22] As standard in probabilistic voting models (see, for example, Lindbeck and Weibull, 1987; Dixit and Londregan, 1995, 1998), both shocks reflect the preferences of the agent for the non-economic policies potentially implemented by the party in office at time t. 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 one.

- 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.

- (Second election). 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 dynamic political game.

4 Characterization of the Equilibrium

4.1 Equilibrium in the Last Period ($t = 1$)

Since we have a finite game, we solve the model by backward induction by computing the political equilibrium in the last period, $t = 1$. In particular, we characterize the equilibrium of the subgame played after elections have been held and a government appointed, for every possible history.

If in office at period 1, party $J \in \{P, R\}$ implements the fiscal policy solving the following problem

$$V_1^J (\theta_1 | \rho_1 = J) \quad = \quad \max_{\tau_1, G_1} \left\{ (1 - \tau_1) a^J (\theta_1) + G_1 + \delta \right\}$$
$$s.t. \ (2),$$

where $V_1^J (\theta_1 | \rho_1 = J)$ denotes the maximized utility of party J from being in power, conditionally on θ_1 that has been chosen by the government in office at time $t = 0$. 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

$$V_1^R (\theta_1 | \rho_1 = R) = \max_{\tau_1} \left\{ (1 - \tau_1) a^r + \tau_1 \left[\lambda a^r + (1 - \lambda) \theta_t a^p \right] + \delta \right\}.$$

From $a^r < \theta_t a^p$ follows that this function is decreasing in τ_1 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 fiscal revenues in G. Using Proposition 1, the per period utilities of low-skill and high-skill producers from voting for party R are respectively

$$u_1^{i,p}\left(\eta_1, \varepsilon_1^i, \theta_1 | \varphi_1^i = R\right) = \theta_1 a^p, \tag{6}$$

and

$$u_1^{i,r}\left(\eta_1, \varepsilon_1^i, \theta_1 | \varphi_1^i = R\right) = a^r, \tag{7}$$

where these expressions incorporate the normalization to zero of the ideological bias in favor of party R (see (5)).

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 $G_1^P = \hat{\tau} \bar{a}_1\left(\theta_1\right)$.*

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

$$V_1^P\left(\theta_1 | \rho_1 = P\right) = \max_{\tau_1}\left\{\left(1 - \tau_1\right)\theta_1 a^p + \tau_1\left[\lambda a^r + \left(1 - \lambda\right)\theta_1 a^p\right] + \delta\right\}.$$

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

$$u_1^{i,p}\left(\eta_1, \varepsilon_1^i, \theta_1 | \varphi_1^i = P\right) = \theta_1 a^p + \Delta_1^p\left(\theta_1\right) + \varepsilon_1^i + \eta_1, \tag{8}$$

where

$$\Delta_1^p\left(\theta_1\right) = G_1^P - \hat{\tau}\theta_1 a^p = \hat{\tau}\lambda\left(a^r - \theta_1 a^p\right) > 0, \tag{9}$$

denotes the net value of fiscal redistribution for the poor. Similarly, the per period utility of a high-skill producer from voting for party P reads

$$u_1^{i,r}\left(\eta_1, \varepsilon_1^i, \theta_1 | \varphi_1^i = P\right) = a^r + \Delta_1^r\left(\theta_1\right) + \varepsilon_1^i + \eta_1, \tag{10}$$

where

$$\Delta_1^r\left(\theta_1\right) = G_1^P - \hat{\tau}a^r = -\hat{\tau}\left(1 - \lambda\right)\left(a^r - \theta_1 a^p\right) < 0, \tag{11}$$

16

is the net value of fiscal redistribution for the rich.

Expression (8) 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 net fiscal transfer $\Delta_1^p(\theta_1)$, and the overall (idiosyncratic plus aggregate) ideological bias in favor of the Left. Since the Right provides no fiscal redistribution, $\Delta_1^p(\theta_1)$ also represents the relative net economic gain for the low-skill agents from voting for party P rather than for party R. Expression (10) has a similar interpretation for a high-skill worker, with the difference that the net fiscal redistribution for the rich, $\Delta_1^r(\theta_1)$, is negative.

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

$$\partial \Delta_1^p(\theta_1) / \partial \theta_1 \equiv -\hat{\tau} \lambda a^p < 0.$$

This is because as θ_1 increases both the income of each low-skill worker and the average (and aggregate) income increase; however, the average income increases proportionally less since the income of the high-skilled does not depend on θ_1. It follows that the difference between the average and personal income of the poor, and therefore the net fiscal transfer, decreases with θ_1. As we will see, this result is important because it implies that the Left can strategically increase its comparative politico-economic value (relative to the Right) for the lower class, $\Delta_1^p(\theta_1)$, by reducing the pre-tax income of the unskilled, i.e. of its natural constituency.

In addition, the net gain of the upper class from fiscal redistribution is increasing in θ_1 since the difference between average income and income of the rich is increasing in θ_1, i.e.,

$$\partial \Delta_1^r(\theta_1) / \partial \theta_1 \equiv \hat{\tau}(1-\lambda) a^p > 0.$$

Citizen $i \in \{p, r\}$ votes for party P at time 1, $\varphi_1^i = P$, if $U_1^i(\varphi_1^i = P) \geq U_1^i(\varphi_1^i = R)$, i.e.,

$$U_1^i(\eta_1, \varepsilon_1^i, \theta_1 | \rho_1 = P) \geq U_1^i(\eta_1, \varepsilon_1^i, \theta_1 | \rho_1 = R), \tag{12}$$

namely if the maximized utility of citizen i when party P in office is greater than his maximized utility when party R is in power. Using (6) and (8), condition (12) for the low-skill citizens, $i = p$, can be written as[23]

$$\varepsilon_1^{i,p} \geq -\Delta_1^p(\theta_1) - \eta_1. \tag{13}$$

For high-skill citizens, $i = r$, the substitution of (7) into (12) leads to

$$\varepsilon_1^{i,r} \geq -\Delta_1^r(\theta_1) - \eta_1. \tag{14}$$

---

[23] We denote with $\varepsilon_1^{i,p}$ ($\varepsilon_1^{i,r}$) the value of the idiosyncratic ideological shock of a poor (rich) citizen i for party P at time 1 when this is necessary for clarity.

Since $\Delta_1^p(\theta_1) > 0$ and $\Delta_1^r(\theta_1) < 0$, 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 one 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 (13) and (14), and the distributional assumptions on the ideological shocks, the total number of votes that party P obtains in state (η_1, θ_1) is

$$S^P(\eta_1, \theta_1) = (1 - \lambda)[1 - F(-\Delta_1^p(\theta_1) - \eta_1)] + \lambda[1 - F(-\Delta_1^r(\theta_1) - \eta_1)]. \qquad (15)$$

The first term on the right hand side of (15) 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 workers multiplied by the fraction of them, $[1 - F(-\Delta_1^p(\theta_1) - \eta_1)]$, voting for party P. The second term on the right hand side of (15) represents instead the number of votes that party P obtains from the high-skill citizens and, again, this is given by the size of high-skill agents λ multiplied by the fraction of these, $[1 - F(-\Delta_1^r(\theta_1) - \eta_1)]$, voting for the Left.

Conditionally on the realization of η_1 and θ_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 $S^P(\eta_1, \theta_1) < 1/2$.[24] 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 $\eta > \eta_1^*(\theta_1)$, where the cutoff $\eta_1^*(\theta_1)$ is implicitly defined by the equation

$$S^P(\eta_1, \theta_1) = 1/2.$$

Using expression (15), this condition can be rewritten as

$$(1 - \lambda)F(-\Delta_1^p(\theta_1) - \eta_1^*(\theta_1)) + \lambda F(-\Delta_1^r(\theta_1) - \eta_1^*(\theta_1)) = 1/2. \qquad (16)$$

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

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

---

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

3. If $\eta_1 = \eta_1^*(\theta_1)$, party R and party P both 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*

$$\sigma_1^R(\theta_1) \equiv \Pr\left\{S^P(\eta_1, \theta_1) < 1/2\right\} = \Pr\left\{\eta < \eta_1^*(\theta_1)\right\} = \Phi(\eta_1^*(\theta_1)), \tag{17}$$

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

$$\sigma_1^P(\theta_1) \equiv 1 - \Phi(\eta_1^*(\theta_1)). \tag{18}$$

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 $t = 1$, $\eta_1^*(\theta_1) < 0$ 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 $\eta_1^*(\theta_1)$ is the solution to the equation $\psi(\hat{\tau}(a^r - \theta_1 a^p), \eta_1^*(\theta_1)) = 1/2$. Since $\psi(h, .)$ is decreasing in η and $\psi(\hat{\tau}(a^r - \theta_1 a^p), 0) < 1/2$, it must be that $\eta_1^*(\theta_1) < 0$. ∎

The intuition for this result is simple. The low-skill workers are more numerous than the high-skilled, and as the former are on average more inclined to vote for the Left than for the Right (due to the more convenient fiscal policy), the two parties obtain the same number of votes in equilibrium if there is a sufficiently large aggregate ideological shock in favor of the Right. This is the case when the realization of η falls below some negative threshold.

The following lemma clarifies how the threshold $\eta_1^*(\theta_1)$ depends on θ_1.

Lemma 2 *In the equilibrium of the subgame beginning in period $t = 1$, $\eta_1^{*\prime}(\theta_1) \equiv d\eta_1^*(\theta_1)/d\theta_1 > 0$ for all θ_1.*

Proof. Let us define

$$z_P \equiv -\Delta_1^p(\theta_1) - \eta_1^*(\theta_1), \tag{19}$$

and

$$z_R \equiv -\Delta_1^r(\theta_1) - \eta_1^*(\theta_1). \tag{20}$$

From $\eta_1^*(\theta_1) < 0$, (9), (11), and (16), we have $z_R > 0 > z_P > -z_R$. Differentiating (16) with respect to θ_1, we get

$$\eta_1^{*\prime}(\theta_1) = \lambda(1-\lambda)a^p \hat{\tau} \frac{f(z_P) - f(z_R)}{(1-\lambda)f(z_P) + \lambda f(z_R)}. \tag{21}$$

Since $|z_P| < |z_R|$, from Assumptions 1 and 2 it follows that $f(z_P) > f(z_R)$ and, therefore, $\eta_1^{*\prime}(\theta_1) > 0$. ∎

Intuitively, the threshold $\eta_1^*(\theta_1)$ 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. A greater θ_1 also increases the fraction of rich voting 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 natural constituency in order to consolidate its own future political power.

4.2   Equilibrium in the Initial Period ($t = 0$)

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

$$V_0^J(\theta_0) = \max_{\tau_0, G_0, \theta_1} \left\{ \left[(1-\tau_0)\, a^J(\theta_0) + G_0 \right] + \delta + \beta \mathbb{E}_\eta \left[V_1^J(\eta, \theta_1) \right] \right\},$$
$$\text{s.t. } (2),$$

where $\mathbb{E}_\eta \left[V_1^J(\eta, \theta_1) \right]$ is the expected continuation value (with respect to the realization of η) of party J conditionally on θ_1, defined as

$$\mathbb{E}_\eta \left[V_1^J(\eta, \theta_1) \right] = \int V_1^J(\theta_1 | \rho_1(\eta, \theta_1))\, \phi(\eta)\, d\eta.$$

Using Proposition 3, this value can be rewritten as

$$\mathbb{E}_\eta \left[V_1^J(\eta, \theta_1) \right] = \sigma_1^P(\theta_1)\, V_1^J(\theta_1 | \rho_1 = P) + \sigma_1^R(\theta_1)\, V_1^J(\theta_1 | \rho_1 = R). \tag{22}$$

Expression (22) reflects the expectation that in period $t = 1$ party P will be in office with probability $\sigma_1^P(\theta_1)$, in which case party J obtains utility $V_1^J(\theta_1|\rho_1 = P)$, and that party R is in power with the complementary probability $\sigma_1^R(\theta_1) = 1 - \sigma_1^P(\theta_1)$, in which case party J's utility is $V_1^J(\theta_1|\rho_1 = R)$. The expected continuation value of party J is affected by the uncertainty relative to the exogenous future realization of the aggregate ideological shock, η_1, the endogenous choice of θ_1, and the future allocation of political power $\rho_1 \in \{P, R\}$ that also depends on the two aggregate state variables of the model.[25]

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

$$\theta_1^J = \arg \max_{\theta_1 \in [\theta^L, 1]} \mathbb{E}_\eta \left[V_1^J(\eta, \theta_1) \right].$$

Using Propositions 1 and 2, and Corollary 1, the expected continuation value $\mathbb{E}_\eta \left[V_1^J(\eta, \theta_1) \right]$ of party $J = P$ in period 1 can be written as

$$V_1^P(\theta_1) \equiv \mathbb{E}_\eta \left[V_1^P(\eta, \theta_1) \right] = \theta_1 a^p + [1 - \Phi(\eta_1^*(\theta_1))] [\delta + \Delta_1^p(\theta_1)], \tag{23}$$

and it is the case that

$$\frac{\partial V_1^P(\theta_1)}{\partial \theta_1} = a^p - [1 - \Phi(\eta_1^*(\theta_1))] \hat{\tau} \lambda a^p - \phi(\eta_1^*(\theta_1)) \eta_1^{*\prime}(\theta_1) [\delta + \Delta_1^p(\theta_1)]. \tag{24}$$

The sign of this derivative is generally ambiguous. 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, which is partially internalized by party P. 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.[26] The third term is also negative (note that $\eta_1^{*\prime}(\theta_1) > 0$ from Lemma 2), as it represents the loss of party P due to the reduced chances of winning the elections in period 1 when θ_1 increases. This utility loss includes both the office rent (which the left party P does not get if the Right is in power) and the fiscal transfer potentially benefiting the lower class (which party P internalizes due to its partial altruism).

Similarly, the expected value $\mathbb{E}_\eta \left[V_1^J(\eta, \theta_1) \right]$ of party $J = R$ in period 1 is given by

$$V_1^R(\theta_1) = a^r + \Phi(\eta_1^*(\theta_1))\delta + [1 - \Phi(\eta_1^*(\theta_1))] \Delta_1^r(\theta_1),$$

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

$$\frac{\partial V_1^R(\theta_1)}{\partial \theta_1} = [1 - \Phi(\eta_1^*(\theta_1))]\,\hat{\tau}\,(1 - \lambda)\,a^p + \phi(\eta_1^*(\theta_1))\eta_1^{*\prime}(\theta_1)\,[\delta - \Delta_1^r(\theta_1)] > 0. \qquad (25)$$

Equation (25), unlike (24), contains only positive terms, which means that the objective function of the Right party is strictly increasing in θ_1. A higher value of θ_1 increases the utility of party R directly by increasing both the income of the poor and the aggregate income, which leads to lower taxes for the rich when the Left is in power (first term of (25)), and indirectly by increasing the probability of electoral victory of the right-wing party (second term in (25)).

The following proposition summarizes these results.

Lemma 3 *When in power in the initial period $(t = 0)$, party R always sets the value of θ_1 at level $\theta_1^R = 1$, while the optimal value of θ_1 for party P is the solution to the following maximization problem*

$$\theta_1^* \equiv \theta_1^P = \arg \max_{\theta_1 \in [\theta^L, 1]} V_1^P(\theta_1) = \theta_1 a^p + [1 - \Phi(\eta_1^*(\theta_1))]\,[\delta + \Delta_1^p(\theta_1)]. \qquad (26)$$

Proof. In the text. ∎

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_1 as the equilibrium value when the Left is in power.

Remark 2 Since the function (24) is continuous and bounded in θ_1, there exists a finite value of δ independent of θ_1, that we call δ^*, such that $V_1^P(\theta_1; \delta)$ is everywhere decreasing in θ_1 for any $\delta > \delta^*$.[27] In this case, the political rents are sufficiently large to insure that the program of party P has the lower corner solution $\theta_1^* = \theta^L$.

We now discuss how election results are determined in the initial period $(t = 0)$.

Anticipating the policy vector potentially chosen by each party, a low-skill citizen, $i = p$, votes for party P at time 0, $\varphi_0^{i,p} = P$, if $U_0^{i,p}(\varphi_0^{i,p} = P) \geq U_0^{i,p}(\varphi_0^{i,p} = R)$, i.e.,

$$U_0^{i,p}(\eta_0, \varepsilon_0^{i,p}, \theta_0 | \rho_0 = P) \geq U_0^{i,p}(\eta_0, \varepsilon_0^{i,p}, \theta_0 | \rho_0 = R),$$

---

[27] This follows readily upon inspection of equation (24), which is continuous and bounded in θ_1 for the following reasons. First, η_1^* is locally a C^1 function of θ_1 for all θ_1 by the implicit function theorem; therefore, η_1^* is globally a continuous and bounded function of θ_1 on the compact interval $[\theta^L, 1]$. Second, $\phi(\cdot)$ is a continuous and bounded function of η over its support by assumption. Finally, $\eta_1^{*\prime}(\theta_1)$ is also continuous and bounded, as can be immediately verified by inspecting equation (21) (recall that $f(\cdot)$ is a continuos and bounded function over its support by assumption).

where the left (right) hand side is the expected discounted maximized utility of a poor citizen when the left-wing (right-wing) party is in power. This condition can be rewritten as

$$(1 - \hat{\tau})\theta_0 a^p + \hat{\tau}\left[\lambda a^r + (1 - \lambda)\theta_0 a^p\right] + \varepsilon_0^{i,p} + \eta_0 + \beta U_1^p(\theta_1^P) \geq \theta_0 a^p + \beta U_1^p(1),$$

where $U_1^p(\theta_1^P)$ and $U_1^p(1) \equiv U_1^p(\theta_1^R)$ 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 $t = 0$. From Lemma 3, θ_1 is set to $\theta_1^P \equiv \theta_1^*$ given by (26) in the first case, and to $\theta_1^R = 1$ in the second one.

From the analysis of the voting decision of the poor, that we do not repeat in detail as it is similar to the one presented above for period 1, we obtain that a low-skill citizen votes for the Left ($\varphi_0^{i,p} = P$) if

$$\varepsilon_0^{i,p} \geq -\Delta_0^p(\theta_0) - \eta_0,$$

where

$$\Delta_0^p(\theta_0) = \hat{\tau}\lambda(a^r - \theta_0 a^p) + \beta\left[U_1^p(\theta_1^*) - U_1^p(1)\right]. \tag{27}$$

Similarly, a high-skill citizen votes for the party P at time 0 ($\varphi_0^{i,r} = P$) if

$$\varepsilon_0^{i,r} \geq -\Delta_0^r(\theta_0) - \eta_0,$$

with

$$\Delta_0^r(\theta_0) \equiv -\hat{\tau}(1 - \lambda)(a^r - \theta_0 a^p) + \beta\left[U_1^r(\theta_1^*) - U_1^r(1)\right] < 0. \tag{28}$$

Again, $U_1^r(\theta_1^*)$ and $U_1^r(1)$ 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.[28]

It is worth noting that these conditions differ from those derived for the second period of the game since they are *dynamic* voting rules. In fact, citizens anticipate that their future utility depends on the current elected government which chooses the value of the future productivity of the low-skill workers (i.e., the government at time 0 can "manipulate" the electoral outcome in the next period through the strategic choice of the endogenous state variable, θ_1). This choice affects the utility of the citizens directly and indirectly by changing the probability of being in office by the two parties.

In particular, when the Left is also expected to choose $\theta_1^* = 1$ 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 first and second period coincide and citizens discriminate between the two parties depending only on the present welfare that they promise

---

[28] It is worth noting that while $\Delta_0^r(\theta_0)$ is always negative since $U_1^r(\theta_1^*) < U_1^r(1)$, $\Delta_0^p(\theta_0)$ may have an ambiguous sign as $U_1^p(\theta_1^*)$ may be lower than $U_1^p(1)$.

to the voters. This is an equilibrium *with no* political entrenchment which 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 each party is in power at time 1 does not depend on which party is in power in the first period. When instead the Left is expected to choose $\theta_1^* < 1$ if in power, the continuation value promised by the two parties to the rich and to the poor differs. This is an equilibrium *with* political entrenchment and it is characterized by the fact that the left-wing party deliberately reduces the income of its constituency in order to consolidate its power. This equilibrium has two important features. First, it is Pareto inefficient since potential output is not maximized. Second, there is persistence of political power, in the sense that the probability that party P is power at time 1, $\sigma_1^P(\theta_1^*)$, depends on which party is in power at time 0. In particular,

$$\Pr\{\rho_1 = P|\rho_0 = P\} = \sigma_1^P(\theta_1^*) > \Pr\{\rho_1 = P|\rho_0 = R\} = \sigma_1^P(1)$$

for all $\theta_1^* < 1$, as $\sigma_1^P(\theta_1^*) = 1 - \Phi(\eta_1^*(\theta_1^*))$ and $\eta_1^{*\prime}(\theta_1^*) > 0$.

The number of votes that party P obtains in the first period is

$$S^P(\eta_0, \theta_0) = (1 - \lambda)\left[1 - F(-\Delta_0^p(\theta_0) - \eta_0)\right] + \lambda\left[1 - F(-\Delta_0^r(\theta_0) - \eta_0)\right]. \tag{29}$$

As before, party P wins the election with certainty if $S^P(\eta_0, \theta_0) > 1/2$, which is the case when $\eta > \eta_0^*(\theta_0)$, where $\eta_0^*(\theta_0)$ is defined as the value of η_0 satisfying the condition[29]

$$S^P(\eta_0^*(\theta_0), \theta_0) = 1/2, \tag{30}$$

with $S^P(\cdot)$ defined by (29).

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

2. *If $\eta_0 > \eta_0^*(\theta_0)$, party P wins elections and implements the fiscal policy $(\tau_0^P = \hat{\tau}, G_0^P = \hat{\tau}\bar{a}_0(\theta_0))$. Moreover, party P chooses $\theta_1^P \equiv \theta_1^* \in \left[\theta^L, 1\right]$ as defined in Lemma 3.*

---

[29] Note that η_0^* also depends on θ_1 since $\Delta_0^p(\theta_0)$ and $\Delta_0^r(\theta_0)$ (see respectively (27) and (28)) are affected by this state variable. To simplify the notation, we will use $\eta_0^*(\theta_0)$ rather than $\eta_0^*(\theta_0, \theta_1^*)$ when this does not make confusion.

24

3. If $\eta_0 = \eta_0^*(\theta_0)$, 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 one of period 1 contained in Proposition 3. It is also immediate to verify that the threshold $\eta_0^*(\theta_0)$ 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 value of θ, as well as the identity if the party in power, potentially exhibit a form of *path dependence*. This is because the threshold $\eta_0^*(\theta_0)$, and therefore the probability that the Right is in power in period $t = 0$, increases with θ_0. Hence, if θ_0 is relatively high, then the probability that the Right is in power in period $t = 0$ is also high, and so will be the value of θ_1 as this party chooses $\theta_1 = 1$. In turn, a high value of θ_1 implies a relatively higher probability that the Right is in power at time $t = 1$. Similarly, the lower is θ_0, the higher is the probability that the Left will be in office in both periods.

5   Comparative Statics

We now move to analyze how some parameters of the model change the degree of entrenchment of the Left, namely how they affect the level of productivity of the low-skilled chosen by party P 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 exogenous component of the balance of power between Left and Right, and a change in state capacity $\hat{\tau}$.

As the objective function (23) of the Left 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 it leads party P to choose a (weakly) lower level of θ_1. This result comes from the fact that $V_1^P(\theta_1; \delta)$ is submodular in (θ_1, δ), i.e., $\partial^2 V_1^P(\theta_1; \delta)/\partial\theta_1\partial\delta < 0$, so that $\partial\theta_1^*/\partial\delta \leq 0$ by

Topkis' theorem (Topkis, 1998). In fact, from (24) we obtain that

$$\frac{\partial^2 V_1^P(\theta_1; \delta)}{\partial \theta_1 \partial \delta} = -\phi(\eta_1^*(\theta_1))\eta_1^{*\prime}(\theta_1),$$

which is always negative. By Topkis' theorem, this implies that θ_1^* is nonincreasing in δ (i.e., $\partial \theta_1^*/\partial \delta \leq 0$).[30] 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 Inter-Party Balance of Power

We now consider the effect of an exogenous variation in the balance of power in favor of the Left on θ_1^*. 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).[31] An 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 $V_1^P(\theta_1; \nu)$ is supermodular or submodular in (θ_1, ν). From (24), we have that

$$\begin{aligned}
\frac{\partial^2 V_1^P(\theta_1; \nu)}{\partial \theta_1 \partial \nu} &= \phi(\eta_1^*(\theta_1))\frac{\partial \eta_1^*(\theta_1)}{\partial \nu}\hat{\tau}\lambda a^p - \phi'(\eta_1^*(\theta_1))\frac{\partial \eta_1^*(\theta_1)}{\partial \nu}\eta_1^{*\prime}(\theta_1)\left[\delta + \Delta_1^p(\theta_1)\right] \\
&\quad - \phi(\eta_1^*(\theta_1))\frac{\partial \eta_1^{*\prime}(\theta_1)}{\partial \nu}\left[\delta + \Delta_1^p(\theta_1)\right].
\end{aligned}$$

By differentiating (16), it can be easily shown that $\partial \eta_1^*(\theta_1)/\partial \nu = -1$ and $\partial \eta_1^{*\prime}(\theta_1)/\partial \nu = 0$, so that

$$\frac{\partial^2 V_1^P(\theta_1; \nu)}{\partial \theta_1 \partial \nu} = -\phi(\eta_1^*(\theta_1))\hat{\tau}\lambda a^p + \phi'(\eta_1^*(\theta_1))\eta_1^{*\prime}(\theta_1)\left[\delta + \Delta_1^p(\theta_1)\right]. \tag{31}$$

The first term in (31) is negative, which implies that θ_1^* is nonincreasing in ν (i.e., $\partial \theta_1^*/\partial \nu \leq 0$), and it represents the *appropriation effect*. A higher ν means that citizens have a positive bias in favor of the Left, that will then be in power more often. This implies that, ceteris paribus, the poor will get income redistribution more often and, therefore, that the cost of entrenchment (i.e., of a lower pre-tax income) for party P, that internalizes to some extent the utility of the poor, is smaller.[32] This leads that party to entrench more, namely to choose a lower level of

---

[30] The notation used to write the expression of the value function $V_1^P(\theta_1; \delta)$ reflects that θ_1 is an endogenous state variable whereas δ represents an exogenous parameter of the model.

[31] Note that Assumptions 1 and 2 are no longer valid. However, if ν is not too large, the critical property of the equilibrium 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.

[32] It is useful to remind that income redistribution partially offset the income loss of the poor generated by a lower level of their productivity.

θ_1^*. The second term in (31) is instead positive, which leads to θ_1^* being nondecreasing in ν (i.e., $\partial \theta_1^* / \partial \nu \geq 0$), and it represents the *decreasing marginal political return effect*. A positive bias in favor of party P (higher ν) reduces the local density of marginal swing voters. This is due to the fact that $\partial \eta_1^*(\theta_1) / \partial \nu < 0$ and Assumption 3 which implies that $\phi'(\eta_1^*(\theta_1)) > 0$ as $\eta_1^*(\theta_1) < 0$. Given the existence of fewer citizens at the margin (i.e., $\phi'(\eta_1^*(\theta_1))\eta_1^{*\prime}(\theta_1) > 0$), 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.

As the sign of (31), and therefore of $\partial \theta_1^* / \partial \nu$, is generally ambiguous, so will be the effect of a positive bias in favor of the Left on entrenchment. However, one can observe that the second term in (31) 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 ($\partial^2 V_1^P(\theta_1; \nu) / \partial \theta_1 \partial \nu > 0$ and $\partial \theta_1^* / \partial \nu \geq 0$). The intuition is the following. The appropriation effect comes from the "citizen-candidate" part of the politicians' objective function; as individuals or as agents of a social constituency, 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 the reelection probability. Here it tells us that entrenchment yields less political leverage, the higher ν, because people like the Left more and their voting behavior is therefore less elastic to the choice of θ.[33]

5.3 The Effect of State Fiscal Capacity

We now analyze the effect of a higher state capacity $\hat{\tau}$ on the entrenchment of the Left. Again, from (24), we obtain

$$
\begin{aligned}
\frac{\partial^2 V_1^P(\theta_1; \hat{\tau})}{\partial \theta_1 \partial \hat{\tau}} = & -\phi(\eta_1^*(\theta_1))\eta_1^{*\prime}(\theta_1)\lambda(a^r - \theta_1 a^p) - [1 - \Phi(\eta_1^*(\theta_1))]\lambda a^p \\
& +\hat{\tau}\lambda a^p \phi(\eta_1^*(\theta_1))\frac{\partial \eta_1^*(\theta_1)}{\partial \hat{\tau}} - \phi'(\eta_1^*(\theta_1))\eta_1^{*\prime}(\theta_1)\left[\delta + \Delta_1^p(\theta_1)\right]\frac{\partial \eta_1^*(\theta_1)}{\partial \hat{\tau}} \\
& -\phi(\eta_1^*(\theta_1))\frac{\partial^2 \eta_1^*(\theta_1)}{\partial \theta_1 \partial \hat{\tau}}\left[\delta + \Delta_1^p(\theta_1)\right].
\end{aligned}
\tag{32}
$$

This expression includes two derivatives that we need to compute.

---

[33] 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.

To determine $\partial \eta_1^*(\theta_1)/\partial \hat{\tau}$, we derive the implicit function (16) defining $\eta_1^*(\theta_1)$ with respect to $\hat{\tau}$ and use (19) and (20), which leads to

$$\frac{\partial \eta_1^*(\theta_1)}{\partial \hat{\tau}} = -\lambda(1-\lambda)\left(a^r - \theta_1 a^p\right)\frac{f(z_P) - f(z_R)}{(1-\lambda)f(z_P) + \lambda f(z_R)} < 0,$$

as $f(z_P) > f(z_R)$.

Moreover, differentiating the RHS of (21) with respect to $\hat{\tau}$ we get

$$
\begin{aligned}
\frac{\partial^2 \eta_1^*(\theta_1)}{\partial \theta_1 \partial \hat{\tau}} =\ & \lambda(1-\lambda)a^p \frac{f(z_P) - f(z_R)}{(1-\lambda)f(z_P) + \lambda f(z_R)} \\
& -\lambda(1-\lambda)\hat{\tau}a^p\left(a^r - \theta_1 a^p\right)\frac{\left[(1-\lambda)\,f(z_P)f'(z_R) + \lambda f(z_R)f'(z_P)\right]}{\left[(1-\lambda)f(z_P) + \lambda f(z_R)\right]^2} \\
& -\lambda(1-\lambda)\hat{\tau}a^p \frac{\partial \eta_1^*(\theta_1)}{\partial \hat{\tau}}\frac{f(z_R)f'(z_P) - f(z_P)f'(z_R)}{\left[(1-\lambda)f(z_P) + \lambda f(z_R)\right]^2},
\end{aligned}
\tag{33}
$$

which is generally ambiguous as the first term is positive, the second is ambiguous and the last term is positive since $f'(z_P) > 0$ and $f'(z_R) < 0$ (from $z_R > 0 > z_P$ and Assumption 2).[34]

---

[34] Remember that when θ_1 goes up, the benefits of redistribution fall. Therefore, the poor swing voter dislikes the Left less, while the rich one likes it less. Both swing voters become less extreme. (Remember that the poor swing voter dislikes the Left, while the rich one likes it). As is clear from equation (21), the intensity of this effect is proportional to the extent of redistribution, i.e. to state capacity $\hat{\tau}$. Furthermore, the net effect of these shifts on the net propensity to vote for the Left depends on the difference between the marginal density of poor swing voters $f(z_P)$ and the marginal density of rich voters $f(z_R)$. This difference is positive due to our assumptions about the shape of $f(\cdot)$ and the proportion of poor. Finally, the response of the critical η_1^* is larger, the smaller the average density of swing voters, $(1-\lambda)f(z_P) + \lambda f(z_R)$. This is because the smaller this density, the greater the change in η which is needed at the margin to switch the balance of power by a given number of votes. Since the number of votes for the Left lost due to a higher θ_1 is also proportional to these marginal densities, what matters for the response of η_1^* to θ_1 is the *relative* density of swing voters $f(z_P)/f(z_R)$, as implied by the fraction expression in (21).

Therefore we see that the increase in η_1^*, and therefore the reduction in the Left's probability of winning the next election, is larger, the greater the state capacity and the greater the relative density of swing voters.

How does this response of η_1^* to θ_1 change when state capacity becomes larger? The answer is in the RHS of (33). First there is the direct effect of $\hat{\tau}$, which increases the intensity of preferences for redistribution and therefore the magnitude with which the swing voters move to the center. This effect tends to magnify the positive response of η_1^* to θ_1 (which in itself tends to increase entrenchment). Hence the first term on the RHS of (33).

Second, $\hat{\tau}$ has a direct impact on the ideological preferences of the poor and the rich swing voters, z_P and z_R. As more is redistributed from rich to poor whenever the Left is in power, fewer rich people, and more poor people, vote in its favor. Consequently, the poor swing voter dislikes the Left more, while the rich swing voter likes it more. The marginal density of swing voters falls for both the poor and the rich, and therefore the effect on the relative density is ambiguous. This is captured by the second term in the RHS of (33). However, if it were the case that $|f'(x)/f(x)|$ were nonincreasing with $|x|$, then we would have that $-f'(z_R)/f(z_R) \geq f'(z_P)/f(z_P)$ and this second term would be non negative.

Finally, $\hat{\tau}$ has an indirect effect on the identity of the swing voters through its effect on η_1^*. We know that $\partial \eta_1^*/\partial \hat{\tau} < 0$, meaning that the Left wins the election more often when state capacity is larger, due to the fact that the poor are a majority. Thus, the aggregate ideological shock beyond which the Left wins the election is more unfavorable to the Left, meaning that the corresponding swing voters have an increased own taste for the Left. Therefore, z_P and z_R both go up. This makes the poor swing voter less extreme (i.e. disliking the Left less) and the rich swing voter more extreme (i.e. liking the Left more). This raises $f(z_P)$ and reduces $f(z_R)$. Therefore, the relative density $f(z_P)/f(z_R)$ goes up, which increases the votes lost by the Left when θ_1 goes up, thus increasing $\eta_1^{*'}(\theta_1)$. Hence the third term in the RHS of (33), which is positive.

The first three terms in (32) 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 (i.e., $\partial \theta_1^* / \partial \hat{\tau} \leq 0$). 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, the higher is $\hat{\tau}$ and the higher is the incentive for party P to reduce θ_1 in order to increase the probability of being elected. The second term captures the expected increase in redistribution, which takes place when the Left is in power, generated by a higher $\hat{\tau}$. 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 $\eta_1^*(\theta_1)$ 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 in (32) represent the *marginal political return effect* generated by the increase in state capacity. The fourth term is positive and comes from the fact that the increase in $\hat{\tau}$ reduces $\eta_1^*(\theta_1)$ (the Left is more popular), which in turn leads to a reduction in the mass of citizens at the margin. Hence, a lower level of θ_1 would lead to a smaller increase in the votes gained by the Left. This lowers the scope for entrenchment and leads 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 $\eta_1^{*\prime}(\theta_1)$, and it can be positive or negative. In sum, 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 on entrenchment is not ambiguous.

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 results obtained above.

Lemma 4 *The preferred level of productivity of the low-skill producers θ_1^* from the point of view of party P has the following features.*

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 Left 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 level of state capacity $\hat{\tau}$ has in general an ambiguous effect on θ_1^*.*

29

5.4 Some Parametric Cases

We can obtain sharper results concerning the comparative statics properties of the political equilibrium if the following additional restrictions are imposed on the functional form of the densities of the two ideological shocks.

Condition 1. *The density of the aggregate popularity shock η has the following property:* $\phi'(x)/\phi(x) \leq 1/\left[\delta + \Delta_1^p(\theta^L)\right]$.[35]

Condition 2. *The density of the idiosyncratic popularity shock ε is such that $\left|f'(x)/f(x)\right|$ is weakly increasing in $|x|$.*[36]

If Conditions 1 and 2 hold, it is straightforward to verify the following set of results.

Lemma 5 *The preferred level of productivity of the low-skill producers θ_1^* from the point of view of party P has the following features.*

1. *Same as Point 1 in Lemma 4 regardless on distributional assumption.*

2. *A positive bias ν in favor of the Left increases entrenchment (i.e., $\partial\theta_1^*/\partial\nu \leq 0$).*

3. *A higher level of state capacity $\hat{\tau}$ increases entrenchment (i.e., $\partial\theta_1^*/\partial\hat{\tau} \leq 0$).*

Proof. *Point 2.* Note that $\eta_1^{*\prime}(\theta_1)$ is bounded from above by $\lambda\hat{\tau}a^p$; using this result, and the fact that $\Delta_1^p(\theta_1)$ is decreasing in θ_1 and therefore it is maximized at $\theta_1 = \theta^L$, it is straightforward to verify that $\partial^2 V_1^P(\theta_1;\nu)/\partial\theta_1\partial\nu < 0$ everywhere if Condition 1 is satisfied. *Point 3.* Condition 2 ensures that $-f'(z_R)/f(z_R) \geq f'(z_P)/f(z_P)$ since $|z_P| < |z_R|$, which implies that also the second term in equation (32) is positive, so that $\partial^2\eta_1^*(\theta_1)/\partial\theta_1\partial\hat{\tau} > 0$. Condition 1 guarantees that the algebraic sum of the third and fourth terms in equation (32) is negative, again since $\eta_1^{*\prime}(\theta_1)$ is bounded from above by $\lambda\hat{\tau}a^p$ and $\Delta_1^p(\theta_1)$ is maximized at $\theta_1 = \theta^L$, which implies that $\partial^2 V_1^P(\theta_1;\hat{\tau})/\partial\theta_1\partial\hat{\tau} < 0$ everywhere. ∎

Intuitively, Condition 1 implies that the marginal density of the aggregate shock is not too much "jumpy," which means that relatively few voters can be captured by Left by reducing entrenchment. As a result, the appropriation effect of a greater political bias in favor of the Left, which leads to more entrenchment, dominates over the marginal political return effect, which has the opposite effect, and therefore entrenchment increases in the political equilibrium.

---

[35] Note that this condition is satisfied, in particular, if the aggregate popularity shock has a uniform distribution (and therefore $\phi'(x) = 0$ for any x), as it is often assumed in political economy models based on probabilistic voting.

[36] Note that a number of important continuous distribution has this property, including the Gaussian and the double-exponential distribution.

Point 2 in Lemma 5 is particularly interesting since it implies that less political competition (or, more precisely, greater exogenous insulation of the Left when in office) may lead to more distortions, in the form of higher equilibrium entrenchment. In this respect, our result is consistent with the recent finding of Acemoglu, Golosov and Tsyvinski (2012) that politicians facing a higher "effective discount factor" (which will be the case if they expect to stay in power longer) are likely to implement macroeconomic policies featuring more, rather than less, political distortions. This conclusion contradicts one of the main insights of the previous literature (e.g., Barro, 1973; Ferejohn, 1986; Persson, Roland and Tabellini, 1997), according to which more stability of power should be conducive to better policies and allocations of resources, by alleviating the political agency problem arising between voters and politicians.

Condition 1 also implies that the appropriation effect of greater state capacity, which leads to more entrenchment, dominates over the marginal political return effect, which has the opposite effect, and therefore makes sure that higher state capacity boosts the incentive of Left to invest in entrenchment. See also Footnote 34 for the intuition of the role played by Condition 2 in determining how state capacity affects directly the responsiveness of electoral outcomes to entrenchment, and therefore the overall sign of the political marginal return effect which is in general ambiguous.

Another special case of interest is the one obtaining when state capacity is small enough.

Lemma 6 *If $\hat{\tau}$ is small enough, higher state capacity generates more political entrenchments.*

Proof. As $\hat{\tau} \downarrow 0$, $\eta_1^{*\prime}(\theta_1) \downarrow 0$, and z_P and z_R both tend to zero, which means that also $[f(z_P) - f(z_R)]$ tend to zero. This in turn implies that $\partial \eta_1^*(\theta_1)/\partial \hat{\tau} \uparrow 0$ and $\partial^2 \eta_1^*(\theta_1)/\partial \theta_1 \partial \hat{\tau} \downarrow 0$, so that from (24) follows that

$$\lim_{\hat{\tau} \to 0} \frac{\partial^2 V_1^P(\theta_1; \hat{\tau})}{\partial \theta_1 \partial \hat{\tau}} = -\left[1 - \Phi(\eta_1^*(\theta_1))\right] \lambda a^p < 0.$$

Therefore, if state capacity is limited, by submodularity it follows that higher state capacity weakly increases entrenchment by the Left (i.e., $\partial \theta_1^*/\partial \hat{\tau} \leq 0$). ■

Intuitively, as state capacity tends to zero, only the appropriation effect of higher state capacity on entrenchment survives, whereas the decreasing marginal political return effect vanishes entirely; hence, equilibrium entrenchment is unambiguously higher. In particular, higher state capacity reduces the marginal loss that the low-skill people experience as entrenchment increases (conditionally on the electoral success of the Left) due to their lower productivity, by allowing for more redistribution and, therefore, it increases the extent of equilibrium entrenchment.

Lemma 5 (Point 3) and Lemma 6 are of special interests since they imply that higher state capacity is not necessarily beneficial for the society but that, under some conditions, it may actually turn out to reduce its welfare. Specifically, according to our model, higher state capacity may be socially harmful by giving to the government currently in power more incentives to consolidate its own future power. Note also that since a low level of state capacity is characteristic of developing economies, Lemma 6 suggests that an autonomous increase in state capacity, is particularly likely to have harmful consequences for developing countries. While the conventional wisdom view (e.g., Besley and Persson, 2009) is that higher state capacity is generally socially beneficial, our result is more in line with Acemoglu (2010) who makes the similar point that an autonomous increment in state capacity (not accompanied by an appropriate change in political institutions) may be socially harmful.

6   The Incentive of the Left to Limit Entrenchment Policies

In our analysis we have assumed that θ_1 has an exogenous lower bound at $\theta^L > 0$. When this constraint is binding, i.e. $\theta_1^* = \theta^L$, a lower value of θ^L means that the Left potentially has a greater capacity to manipulate the economy for electoral purposes if it is in power in the initial period; remember that the higher the reduction of income of the poor, the higher the probability that the Left wins the elections in the following period. Nevertheless, lower values of θ^L may also turn against the left-wing party because the low-skill citizens have lower incentives to vote for this party in the first period since entrenchment reduces their future income. Therefore, if the Left could freely set the value of θ^L before the first elections are held, it would not necessarily choose a low value or the minimum possible one ($\theta^L = 0$). In other words, there can be situations where it is optimal for the left-wing party to commit not to reduce the welfare of its own constituency beyond a certain point in the future in order not to become too much unpopular at the first electoral round. In this Section we analyze when this is the case.

To make this point, we consider the case where θ^L is endogenous and chosen by the Left before the first elections are held, and analyze under what conditions this party prefers to constrain itself by choosing relatively high values of θ^L. In order to simplify the analysis and make it more transparent, we focus on the case where the left-wing party prefers committing to not undertaking entrenchment policies in the future, which corresponds to choosing a value of θ^L equal to 1.

The maximized expected utility of the Left at $t = 0$ is given by

$$V_0^P(\theta_0, \theta_1^*) = [1 - \Phi(\eta_0^*(\theta_0, \theta_1^*))] \left[\theta_0 a^p + \hat{\tau}\lambda(a^r - \theta_0 a^p) + \delta + \beta V_1^P(\theta_1^*)\right] \tag{34}$$
$$+ \Phi(\eta_0^*(\theta_0, \theta_1^*)) \left[\theta_0 a^p + \beta V_1^P(1)\right],$$

where we have emphasized that η_0^* defined by (30) also depends from the optimal value of the θ_1 chosen by the party in power in the first period and reported in Lemma 3. The first term of (34) represents the expected utility that the Left obtains if it wins the first round of elections; this event occurs with probability $1 - \Phi(\eta_0^*(\theta_0, \theta_1^*))$, θ_1 is set at the level $\theta_1^* \in [\theta^L, 1]$ and $V_1^P(\theta_1^*)$ is the expected utility of the Left at $t = 1$ as defined by (23). The second term of (34) is the expected utility of this party if the Right wins the first round of elections, which happens with probability $\Phi(\eta_0^*(\theta_0, \theta_1^*))$; this party sets $\theta_1 = 1$ and $V_1^P(1)$ is the expected utility of the Left at $t = 1$ as defined by (23) with $\theta_1 = 1$.

We consider the case where the constraint $\theta_1^* \geq \theta^L$ is binding, so that the Left sets $\theta_1^* = \theta^L$ when in power in the first period. A sufficient condition ensuring this is that $\partial V_1^P(\theta_1)/\partial\theta_1 < 0$ for all $\theta_1 \in [\theta^L, 1]$; Remark 2 clarifies that this is always the case when the office rents are sufficiently high ($\delta \geq \delta^*$). We then analyze under what conditions the Left prefers setting $\theta^L = 1$ before elections are held in the first period, so giving up the possibility of pursuing entrenchment policies once in power. A sufficient condition for this result is that $V_0^P(\theta^L)$ is monotonically increasing in θ^L for all $\theta^L \in [0, 1]$.

If $\theta_1^* = \theta^L$, then the value in (34) becomes

$$V_0^P(\theta_0, \theta^L) = [1 - \Phi(\eta_0^*(\theta_0, \theta^L))] \left[\theta_0 a^p + \hat{\tau}\lambda(a^r - \theta_0 a^p) + \delta + \beta V_1^P(\theta^L)\right]$$
$$+ \Phi(\eta_0^*(\theta_0, \theta^L)) \left[\theta_0 a^p + \beta V_1^P(1)\right],$$

which implies that

$$\frac{\partial V_0^P(\theta^L)}{\partial\theta^L} = [1 - \Phi(\eta_0^*(\theta^L))] \beta \frac{\partial V_1^P(\theta^L)}{\partial\theta^L} \tag{35}$$
$$- \phi(\eta_0^*(\theta^L)) \frac{\partial\eta_0^*(\theta^L)}{\partial\theta^L} \left\{\hat{\tau}\lambda(a^r - \theta_0 a^p) + \delta + \beta\left[V_1^P(\theta^L) - V_1^P(1)\right]\right\},$$

where to simplify the notation we have omitted the dependency of η_0^* and V_0^P from θ_0. The first component of (35) is negative since $\partial V_1^P(\theta_1)/\partial\theta_1 < 0$ for all $\theta_1 \in [\theta^L, 1]$, and it represents the reduction of the expected utility of the Left that an increase in θ^L induces through the reduction of the probability of winning the elections at time $t = 1$; a higher θ^L lowers the future utility of the party because it reduces its probability to be in power in the future by limiting the size of the entrenchment policies. This effect leads the party to choose low values

of θ^L. The second component represents the variation in the expected utility of the Left that an increase in θ^L generates through the change in the probability of winning the elections at time $t = 0$; this change has the same sign as $\partial \eta_0^*(\theta^L)/\partial \theta^L$, since $V_1^P(\theta^L) > V_1^P(1)$ from the fact that $\partial V_1^P(\theta_1)/\partial \theta_1 < 0$ for all $\theta_1 \in \left[\theta^L, 1\right]$. The term $-\phi(\eta_0^*(\theta^L))(\partial \eta_0^*(\theta^L)/\partial \theta^L)$ represents the variation in the probability that the Left wins the elections at time $t = 0$ when θ^L increases, while the expression in brace brackets is the expected utility gain of this party from being in power in the initial period.

If $\partial \eta_0^*(\theta^L)/\partial \theta^L$ is nonnegative for all $\theta^L \in [0, 1]$, then $\partial V_0^P(\theta^L)/\partial \theta^L$ is always negative and the left-wing party prefers the lowest possible value of θ^L also before the first elections are held. In the Appendix, we show that $\partial \eta_0^*(\theta^L)/\partial \theta^L < 0$ when state capacity is sufficiently small. This implies that $-\phi(\eta_0^*(\theta^L))(\partial \eta_0^*(\theta^L)/\partial \theta^L) > 0$, so that lower values of θ^L reduce the probability that the Left wins the elections in the first period as citizens anticipate the higher losses from the entrenchment policies implemented by this party when in power. When this effect is sufficiently strong, then it is possible that $\partial V_0^P(\theta^L)/\partial \theta^L > 0$ and that the left-wing party prefers choosing a high value of θ^L, so limiting its future possibilities of adopting entrenchment policies.

While the sign of $\partial V_0^P(\theta^L)/\partial \theta^L$ is in general ambiguous, we can prove the following result.

Lemma 7 *If $\hat{\tau}$ is small enough and δ is higher than a certain threshold (i.e., $\delta > \delta^{**} \equiv \max\{\delta^*, \tilde{\delta}\}$, where $\tilde{\delta} = \beta a^p$ and δ^* is defined in Remark 2), the Left prefers setting $\theta^L = 1$ (i.e., $\partial V_0^P(\theta^L)/\partial \theta^L > 0$ for all $\theta^L \in [0, 1]$), which means that party P would ex-ante prefer to give up the possibility of implementing entrenchment policies once in power.*

Proof. See Appendix. ∎

The intuition for the result in Lemma 7 is the following. Under some conditions, the entrenchment policies may be so harmful for the low-skill individuals that they have strong incentives to voting for the Right in order to avoid them. The consequent reduction in the probability of winning the elections in the first period may be so high for the Left that it may prefer to committing not to implement such policies once in power, which means setting $\theta^L = 1$. This is likely to be case when state capacity is low, as in this case the low-skill agents have less to gain from the Left in power and are therefore more inclined to voting for the Right if they anticipate that the left-wing party will implement entrenchment policies.

7 Extension: The Possibility of Entrenchment by the Right

In this section, we extend the baseline version of the model by assuming that the policy variable θ also affects the income of the rich, and show that also the Right might pursue entrenchment policies damaging their own constituency in order to obtain electoral gains. In particular, we now assume that the pre-tax income of each high-skill agent is

$$a^r(\theta; \gamma) = (1 + \gamma\theta)a^r,$$

where γ can be positive or negative. To make sure that the rich have a pre-tax income higher than the poor for any level of θ, we make the following assumption.

Assumption 4 $\gamma > (a^p/a^r) - 1 \equiv \underline{\gamma}$.

The aggregate and average level of income can now be written as

$$\bar{a}(\theta; \gamma) \equiv \lambda a^r + [\lambda\gamma a^r + (1 - \lambda)a^p]\theta.$$

Note that when $\gamma > 0$ the income of both groups, the high-skilled and the low-skilled, is increasing in θ. Therefore, a reduction of θ harms both groups, thus generating an inefficient outcome. Conversely, when $\gamma < 0$ the pre-tax income of the rich is decreasing in θ and a reduction of θ harms the poor and benefit the rich, but it may increase or decrease aggregate output; this case may capture those policies that damage directly the unskilled but benefit directly the skilled, such as permitting the immigration of unskilled labor, which corresponds to a reduction of θ when $\gamma < 0$. The following Corollary summarizes these points.

Corollary 2 *A change in θ has the following effects on individual and aggregate productivity.*

1. *If $\gamma > 0$, a fall in θ harms both groups, and reduces aggregate output; hence, the socially efficient value of θ is 1.*

2. *If $\gamma < 0$, a fall in θ harms the poor and benefits the rich. Moreover,*

2a. *if $-(1 - \lambda)a^p/\lambda a^r < \gamma < 0$, a fall in θ leads to aggregate productivity losses; hence, as in the baseline model, the socially efficient value of θ is 1;*

2b. *if $\gamma < -(1 - \lambda)a^p/\lambda a^r$, a fall in θ leads to aggregate productivity gains; hence, the socially efficient value of θ is θ^L.[37]*

---

[37] This last range of parameters in non-empty if and only if $a^p/a^r < \lambda$, which is the case when $\underline{\gamma} < -(1 - \lambda)a^p/\lambda a^r$.

The analysis of the equilibrium of the subgame played at time $t = 1$ is basically unchanged and will not be repeated in detail. In particular, since the value of θ_1 has been already determined at this point, the only action happening in the second stage of the game is the counting of votes. Again, there will exist a threshold value $\eta_1^*(\theta_1)$ such that the Left wins when $\eta > \eta_1^*(\theta_1)$, and vice versa.

To understand the incentives of the two parties to manipulate strategically θ_1 in this new environment, it is useful to look at net value of fiscal redistribution for the two income groups, which is the endogenous determinant of the period 1 voting rules (see expression (13) and (14)). The net value of the fiscal transfer obtained by the poor and the rich when the Left is in power is respectively given by

$$\Delta_1^p(\theta_1; \gamma) = \hat{\tau}\lambda\left[(1 + \gamma\theta_1)a^r - \theta_1 a^p\right],\tag{36}$$

and

$$\Delta_1^r(\theta_1; \gamma) = -\hat{\tau}(1 - \lambda)\left[(1 + \gamma\theta_1)a^r - \theta_1 a^p\right].\tag{37}$$

Differentiating these expressions with respect to θ_1, we obtain that

$$\partial\Delta_1^p(\theta_1; \gamma)/\partial\theta_1 = \hat{\tau}\lambda(\gamma a^r - a^p),$$

and

$$\partial\Delta_1^r(\theta_1; \gamma)/\partial\theta_1 = -\hat{\tau}(1 - \lambda)(\gamma a^r - a^p).$$

The main difference with the baseline model is that $\Delta_1^p(\theta_1; \gamma)$ and $\Delta_1^r(\theta_1; \gamma)$ may now increase or decrease with θ_1 depending on the value of γ. As we will see, there exists now conditions under which the Right may have the incentive to entrench itself, i.e., to reduce the future income of its constituency in order to consolidate its political power. Moreover, if the political rents δ appropriated by incumbent politicians are sufficiently high, social welfare (measured by aggregate output) will not necessarily be reduced by entrenchment policies and may in some cases even increase, in contrast to the baseline model where higher levels of δ never raise aggregate output.[38]

It is useful to remember that the value of θ_1 chosen by the party in office in the first period is affected by two forces. As parties are partially benevolent, they tend to choose the value of θ_1 that maximizes the income of their own natural constituency. However, as they are partially selfish, they also take into account the effect of θ_1 on their probability of winning the elections

---

[38] We recall that, in the baseline model, higher political rents have either no effect on the choice of θ_1, and therefore on social welfare, if the Right is in office, or they are potentially welfare-reducing when the Left is in power since they make entrenchment more valuable.

in the following period. In particular, from (17) follows that the effect of θ_1 on the probability of the Right of winning the elections at time 1 is

$$\frac{\partial \sigma_1^R(\theta_1)}{\partial \theta_1} = \phi(\eta_1^*(\theta_1))\eta_1^{*\prime}(\theta_1), \tag{38}$$

which has the same sign of $\eta_1^{*\prime}(\theta_1)$. By differentiating equation (16) with respect to θ_1, we obtain that

$$\eta_1^{*\prime}(\theta_1) = \lambda(1 - \lambda)\hat{\tau} \, (a^p - \gamma a^r) \, \frac{f(z_P) - f(z_R)}{(1 - \lambda)f(z_P) + \lambda f(z_R)}, \tag{39}$$

where z_P and z_R are defined as in Lemma 2 and where $\Delta_1^p(\theta_1; \gamma)$ and $\Delta_1^r(\theta_1; \gamma)$ in (36) and (37) replace $\Delta_1^p(\theta_1)$ and $\Delta_1^r(\theta_1)$. From (39) it is immediate that the sign of $\eta_1^{*\prime}(\theta_1)$ is the same as of $(a^p - \gamma a^r)$. Therefore, if $a^p - \gamma a^r > 0$, then $\eta_1^{*\prime}(\theta_1) > 0$, and higher levels of θ_1 increase the probability of the Right of winning the elections at time 1, and vice versa.[39]

The optimal level of θ_1 for the Right can be derived from the value of party R at $t = 1$

$$V_1^R(\theta_1; \delta) = (1 + \gamma\theta_1)a^r + \Phi(\eta_1^*(\theta_1))\delta + [1 - \Phi(\eta_1^*(\theta_1))] \, \Delta_1^r(\theta_1; \gamma),$$

which implies that

$$\frac{\partial V_1^R(\theta_1; \delta)}{\partial \theta_1} = \gamma a^r + [1 - \Phi(\eta_1^*(\theta_1))] \, \hat{\tau} \, (1 - \lambda) \, a^p + \phi(\eta_1^*(\theta_1))\eta_1^{*\prime}(\theta_1) \, [\delta - \Delta_1^r(\theta_1; \gamma)]. \tag{40}$$

Finally notice that the expressions in (23) and (24) to determine the optimal level of θ_1 for the Left are still valid with the caveat that the net transfer is now given by (36).

7.1   Case 1a. $\gamma > a^p/a^r$

In this case, a higher level of θ also increases the income of the high-skill individuals, so increasing aggregate output. Therefore, both parties should choose $\theta_1 = 1$ to maximize the income of their own constituencies. However, from $\eta_1^{*\prime}(\theta_1) < 0$ (as $a^p - \gamma a^r < 0$) it follows that the Right can increase the probability of winning the elections at time 1 by choosing values of θ_1 lower than 1. In other words, the entrenchment strategy may be optimal for the Right, while the Left will now choose the efficient level of $\theta_1 = 1$ as this also maximizes its probability of winning future elections.

In particular, as the first two components of (40) are positive, $\Delta_1^r(\theta_1; \gamma) < 0$, $\eta_1^{*\prime}(\theta_1) < 0$, and $\eta_1^*(\theta_1)$ is independent on δ, then $\partial V_1^R(\theta_1; \delta)/\partial \theta_1$ is linearly increasing in δ and positive at $\delta = 0$. This means that $V_1^R(\theta_1; \delta)$ is monotonically increasing in θ_1 for all $\theta_1 \in [\theta^L, 1]$ if δ is lower than a certain threshold $\hat{\delta}$. This implies that there is no entrenchment by the Right,

---

[39] The opposite result holds for the probability of the Left (18) of winning the elections at time $t = 1$.

which will choose $\theta_1 = 1$, when the rents from office are relatively low. However, there exists a level of rents, $\bar{\delta}$, such that $\partial V_1^R(\theta_1; \delta)/\partial \theta_1 < 0$ for all $\theta_1 \in [\theta^L, 1]$ when $\delta \geq \bar{\delta}$. This means that $V_1^R(\theta_1; \delta)$ is monotonically decreasing in θ_1 and party R will choose $\theta_1^R = \theta^L$. In other words, high levels of rents may lead the Right to choose the entrenchment strategy, i.e., lower values of θ in order to increase its probability of winning future elections at the expense of its constituency. The entrenchment strategy may also be optimal for intermediate values of δ. In this case, the third term in (40) is negative but not too high, and the optimal value of θ_1 for party R could be interior, i.e., $\theta_1^R \in (\theta^L, 1)$. Finally notice that the negative effect of δ on θ_1^R can also be shown from the fact that

$$\frac{\partial^2 V_1^R(\theta_1; \delta)}{\partial \theta_1 \partial \delta} = \phi(\eta_1^*(\theta_1))\eta_1^{*\prime}(\theta_1) < 0, \tag{41}$$

meaning that the objective function of the Right is submodular in θ_1 and δ, and therefore that $\partial \theta_1^R/\partial \delta \leq 0$.

7.2   Case 1b. $0 < \gamma < a^p/a^r$

Again, as $\gamma > 0$, higher levels of θ increase the income of the high-skill citizens and aggregate output, so that $\theta_1 = 1$ maximizes the income of both constituencies. However, as $a^p - \gamma a^r > 0$, then $\eta_1^{*\prime}(\theta_1) > 0$ and the situation is similar to the baseline model when only the Left has the incentive to choose inefficient values of θ_1 in order to increase its electoral advantage in the following period.

7.3   Case 2. $\gamma < 0$

When γ is negative, $\theta_1 = \theta^L$ is the level that maximizes the income of the high-skill individuals. Moreover, $\gamma < 0$ implies that $a^p - \gamma a^r > 0$ and $\eta_1^{*\prime}(\theta_1) > 0$, so that the probability of the Right of winning the elections at time 1, $\sigma_1^R(\theta_1)$, is increasing in θ_1 (see (38)). This means that the Right may have the incentive to choose high levels of θ_1, and in the limit case $\theta_1 = 1$, even though this reduces the income of the rich. The expression in (40) contains both positive and negative components. However, when the rents from office δ are sufficiently high, the third term in (40) is likely to be so high and positive that $\partial V_1^R(\theta_1; \delta)/\partial \theta_1 > 0$ for all $\theta_1 \in [\theta^L, 1]$. This means that $V_1^R(\theta_1; \delta)$ is monotonically increasing in θ_1 and the Right will find optimal the entrenchment strategy by choosing $\theta_1^R = 1$. In this respect, it useful to observe that $V_1^R(\theta_1; \delta)$ is supermodular in θ_1 and δ as from (41) follows that $\partial^2 V_1^R(\theta_1; \delta)/\partial \theta_1 \partial \delta > 0$, since $\eta_1^{*\prime}(\theta_1) > 0$, and therefore that θ_1^R is nondecreasing in δ (i.e., $\partial \theta_1^R/\partial \delta \geq 0$). This is another case where

Right has the incentive to entrench itself by choosing a policy that maximizes its benefits at the expense of its constituency.

As in the baseline model, $\theta_1 = 1$ is the level of θ that maximizes the income of the low-skill producers and the Left has still the incentive to choose low levels of θ as this increases its chances of winning future elections and getting office rents.

Finally notice that we can distinguish two cases when $\gamma < 0$, depending on whether a fall in θ_1 raises or lower aggregate output. Hence, when $-(1-\lambda)a^p/\lambda a^r < \gamma < 0$ (Case 2a in Corollary 2), a reduction of θ_1 raises the productivity of the rich but it is socially inefficient. Therefore, a "selfish" right-wing party that cares mainly of office rents (i.e., when δ is large) may choose for electoral reasons a high level of θ_1 that harms its own constituency but, nonetheless, it pursues the efficient policy. Conversely, when $\gamma < -(1-\lambda)a^p/\lambda a^r$ (Case 2b in Corollary 2), a fall in θ_1 leads to aggregate productivity gains. In this case, the entrenchment by the Right, i.e. high levels of θ_1, is socially inefficient, while the entrenchment by the Left, which chooses low values of θ_1 is socially efficient. In other words, we obtain the somewhat paradoxical result that greater political rents may now be potentially welfare improving, in the sense that they induce the Left to entrench itself more and raise aggregate output. This case is particularly interesting because a reduction of θ_1 corresponds to those policies that benefit the rich, increase aggregate output, but have a direct negative effect on the income of the unskilled, such as a trade liberalization in a capital-rich economy or liberal immigration policies.

The following lemma summarizes the main results of this section.

Lemma 8 *If the level of rents from office δ is sufficiently high, then party R may choose the entrenchment strategy when $\gamma > a^p/a^r$ or $\gamma < 0$. In the former case, the income of the rich is maximized at $\theta_1 = 1$, the Right chooses $\theta_1^R < 1$ and entrenchment always produces aggregate productivity losses. In the latter case, the income of the rich is maximized at $\theta_1 = \theta^L$, the Right chooses $\theta_1^R > \theta^L$ and entrenchment produces aggregate productivity gains when $-(1-\lambda)a^p/\lambda a^r < \gamma < 0$. When $\gamma < -(1-\lambda)a^p/\lambda a^r$, the entrenchment strategy by the Left, which chooses $\theta_1^P < 1$ is socially efficient.*

8 The Importance of the Theory for Term Limits

According to the standard models of Barro (1973) and Ferejohn (1986) elections allow the citizens to (partially) solve the potential moral hazard problem of incumbent politicians, who can be induced not to appropriate for themselves too much resources by the implicit threat of non being reelected 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 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 is that a term limit may be beneficial for society (i.e., the sum of rich and poor citizens) as it reduces the scope for political entrenchment.[40]

Let us now assume that the 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 *never* in office due to the presence of a term limit) is equal to a fraction $\epsilon \in [0,1]$ of the rents δ potentially appropriated by their party at that point. 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 may 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 follows

$$\mathbf{v}_0^{J,TL} = \mathbb{E}_0 \sum_{t=0}^{1} \beta^t v_t^{J,TL},$$

with

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

for $t \in \{0, 1\}$.

The analysis of the equilibrium in presence of terms limits is straightforward since the

---

[40] We do not provide here a complete positive theory of term limits (which is beyond the purpose of this paper) since in our framework term limits will have only beneficial effects for society (i.e., they reduce the scope for entrenchment at zero cost). 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.

objective function of party J at time $t = 0$ is the same as in the baseline model, while it contains the term $\epsilon\delta$ rather than δ at time $t = 1$. Therefore, 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 (i.e., higher ϵ) 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 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, as in the case were there are 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.

9   Conclusions

This paper has posed a simple but, to the 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 proposed a simple dynamic model of voting with redistributive fiscal policy in a two-parties system. The Left party alone can credibly be expected 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 in the future. 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 these policies make the natural constituency of a left-wing party endogenously more dependent on it and, therefore, increase the support 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. Such an equilibrium is not based on

any form of myopia or irrationality of politicians or voters and 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 (or in presidential governments with limited checks and balances).

Our comparative statics analysis has clarified how some features, such as state capacity, the bias in favor of a party and income inequality affect the incentive of parties to pursue entrenchment policies. We have discussed the case where the left-wing party may have the incentive to ex-ante commit to not pursue entrenchment policies once in power. And we have shown that, in a more general framework, also the right-wing party may adopt policies damaging its own constituency in order to increase its electoral advantage and that entrenchment policies are not necessarily inefficient from a social perspective. Moreover, it is important to remark that the scope for political entrenchment is much wider than the one suggested by our simple model focusing only on the politics of pure income redistribution and can be applied to other scenarios, involving the provision of public goods such as, for example, national defence.

An interesting application of the theory in the area of the positive analysis of institutional design is the rationale for the existence of term limits in democratic constitutions. Our analysis suggests that term limits are potentially beneficial for society to the extent that they reduce the incentives of incumbent politicians in promoting socially inefficient entrenchment policies by reducing their political time-horizon. Nevertheless, the benefits from term limits may be related to other features of the political system, such as the degree of party discipline. The endogenous determination of the politics of entrenchment, party discipline and term limits in a dynamic political setup seems to be a potentially interesting topic for future research.

10 Appendix

10.1 The Effect of Income Inequality

In this appendix, we analyze the effect of income inequality on entrenchment. We rewrite

$$a^r = \frac{\mu}{\lambda}a \quad \text{and} \quad a^p = \frac{1-\mu}{1-\lambda}a, \tag{42}$$

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

Using (24), we obtain that

$$
\frac{\partial^2 V_1^P(\theta_1;\mu)}{\partial\theta_1\partial\mu} = -\frac{a}{1-\lambda}\left[1-(1-\Phi(\eta_1^*(\theta_1)))\hat{\tau}\lambda\right] \tag{43}
$$

$$
-\phi(\eta_1^*(\theta_1))\eta_1^{*\prime}(\theta_1)\left[1+\hat{\tau}\lambda\left(\frac{a}{\lambda}+\frac{\theta_1 a}{1-\lambda}\right)\right]+\phi(\eta_1^*(\theta_1))\frac{\partial\eta_1^*(\theta_1)}{\partial\mu}\hat{\tau}\lambda a^p
$$

$$
-\phi'(\eta_1^*(\theta_1))\frac{\partial\eta_1^*(\theta_1)}{\partial\mu}\eta_1^{*\prime}(\theta_1)\left[\delta+\Delta_1^p(\theta_1)\right]-\phi(\eta_1^*(\theta_1))\frac{\partial\eta_1^{*\prime}(\theta_1)}{\partial\mu}\left[\delta+\Delta_1^p(\theta_1)\right].
$$

From the implicit differentiation of (16) with respect to μ and (42) we have that

$$
\frac{\partial\eta_1^*(\theta_1)}{\partial\mu} = -\hat{\tau}\lambda(1-\lambda)\left(\frac{a}{\lambda}+\frac{\theta_1 a}{1-\lambda}\right)\frac{f(z_P)-f(z_R)}{(1-\lambda)f(z_P)+\lambda f(z_R)} < 0,
$$

and the derivative of $\eta_1^{*\prime}(\theta_1)$ in (21) with respect to μ leads to

$$
\frac{\partial\eta_1^{*\prime}(\theta_1)}{\partial\mu} = -\lambda\hat{\tau}a\frac{f(z_P)-f(z_R)}{(1-\lambda)f(z_P)+\lambda f(z_R)}
$$

$$
-\lambda(1-\lambda)\hat{\tau}^2 a^p\left(\frac{a}{\lambda}+\frac{\theta_1 a}{1-\lambda}\right)\frac{[\lambda f(z_R)f'(z_P)-(1-\lambda)f(z_P)f'(z_R)]}{[(1-\lambda)f(z_P)+\lambda f(z_R)]^2}
$$

$$
-\lambda(1-\lambda)\hat{\tau}a^p\frac{\partial\eta_1^*(\theta_1)}{\partial\mu}\frac{f(z_R)f'(z_P)-f(z_P)f'(z_R)}{[(1-\lambda)f(z_P)+\lambda f(z_R)]^2},
$$

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

From (43) follows that the effect of higher income inequality on entrenchment is ambiguous and it is very similar to the effect of state capacity. In fact, the first three terms in (43) are negative and they represent the *appropriation effect* generated by higher inequality; these components push towards a reduction of θ_1^* as income inequality 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 comes 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 since this is case for $\partial\eta_1^{*\prime}(\theta_1)/\partial\mu$. 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.

10.2 Proof of Lemma 7

We first show that $\partial\eta_0^*(\theta^L)/\partial\theta^L < 0$ when state capacity is sufficiently small. From the implicit differentiation of equation (30) defining $\eta_0^*(\theta_0,\theta^L)$ when $\theta_1^*=\theta^L$, we obtain

$$
\frac{\partial\eta_0^*(\theta^L)}{\partial\theta^L} = -\frac{(1-\lambda)f(\hat{z}_P)(\partial\Delta_0^p(\theta_0,\theta^L)/\partial\theta^L)+\lambda f(\hat{z}_R)(\partial\Delta_0^r(\theta_0,\theta^L)/\partial\theta^L)}{(1-\lambda)f(\hat{z}_P)+\lambda f(\hat{z}_R)},
$$

43

where $\hat{z}_P \equiv -\Delta_0^p(\theta_0, \theta^L) - \eta_0^*(\theta_0, \theta^L)$ and $\hat{z}_R \equiv -\Delta_0^r(\theta_0, \theta^L) - \eta_0^*(\theta_0, \theta^L)$. The differentiation of $\Delta_0^p(\theta_0, \theta^L)$ and $\Delta_0^r(\theta_0, \theta^L)$, defined respectively in (27) and (28), with respect to θ^L leads to

$$\frac{\partial \Delta_0^p(\theta_0, \theta^L)}{\partial \theta^L} = \beta \frac{\partial U_1^p(\theta^L)}{\partial \theta^L} = \beta \left\{ a^p - \left[1 - \Phi(\eta_1^*(\theta^L))\right] \hat{\tau} \lambda a^p - \phi(\eta_1^*(\theta^L)) \eta_1^{*\prime}(\theta^L) \Delta_1^p(\theta^L) \right\},$$

and

$$\frac{\partial \Delta_0^r(\theta_0, \theta^L)}{\partial \theta^L} = \beta \frac{\partial U_1^r(\theta^L)}{\partial \theta^L} = \beta \left\{ \left[1 - \Phi(\eta_1^*(\theta^L))\right] (1 - \lambda) \hat{\tau} a^p + \phi(\eta_1^*(\theta^L)) \eta_1^{*\prime}(\theta^L)(1 - \lambda) \hat{\tau}(a^r - \theta_1 a^p) \right\}.$$

From Lemma 2 follows that $\eta_1^{*\prime}(\theta^L) > 0$ and, therefore, that $\partial \Delta_0^r(\theta_0, \theta^L)/\partial \theta^L$ is always positive. The sign of $\partial \Delta_0^p(\theta_0, \theta^L)/\partial \theta^L$ is instead generally ambiguous. However, as $\hat{\tau} \downarrow 0$, $\eta_1^{*\prime}(\theta_1) \downarrow 0$, and $\partial \Delta_0^p(\theta_0, \theta^L)/\partial \theta^L > 0$, which in turn implies that $\partial \eta_0^*(\theta^L)/\partial \theta^L < 0$.

We now show that $\partial V_0^P(\theta^L)/\partial \theta^L > 0$ for all $\theta^L \in [0,1]$ when $\hat{\tau}$ is small enough and δ is relatively high.

From (23), it follows that

$$V_1^P(\theta^L) - V_1^P(1) = \theta_1 a^p + [1 - \Phi(\eta_1^*(\theta_1))] \left[\delta + \Delta_1^p(\theta^L)\right] - a^p - [1 - \Phi(\eta_1^*(1))] \left[\delta + \Delta_1^p(\theta^L)\right],$$

and substituting this expression and (24) into (35), we obtain that

$$
\begin{aligned}
\partial V_0^P(\theta^L)/\partial \theta^L = \ & \left[1 - \Phi(\eta_0^*(\theta^L))\right] \beta a^p \left\{ 1 - \left[1 - \Phi(\eta_1^*(\theta^L))\right] \hat{\tau} \lambda \right\} \\
& - \left[1 - \Phi(\eta_0^*(\theta^L))\right] \beta \phi(\eta_1^*(\theta^L)) \eta_1^{*\prime}(\theta^L) \left[\hat{\tau}\lambda(a^r - \theta_1 a^p) + \delta\right] \\
& - \phi(\eta_0^*(\theta^L)) \frac{\partial \eta_0^*(\theta^L)}{\partial \theta^L} \beta \left\{ \left[1 - \Phi(\eta_1^*(\theta^L))\right] \left[\delta + \Delta_1^p(\theta^L)\right] - [1 - \Phi(\eta_1^*(1))] \left[\delta + \Delta_1^p(1)\right] \right\} \\
& - \phi(\eta_0^*(\theta^L)) \frac{\partial \eta_0^*(\theta^L)}{\partial \theta^L} \left[\delta - \beta(1 - \theta^L)a^p\right].
\end{aligned}
\tag{44}
$$

The first and second term of (44) are respectively positive and negative since $\eta_1^{*\prime}(\theta^L) > 0$. $\partial \eta_0^*(\theta^L)/\partial \theta^L < 0$ ensures that the third term is positive; in fact, from $\theta^L < 1$ follows that $\Delta_1^p(\theta^L) > \Delta_1^p(1)$, and $\eta_1^{*\prime}(\theta^L) > 0$ implies that $\eta_1^*(\theta^L) < \eta_1^*(1)$ and that $\Phi(\eta_1^*(\theta^L)) < \Phi(\eta_1^*(1))$. The fourth term has an ambiguous sign depending on the level of the rents from office; however, this term is positive when the rents are sufficiently high, i.e. $\delta > \tilde{\delta} \equiv \beta a^p$.

As $\hat{\tau} \downarrow 0$, $\eta_1^{*\prime}(\theta_1) \downarrow 0$, and the second term of (44) tends to zero. If $\delta > \delta^{**} \equiv \max\{\delta^*, \tilde{\delta}\}$, then $\partial V_0^P(\theta^L)/\partial \theta^L$ is positive for all $\theta^L \in [0,1]$.[41] This means that the ex-ante welfare of the Left is globally strictly increasing in θ^L and, therefore, it is maximized at $\theta^L = 1$. This result completes the proof of the Lemma.

---

[41] We remind that δ^* guarantees that $V_1^P(\theta_1; \delta)$ is everywhere decreasing in θ_1 for any $\delta > \delta^*$ (see Remark 2) so that $\theta_1^* = \theta^L$.

11 References

Acemoglu, Daron (2010) "Institutions, Factor Prices, and Taxation: Virtues of Strong States?" *American Economic Review: Papers & Proceedings*, 100(2), pp. 115–119.

Acemoglu, Daron, Michael Golosov and Aleh Tsyvinski (2012) "Power Fluctuations and Political Economy," *Journal of Economic Theory*, forthcoming.

Acemoglu, Daron and James A. Robinson (2001) "Inefficient Redistribution," *American Political Science Review*, 95(3), pp. 649–661.

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," *Journal of the European Economic Association*, April, 9(2), pp. 177–208.

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(4), pp. 796–805.

Alesina, Alberto, and Stephen Spear (1988) "An Overlapping Generations Model of Electoral Competition," *Journal of Public Economics*, 37(3), pp. 359–379.

Alesina, Alberto and Guido Tabellini (1990) "A Positive Theory of Fiscal Deficits and Government Debt," *The Review of Economic Studies*, 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*, 112(1), 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.

Besley, Timothy and Torsten Persson (2009) "The Origins of State Capacity: Property Rights, Taxation, and Policy," *American Economic Review*, 99(4), pp. 1218–1244.

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

Borjas, George J. (2003) "The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market," *Quarterly Journal of Economics*, 118(4), pp. 1335–1374.

Borjas, George J. and Abdurrahman Aydemir (2007) "Cross-Country Variation in the Impact of International Migration: Canada, Mexico, and the United States," *Journal of the European Economic Association*, 5(4), pp. 663–708.

Borjas, George J. and Abdurrahman Aydemir (2011) "Attenuation Bias in Measuring the Wage Impact of Immigration," *Journal of Labor Economics*, pp. 69–112.

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.

Card, David (2001) "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts," *Journal of Labor Economics*, 19(1), pp. 22–64.

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(2), 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.

Hamermesh, Daniel (2003) *Labor Demand*, Princeton, NJ: Princeton University Press.

Harrington, Joseph E. Jr. (1993) "Economic Policy, Economic Performance, and Elections," *American Economic Review*, 83(1), pp. 27–42.

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.

Hsieh, Chang-Tai, Edward Miguel, Daniel Ortega and Francisco Rodriguez (2011) "The Price of Political Opposition: Evidence from Venezuela's *Maisanta*," *American Economic Journal: Applied Economics*, 3(2), pp. 196–214.

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.

Krusell, Per and Jose-Victor Rios-Rull (1996) "Vested Interests in a Positive Theory of Stagnation and Growth," *Review of Economic Studies*, 63(2), pp. 301–329.

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.

Lindbeck, Assar and Jörgen W. Weibull (1987) "Balanced-budget Redistribution as the Outcome of Political Competition," *Public Choice*, 52(3), pp. 272–297.

Mainwaring, Scott and Matthew Soberg Shugart, Editors (1997) *Presidentialism and Democracy in Latin America*, Cambridge: Cambridge University Press.

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.

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

Milesi-Ferretti, Gian Maria (1995) "The Disadvantage of Tying Their Hands: On the Po-

litical Economy of Policy Commitments," *Economic Journal*, 105(433), pp. 1381–1402.

Milesi-Ferretti, Gian Maria and Enrico Spolaore (1994) "How Cynical Can an Incumbent Be? Strategic policy in a Model of Government Spending," *Journal of Public Economics*, 55(1), pp. 121-140.

Milgrom, Paul and Shannon, Chris (1994) "Monotone Comparative Statics," *Econometrica*, 62(1), pp. 157–180.

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

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(2–3), pp. 197–210.

Romer, Thomas (1975) "Individual Welfare, Majority Voting, and the Properties of a Linear Income Tax," *Journal of Public Economics*, 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 (2002) "The Political Economy of Employment Protection," *Journal of Political Economy*, 110(3), pp. 672–704.

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

Scheve, Kenneth F. and Matthew Slaughter (2001) "Labor–Market Competition and Individual Preferences Over Immigration Policy," *Review of Economics and Statistics*, 83(1), pp. 133–145.

Shleifer, Andrei and Robert W. Vishny (1989) "Management Entrenchment. The Case of Manager-Specific Investments," *Journal of Financial Economics*, 25(1), pp. 123–139.

Stokes, Susan C. (2001) *Mandates and Democracy: Neoliberalism by Surprise in Latin America*, Cambridge: Cambridge University Press.

Stokes, Susan C. (2009) "Globalization and the Left in Latin America," Yale University, 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: Princeton University Press.

Zolberg, Aristide R. (2008) *A Nation by Design: Immigration Policy in the Fashioning of America*, Cambridge, MA: Harvard University Press.

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