EMERGENCE AND PERSISTENCE OF INEFFICIENT STATES

Daron Acemoglu
Massachusetts Institute of Technology

Davide Ticchi
University of Urbino

Andrea Vindigni
Princeton University

Abstract
We present a theory of the emergence and persistence of inefficient states based on patronage politics. The society consists of rich and poor. The rich are initially in power, but expect to transition to democracy, which will choose redistributive policies. Taxation requires the employment of bureaucrats. By choosing an inefficient state structure, the rich may be able to use patronage and capture democratic politics, so reducing the amount of redistribution in democracy. Moreover, the inefficient state creates its own constituency and tends to persist over time. Intuitively, an inefficient state structure creates more rents for bureaucrats than would an efficient one. When the poor come to power in democracy, they will reform the structure of the state to make it more efficient so that higher taxes can be collected at lower cost and with lower rents for bureaucrats. Anticipating this, when the society starts out with an inefficient organization of the state, bureaucrats support the rich, who set lower taxes but also provide rents to bureaucrats. We obtain that the rich–bureaucrats coalition may also expand the size of bureaucracy *excessively* so as to generate enough political support. The model shows that an equilibrium with an inefficient state is more likely to arise when there is greater income inequality, when bureaucratic rents take intermediate values, and when individuals are sufficiently forward-looking. (JEL: P16, H11, H26, H41)

1. Introduction

There are large cross-country differences in the extent of bureaucratic corruption and the efficiency of the state organization (see for example World Bank 2004). An influential argument, dating back at least to Tilly (1990), maintains that differences in

The editor in charge of this paper was Roberto Perotti

Acknowledgments: We thank Timothy Besley, Bruno Contini, Miriam Golden, Roberto Perotti, Gilles Saint-Paul, Guido Tabellini, three anonymous referees, and seminar participants at IGIER-Bocconi, the Latin American Meetings of the Econometric Society, National Bureau of Economic Research Summer Institute, Toulouse University, and Yale University for useful comments and suggestions. Acemoglu gratefully acknowledges financial support from the National Science Foundation and the hospitality of the Yale Economics Department. Vindigni gratefully acknowledges the hospitality of Collegio Carlo Alberto and of the Yale Leitner Program in International Political Economy. Acemoglu is a member of CIFAR.

E-mail: daron@mit.edu (Acemoglu); davide.ticchi@uniurb.it (Ticchi); vindigni@princeton.edu (Vindigni)

Journal of the European Economic Association April 2011 9(2):177–208

DOI: 10.1111/j.1542-4774.2010.01008.x

"state capacity" are an important determinant of economic development.[1] The evidence that many less-developed economies in sub-Saharan Africa, Asia, and Latin America only have a small fraction of their GDP raised in tax revenue and invested by the government (see for example Acemoglu 2005) and the correlation between measures of state capacity and economic growth (see for example Rauch and Evans 2000) are consistent with this view as well. Societies with limited state capacity also tend to be those that invest relatively little in public goods and do not adopt policies that redistribute resources to the poor.[2]

In this paper, we construct a political economy model linking the emergence and persistence of inefficient states to the strategic use of *patronage politics* by the elite as a means of capturing democratic politics. Democratic capture enables the elite to limit the provision of public goods and redistribution, but at the cost of aggregate inefficiencies. Our approach therefore provides a unified answer to the questions of:

(i) why inefficient states emerge in some societies;
(ii) what explains the prevalence of patronage politics and its association with greater inefficiencies in bureaucracy; and
(iii) why many democracies pursue relatively pro-elite policies.

Our model also suggests a potential reason why certain democracies may exhibit relatively poor economic performance and adopt various inefficient policies.

We consider an infinite-horizon economy consisting of two groups, the rich elite and poor citizens. Linear taxes can be imposed on both groups, with the proceeds used to finance public good investments. The rich are opposed to high levels of taxes and public good investments. Tax collection requires that the state employs bureaucrats to prevent individuals from evading taxes, but bureaucrats themselves also need to be given incentives so that they exert effort (or do not accept bribes). The efficiency with which a central authority can monitor the bureaucrats is our measure of the *organization of the state*. Political competition is modeled either by assuming the existence of two parties, respectively aligned with the rich and the poor, or by allowing free entry into the political arena by citizen-candidates (Osborne and Slivinski 1996; Besley and Coate 1997). In both cases, there is no commitment to policies before elections and the party that comes to power chooses the policy vector, consisting of taxes, public good provision, and bureaucratic wages, and whether to reform the efficiency of the state institutions. Democratic political competition is made interesting by the fact that bureaucrats may support either the rich or poor parties (candidates) and their support may be pivotal in the outcome of elections.[3]

1. See, for example, Evans 1995; Levi 1989; Migdal 1988; Epstein 2000; Herbst 2000; Centeno 2002; Kohli 2004.

2. See, for example, Etzioni-Halevy (1983) on the importance of state capacity and bureaucratization for the development of the welfare state in the West, and Rothstein and Uslaner (2005) on the importance of state capacity for income redistribution.

3. In the main analysis we focus on Markov Perfect Equilibria (where strategies only depend on payoff-relevant state variables). This equilibrium concept implies that there is no commitment to policies before

There are two possible organizations of the state: the first is an *efficient* organization, in which bureaucrats will be detected easily if they fail to exert effort, while the second is an *inefficient* one in which monitoring bureaucrats is difficult. In equilibrium, when the state is inefficient, bureaucrats need to be paid rents in order to induce them to perform their roles of tax collection and inspection. The presence of rents creates the possibility of patronage politics, whereby bureaucrats may support the party that will maintain the inefficient structure.

In a society that is always dominated by the rich elite or that is permanently in democracy (with a poor citizen as the median voter), the political process produces an efficient organization of bureaucracy, since an inefficient state creates additional costs and no benefits for those holding power. Our main result is that when the society starts out as nondemocratic (under the control of the rich elite) and is expected to transition to democracy, the rich may find it beneficial to choose an inefficient organization of the state so as to exploit patronage politics to limit redistribution. In particular, bureaucrats realize that once the poor come to power in democracy, there will be bureaucratic reform, reducing their rents. Therefore, if the rich elite, when in power, choose an inefficient organization of the state, the current bureaucrats—who are receiving rents—prefer to support the rich rather than vote with the poor. Consequently, an inefficient state organization emerges as a political instrument for the rich elite to *capture* the democratic decision-making process by fostering a coalition between themselves and the bureaucrats. It is also noteworthy that the inefficient state not only emerges in equilibrium, but also persists; when the state is inefficient, the bureaucrats vote for the party of the rich, which chooses not to reform the bureaucracy and continues to maintain the support of the existing bureaucrats and thus its political power.

Our analysis shows that patronage politics can lead not only to the emergence and persistence of an inefficient state but also to the *overemployment* of bureaucrats. This is because the rich may prefer to hire additional (unnecessary) bureaucrats so as to boost their party's votes. Consequently, a captured democracy will typically feature an inefficient state (bureaucracy), provide relatively few public goods, and employ an excessive number of bureaucrats. This pattern of bureaucratic inefficiency is consistent with the stylized view of corrupt and low-capacity bureaucracies in many less-developed countries (Geddes 1991; Rauch and Evans 2000).[4]

Two factors are important in the result that the rich distort the efficiency of the state organization in order to gain political advantage in democracy. The first is that the efficiency of the state (bureaucracy) is a *state variable*, meaning that the rules

elections and thus enables us to illustrate the logic of patronage politics in the clearest fashion. Section 4.1 shows that similar results hold when we focus on Subgame Perfect Equilibria.

4. Even with the overemployment of bureaucrats, bureaucrats and the rich elite are unlikely to have an absolute majority in the electorate, and we do not imply that this channel alone will be sufficient for the elite to control democratic politics. Instead, our objective is to focus on a specific mechanism of democratic capture by the elite to highlight how it functions and what its implications are. In practice, as discussed further in what follows, the elite may be able to gain additional influence using other, complementary methods such as lobbying, vote buying, or use of paramilitaries. In addition, the lower turnout rates of relatively poor individuals may further limit the political power of parties favoring redistributive policies (Bénabou 2000).

and regulations governing bureaucrats' behavior and monitoring cannot be reformed immediately. Instead, it will take some time (in our model one period) for these reforms to take effect. We view this as a good approximation to reality, where institutional variables are endogenously determined and can be changed, but often with some delay or sluggishness.[5] The second and more important ingredient for this result is that the rich elite can credibly commit to an inefficient state, which provides rents to current bureaucrats. A party representing the preferences of the median voter (a relatively poor agent) cannot make such a credible commitment, however. This feature emerges as part of the equilibrium; when a party representing the preferences of the poor comes to power it will be in its interest to reform the bureaucracy in order to be able to increase tax revenues, thus it cannot make a credible commitment to an inefficient state. In contrast, the party representing the preferences of the rich can do so because its disproportionate effect on democratic politics crucially depends on the persistence of the inefficient state. It is this differential (equilibrium) commitment power of the rich and the poor that leads to the emergence of an inefficient state and to the coalition between the rich and the bureaucrats, limiting redistribution in democracy.

The comparative static results of our model shed light on the conditions under which an inefficient state may emerge. Most importantly, greater (pre-tax) inequality makes the emergence and persistence of an inefficient state more likely. This is because greater inequality raises the equilibrium tax rate in democracy and makes it more appealing for the rich to create an inefficient state apparatus to prevent democratic outcomes. An inefficient state also requires intermediate levels of *rents/efficiency wages* for bureaucrats; when rents are limited, bureaucrats would not support the rich, while too high rents would make the inefficient state equilibrium prohibitively costly for the rich elite. Finally, an inefficient state is more likely to arise when agents are more forward-looking, because bureaucrats support the inefficient state in order to obtain future rents.

It is worth emphasizing that in our model patronage politics and the creation of an inefficient state are the only way for the rich to increase their votes and influence in politics. We certainly do not argue or believe that there are no other, complementary methods of democratic capture. Instead, our purpose is to isolate a particular mechanism that appears to be important in practice and investigate its implications. Other methods via which a rich elite may have disproportionate effect in a democracy include lobbying, vote buying, co-optation of a subgroup of the population, and use of force and paramilitaries. Co-optation in general plays a role similar to that of patronage politics in our model and may be important in practice, but only a few papers have investigated how co-optation affects political equilibria. Our mechanism can be interpreted as an example of co-optation, though

5. This feature plays a role in the analysis, since strategies in Markov Perfect Equilibria can be conditioned on the current organization of the state. It is no longer necessary when we focus on Subgame Perfect Equilibria.

it is not a specific social or ethnic group that is being co-opted and equilibrium co-optation is only possible when the structure of the state is distorted in an inefficient manner.[6]

Brazil provides a typical example of a society where the state sector has been relatively inefficient and democratic politics has generated only limited public goods and benefits for the poor, and illustrates the forces emphasized by our analysis. It is generally agreed that the distribution of large numbers of public jobs, both in the public administration and in parastatal organizations, has created a pattern of patronage politics in Brazil (see for example Gay 1990; Evans 1992; Weyland 1996; Roett 1999). The control over these jobs appears to have enabled traditional elites to preserve their political power and limit the amount of public good provision and redistribution. Cohen (1987) argues that the origins of state expansion in Brazil lie precisely with the attempts of the elite to control politics following World War II, when transition to democracy in Brazil became inevitable. He describes this episode and its aftermath as follows: "During the relatively peaceful transition to democracy between 1943 and 1945, the political elite used the resources of this greatly expanded state to forge the broad electoral coalition that would allow it to perpetuate its power in the future regime" (p. 49). Consistent with this picture of Brazilian politics, patronage relations have also ensured that even those in the poorest neighborhoods of Rio have supported the traditional parties rather than socialist or social democratic parties running on platforms of greater public good provision and redistribution (Gay 1990). Roett (1999) emphasizes the role of public sector employees in this process and writes that "state company employees emerged as being among the strongest supporter of the patrimonial order" . In return, successive governments have withstood external pressures from the IMF and have not reformed the public sector, despite the "public perception that public-sector workers were overpaid and underworked" (Roett, p. 97). The process of reforming the public sector in Brazil has started only recently and progressed slowly.

Our paper is related to a number of different literatures, though we are not aware of other papers that derive an equilibrium explanation for the various inefficiencies in the organization of bureaucracies and provide a range of comparative statics showing when such inefficiencies are more likely. The political science and sociology literature on the organization of the state and the bureaucracy mentioned previously discusses related issues, but does not provide formal models or emphasize the mechanisms we focus on in this paper. The small economics literature on the internal organization of the state and bureaucracy, for example Acemoglu and Verdier (1998), Dixit (2002), Egorov and Sonin (2005) and Debs (2006), does not investigate the relationship between patronage politics and the emergence of the inefficient state as a method of limiting redistribution.

6. We are only aware of two papers on co-optation in political economy, Gershenson and Grossman (2001) and Bertocchi and Spagat (2001), neither of which focuses on the questions or the mechanisms emphasized here. Lobbying is discussed extensively in the literature, but typically not in the context of democratic capture and inefficient state structures.

Acemoglu (2005) and Besley and Persson (2009) are complementary to our approach, since they emphasize the importance of state capacity.

The literature on the inefficiency of the form of redistribution is also related to our work. Wilson (1990) and Becker and Mulligan (2003) argue that inefficient methods of redistribution are chosen as a way of limiting the amount of redistribution (see also Coate and Morris 1995; Rodrik 1995; Saint-Paul 1996; Acemoglu and Robinson 2001).[7] However, these papers do not model or explain how an existing elite can capture democratic politics by creating an inefficient state structure or provide a mechanism for patronage politics.[8]

Most closely related to our work is the small literature on how politicians may distort policies for strategic reasons. Papers in this literature include models where inefficient policies are chosen in order to gain votes (see for example Fiorina and Noll 1978; Geddes 1991; Shleifer and Vishny 1994; Lizzeri and Persico 2001; Robinson and Torvik 2005). Other papers focus on how certain inefficient choices (including wasteful investments, large budget deficits, and inefficient fiscal systems) could be made in order to constrain future politicians (see for example Glazer 1989; Persson and Svensson 1989; Tabellini and Alesina 1990; Aghion and Bolton 1990; Cukierman, Edwards, and Tabellini 1992; Biais and Perotti 2002). Cukierman, Edwards, and Tabellini (1992) has the most closely related focus to our paper, since, as in the papers by Wilson and Becker and Mulligan mentioned above, they also emphasize the role of an inefficient tax system in limiting future redistribution. As in those papers, however, there is no explanation for why there can be commitment to the tax system but not to the level of taxes. The main contribution of our model relative to all of these papers is the idea that the rich elite, who need the support of the bureaucracy, can make a credible commitment to keeping an inefficient state, while the poor, who wish to raise greater tax revenues, cannot credibly commit to such a policy. More generally, none of these papers derive a rationale for patronage politics or feature the mechanism of an elite creating an inefficient state structure to maintain their political power in the face of an emerging democracy.

The rest of the paper is organized as follows. Section 2 outlines the basic economic and political environment. Section 3 presents the main results. Section 4 briefly discusses a number of extensions. Section 5 concludes, while the Appendix contains all the proofs.

7. Note, however, that there is an important distinction between our theory and the basic Becker–Mulligan–Wilson story. In the latter, it is not clear why the society can commit to the form of redistribution and *not* to the amount of redistribution. In contrast, in our model the choice of an inefficient bureaucracy is an equilibrium strategy for affecting the future political equilibrium so as to bring the party aligned with the interests of the rich to political power, and via this channel, to limit the provision of public goods and taxation.

8. Another related paper is Alesina and Drazen (1991), who suggest an explanation for delayed reform due to a war-of-attrition between different groups. Our theory provides an explanation for lack of reform based on democratic capture by the elite.

2. Basic Model

2.1. Description of the Economic Environment

Consider the following discrete-time infinite-horizon economy populated by a continuum 1 of agents, each of which has the following risk-neutral preferences:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(c_t^j + G_t - h e_t^j \right),$$

at time $t = 0$, where $\mathbb{E}_0$ is the expectation at time $t = 0$, $\beta \in (0, 1)$ is the discount factor, $c_t^j \geq 0$ denotes the consumption of the agent in question (agent j), $G_t \geq 0$ is the level of public good enjoyed by all agents, $e_t^j \in \{0, 1\}$ is the effort decision of the agent (which will be necessary in some occupations), and $h > 0$ is the cost of effort.

There are two types of agents: $n > 1/2$ are poor (low-skill), while $1 - n$ are rich (high-skill). We denote poor agents by the symbol L (corresponding to low-productivity), and rich agents by H, and also use $\mathcal{L}$ and $\mathcal{H}$ to denote the set of poor and rich agents.

There are two occupations: producer and bureaucrat. In each period, as long as some amount of investment in infrastructure, $K > 0$, is undertaken, each producer generates an income depending on his skill; A^L for poor agents and $A^H > A^L$ for rich agents. If the investment in infrastructure K is not undertaken at time t, then no agent can produce within that period.[9] Producers receive and consume their income net of taxes.

A set of agents denoted by $\mathcal{X}_t$ are bureaucrats at time t. These agents do not produce, but receive a *net* wage of $w_t \geq 0$ from the government (that is, they do not pay taxes on their wage income, which is simply a normalization). The role of bureaucrats is tax collection. In particular, we will allow for a linear tax rate $\tau_t \in [0, 1]$ on earned incomes in order to finance the infrastructure investment K, additional spending on the public good G_t, and the wages of bureaucrats. This tax rate is the same regardless of whether the individual is rich or poor. To simplify the discussion, we assume that only poor agents can become bureaucrats. This assumption is not necessary for the results, since it will be evident in what follows that low-productivity poor agents always prefer bureaucracy more than do high-productivity rich agents.

Both rich and poor agents can try to evade taxes. We assume that if an individual tries to evade taxes, he gets caught with probability $p(x_t)$, where $p : [0, 1] \to [0, 1]$ is an increasing, twice continuously differentiable, and strictly concave function with $p(0) = 0$, and x_t denotes the number of bureaucrats exerting positive effort at time

9. While the level of K is not important for our results (and we could set $K = \varepsilon > 0$ for ε arbitrarily small), it is important that bureaucrats undertake productive activities, since otherwise the rich cannot commit to employing bureaucrats once they are in power.

t. More formally, this is defined as $x_t = \int_{j \in \mathcal{X}_t} e_t^j dj$. This expression incorporates the fact that bureaucrats who do not exert effort are not useful.[10]

If an individual is caught evading taxes, all of his income during that period is lost. For simplicity, we assume that this income does not accrue to the government either.[11] We also assume that there is *limited liability*, so that $c_t^j \geq 0$, and *anonymity*, so that the past history of individual producers is not observed and future punishments on tax evaders are not possible.

Since effort is costly, bureaucrats will exert effort only if their compensation depends on their effort decision. We assume that if they do not exert effort, bureaucrats are caught with probability q_t at time t. If they are not caught, they receive the wage w_t, and if they are caught shirking, they lose their wage, but are not fired from the bureaucracy.[12] This assumption simplifies the exposition and is relaxed in Section 4.2.

The probability of detection q_t depends on the *organization of the state*. We allow two types of organizations, represented by $I_t \in \{0, 1\}$, which correspond to different types/degrees of monitoring within the bureaucracy. In particular, we assume $q\,(I = 1) = 1$, so that one of the possible organizations of the state allows for perfect monitoring of bureaucrats. In contrast, $q\,(I = 0) = q_0 < 1$, so that the alternative organization involves an imperfect monitoring technology and shirking bureaucrats are not necessarily detected. Clearly, imperfect monitoring will lead to equilibrium distortions, and for this reason we refer to $I = 0$ as an *inefficient organization of the state* and to $I = 1$ as *efficient state*. To simplify the analysis we assume that $I = 1$ has no cost relative to $I = 0$.[13]

At each date, the political system chooses the following policies: (i) a tax rate on all earned income $\tau_t \in [0, 1]$; (ii) the wage rate for bureaucrats $w_t \in \mathbb{R}_+$; (iii) a level of public good $G_t \in \mathbb{R}_+$; (iv) the number of bureaucrats hired, $X_t \in [0, 1]$; and (v) the efficiency of the state for the next date, $I_t \in \{0, 1\}$; the efficiency of the state at the current date, I_{t-1}, is part of the state variable, determined by choices in the previous period. These policies must satisfy the government budget constraint (specified in what follows), and we also make the following assumption to simplify the exposition: if $X_t \geq X_{t-1}$, then existing bureaucrats cannot be fired (but each bureaucrat can decide to quit if he finds this beneficial); if $X_t < X_{t-1}$, then no new bureaucrats are hired and a fraction $(X_t - X_{t-1})\,/\,X_t$ of the bureaucrats is fired (those fired being randomly

10. Alternatively, instead of inducing bureaucrats to exert effort, it may be important to ensure that they do not accept bribes from the individuals supposed to pay taxes (see for example Acemoglu and Verdier 1998, 2000). We investigate a variant of our model with corruption in Section 4.4.

11. This is not an important assumption, since there is no tax evasion in equilibrium.

12. Bureaucrats also do not receive their wage if the investment in infrastructure, K, is not undertaken, since in this case there is no production and thus no government revenue. This event does not take place along the equilibrium path. Moreover, since there is a continuum of agents, no individual takes into account whether his decision to enter or quit bureaucracy will have an effect on government revenues and the financing of the investment in infrastructure.

13. In general, one can imagine that setting up a more efficient state apparatus may involve additional expenditures. We ignore those both to simplify the algebra and also to highlight that inefficient states can arise *even when* an efficient organization is costlessly available.

chosen irrespective of their past history).[14] We denote a vector of policies satisfying these restrictions by $\rho_t \equiv (\tau_t, w_t, G_t, X_t, I_t) \in \mathcal{R}$.

2.2. *Description of the Political System*

We will consider three different political environments: (i) Permanent Nondemocracy: the rich elite are in power at all dates, meaning that only the rich can vote, and since all rich agents have the same policy preferences over the available set of policies, the policy vector most preferred by a representative rich elite will be implemented. (ii) Permanent Democracy: the citizens, who form the majority, are in power at all dates starting at $t = 0$ (or at all dates there are elections as described in what follows). (iii) Emerging Democracy: the rich elite are in power at $t = 0$, and in all future dates, the regime will be democratic with majoritarian elections.

The first two environments are for comparison. The third one is our main focus in this paper. It is a simple way of capturing the idea that some decisions are originally taken by elites, anticipating that democracy will arrive at some point—in this case right at date $t = 1$.[15]

To start with, we model the democratic system in a very simple way, by assuming that there are two parties, one run by a poor agent and one run by an elite agent, and that bureaucrats cannot run for office. We use the symbols P and R to denote these parties and $d_t = P$ denotes that party P is elected to office at date t. Parties are unable to make commitments to the policies they will implement once they come to power. Thus whichever party receives the majority of the votes comes to power and the agent in control of the party chooses the policy vector that maximizes his own utility. This last assumption departs from the standard Downsian models of political competition where parties commit to their policy platform before the election (see Section 4.3 for further discussion).

2.3. *Timing of Events*

To recap, the timing of events within each date is as follows. The society starts with some political regime, nondemocracy or democracy, that is, $s_t \in \{N, D\}$, a set $\mathcal{X}_{t-1} \subset \mathcal{L}$ of agents who are already bureaucrats (since, by assumption, the set of bureaucrats $\mathcal{X}_{t-1}$ must be a subset of the set of poor agents), and a level of efficiency of the state, $I_{t-1} \in \{0, 1\}$. Then, the following occurs.

14. We adopt this restriction to simplify the discussion and the notation. The same results apply if the party in power can choose, without any restrictions, who the bureaucrats will be, but in this case we have to specify the exact identity of those in bureaucracy at each date as part of the policy vector.

15. In this case, the society is nondemocratic at date $t = 0$, and we assume that it will become democratic for exogenous reasons at date $t = 1$. It is possible to model democratization as equilibrium institutional change along the lines of the models of endogenous democratization in the literature (for a discussion and references see Acemoglu and Robinson 2006), but doing so would complicate the analysis without generating additional economic insights in the current context.

(1) In democracy, all individuals $j \in [0, 1]$ vote for either party P or party R, that is, individual j decides $v_t^j \in \{P, R\}$. Whichever party receives the majority of the votes is elected to office. To simplify the discussion, we assume that if both parties receive exactly half of the votes, party R is elected.
(2) The elected party (in democracy) or the representative elite agent (in nondemocracy) decides the policy vector $\rho_t \equiv (\tau_t, w_t, G_t, X_t, I_t) \in \mathcal{R}$.
(3) Observing this vector, each individual $j \notin \mathcal{X}_{t-1}$ decides whether to apply to become a bureaucrat, $\chi_t^j \in \{0, 1\}$, and each individual $j \in \mathcal{X}_{t-1}$ decides whether to quit bureaucracy, $\chi_t^j \in \{0, 1\}$ (which is denoted by the same symbol without any risk of confusion). Naturally, by assumption, $\chi_t^j = 0$ for all the rich agents. The number of bureaucrats at time t is then $\min\{X_t, \int_0^1 \chi_t^j dj\}$, that is, the minimum of the number of bureaucrats chosen by the polity in power and the number of people applying to or remaining in bureaucracy. This also determines the current set of bureaucrats, $\mathcal{X}_t$.
(4) Each bureaucrat decides whether to exert effort, $e_t^j \in \{0, 1\}$, which determines $x_t = \int_{j \in \mathcal{X}_t} e_t^j dj$, and thus the probability of detection of individuals evading taxes.
(5) Production takes place and each producer decides whether to evade taxes or not, denoted by $z_t^j \in \{0, 1\}$.
(6) A fraction $p(x_t)$ of producers evading taxes are caught.
(7) A fraction $q_t = q(I_{t-1})$ of shirking bureaucrats are caught and punished.
(8) Taxes are collected, remaining bureaucrats are paid their wage, w_t, and the public good G_t is supplied.

Naturally, the society starts with $\mathcal{X}_{-1} = \varnothing$, that is, at the initial date there are no incumbent bureaucrats. We also suppose that $I_{-1} = 0$ (though this has no bearing on any of our results except the actions at time $t = 0$, since the choice of $I_t \in \{0, 1\}$ is without any costs).

3. Characterization of Equilibria

3.1. Definition of Equilibrium

In this section, we focus on pure strategy Markov perfect equilibria (MPE). In our model, MPE will be unique and relatively straightforward to characterize. In addition, the focus on MPE makes the emergence of a coalition between the rich and the bureaucrats more difficult (since there cannot be *commitment* to future rents for bureaucrats). Subgame perfect equilibria are discussed in Section 4.1.

Recall that Markovian strategies condition only on the payoff-relevant state variables (and on the prior actions within the same stage game). A MPE is defined as a set of Markovian strategies that are best responses to each other given every history. In the current game, the aggregate state vector can be represented as $S_t \equiv (s_t, I_{t-1}, X_{t-1}) \in \mathcal{S}$, where $s_t \in \{N, D\}$ is the political regime at time t, $I_{t-1} \in \{0, 1\}$ is the efficiency of the bureaucracy inherited from the previous period, and X_{t-1} is the size of the bureaucracy inherited from the previous period.[16] Individual actions will be a function of the aggregate state vector S_t and the individual's identity, in particular, $a_t \in \{L, H, B\}$ representing whether the individual is a poor producer, rich producer or a bureaucrat. Thus as a function of S_t and a_t, each individual will decide which party to vote for, that is, $v_t^j \in \{P, R\}$, whether to apply (or to remain) in bureaucracy, $\chi_t^j \in \{0, 1\}$, whether to evade taxes, $z_t^j \in \{0, 1\}$, if the individual is a producer, and whether to exert effort, $e_t^j \in \{0, 1\}$, if the individual is a bureaucrat. Finally, strategies also include the choice of $I_t \in \{0, 1\}$, $\tau_t \in [0, 1]$, $X_t \in [0, n]$, and $G_t \in \mathbb{R}_+$ when the individual is the party leader. Thus Markovian strategies can be represented by the following mapping:

$$\sigma : \mathcal{S} \times \{L, H, B\} \to \{P, R\} \times \{0, 1\}^4 \times [0, 1] \times [0, n] \times \mathbb{R}_+.$$

A MPE is a mapping σ^* that is best response to itself at every possible history. In addition, we impose the natural restriction that in the voting stage no agent uses weakly dominated strategies and this qualification is implicit in our notion of MPE.[17]

We often refer to subcomponents of σ rather than the entire strategy profile, and with a slight abuse of notation, we use $v(I|a)$ to denote the voting strategy of an individual of group $a \in \{L, H, B\}$ as a function of the efficiency of the state institutions. Moreover, when there is no risk of confusion, we use the index j to denote individuals or groups interchangeably.

3.2. Preliminary Results

Let us first note that if $p(x_t) < \tau_t$, then all producers will evade taxes at time t, that is, $z_t^j = 0$ for all $j \notin \mathcal{X}_t$. This follows from the fact that by anonymity, the continuation value of a producer caught evading taxes is the same as that of a producer paying taxes. This observation combined with limited liability implies that the decision to evade taxes depends on whether post-tax income $(1 - \tau_t) A^j$ is greater than expected income from evasion, $(1 - p(x_t)) A^j$ (taking into account that tax evaders are caught with probability $p(x_t)$). Since with tax evasion there is no government revenue, this implies that in equilibrium we need to have the following *incentive compatibility constraint* for producers,

$$p(x_t) \geq \tau_t,$$

16. In addition, for each individual we could specify whether the individual is currently a bureaucrat, that is, whether $j \in \mathcal{X}_{t-1}$ and whether he is a party leader as part of the individual-specific state vector. Nevertheless, Markovian strategies can be defined without doing this, which simplifies the notation.

17. Without this restriction, as in other voting games, there will exist equilibria in which all agents (in particular, poor agents) use weakly dominated strategies and vote for the party that gives them lower utility. Focusing on strategies that are not weakly dominated is standard practice in such games and is without loss of any economic insight.

to be satisfied. Alternatively, defining $\pi(\tau) \equiv p^{-1}(\tau)$, producers' incentive compatibility constraint can be expressed as[18]

$$x_t \geq \pi(\tau_t). \tag{1}$$

This condition requires the number of bureaucrats exerting effort to be greater than $\pi(\tau_t)$. This constraint is sufficient to ensure that all individuals choose not to evade taxes. Since $p(\cdot)$ is strictly increasing, continuously differentiable and strictly concave, $\pi(\cdot)$ is strictly increasing, continuously differentiable and strictly convex.

Next, a similar argument gives the following *incentive compatibility constraint* for the bureaucrats:

$$w_t \geq \frac{h}{q_t}, \tag{2}$$

where $q_t = q(I_{t-1})$. This constraint is necessary and sufficient to ensure that all bureaucrats choose to exert effort.[19] Moreover, since bureaucrats are necessary to prevent tax evasion and some amount of taxation is necessary for the investment in infrastructure, every allocation must satisfy (2).

Finally, in equilibrium (poor) individuals must prefer to become bureaucrats. This requires the following *participation constraint*:

$$w_t \geq (1-\tau_t) A^L + h, \tag{3}$$

which imposes that bureaucrats receive at least as much as they would obtain in private production.[20]

This discussion immediately establishes the following lemma (proof omitted).

LEMMA 1. *In any MPE, conditions (1), (2) and (3) must hold and $e_t^j = 1$ for all $j \in \mathcal{X}_t$ and all t, and $z_t^j = 1$ for all $j \notin \mathcal{X}_t$ and all t.*

In other words, in any equilibrium the incentive compatibility constraints of producers and bureaucrats and the participation constraint of bureaucrats are satisfied, and no producer evades taxes and all bureaucrats exert effort. This also implies that as long as the constraints (1) and (2) are satisfied, the government budget constraint can be written as

$$K + G_t + w_t X_t \leq (1-n)\tau_t A^H + (n - X_t)\tau_t A^L, \tag{4}$$

18. This condition can also be interpreted as a *state capacity constraint* since, given the effective size of the bureaucracy, it determines the maximum tax rate.

19. This incentive compatibility constraint exploits the fact that bureaucrats caught shirking are not fired. Then by shirking a bureaucrat saves the effort cost h and incurs the expected cost of $q_t w_t$. We show in Section 4.2 that our main results do not depend on the assumption that bureaucrats cannot be fired.

20. If rich agents could become bureaucrats, the equivalent participation constraint, corresponding to (3), for rich agents would be $w_t \geq (1-\tau_t)A^H + h$. Clearly, poor agents are always more willing to enter bureaucracy than rich agents. Our assumption that rich agents cannot become bureaucrats therefore enables us to avoid imposing explicit conditions to ensure that this inequality is not satisfied and (3) is.
In addition, because as noted in footnote 12, individuals are infinitesimal and thus do not take into account the implications of their occupational choice on government revenue, no individual would accept to work in bureaucracy at a wage that does not satisfy (3).

where the left-hand side is government expenditures, consisting of the investment in infrastructure, spending on public goods and bureaucrats' wages, while the right-hand side is total government tax receipts. This expression takes into account that all bureaucrats exert effort and no producer evades taxes. Moreover, (4) highlights that in our model, taxation reduces output through a particular general equilibrium mechanism; the government can raise taxes only by hiring bureaucrats and bureaucrats themselves do not produce any output.

Finally, the following lemma is immediate and is also stated without proof.

LEMMA 2. *Rich agents always vote for party R, that is, for all $j \in \mathcal{H}$, $v_t^j = R$, and poor producers always vote for party P, that is, for all $j \in \mathcal{L}$ and $j \notin \mathcal{X}_{t-1}$, $v_t^j = P$.*

3.3. Equilibria under Permanent Democracy and Nondemocracy

Equilibria under permanent democracy and permanent nondemocracy are of interest as a comparison to our main political environment, which involves the society starting as nondemocratic and then transitioning to democracy. The following results are straightforward and as with all remaining lemmas and propositions, the proofs are provided in the Appendix.

PROPOSITION 1. *Under permanent democracy, there exists a unique MPE. In this equilibrium, $d_t = P$ at each $t \geq 0$, and the following policy vector is implemented at each $t > 0$:*

$$\begin{aligned} &I_t = 1, \quad w_t = (1-\tau^D)A^L + h, \quad X_t = \pi(\tau^D), \quad \text{and} \\ &G_t = G^D \equiv (1-n)\tau^D A^H + [n - \pi(\tau^D)]\tau^D A^L - [(1-\tau^D)A^L + h]\pi(\tau^D) - K, \end{aligned} \tag{5}$$

where τ^D is the unique solution to the maximization problem,

$$\begin{aligned} &\max_{\tau, G} \ (1-\tau)A^L + G \\ &\text{subj. to:} \ G = (1-n)\tau A^H + [n-\pi(\tau)]\tau A^L - [(1-\tau)A^L + h]\pi(\tau) - K. \end{aligned} \tag{6}$$

The next proposition provides analogous results and shows that the organization of the state will be efficient under permanent nondemocracy.

PROPOSITION 2. *Under permanent nondemocracy, there exists a unique MPE. In this equilibrium, the following policy vector is implemented at each $t > 0$:*

$$I_t = 1, \quad w_t = (1-\tau^N)A^L + h, \quad X_t = \pi(\tau^N), \quad G_t = G^N \equiv 0,$$

and τ^N is the unique solution to the equation

$$[(1-\tau)A^L + h]\pi(\tau) - (1-n)\tau A^H - [n - \pi(\tau)]\tau A^L + K = 0. \tag{7}$$

The main conclusion from both of these benchmark political environments is that the politically decisive agents choose a policy vector consistent with their own interests, and this always involves an *efficient* organization of the state, that is, $I_t = 1$ for all $t \geq 0$. There is no reason to make the state inefficient. Consequently, both consolidated democratic and nondemocratic regimes involve $I = 1$. Moreover, in both regimes the capacity of the state is fully utilized in the sense that constraint (1) holds as equality and the minimum number of bureaucrats necessary to prevent tax evasion are employed.

It is straightforward to see that the unique solution (τ^D, G^D) in (6) involves $\tau^D > 0$, since infrastructure spending, $K > 0$, has to be financed (and for the same reason, $\tau^N > 0$ in Proposition 2). However, because raising further revenues involves the employment of bureaucrats which is costly, it is possible that the solution to (6) involves $G^D = 0$. If this were the case, there would be no difference between the political bliss points of poor and rich agents given in Propositions 1 and 2 and thus no interesting political conflict. Consequently, we are more interested in the case where the following condition is satisfied.

CONDITION 1. *The solution to (6) involves $G^D > 0$.*

It can be verified that if the gap between A^H and A^L is small and $\pi'(\tau)$ is large, this condition will be violated. Therefore, this condition imposes that there is a certain degree of inequality in society and raising taxes is not excessively costly, so that the poor would like a higher level of public good provision than the rich. When Condition 1 is satisfied, it also follows that $\tau^D > \tau^N$, and since $\pi(\cdot)$ is strictly increasing, $\pi(\tau^D) > \pi(\tau^N)$ and the size of the bureaucracy is larger in permanent democracy than in permanent nondemocracy. Throughout the rest of the paper, we assume that Condition 1 is satisfied.

3.4. Political Equilibrium with Regime Change

We now look at the more interesting case with regime change—that is, where at date $t = 0$, the rich are in power and from then on there will be elections. We start with a series of lemmas. The following lemma establishes three useful results: (i) with efficient state institutions, the rich will choose their political bliss point as in Proposition 2; (ii) the party representing the poor, party P, being elected to office is an *absorbing state*, so that once the party of the poor is elected, the results of Proposition 1 apply subsequently; and (iii) the structure of equilibrium policies starting with an inefficient state, that is, $I_{t-1} = 0$.

LEMMA 3.

1. *In a MPE, if $d_t = R$ and $I_{t-1} = 1$, then $w_t = (1 - \tau^N)A^L + h$, $X_t = \pi(\tau^N)$, $G_t = G^N \equiv 0$, and τ^N is given by (7).*
2. *If $d_t = P$, then $d_{t'} = P$ for all $t' \geq t$, and equilibrium policies at all dates $t' > t$ are given by (5).*
3. *Suppose that $I_{t-1} = 0$, then $w_t = h/q_0$. Moreover, if $d_t = R$, then $G_t = G^E \equiv 0$, and if $d_t = P$, then $G_t = \hat{G}^D$ given by the solution to the following maximization*

program:

$$\max_{\tau,G} \ (1-\tau)\,A^L + G$$
$$\text{subj. to: } G = (1-n)\,\tau A^H + [n - \pi(\tau)]\,\tau A^L - \frac{h}{q_0}\pi(\tau) - K. \tag{8}$$

The most important result is part 2 of this lemma, which establishes that party P being elected to office is an absorbing state. The intuition for this result is as follows. Once the party of the poor wins an election, they will choose their preferred policy vector, which includes $I_t = 1$, and given an efficient state, bureaucrats will have no reason to support the rich party and the poor will continue to win elections in all future periods and the organization of the state will continue to be efficient. An efficient organization of the state ensures that bureaucrats receive no rents and receive the same payoff as poor producers. Thus they will also support party P, and the political bliss point of the poor will be implemented in all future periods. This lemma also implies that when $I_{t-1} = 1$, that is, when the state is efficient, the rich will not be able to win a majority. This is related to the basic idea of our approach: the rich can only convince bureaucrats to vote for their party by *committing* to giving them rents and this can only be achieved when the organization of the state is inefficient, that is, $I_{t-1} = 0$.

Given the characterization of continuation equilibria in Lemma 3, the key question is whether the party of the rich, party R, can ever win an election starting with $I_{t-1} = 0$. The following lemma answers this question.

LEMMA 4. *In a MPE, $d_t = R$, that is, the rich will win the election at time t, if $I_{t-1} = 0$,*

$$(1-q_0)\frac{h}{q_0} > (1-\tau^D)A^L + G^D + \frac{1-\beta}{\beta}\hat{G}^D, \tag{9}$$

and

$$X_t \geq n - \frac{1}{2}, \tag{10}$$

where G^D is given by (5), $\hat{G}^D$ is given by (8), and τ^D is given by (6).

Lemma 4 determines the conditions under which the bureaucrats will support party R (a rich agent running for office) and will be numerous enough to give them the majority. Condition (10) requires the size of the bureaucracy to be sufficient to give the majority to party R when all bureaucrats vote with the rich. Nevertheless, $n - 1/2$ may not be the actual size of bureaucracy. In particular, at $X = n - 1/2$, the government budget may not balance. To ensure that it does, we need to consider two cases separately.

Let us first define τ^E as the tax rate that party R would choose as its unconstrained optimal policy to finance the investment in infrastructure, K, given that bureaucratic wages are equal to $w = h/q_0$. Clearly, τ^E is given by the unique solution to the equation

$$\pi(\tau^E)\frac{h}{q_0} - (1-n)\tau^E A^H - [n - \pi(\tau^E)]\tau^E A^L + K = 0. \tag{11}$$

In other words, τ^E balances the government budget when the minimum number of bureaucrats necessary to avoid tax evasion, $X = \pi(\tau^E)$, are employed.

The first case corresponds to the one where $\pi(\tau^E) \geq n - 1/2$, so that the unconstrained optimal size of bureaucracy for party R is also sufficient to make sure that (10) is satisfied and the rich have a majority.

The second case applies when this inequality does not hold, that is, when $\pi(\tau^E) < n - 1/2$. In this case, the unconstrained optimal policy for the rich would *not* satisfy (10), and party R cannot win the election with the minimum number of bureaucrats. Instead, party R can win an election only if $X \geq n - 1/2$, and with this larger size of bureaucracy, budget balance requires the greater tax rate $\hat{\tau}^E$ given by the solution to

$$\left(n - \frac{1}{2}\right)\frac{h}{q_0} - (1-n)\hat{\tau}^E A^H - \frac{1}{2}\hat{\tau}^E A^L + K = 0. \tag{12}$$

It can be verified that whenever $n - 1/2 > \pi(\tau^E)$, we also have $\hat{\tau}^E > \tau^E$, and whenever $n - 1/2 \leq \pi(\tau^E)$, $\hat{\tau}^E \leq \tau^E$. This implies that the size of the bureaucracy necessary for the rich to form a winning coalition is the maximum of $\pi(\tau^E)$ and $n - 1/2$, and correspondingly, the tax rate that party R needs to set is $\max\{\tau^E, \hat{\tau}^E\}$.

The results so far have provided the necessary conditions for the rich to be able to generate sufficient votes from the bureaucrats to remain in power. It remains to check whether the rich prefer to pursue this strategy and commit to an inefficient state in order to maintain political power in democracy. The following lemma answers this question.[21]

LEMMA 5. *Suppose that (9) holds. Then the rich prefer to set $I_t = 0$ for all t if one of the following hold:*

$$\begin{aligned} &\tau^E \geq \hat{\tau}^E \text{ and } (\tau^D - \tau^E)A^H > G^D; \text{ or} \\ &\tau^E < \hat{\tau}^E \text{ and } (1-\hat{\tau}^E)A^H > (1-\beta)(1-\tau^E)A^H + \beta[(1-\tau^D)A^H + G^D], \end{aligned} \tag{13}$$

where G^D is given by (5), τ^D is given by (6), τ^E is given by (11), and $\hat{\tau}^E$ is given by (12).

Now putting all these lemmas together we obtain the main result of this section.[22]

21. If Condition 1 were not satisfied, the conditions in Lemma 5 could never be satisfied. In particular, when Condition 1 does not hold, we have $G^D = 0$ and $\tau^D = \tau^E$, so that neither part of condition (13) could hold. This is a direct consequence of the fact that a significant conflict in policies between the rich and the poor is necessary for the rich to set up an inefficient system of patronage politics.

22. It can be verified that the set of parameter values where $I_t = 0$ emerges as an equilibrium in Proposition 3 is nonempty. A straightforward way of doing this is to consider high values of β as in Proposition 5. Note also that Proposition 3 does not cover the case in which one of (9) and (13) holds as equality; in this case the MPE is no longer unique. It is straightforward to see that in such a case, either the rich or the poor party could receive the majority of the votes, or the rich could be indifferent between maintaining an inefficient and an efficient state. We do not describe the equilibrium in these cases to save space.

PROPOSITION 3. *Consider the political environment with emerging democracy. If (9) and (13) hold, then there exists a unique MPE. In this equilibrium, the rich elite choose* $I_t = 0$ *for all* $t \geq 0$*, the rich party R always remains in power and the following policies are implemented:*

$$w_t = \frac{h}{q_0}, \quad X_t = \max\{\pi(\tau^E), n - 1/2\}, \quad G_t = G^E \equiv 0, \quad \text{and} \quad \tau_t = \max\{\tau^E, \hat{\tau}^E\},$$

where τ^E *is given by (11) and* $\hat{\tau}^E$ *is given by (12).*

If, on the other hand, one or both of (9) and (13) hold with the reverse inequality, the unique MPE involves $I_t = 1$ *in the initial period, and for all* $t \geq 1$, $d_t = P$ *and the unique policy vector is given by (5).*

Proposition 3 is our first major result. It establishes the possibility that the rich elite, who are in power at time $t = 0$, may choose an inefficient state organization and a large (inefficient) bureaucracy as a way of credibly committing to providing rents to bureaucrats. This enables them to create a majority coalition consisting of themselves and the bureaucrats, and thus capture democratic politics. This coalition implements policies that support low redistribution and low provision of public goods, but creates high rents for bureaucrats. Perhaps more interestingly, after $t = 1$, even when the society is democratic, the inefficient state institutions *persist* and the rule of the rich continue. This is in spite of the fact that at any date these inefficient institutions can be reformed at no cost and made more efficient. The reasoning is related to the formation of the coalition between the rich and the bureaucrats in the first place. The rich realize that they will be able to maintain power only by keeping an inefficient state structure and creating sufficient rents for bureaucrats. If these rents disappear, bureaucrats will ally themselves with the poor, since their net income will be the same as the net income of poor producers (recall parts 1 and 2 of Lemma 3). It is precisely the presence of inefficient state institutions creating rents for the bureaucrats that induces them to support the policies of the rich. Recognizing this, when in power the rich choose to maintain the inefficient state structure. At the next date, the party representing the rich receives the support of the bureaucrats and the rich; consequently, it remains in power and the cycle continues. The model therefore generates a political economy theory for both the emergence and the persistence of inefficient state institutions.[23]

It is also noteworthy that even though taxes are lower in the equilibrium with inefficient state than they would have been under permanent democracy (recall Proposition 1 and Lemma 5), the size of the bureaucracy can be greater than under permanent democracy. This could be the case when the rich elite hire more bureaucrats than necessary for preventing tax evasion in order to create a majority in favor of the persistence of the inefficient state—that is, in the case where $X > \pi(\tau^E)$. In particular, note that bureaucracy will be more numerous under the control of the elite than in

23. The nature of persistence here is different from the persistence of policies arising in Coate and Morris (1999), Hassler et al. (2003), or Acemoglu and Robinson (2008), because the focus is not on persistence of a certain set of collective decisions within a given institutional framework, but on the persistence of the inefficiency of state institutions.

democracy whenever $\pi(\tau^D) < n - 1/2$. Since in this case equation (13) implies that $\tau^E < \tau^D$, we must also have $\pi(\tau^E) < \pi(\tau^D) < n - 1/2$ and thus

$$X > \pi(\tau^E).$$

Consequently, the rich not only choose an inefficient state organization, but they also choose *overemployment* of bureaucrats, in the sense that bureaucracy is now unnecessarily large and the number of bureaucrats is strictly greater than that necessary for tax inspection. The capture of democratic politics by the rich elite therefore creates an inefficient state, with poorly monitored and overpaid bureaucrats, and also leads to a situation in which the capacity of the state is not fully utilized. These inefficiencies imply that the allocation of resources in a captured democracy is worse than in a nondemocracy (or than in a perfectly functioning democracy). Naturally, these inefficiencies have a political rationale, which is to increase the number of bureaucrats that will vote for the party aligned with the rich, so that the rich can maintain political power in the future.

Interestingly, because creating an inefficient bureaucracy is more costly than creating an efficient one (which is smaller and gives bureaucrats no rents), the citizens are *worse off* in a nonconsolidated (emerging) democracy, where they are taxed at rate $\max\{\tau^E, \hat{\tau}^E\}$, than they would be under a consolidated nondemocracy, where they are only taxed at rate $\tau^N < \max\{\tau^E, \hat{\tau}^E\}$. Moreover, the rich are also worse off in this equilibrium than they would be in a permanent nondemocracy, since they are paying higher wages to bureaucrats and possibly employing an excessive number of them.

3.5. Comparative Statics

We next investigate the conditions under which the equilibrium involves the emergence and persistence of inefficient state institutions. The following proposition establishes that a certain degree of inequality between the poor and the rich (that is, a high level of A^H/A^L), a sufficiently high discount factor, β, and intermediate bureaucratic rents, $(1 - q_0)h/q_0$, are necessary for the emergence of inefficient state institutions.

PROPOSITION 4. *Consider an economy characterized by the parameters* $(\beta, n, A^L, A^H, K, h, q_0)$ *and the function* $p(\cdot)$. *Holding all other parameters constant, we have*

(1) *there exists* $a > 1$ *such that if* $A^H/A^L \leq a$, *then the state is always efficient, that is,* $I_t = 1$;
(2) *there exist* $a' > 1$ *and* $\bar{\beta} \in (0, 1)$ *such that as long as* $A^H/A^L \geq a'$, $\beta \leq \bar{\beta}$ *implies* $I_t = 1$;
(3) *there exists* $\underline{\theta} > 0$ *and* $\bar{\theta}$ *such that if* $(1 - q_0)h/q_0 \notin (\underline{\theta}, \bar{\theta})$, *then* $I_t = 1$.

The first part of the proposition implies that a certain level of inequality is necessary for the emergence of an inefficient state. This is intuitive; with limited inequality, democracy will not be redistributive and it will not be worthwhile for the rich to set up an inefficient bureaucracy in order to keep the poor away from power. The second part implies that the high discount factor is also necessary for the emergence of the inefficient state. This follows because bureaucrats vote for party R as an *investment*, that is, to obtain higher returns in the future. Instead, if they deviate and vote for party P, in the current period they receive both the same high wages (since $I_t = 0$) and the positive level of public good provided by party P, $\hat{G}^D > 0$. If their discount factor were very small, it would be impossible for rich agents to convince bureaucrats to support their party.[24] Finally, the third part of the proposition implies that bureaucratic rents need to take intermediate values. If bureaucratic rents are very small, bureaucrats would not support the party of the rich. If they are very large, it is prohibitively costly for the rich to control democratic politics.

While Proposition 4 shows that a certain degree of inequality is necessary for $I_t = 0$, it does not establish that inequality has a monotonic effect on the likelihood of an inefficient state. The next proposition establishes this result under somewhat more restrictive assumptions. In this proposition, by greater inequality we mean a mean-preserving spread of the income distribution in the economy, that is, a simultaneous increase in A^H and decrease in A^L such that mean *potential* income, $Y = (1 - n)A^H + nA^L$, remains constant.

PROPOSITION 5. *Suppose that $\pi(\tau)$ is log-concave in τ and τ^D given by (6) satisfies $\tau^D < 1 - \pi(\tau^D) < 1$. Then there exists $\tilde{\beta} \in (0, 1)$ such that for all $\beta \geq \tilde{\beta}$, greater inequality makes the inefficient state equilibrium, that is, $I_t = 0$, more likely.*

The proof of this proposition is long and somewhat involved. It is provided in Acemoglu, Ticchi, and Vindigni (2006) and we omit it here to save space. Note that the condition that $\pi(\tau)$ is log-concave is not very restrictive. For example, any $p(x)$ that takes the power function form, that is, $p(x) = P_0 x^\alpha$ for $P_0 > 0$ and $\alpha \in (0, 1)$, satisfies this condition. The condition that $\tau^D < 1 - \pi(\tau^D) < 1$ is also natural; if this condition were violated, we would have that the utility of the poor in democracy $(1 - \tau^D)A^L + G^D$ would be non-increasing in A^L.

In addition to generalizing the first part of Proposition 4, Proposition 5 implies that taxes (and public spending) can be higher in more equal societies, because unequal societies are more likely to create inefficient bureaucracies to limit taxation and public spending. This result therefore presents an alternative explanation to the often-discussed negative cross-sectional correlation between inequality and redistribution (see for example Perotti 1996; Bénabou 2000).

24. Robinson (2001) and Acemoglu and Robinson (2008) also obtain the result that higher discount factors may lead to greater inefficiencies. However, in these models the source of inefficiency is very different. In particular, inefficient political equilibria arise when pivotal agents—elites or rulers—are sufficiently patient and thus take inefficient actions in order to secure their future political survival.

4. Extensions

In this section, we discuss the robustness of the main conclusions of our benchmark model and a number of extensions.

4.1. Subgame Perfect Equilibria

We have so far focused on MPE. A natural question is whether similar insights apply without the restriction to Markovian strategies. This is important since commitment and credibility problems are at the center of our theory of emergence and persistence of inefficient states and we like to ensure that these are not imposed artificially by the concept of MPE.

Our brief analysis in this section shows that similar insights apply when we focus on the subgame perfect equilibria (SPE). The reason is that the elite have an equilibrium (credible) commitment to redistributing to bureaucrats that the poor lack; the poor cannot commit to keeping an inefficient state and paying efficiency wages to bureaucrats, because, when in power, they will want to increase tax revenues, and reform or downsize bureaucracy. As in the analysis so far, we maintain the assumption of individual anonymity, which implies that individual histories are not observed and thus future hiring decisions for bureaucracy cannot be conditioned on whether an individual was previously a bureaucrat.

To state the main result of this section, recall the notation in Section 3.1 and let $\mathcal{H}^t$ denote the set of all possible histories of the game up to stage t (which includes the state vector S up to time t). Then a (possibly non-Markovian) strategy profile can be represented as

$$\tilde{\sigma} : \mathcal{H}^t \times \{L, H, B\} \to \{P, R\} \times \{0, 1\}^4 \times [0, 1] \times [0, n] \times \mathbb{R}_+,$$

and specifies the behavior of each type of agent (poor producer, rich producer, and bureaucrat) as a function of history $h^t \in \mathcal{H}^t$. An SPE is a mapping $\tilde{\sigma}^*$ that is the best response to itself at every possible history $h^t \in \mathcal{H}^t$. As with MPE, we rule out weakly dominated strategies, so that an SPE refers to a strategy profile that is a best response to itself and does not involve the play of weakly dominated strategies. The following proposition characterizes the SPE that maximizes the date $t = 0$ utility of the elite when condition (9) and (13) hold and when $\pi(\tau^E) > n - 1/2$. It shows that the best SPE from the viewpoint of the elite is qualitatively similar to the MPE in Proposition 3.

PROPOSITION 6. *Consider the political environment with emerging democracy, and suppose that conditions (9) and (13) hold and that $\pi(\tau^E) < n - 1/2$, where τ^E is defined as in (11). Then the SPE that maximizes the date $t = 0$ utility of the elite involves a bureaucracy of size $X = n - 1/2$, no redistribution ($G = 0$), and party R winning the election at each date.*

This proposition states that, though some details of the best SPE from the viewpoint of the elite may be different than in the MPE of Proposition 3, the qualitative features of the two equilibrium allocations are very similar. In both cases, the elite are able to retain their political power in democracy and avoid redistribution. In particular, since $\pi(\tau^E) < n - 1/2$, in both cases they achieve this by creating an oversized bureaucracy. The key intuition is once again that the poor (party *P*) do not have a credible commitment to paying high salaries to bureaucrats. More specifically, since, $\pi(\tau^E) < n - 1/2$, the elite requires an *oversized* bureaucracy, that is, $X = n - 1/2 > \pi(\tau^E)$, to win elections. This implies that as soon as party *P* comes to power, its best response is to reduce the size of bureaucracy (which it can do without violating (1)) and increase redistribution—thus increasing the utility of the representative poor agent. Individual anonymity implies that poor agents who were previously bureaucrats do not have a higher probability of being hired back into bureaucracy if party *R* ever comes back to power. Consequently, after the downsizing of bureaucracy, all poor agents who are currently not bureaucrats will vote for party *P*, which will then reform the organization of the state ($I_t = 1$) and win all future elections. Consequently, there is no SPE in which party *P* can pay efficiency wages to bureaucrats (except possibly in the first period in which it is in power if $I = 0$ in that period).

Turning to the strategies of the elite, part 3 of Lemma 3 still applies and shows that when $I_{t-1} = 0$, any party that is in power must pay at least the minimum efficiency wage $w_t = h/q_0$ (otherwise there will be no production). Moreover, Lemma 5 also applies and shows that when (9) and (13) hold, the elite prefer the inefficient state equilibrium to living under democracy. These observations imply that the elite can, and would like to, build a large bureaucracy and hold on to power. In fact, the proof in the Appendix shows that the elite may be able do this more cheaply than in the MPE of Proposition 3, and the relevant condition for this is provided in the Appendix. Regardless of whether this condition holds the important conclusion is that the SPE has a similar structure to the MPE in Proposition 3; in particular, it involves the elite maintaining political power, setting up an oversized bureaucracy, and preventing redistribution.

For brevity, Proposition 6 is stated for the case in which (9) and (13) hold and $\pi(\tau^E) < n - 1/2$. If $\pi(\tau^E) \geq n - 1/2$, it may be possible for party *P* to commit to bureaucratic wages greater than $(1 - \tau^D)A^L + G^D + h$, but in any such allocation, there will still be redistributive taxation and the elite will be worse off than in the equilibrium of Proposition 3. Therefore, the best SPE from the viewpoint of the elite again involves party *R* being in power and $G = 0$, though now the size of bureaucracy may be smaller than $\pi(\tau^E)$. It can also be verified that if these conditions fail to hold, there may still exist an SPE with elite control, because now the rich can promise even a higher wage to bureaucrats. Thus, SPE may in fact involve a greater likelihood of the emergence and persistence of an inefficient state.

4.2. Equilibrium When Bureaucrats Can Be Fired

The main result of the previous section, Proposition 3, was derived under the assumption that bureaucrats cannot be fired when they are caught shirking. This simplified the analysis by enabling us to write the incentive compatibility constraint of bureaucrats in the form of condition (2). We now allow bureaucrats to be fired when they are caught shirking. It is clear that from the viewpoint of discouraging shirking, a contract which commits to firing bureaucrats when they are caught shirking is optimal.

To study the structure of equilibria when bureaucrats can be fired, we focus on a stationary equilibrium, where today and in all future periods the tax rate is equal to $\hat{\tau}$, the wage rate for bureaucrats is $\hat{w}$, and the probability of getting caught is $\hat{q}$. In this case, if

$$\frac{\hat{w}-h}{1-\beta} < \hat{q}\beta\frac{(1-\hat{\tau})A^L}{1-\beta} + (1-\hat{q})\left(\hat{w} + \beta\frac{\hat{w}-h}{1-\beta}\right),$$

then bureaucrats would prefer to shirk. This is intuitive since the left-hand side of this expression is what the individual would receive by exerting effort at every date, whereas the right-hand side is the payoff to deviating for one period, and then switching to exerting effort from then on (implicitly using the one-step ahead deviation principle, see Fudenberg and Tirole, 1991, Chapter 4). In particular, the right-hand side has the individual getting caught with probability $\hat{q}$, receiving nothing today and the wage of a low-skill producer from then on, and not getting caught with probability $1-\hat{q}$, in which case he receives $\hat{w}$ today and then receives the discounted version of the left-hand side (as he switches back to exerting effort). A bureaucrat who loses his job always receives the wage of a low-skill producer from then on, since along the equilibrium path, there will be no further hiring into bureaucracy. Rearranging the previous expression, we conclude that whenever the following incentive compatibility constraint is violated, bureaucrats will shirk. The relevant incentive compatibility constraint is

$$\hat{w} \geq \beta(1-\hat{\tau})A^L + \frac{(1-\beta(1-\hat{q}))h}{\hat{q}}. \tag{14}$$

Given this modified incentive compatibility condition, all of the results from the previous section apply with appropriate modifications. The following proposition can be proved with identical arguments to those in Section 3 (see Acemoglu, Ticchi, and Vindigni, 2006).

PROPOSITION 7. *Consider the political environment with emerging democracy and suppose that bureaucrats can be fired if caught shirking. Let G^D and τ^D be given by (5) and (6), $\tilde{\tau}^E$ be such that*

$$\lambda_m\left[\beta(1-\tilde{\tau}^E)A^L + \frac{(1-\beta(1-q_0))h}{q_0}\right] - (1-n)\tilde{\tau}^E A^H - (n-\lambda_m)\tilde{\tau}^E A^L + K = 0,$$

where $\lambda_m \equiv \max\{\pi(\tilde{\tau}^E), n-1/2\}$, and $\hat{G}^D$ be given by (8) with

$$\beta(1-\tilde{\tau}^E)A^L + (1-\beta(1-q_0))h/q_0$$

replacing h/q_0. Suppose that

$$\beta(1-\tilde{\tau}^E)A^L + (1-\beta)(1-q_0)h/q_0 > (1-\tau^D)A^L + G^D + (1-\beta)\hat{G}^D/\beta$$

and $(\tau^D - \tilde{\tau}^E)A^H > G^D$*. Then the unique MPE is as follows: the rich elite choose* $I_t = 0$ *in the initial period and for all t thereafter, the rich party always remains in power and the following policies are implemented at all dates:*

$$\begin{aligned} w_t &= \beta(1 - \tilde{\tau}^E)A^L + (1 - \beta(1 - q_0))h/q_0, \\ X_t &= \max\{\pi(\tilde{\tau}^E), n - 1/2\}, \\ G_t &= G^E \equiv 0, \quad \text{and} \quad \tau_t = \tilde{\tau}^E. \end{aligned}$$

Proposition 7 demonstrates that the results in Proposition 3 generalize to the environment where bureaucrats can be fired if caught shirking. One important difference is worth noting, however. In our main analysis, Proposition 4 showed that a higher discount factor, β, makes the emergence of an inefficient state more likely. Instead, when bureaucrats can be fired, the relationship between the discount factor and the emergence of inefficient states is more complex. Higher β again increases the importance that bureaucrats attach to future rents, but it also reduces the level of rents, because being fired from bureaucracy becomes more costly.

4.3. Political Equilibrium Citizen-Candidates

We have so far assumed that democracy involves competition between the two parties representing the interests of the poor and the rich. If a party representing the interests of bureaucrats forms, bureaucrats might vote for that party and the coalition between the rich and the bureaucrats, choosing low public good provision and low taxes, might not materialize.

In Acemoglu, Ticchi, and Vindigni (2006), we showed that all of the results in Section 3 continue to apply when the set of parties (citizen-candidates) running for office is endogenized, so that the bureaucrats can also form a party and contest elections. We showed that in this extended environment the poor vote for a poor candidate and the rich vote for a rich candidate. Moreover, when the party of the poor cannot win an election, the poor never support the bureaucrats; instead, they vote for the party of the rich, because a party representing the bureaucrats would impose higher taxes but would not provide public goods (whereas the party of the rich at least sets lower taxes). This analysis overall confirms that our theory of emergence and persistence of inefficient states does not depend on artificially restricting the set of party platforms.

4.4. Bureaucratic Corruption

A final extension involves modifying the basic model so that the moral hazard problem on the side of bureaucrats is not related to their effort but to whether or not they accept bribes from producers evading taxes. This source of moral hazard problem is arguably as important as the effort choice of bureaucrats. More importantly, this type of moral hazard problem leads to an interesting pattern of *de facto regressive taxation* as a result of successful patronage politics by the rich elite.

Because of space restrictions, we will only sketch this extension, referring the reader to Acemoglu, Ticchi, and Vindigni (2006) for details. The main difference from our benchmark model in Section 2 is that the bureaucrats no longer have an effort choice, but instead, they can accept bribes from producers that have evaded taxes. The efficiency of the state, $I_t \in \{0, 1\}$, now determines whether bureaucrats accepting bribes are caught or not. In particular, when $I = 1$, there is an efficient organization of the state and corruption is detected with probability $q(I = 1) = 1$. When $I = 0$, the state organization is inefficient and corruption is detected with probability $q(I = 0) = q_0 < 1$. We also assume that each bureaucrat can be matched with at most one producer, and thus receives at most the bribe payments from a single producer, denoted by $b_t \geq 0$. The incentive compatibility constraint for bureaucrats (2) is now replaced by the following "no bribe constraint":

$$w_t \geq (1 - q_0)(w_t + b_t), \tag{15}$$

where b_t is the bribe offered to the bureaucrat by a producer. Intuitively, the right hand side of (15) represents the expected return of a bureaucrat that accepts a bribe b_t, given by the sum of the wage and the bribe, weighted by the probability of not being detected. If condition (15) does not hold, it is not possible to prevent the corruption of bureaucrats by producers.

Given this setup, the formal analysis parallels that in Section 3 and leads to the following result (see Acemoglu, Ticchi, and Vindigni 2006 for details).

PROPOSITION 8. *Consider the political environment with emerging democracy. If q_0 is sufficiently small and inequality is sufficiently large, then the unique MPE is one in which the rich elite choose $I_t = 0$ in the initial period and for all t thereafter, the rich party R always remains in power. They provide no public goods and pay relatively low wages to bureaucrats, so that in equilibrium the rich evade taxes and pay a bribe equal to $b = A^L$ when inspected. The poor do not evade taxes.*

The most interesting result in Proposition 8 is that, when they are able to capture democratic politics, the rich do not pay any taxes at all. Instead, they (sometimes) pay bribes equal to the tax burden on poor agents, A^L. This implies that patronage politics turns de jure proportional taxation into a de facto regressive one. In other words, when the rich elite are able to set up an inefficient state and receive the support of bureaucrats, they are not only able to limit redistribution and public good provision, but they are also able to shift most of the burden of taxation to the poor. Consequently, the tax rate faced by the poor may be higher when corruption is possible than in the baseline model where both rich and poor pay taxes.

5. Concluding Remarks

Inefficiencies in the bureaucratic organization of the state are often viewed as an important factor in retarding economic development. Many sociological accounts of comparative development emphasize the role of state capacity (or lack thereof) in explaining why some societies are able to industrialize and modernize (see for example Evans 1995; Migdal 1988). In addition, inefficient state organizations appear to coincide with limited amounts of public good provision and redistribution towards the poor. Existing approaches do not address the question of why certain societies choose or end up with such inefficient organizations and do not clarify the relationship between inefficient state organizations and limited redistribution.

We presented a simple theory of the emergence and persistence of inefficient states, in which the organization of the public bureaucracy is manipulated by the rich elite in order to influence redistributive politics. In particular, by instituting an inefficient state structure, the elite are able to use patronage and capture democratic politics. This enables them to limit the extent of redistribution and public good provision. Captured democracies not only limit redistribution, but also create a number of major distortions: the structure of the state is inefficient, there is too little public good provision and there may be overemployment of bureaucrats.

We also showed that an inefficient state creates its own constituency and tends to persist over time. Intuitively, an inefficient state structure creates more rents for bureaucrats than would an efficient state structure. When the median (poor) agent comes to power in democracy, he will reform the structure of the state to make it more efficient so that the higher taxes can be collected at lower cost (especially in terms of lower rents for bureaucrats). Anticipating this, when the organization of the state is inefficient, bureaucrats support the rich, who set lower taxes but pay high wages to bureaucrats. In order to generate enough political support, the coalition of the rich and the bureaucrats may not only choose an inefficient organization of the state, but they may further expand the size of bureaucracy so as to gain additional votes.

The model shows that an equilibrium with an inefficient state is more likely when there is greater income inequality and when democratic taxes are anticipated to be higher. An interesting implication of this result is that inequality and redistribution may be negatively correlated because higher inequality makes the capture of democratic politics more likely.

The general message from our analysis is that "not all democracies are created equal"; while some democracies will adopt policies that redistribute to poorer segments of the society, others may become captured by traditional elites. These captured democracies not only choose low levels of redistribution, but, as part of their political rationale for survival, they also typically create a range of inefficiencies. While we focused on the emergence and persistence of inefficient state structures in democratic or quasi-democratic polities, such inefficiencies are also present in other systems, most notably, in socialist economies and other dictatorships. We conjecture that the unwillingness of leaders to reform inefficient state organizations in these regimes may also be related to their attempts to forge coalitions with bureaucrats and state employees, who are often more powerful in these regimes than in democracies. Other mechanisms for democratic capture and reasons for maintaining inefficient bureaucracies in nondemocratic systems constitute fruitful areas for future research.

Appendix Proofs

Proof of Proposition 1. By Lemma 2, for all $j \in \mathcal{L}$, $v_t^j = P$. Under permanent democracy, the poor can vote and form the majority starting at $t = 0$, thus $d_t = P$ for all t. Then the payoff to the decisive voter $j' \in \mathcal{L}$ can be written as $V_t^{j'} = (1-\tau)A_t + G_t + \beta V_{t+1}^{j'}(\sigma^*)$, where σ^* is the optimal policy and $\beta V_{t+1}^{j'}(\sigma^*)$ is the discounted optimal continuation value. The continuation value $\beta V_{t+1}^{j'}(\sigma^*)$ is unaffected by current policies, thus the optimal policy can be determined as a solution to the following program:

$$\begin{aligned}
\max_{\tau,w,X,I,G} \quad & (1-\tau)A^L + G \\
\text{subj. to:} \quad & \pi(\tau) \leq X, \\
& \max\left\{\frac{h}{q(I_t)}, (1-\tau)A^L + h\right\} \leq w, \\
& G \leq (1-n)\tau A^H + (n-X)\tau A^L - wX - K, \\
& 0 \leq G.
\end{aligned} \tag{A.1}$$

It is evident that $I_t = 1$ relaxes the second constraint relative to $I_t = 0$, so will always be chosen for all $t \geq 0$. Moreover, there cannot be a solution in which any one of the first three constraints is slack (since this would allow an increase in G, raising the value of the objective function), so $X = \pi(\tau)$ and $w = \max\{h, (1-\tau)A^L + h\} = (1-\tau)A^L + h$. Substituting these equalities yields (6) for all periods where $I_t = 1$, that is, for all $t > 0$. Strict convexity of $\pi(\cdot)$ then ensures that τ^D is uniquely defined. □

Proof of Proposition 2. Under permanent nondemocracy, the rich always retain political power and the payoff to a rich individual $j' \in \mathcal{H}$ can be written as $V_t^{j'} = (1-\tau)A^H + G_t + \beta V_{t+1}^{j'}(\sigma^*)$, where σ^* is the optimal policy and $\beta V_{t+1}^{j'}(\sigma^*)$ is the continuation value, which is again unaffected by current policies. Therefore, the optimal policy can be determined as a solution to maximizing $(1-\tau)A^H + G$ subject to the same set of constraints as in (A.1). Once again $I_t = 1$ relaxes the second constraint, so will always be chosen. Moreover, the first three constraints must hold as equalities, so $X = \pi(\tau)$ and $w = \max\{h, (1-\tau)A^L + h\} = (1-\tau)A^L + h$. Finally, in this case $G = 0$, and the strict convexity of $\pi(\cdot)$ again ensures the uniqueness of the solution to (7). □

Proof of Lemma 3. Parts 1 and 3 of this lemma are straightforward and we omit the proofs.

Part 2. The policy vector in (5) is the optimal policy of the citizens in permanent democracy (Proposition 1). Now suppose that party P is in power at time t, and suppose that it chooses the policy vector specified in the lemma. Since this includes $I_t = 1$, the following period, we start with $I_t = 1$ as part of the payoff-relevant state vector. Suppose that σ^* is such that $v(I = 1|B) = P$. Then party P wins the majority at time $t + 1$. Alternatively suppose that $v(I = 1|B) \neq P$, but $X < n - 1/2$. Then, party P

again wins the majority at time $t+1$. In both cases, repeating this argument for the next period shows that party P keeps power at all dates and establishes the lemma.

To complete the proof we only need to rule out the case where $v(I=1|B)=R$ and $X \geq n-1/2$ (the proof to eliminate the case where bureaucrats randomize between the two parties in a way to bring party R to power is identical). Since $v(I=1|B)=R$ and $I=0$ is costly for the rich, party R will choose $I_t=1$. Then from part 1, we obtain that $w_t=(1-\tau^N)A^L+h$, $X_t=\pi(\tau^N)$, and $G_t=G^N\equiv 0$. This implies that the utility of the bureaucrat is the same as a poor producer. Then denoting the utility of a bureaucrat supporting party d by $V^B(d)$, we have

$$V^B(R)=(1-\tau^N)A^L+\beta V^j(\sigma^*)<(1-\tau^D)A^L+G^D+\beta V^j(\sigma^*)=V^B(P),$$

where the inequality follows because the last term is the maximal utility of a poor agent and also the utility of a bureaucrat when party P is in power. Thus $v(I=1|B)=R$ cannot be a best response. □

Proof of Lemma 4. Part 2 of Lemma 3 establishes that $I_{t-1}=0$ is necessary. Now suppose that $I_{t-1}=0$ and consider the scenario in which party R chooses $I_t=0$ and $X_t \geq X_{t-1}$ (so that no current bureaucrat is fired). Consider the case in which individual $j \in \mathcal{X}_t$ is pivotal and chooses $v_t^j=R$ in all future periods. Then, his net per-period payoff and his lifetime utility are $w_t-h=(1-q_0)h/q_0$ and

$$V_t^j=\frac{1}{1-\beta}\frac{(1-q_0)h}{q_0}. \tag{A.2}$$

In contrast, if $j \in \mathcal{X}_t$ were to choose $v_t^j=P$ when pivotal, his value would be

$$V_t^j=\frac{h}{q_0}-h+\hat{G}^D+\beta\hat{V}_{t+1}^j, \tag{A.3}$$

where $\hat{V}_{t+1}^j$ is the continuation value when party P is in power from then on, given by $\hat{V}_{t+1}^j=[(1-\tau^D)A^L+G^D]/(1-\beta)$. This last expression incorporates the fact that if the poor are in power, they reform the bureaucracy, setting $I_t=1$, and that $I=1$ is an absorbing state.

The comparison of (A.2) and (A.3) gives (9)—as a weak inequality—as a necessary condition for bureaucrats to support party R when they are pivotal. Condition (10) is also necessary since, if it were violated, bureaucrats would not be pivotal and party R would receive less than half of the votes even with all of bureaucrats voting $v_t^j=R$. This argument establishes that both (9) and (10) are necessary. Moreover, (9)—as a strict inequality—and (10) are also sufficient to ensure $d_t=R$, since when both of these conditions hold, it is a weakly dominant strategy for bureaucrats to vote for party R whenever $I_{t-1}=0$ and the coalition of bureaucrats and the rich have a majority. □

Proof of Lemma 5. Suppose $v(I=0|B)=R$ (that is, bureaucrats vote for party R whenever the state is inefficient). Under party P, the per period return of the rich is $(1-\tau^D)A^H+G^D$. When $\tau^E \geq \hat{\tau}^E$, party R can remain in power by choosing $I=0$ with the per period return $(1-\tau^E)A^H$. This gives the first part of (13).

For the second part, note that, when in power, party R can always choose its myopic optimum, giving each rich agent utility,

$$V^R = (1-\tau^E)A^H + \beta[(1-\tau^D)A^H + G^D]/(1-\beta).$$

The continuation value $\beta[(1-\tau^D)A^H + G^D]/(1-\beta)$ follows from the observation that since, by assumption, $\tau^E < \hat{\tau}^E$, we have $n - 1/2 > \pi(\tau^E)$ and thus party R will lose the election at the next date. Then Lemma 3 implies that party P will win all elections in all future dates. Alternatively, party R can choose $X = n - 1/2$ and guarantee to be in power forever, but at the expense of taxing the rich at the higher rate $\hat{\tau}^E$. This will give a representative rich agent utility $\hat{V}^R = (1-\hat{\tau}^E)A^H/(1-\beta)$. Comparison of $\hat{V}^R$ with V^R in the previous expression gives the second part of (13). □

Proof of Proposition 3. The first part of the proposition follows immediately from combining Lemma 4 and Lemma 5. When (9) or (13) does not hold, then party P is in power and the second part of the proposition follows immediately from Lemma 3 and Proposition 1. □

Proof of Proposition 4. Inspection of the maximization problem (6) immediately shows that as $A^L \to \mathrm{A}^{\mathrm{H}}$, Condition 1 will be violated and the conditions in (13) cannot hold. Then the first result follows from Lemma 5 and Proposition 3.

For the second part, recall from the discussion of Lemma 5 that some minimal level of inequality, say $A^H/A^L \geq a'$, is necessary for $\hat{G}^D > 0$. Suppose this is the case. From Proposition 3, condition (9) is necessary for $I_t = 0$. Since $\hat{G}^D > 0$, there exists $\beta_0 \in (0, 1)$ such that $(1-q_0)h/q_0 = \beta_0\hat{G}^D/(1-\beta_0)$. Since the sum of the other terms on the right-hand side of (9) is positive, this implies that there exists $\bar{\beta} < \beta_0$ such that for all $\beta \leq \bar{\beta}$ (9) will be violated and thus $I_t = 1$.

For the third part, note that bureaucratic rents are equal to $h/q_0 - h = (1-q_0)h/q_0$, which needs to be greater than or equal to the right-hand side of (9). Let this right-hand side be denoted by $\underline{\theta}$ (and note that $\underline{\theta} > 0$). If $(1-q_0)h/q_0 < \underline{\theta}$, then (9) will be violated and $I_t = 1$. This implies that we need $(1-q_0)h/q_0 \geq \underline{\theta} > 0$. Next observe from (11) that there exists a value of $(1-q_0)h/q_0$, say $\bar{\theta}_0$, such that $\hat{\tau}_E = 1$. It is evident that when $\hat{\tau}_E = 1$, condition (13) cannot be satisfied, thus $I_t = 1$. This implies that for $I_t = 0$, we need $h/q_0 \leq \bar{\theta}_0$ and thus $(1-q_0)h/q_0 \leq \bar{\theta}$. □

Proof of Proposition 6. Consider a history $h^t \in \mathcal{H}^t$ where party P is currently in power and observe that the unique best response of party P is to choose its unconstrained optimal policy. In particular, if $I = 1$, its unique best response is to choose the policy vector $(\tau^D, w^D, G^D, \pi(\tau^D), I = 1)$, where τ^D is as defined in Proposition 1 and $w^D \equiv (1-\tau^D)A^L + h$. If $I = 0$, then its unique best response is to choose the policy vector $(\hat{\tau}^D, w = h/q_0, \hat{G}^D, \pi(\hat{\tau}^D), I = 1)$, where $(\hat{\tau}^D, \hat{G}^D)$ are the solution to the maximization problem in part 3 of Lemma 3. To see why no other policy vector could be a best response, note that both of these policies involve a bureaucracy of size less than $n - 1/2$ and that the unique continuation best response of any poor non-bureaucrat is to support party P (since even with the random likelihood of being hired into bureaucracy, the expected utility from switching to an elite controlled regime for a poor agent is less than the utility of permanent democracy in Proposition 1, and individual anonymity implies that specific individuals, for example those who were previously in bureaucracy, cannot be treated differently). Consequently, in any SPE continuation play following this history h^t, party P will remain in power, and thus there is no reason for it to deviate from the policies that maximize its preferences (or the utility of a representative poor agent).

Next, note that since (9) and (13) are satisfied, it is feasible to have an elite-controlled regime making the elite better off than in permanent democracy. This implies that the best SPE from the viewpoint of the elite will involve party R being in power and choosing no redistribution ($G = 0$), and the size of bureaucracy must be $X = n - 1/2$ in order to ensure electoral victory to party R. To complete the characterization of the best SPE from the viewpoint of the elite it remains to determine the organization of the state, I, and bureaucratic wage, w, in this equilibrium. From part 3 of Lemma 3, if $I = 0$, the bureaucratic wage must be $w = h/q_0$, and (9) and (13) imply that at this wage, bureaucrats are happy to support party R and the elite are better off in this regime. However, the elite may possibly obtain higher utility with $X = n - 1/2$, $I = 1$ and some wage $w^* < h/q$. To see whether this is possible, note that the minimum wage that the elite need to pay to bureaucrats is $w^* = (1 - \tau^D)A^L + G^D + h$, since otherwise bureaucrats would be better off in democracy than under elite-controlled democracy and would vote for party P. Condition (9) implies that w^* is indeed strictly less than h/q. Consequently, if it is a subgame perfect strategy for the elite to choose ($I = 1, w^*$), then the best SPE will involve the policy vector ($\tilde{\tau}^E, w^*, G = 0, I = 1$), where $\tilde{\tau}^E$ is the tax rate that satisfies the government budget constraint and is necessarily strictly less than $\hat{\tau}^E$ as defined in (12) since $w^* < h/q_0$.

We next need to check that this strategy profile is subgame perfect. Suppose bureaucrats use the most severe punishment against elite deviation, which we next describe. Consider a history h^t where party R has always been in power (otherwise, simply consider the part of the history where party R has been in power). Then the vote of each bureaucrat after this history, $v(B|h^t)$, is as follows: if h^t contains $w_{t'} < w^*$ for any $t' \leq t$, then $v(B|h^t) = P$, and if h^t contains $w_{t'} = w^*$ for all $t' \leq t$, then $v(B|h^t) = R$ (that is, if the wage w is ever less than w^*, then bureaucrats vote for party P in all future periods). Given this strategy, the utility of the elite from the policy vector ($\tilde{\tau}^E, w^*, G = 0, I = 1$) is

$$\frac{(1 - \tilde{\tau}^E)A^H}{1 - \beta}.$$

Instead, if party R deviates, the best deviation is to pay the lowest wage to bureaucrats consistent with bureaucrats not leaving the public sector (the policy vector in Proposition 2 with the tax rate τ^N). After this deviation, bureaucrats vote for party P and the allocation in Proposition 1, with the policy vector ($\tau^D, w^D, G^D, \pi(\tau^D), I = 1$), is implemented thereafter, giving the elite utility

$$(1 - \tau^N)A^H + \beta \frac{(1 - \tau^D)A^H + G^D}{1 - \beta}.$$

Therefore, choosing the policy vector $(\tilde{\tau}^E, w^*, G = 0, X = n - 1/2, I = 1)$ is subgame perfect if

$$(1 - \tilde{\tau}^E)A^H \geq (1 - \beta)(1 - \tau^N)A^H + \beta[(1 - \tau^D)A^H + G^D].$$

This condition is similar to the second part of (13). If it is not satisfied, then the elite cannot credibly commit to the bureaucratic wage w^* with $I = 1$ and therefore need to choose the policy vector $(\hat{\tau}^E, w = h/q, G = 0, X = n - 1/2, I = 0)$ as in Proposition 3 (which is preferred to transitioning to permanent democracy since condition (13) holds). Therefore, the best SPE (from the viewpoint of the elite) may involve either the policy vector $(\tilde{\tau}^E, w^*, G = 0, X = n - 1/2, I = 1)$ or $(\hat{\tau}^E, w = h/q, G = 0, X = n - 1/2, I = 0)$. But in both cases, it features party R being in power, $G = 0$, and $X = n - 1/2$ as claimed in the proposition. □

References

Acemoglu, Daron (2005). "Politics and Economics in Weak and Strong States." *Journal of Monetary Economics*, 52, 1199–1226.

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

Acemoglu, Daron and James A. Robinson (2006). *Economic Origins of Dictatorship and Democracy*. Cambridge University Press.

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

Acemoglu, Daron, Davide Ticchi, and Andrea Vindigni (2006). "Emergence and Persistence of Inefficient States." NBER Working Paper No. 12748.

Acemoglu, Daron and Thierry Verdier (1998). "Property Rights, Corruption and the Allocation of Talent: A General Equilibrium Approach." *Economic Journal*, 108, 1381–1403.

Acemoglu, Daron and Thierry Verdier (2000). "The Choice Between Market Failures and Corruption." *American Economic Review*, 90, 194–211.

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*," edited by R. Dornbusch and M. Draghi. MIT Press.

Alesina, Alberto and Allan Drazen (1991). "Why Are Stabilizations Delayed?" *American Economic Review*, 81, 1170–1188.

Becker, Gary S. and Casey B. Mulligan (2003). "Deadweight Costs and the Size of Government." *Journal of Law and Economics*, 46, 293–340.

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

Bertocchi, Garziella and Michael Spagat (2001). "The Politics of Co-Optation." *Journal of Comparative Economics*, 29, 591–607.

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

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

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

Centeno, Miguel (2002). *Blood and Debt: War and Nation-State in Latin America*. Pennsylvania State University Press.

Coate, Stephen and Stephen Morris (1995). "On the Form of Transfers to Special Interests." *Journal of Political Economy*, 103, 1210–1235.
Coate, Stephen and Stephen Morris (1999). "Policy Persistence." *American Economic Review*, 89, 1327–1336.
Cohen, Youssef (1987). "Democracy from Above: The Political Origins of Military Dictatorship in Brazil." *World Politics*, 40, 30–54.
Cukierman, Alex, Sebastian Edwards and Guido Tabellini (1992). "Seigniorage and Political Instability." *American Economic Review*, 82, 537–555.
Debs, Alexandre (2006). "The Wheel of Fortune: Agency Problems in Dictatorships." Working paper, Massachusetts Institute of Technology.
Dixit, Avinash K. (2002). "Incentives and Organizations in the Public Sector: An Interpretative Review." *Journal of Human Resources*, 37, 696–727.
Egorov, Georgy and Konstantin Sonin (2005). "Dictators and Their Viziers: Agency Problems in Dictatorships." New Economic School, Moscow, mimeo.
Epstein, Stepan (2000). *Freedom and Growth. The Rise of States and Markets in Europe, 1300–1750*. Routledge.
Etzioni-Halevy, Eva (1983). *Bureaucracy and Democracy*. Routledge.
Evans, Peter B. (1992). "The State as Problem and Solution: Predation, Embedded Autonomy, and Structural Change." In "*The Politics of Economic Adjustment*," edited by S. Haggard and R. Kaufman. Princeton University Press.
Evans, Peter B. (1995). *Embedded Autonomy*. Princeton University Press.
Fiorina, Morris P. and Roger G. Noll (1978). "Voters, Legislators and Bureaucracy: Institutional Design in the Public Sector." *American Economic Review*, 68, 256–260.
Fudenberg, Drew, and Jean Tirole (1991). *Game Theory*. MIT Press.
Gay, Robert (1990). "Popular Incorporation and Prospects for Democracy: Some Implications of the Brazilian Case." *Theory and Society*, 19, 447–463.
Geddes, Barbara (1991). "A Game Theoretic Model of Reform in Latin American Democracies." *American Political Science Review*, 85, 371–392.
Gershenson, Dmitriy and Herschel I. Grossman (2001). "Cooption and Repression in the Soviet Union." *Economics and Politics*, 13, 31–47.
Glazer, Amihai (1989). "Politics and the Choice of Durability." *American Economic Review*, 79, 1207–1213.
Hassler, John, José V. Rodrıguez Mora, Kjetil Storesletten, and Fabrizio Zilibotti (2003). "The Survival of the Welfare State." *American Economic Review*, 93, 87–112.
Herbst, Jeffery I. (2000). *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton University Press.
Kohli, Atul (2004). *State-Directed Development: Political Power and Industrialization in the Global Periphery*. Cambridge University Press.
Levi, Margaret (1989). *Of Rule and Revenue*. University of California Press.
Lizzeri, Alessandro and Nicola Persico (2001). "Provision of Public Goods under Alternative Electoral Systems." *American Economic Review*, 91, 225–239.
Migdal, Joel (1988). *Strong Societies and Weak States: State-Society Relations and State Capabilities in the Third World*. Princeton University Press.
Osborne, Martin J. and Al Slivinsky (1996). "A Model of Political Competition with Citizen-Candidates." *Quarterly Journal of Economics*, 111, 65–96.
Perotti, Roberto (1996). "Growth, Income Distribution, and Democracy: What the Data Say." *Journal of Economic Growth*, 1, 149–187.
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, 325–345.
Rauch, James and Peter B. Evans (2000). "Bureaucratic Structure and Bureaucratic Performance in Less Developed Countries." *Journal of Public Economics*, 75, 49–71.
Robinson, James A. (2001). "When is a State Predatory?" Working paper, Harvard University.
Robinson, James A. and Ragnar Torvik (2005). "White Elephants." *Journal of Public Economics*, 89, 197–210.

Rodrik, Dani (1995). “The Political Economy of Trade Policy.” In “*Handbook of International Economics*,” Vol. 3, edited by G. Grossman and K. Rogoff. North-Holland.

Roett, Riordan (1999). *Brazil: Politics in a Patrimonial Society*. Praeger.

Rothstein, Bo and Eric M. Uslaner (2005). “All for All: Equality, Corruption, and Social Trust.” *World Politics*, 58, 41–72.

Saint-Paul, Gilles (1996). “Exploring the Political Economy of Labour Market Institutions.” *Economic Policy*, 11, 265–315.

Shleifer, Andrei and Robert W. Vishny (1994). “Politicians and Firms.” *Quarterly Journal of Economics*, 109, 995–1025.

Tabellini, Guido and Alberto Alesina (1990). “Voting on the Budget Deficit.” *American Economic Review*, 80, 37–49.

Tilly, Charles (1990). *Coercion, Capital and European States, AD 990–1990*. Blackwell.

Weyland, Kurt (1996). “Obstacles to Social Reform in Brazil’s New Democracy.” *Comparative Politics*, 29, 1–22.

World Bank (2004). *Doing Business in the World: Understanding Regulation*. Oxford University Press.

Wilson, John D. (1990). “Are Efficiency Improvements in Government Transfer Policies Self-Defeating in Political Equilibrium.” *Economics and Politics*, 2, 241–258.