Weak, Despotic, or Inclusive? How State Type Emerges from State vs. Civil Society Competition*

Daron Acemoglu[†] **James A. Robinson**[‡]

February 5, 2022

Abstract

We develop a theory of the accumulation of state capacity as the outcome of a political competition between elites and (civil) society. State capacity is accumulated by elites and is both productive but also useful in controlling society. Society, however, can fight back and accumulate its own capacity, facilitating collective action. The theory leads to three distinct equilibria depending on initial conditions. One type, a weak state, emerges when society is strong relative to the elite. Another, a despotic state, originates where the elite is initially relatively powerful. A third type, an inclusive state, emerges when the elite and society are more evenly matched. The theory has several key implications; first, variation in state capacity does not require large structural differences; second, inclusive states have the highest levels of state capacity in the long run; third, the effects of shocks or external threats like wars are conditional on the balance of power between elites and society.

Keywords: civil society, contest, political divergence, state capacity, weak states.

JEL classification: H4, H7, P16. **Word Count:** 11533 (Using the APSR recommended tool).

*We are grateful to two sets of editors and three anonymous referees for their insightful and constructive comments. We thank Pooya Molavi for exceptional research assistance, Marco Battaglini, Roland Benabou, Kim Hill, Josh Ober, Paul Poast and Mark Pyzyk for discussions and seminar participants at the ASSA 2017, CIFAR, Chicago, NBER, University of Illinois, Lund, Northwestern and Yale for useful comments and suggestions. Daron Acemoglu gratefully acknowledges financial support from the Carnegie Foundation and ARO MURI Award No. W911NF-12-1-0509.

[†]Massachusetts Institute of Technology, Department of Economics, E52-380, 50 Memorial Drive, Cambridge MA 02142; E-mail: daron@mit.edu.

[‡]University of Chicago, Harris School of Public Policy and Department of Political Science, 1155 East 60th Street, Chicago, IL60637; E-mail: jamesrobinson@uchicago.edu.

Weak, Despotic, or Inclusive? How State Type Emerges from State vs. Civil Society Competition

Abstract

We develop a theory of the accumulation of state capacity as the outcome of a political competition between elites and (civil) society. State capacity is accumulated by elites and is both productive but also useful in controlling society. Society, however, can fight back and accumulate its own capacity, facilitating collective action. The theory leads to three distinct equilibria depending on initial conditions. One type, a weak state, emerges when society is strong relative to the elite. Another, a despotic state, originates where the elite is initially relatively powerful. A third type, an inclusive state, emerges when the elite and society are more evenly matched. The theory has several key implications; first, variation in state capacity does not require large structural differences; second, inclusive states have the highest levels of state capacity in the long run; third, the effects of shocks or external threats like wars are conditional on the balance of power between elites and society.

Keywords: civil society, contest, political divergence, state capacity, weak states.

JEL classification: H4, H7, P16.

1 Introduction

There is a great deal of variation in state capacity around the world. Simple statistical measures, such as the ratio of government tax revenues to national income, vary from close to 50% in western and northern Europe to less than 10% in many countries in sub-Saharan Africa. These differences in resources are reflected in large differences in the provisions of public goods, infrastructure, and the ability to deliver justice or hold clean elections.

Such dispersion has enormous consequences for politics. The absence of state capacity has been argued to be the main reason that societies fall into civil war (Fearon and Laitin, 2003); the primary explanation for the absence of accountability and institutions of participation (Levy, 1989, Herbst, 2000, Tilly, 2007); and an important cause of the inability of communities to collectively govern resources in desirable ways (Ostrom, 1990). The presence of state capacity is argued to be the main driving force behind experiences of rapid economic growth (Amsden, 1989, Wade, 1990, Evans, 1995) and more broadly to guarantee that the state works in the collective interest (Geddes, 1994).

In this paper we propose a new theory to explain this variation. Existing explanations account for it either by structural factors, such as population density, topography, or culture or via more contingent influences such as histories of warfare, colonialism or trade. Yet, as we illustrate below, polities with very similar structural features and histories of warfare have experienced dramatic divergences in state capacity.

Our theory conceptualizes the construction of state capacity as a game between those in control of the state, who we call the elite, and civil society—or henceforth "society" for short. State capacity, accumulated by the elite, potentially has benefits for all in the sense that it allows for the provision of public goods, but at the same time it can be used to oppress citizens and thus slant all the benefits to the elite. This trade-off is recognized in the western world at least since the time of Locke (2003). As Scott (2010) and Acemoglu and Robinson (2019) discuss, it is well understood in non-western societies too. But citizens also have capacity in this game with the elite—we conceptualize this as the ability to organize and engage in collective action. Like state capacity accumulated by elites, greater societal capacity adds to productivity because it allows greater coordination and public good provision (Ostrom, 1990). But it simultaneously gives society greater ability to contest with the elite (emphasized by Wood, 2003).

Specifically, we assume that state and society's capacities are used as inputs into a contest for power. Elites accumulate state capacity, which enables them to impose their wishes on and dominate society, while society's capacity empowers it to resist elite schemes (which may take the form of democratic political participation, protests, collective action, strikes, and even violence). Our notion of a contest for power and the model we use is common in the literature on interest group politics, civil wars and international relations (see Tullock, 1980; Skaperdas, 1992; Powell, 1999). As in these literatures we model the outcome as uncertain; for given capacities which side wins depends on a variety of contingent factors that cannot be anticipated in advance. Whoever wins is then able to set

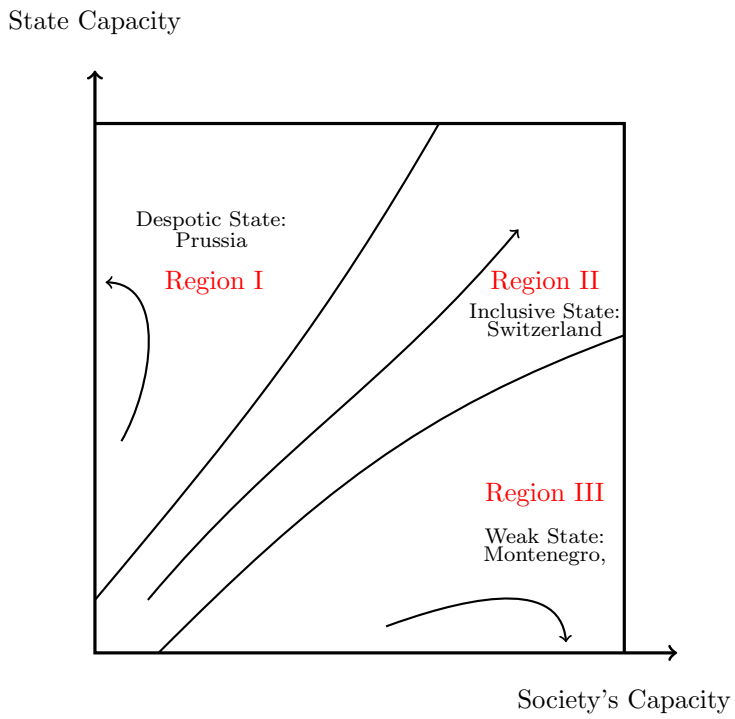

Figure 1: The emergence and dynamics of weak, despotic and inclusive states.

policy and allocate resources in their own interests.

We show that there are three long-run equilibrium outcomes in this contest.[1] Either, elites manage to accumulate far more capacity than society and dominate it. This leads to a *despotic state*. Or, in the opposite case, society dominates elites and very little state capacity develops, and a *weak state* emerges. In the middle, where the capacity of the state and society is balanced a third equilibria emerges where both the state and society end up with large amounts of capacity. We call this an *inclusive state*.

These long-run equilibrium outcomes and the dynamics leading up to them are summarized in Figure 1. We depict society's capacity on the horizontal axis and state capacity on the vertical one. Region I illustrates the political dynamics leading to a despotic state, similar to those of early modern Prussia, a case we discuss in detail in Section 5. Region III is the case of a weak state, which we illustrate in Section 5 with the history of Montenegro contrasted to Prussia's. Finally, Region II depicts the middle ground where the elite and society are initially in balance, and this triggers an ongoing competition between the two, leading to an inclusive state, which we illustrate with Switzerland.

This comparison is not intended to indicate the scope conditions of the model but is simply a use of Mill's (1972) most similar research design. More broadly, our claim in that the mechanisms underlying the dynamics and equilibria of the model help to account for the patterns we see in the

[1]Technically, these are steady-state Nash equilibria of the dynamical system, but to avoid proliferating the use of the word "state" we simply refer to these as equilibria.

world with the simultaneous persistent existence of despotic states (like China), weak states (as in many parts of sub-Saharan Africa) and inclusive states (in western and northern Europe or north America). Existing theories can explain this pattern only by appealing to significant regional or continental differences in underlying conditions. Without denying the role for such differences, our ambition is to develop a global theory that can account for these rich patterns without huge structural differences across countries and regions.

Our theory and Figure 1 clarify that small differences in initial conditions can put a polity into the basin of attraction of one equilibrium rather than another. We believe this feature of the model is consistent with historical evidence (see Section 5).

An important implication of our theory is visible in Figure 1: state capacity is greater under inclusive states than despotic states. This is because when the state dominates society, it has less incentive to accumulate further capacity. It is competition between state and society that triggers greater investments by the elites controlling it. Our model thus highlights that not only do weak, despotic and inclusive states have different amounts of state capacity, but they also have divergent societal capacities.

Another feature of our theory is clear from Figure 1: the effects of a rich array of structural factors are *conditional*. For example, factors that (exogenously) increase the power of the state could move a polity from Region III to Region II, initiating a powerful process of state capacity building. But as Figure 2 illustrates the same factors may also push a polity previously in Region II into Region I, reducing its long-run potential to achieve high state capacity. Similarly, structural factors shift the boundaries of the basins of attraction of the three different types of states as well, but the effects of such changes are also conditional on the prevailing balance between state and society. These conditional comparative statics highlight that there is no simple version of Tilly's "war made the state and the state made war" in our theory. Though warfare may create incentives to strengthen state capacity, its ultimate implications depend on the balance of power between state and society. This provides one explanation for why for each structural factor argued to underpin the development of state capacity, there are always counterexamples going in the opposite direction.[2]

Although our theory allows for diverse long-term outcomes, it also generates several falsifiable predictions. First, the greatest state capacity emerges not when the state is overwhelmingly dominant and strong relative to society, but when there is some sort of balance between the two. This would imply that one should find a non-monotonic relationship between initial state capacity (given society's organization) and ultimate state capacity. Second, the longer a society remains in the basin of attraction of the despotic state (or weak state), the more difficult it is for it to transition to an inclusive state. This perspective implies, for example, that in contrast to optimistic predictions about

[2]Most studies of Tilly's hypothesis, even in Europe find it does not apply without some meditating variables or the addition of various complementary channels (e.g., Abramson, 2017; Hoffman, 2015). Thies (2005) adds to Tilly the notion of "inter-state rivalry". In the Asian context, Taylor and Botea (2008) provide a discussion of how mediating factors like ethnicity may interact with warfare. For Latin America, see Centeno's (1997) discussion of mediating factors, for China, see Hui (2005) and Dincecco and Wang (2018), and for Africa, see Dincecco, Fenske and Onorato (2019).

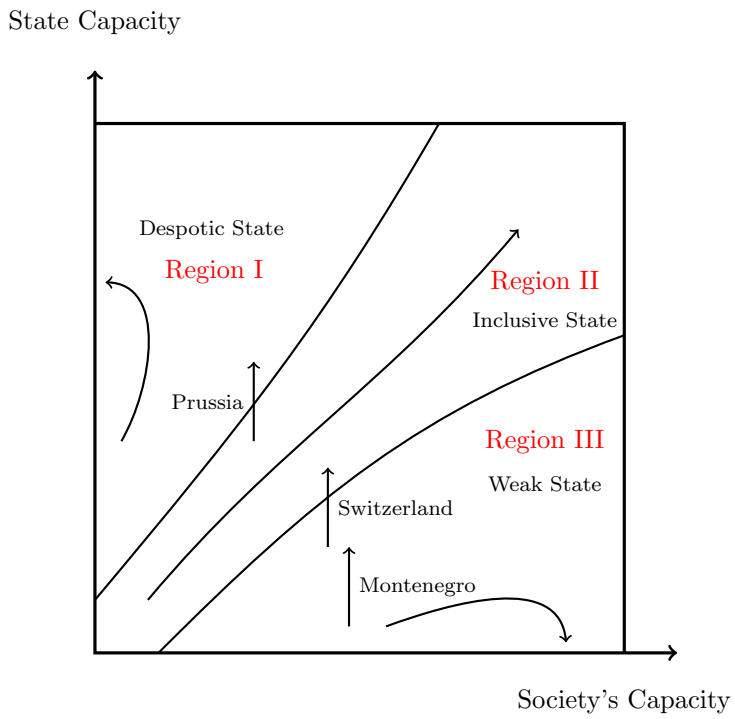

Figure 2: Conditional comparative statics: The same impulse leads to very different trajectories of political development.

China's political development and seamless transition to democracy based on its modernization, building inclusive state-society relationship in China will be very hard. Third, the model implies a type of weak monotonicity in initial conditions in response to shocks, which can be investigated empirically. For example, of two polities that are subject to exactly the same pro-state shock and have the same societal capacity, the one with greater state capacity is more likely to transition to a despotic state than the other, and less likely to remain in the weak state's basin of attraction.

Our theory builds on a few key ideas. First, that state capacity can be used productively, but also to control society and that it can be accumulated. By state capacity most scholars build on Mann's (1993) notion of "infrastructural power" which he defines as the capacity of a central state "penetrate its territories and logistically implement decisions" (p. 59). Such capacity is multi-dimensional and includes bureaucratic, legal and fiscal capacities and even the latter is itself complex, as Fauvelle-Aymar (1999) points out, since it involves not just the construction of state institutions, but also people's willingness to pay taxes which may depend on the extent to which they see the state as legitimate or efficient. Though our model is reduced form and consistent with different interpretations of state capacity, fiscal capacity does satisfy our three conditions; it can be used to provide public goods, to equip and pay state agents of control and more of it can be acquired via institution building, conducting surveys of assets and wealth etc. (even if people may resist, a la Fauvelle-Aymar, 1999). It is also the dominant measure in the literature (e.g. Herbst, 2000, Slater, 2010) and lends itself to measurement.

We pit this against society's capacity. Our notion of societal capacity emphasizes society's ability to solve collective action problems and define and achieve collective goals. Like our definition of state capacity this is productive, can be used in contestation and it can be built. Tilly (1995) studied the emergence of this capacity in 18th century Britain showing how initially "contention" was about "local people and local issues, rather than nationally organized programs and parties" (p. 5). However, "between 1758 and 1833 a new variety of claim making had taken shape ... Mass popular politics had taken hold on a national scale" (p. 13). Critically, this new politics was not just about new and collective issues (instead of parochial ones); it was also organized in very different ways, frequently in the form of "a special purpose association, society or club" (p. 10). Our model captures what Tilly thought was the force driving this reorganization of society: "the contemporary reorganization ... of the state, that re-shaped the state's repressive apparatus" (p. 23), along with expansions of "revenue, expenditure and personnel" (p. 49), which induced on the part of society "a shift towards collective action that was large in scale and national in scope" (p. 49).

Ostrom's work identified aspects of social institutions that are critical for solving collective action problems and building capacity (1990, p. 90). These include: autonomous organization, monitoring capability, availability of sanctions and effective conflict resolution mechanisms. Society lacks capacity when there is "no capacity to communicate with one another, no way to develop trust, and no sense that they share a common future". Ostrom also stressed that "Such groups may need some form of external assistance to break out of the perverse logic of their situation" (p. 21), which often comes from interactions with the state, or as she puts it "are affected by the surrounding political regime" (p. 141). States often influence the creation of institutions to solve collective action problems "by creating and limiting the powers that can be exercised with collective-choice arrangements (creating legislative and judicial bodies, protecting rights of free speech and property etc.)" (p. 192). Her case studies confirmed that "the orientation of the ruling political regime can make a substantial difference" (p. 212). Recent work on collective action in civil war echoes many of these themes. Steele shows that communities in Colombia have managed to organize collectively to resist armed groups when they developed "A combination of a strict internal hierarchy and external support" (2017, p. 173). Arjona's study of local order during the Colombian civil war similarly illustrates how "The quality of dispute institutions plays a central role in fostering the community's capacity for collective action" (2016, p. 71). Institutions further facilitate cohesion if they can target goods and services as emphasized by Cammett (2014) and Cammett and McLean (2014).

Societal capacity also depends on cultural and behavioral factors identified by Wood (2003). In her study of the Salvadorean Civil War, collective action depends on people's "moral commitments and emotional engagements" (p. 18). People acted collectively "as an act of defiance of long-resented authorities and a repudiation of perceived injustices" (p. 18). As in our theory, this type of capacity building is dynamic, since "political culture — the values, norms, practices, beliefs, and collective identity of insurgents — was not fixed but evolved in response to the experiences of the conflict itself" (p. 41), and defiance "... motivated further collective action through a recursive process" (p. 238).

Wood's notion of society's capacity is complemented by Viterna's (2013) theory of how people can assume a "participation identity", selecting from the available repertoire of identities (Swidler, 1986) one that is highly congruent with collective action. As in Wood's analysis, this can be precipitated by the influence of political institutions. "In El Salvador, the military's attacks on rural civilians gave new meanings to existing identities" (p. 54), which tipped people towards participation in the conflict.

A key assumption in our model is that the accumulation of both state and societal capacity is subject to (dynamic) increasing returns to scale. In particular we assume that the marginal cost of building capacity is higher below a certain threshold than afterwards. Intuitively, once one has sufficient capacity, it becomes easier to acquire more. For example, building a nascent fiscal system involves large fixed costs. Detailed studies of this, such as Brewer (1989), show that to build an effective excise tax system in Britain many elements had to be in the place so that personnel can be trained, paid, monitored and stationed throughout Britain. Our assumption is that one has these elements in place, then it becomes easier and cheaper to build further fiscal capacity (see also Dharmapala, Slemrod and Wilson, 2011). Our argument for society is similar: as Tilly's study shows, more effective contention required organization and the creation of new institutions. Once these had been put in place, it subsequently becomes easier to accumulate greater capacity. Both Wood's and Viterna's theories imply increasing returns to the accumulation of what we are calling society's capital. This is evident in Wood's formal model (2003, pp. 267-274), and from the peer effects connected to the formation of a collective participation identity in Viterna (2013). Increasing returns to collective action are also emphasized in Marwell and Oliver (1993) and Pearson (2000).

Our paper is closely related to several important lines of research within comparative politics. Though our findings do not support the idea that divergence in state capacity necessitates large structural differences or is a consequence of the incidence of warfare, they are consistent with other key ideas. Many scholars have studied the idea that elites and society compete and that state capacity can emerge from what Wood calls this "recursive process". This is evident from our discussion of Tilly (1995) and central to Skocpol's (1979) theory of social revolutions and their consequences. She defines social revolutions as "rapid, basic transformations of a society's state and class structures, accompanied and in part carried through by class-based revolts from below" (p. 33) and emphasizes not just "the changes that social revolutions make in the structure and function of states" (p. 164) but also the mutual feedback such that "The changes in state structures that occur during social revolutions typically both consolidate, and themselves entail, socioeconomic changes" (p. 164). [3] Our model formalizes and extends these ideas and mutual feedback loops and shows that mechanisms typically studied in different contexts have common roots and major implications for the distribution of state capacity. For example, our weak state configuration captures Migdals (1988) and Scott's (2010) ideas that week states emerge from society's strength. Yet, in contrast to their emphasis, our

[3] Anderson (1974), Hechter and Brustein (1980), Slater (2010) and Saylor (2014) have also emphasized the idea that the state accumulates capacity in a contest with society.

theory also reveals that, in the inclusive state configuration, society is even "stronger" (using Migdal's terminology) when the state is stronger. Our theory's implications also contrast with Huntington's (1968) and Fukuyama's (2011) claims that state capacity emerges under the auspices of powerful leaders and groups and that there is a specific sequence towards a democratic strong state — state strength first, democracy later. This is not a prediction of our model; when the state and elites become too strong, the development of state capacity and democracy (see below) is arrested. Our theory also nests Grzymala-Busse's (2007) theory that state capacity emerged in the post-Soviet world when there was competition between political parties (as in Hungary and Poland), but did not when one party dominated (as in the Czech Republic or Slovakia).[4]

Much scholarship in comparative politics has attempted to explain not just patterns of state building, but their consequences for regimes. Herbst (2000), building on ideas initially developed by Levy (1989), argues that state weakness in Africa leads to nondemocracy because states that do not raise taxes are not forced to negotiate with their citizens. Ertman (1997) develops a dichotomy of state capacity (bureaucratic or patrimonial) combining it with regime types (constitutional or autocratic rule). Though we do not formally model the process of democratization, our theory implies that democracy cannot easily arise under despotic or weak states. In the former, powerful elites can block the participation of society (as in many standard theories of democracy, e.g. Rueschemeyer, Stephens and Stephens, 1992, and Acemoglu and Robinson, 2006). In the latter, the weakness of the central state does not allow for the emergence of effective democratic institutions. It is therefore in the middle, where there is state capacity, but it is forced to be responsive to a capacitated society that one would expect democracy to emerge.

In Section 5, we also develop an Early Modern European case study in the spirit of Mill's (1872) "most similar" research design (Skocpol, 1979, discusses the strengths and weaknesses of this approach). Our case selection focuses on the large differences within Europe in terms of our main dependent variable. Why did Switzerland develop an inclusive state which had capacity, but which was very accountable to its people, while Prussia created a despotic state with less capacity? Why did Montenegro never really create a state, except a powerless theocracy, until the 20th century?

We develop this case study as two linked pairs, Prussia-Switzerland and Switzerland-Montenegro. This design is "most similar" in the sense that we emphasize both common institutional histories (Prussia-Switzerland), social structures and ecologies (Switzerland-Montenegro). Prussia and Switzerland in 1600 (before Brandenburg had merged with Prussia) were both part of the Holy Roman Empire and had inherited several political and economic institutions in common from the Germanic tribes, the Carolingians, and feudalism. Switzerland and Montenegro had deep institutional roots in common (Roman Empire), but our focus is on social structures (clans) and also mountainous ecology. The role of clans is well established in Montenegro, of course, but it is also a common theme in Swiss studies. Steinberg (1996, p. 17), for example, notes: "[Swiss] medieval clan structures had little to

[4]See Berwick and Christia (2011), Blaydes (2017) and Grzymala-Busse (2020) for three recent surveys of important aspects of this literature, and the essays in Centeno, Kohli, Yashar and Mistree (2017).

do with our images of democratic forms but these peasants were 'free" and he notes the extent of communal activities and organization. It also interesting the extent to which the early development of the Old Swiss Confederacy was focused on conflict resolution and the management of feuds, which were an incessant problem in Montenegro as well. The key difference between the cases is the initial strength of the state: no state in Montenegro, a nascent, non-centralized state in Swiss cantons, and a stronger and more centralized state in Brandenburg (Prussia). In our theory it is these relatively small differences in the initial strength of the state that led to the very different dynamics and outcomes.

Finally, our theory is built on the theory of dynamic contests (e.g., Hirshleifer, 1989). Our results on stronger incentives to build capacity when elites and citizens are evenly matched is related to the discouragement effect in contests emphasized in Harris and Vickers (1985).

The rest of the paper is organized as follows. In the next section, we introduce our main model. In Section 3, we characterize the dynamic equilibrium and steady states of this model. To maximize transparency, this section uses a number of simplifying assumptions, many of which are relaxed in the Appendix. Section 4 outlines how the same model with forward-looking players leads to similar results. Section 5 discusses our case study of early modern European state divergence. Section 6 concludes, while the online Appendix presents the proofs of the results stated in the text, details of several generalizations mentioned in the text, some additional technical material, and also a generalization of our model in which the inclusive state becomes feasible only after the capacities of both the state and society are above certain thresholds.

2 Basic Model

In this section, we introduce our basic model which is then analyzed in the next several sections.

2.1 Preferences and Conflict

We start with a discrete time setup, where period length is $\Delta > 0$ and will later be taken to be small, so that we work with differential rather than difference equations in characterizing the dynamics. At time t, the stock variables inherited from the previous period are $(x_{t-\Delta}, s_{t-\Delta}) \in [0,1]^2$, where the first element corresponds to the capacity of society and the second to state capacity controlled by the elite.[5]

At each point in time, the elite is represented by a single player, and society is also represented by a single player. In the next two sections, we study both the case in which these players are short-lived and are immediately replaced by another player (so that we have a non-overlapping generations model with myopic players), and the case in which players are long-lived and maximize their discounted sum of utilities.

At time t, players simultaneously choose their investments, $i_t^x \geq 0$ and $i_t^s \geq 0$, which determine

[5]Normally in dynamical systems these variables would be referred to as the state variables. We use the terminology stock instead to avoid confusion.

their current capacity according to the equations:

$$x_t = x_{t-\Delta} + i_t^x \Delta - \delta\Delta, \tag{1}$$

and

$$s_t = s_{t-\Delta} + i_t^s \Delta - \delta\Delta, \tag{2}$$

where $\delta > 0$ is the depreciation of the capacities of both parties between periods. Both investment and depreciation are multiplied by the period length, Δ, since they represent "flow" variables, and when period length is taken to be small, they will be suitably downscaled.[6]

The cost of investment for society during a period of length Δ is given as $\Delta \cdot \tilde{C}_x(i_t^x, x_{t-\Delta})$ where

$$\tilde{C}_x(i_t^x, x_{t-\Delta}) = \begin{cases} c_x(i_t^x) & \text{if } x_{t-\Delta} > \gamma_x, \\ c_x(i_t^x) + (\gamma_x - x_{t-\Delta})\, i_t^x & \text{if } x_{t-\Delta} \leq \gamma_x. \end{cases}$$

This cost function is multiplied by Δ, since it is the cost of investing an amount i_t^x during the period of length Δ (as captured by equation (1)). The presence of the term $\gamma_x > 0$, on the other hand, captures the "increasing returns" nature of capacity accumulation mentioned in the Introduction: starting from a low level of capacity, it is more costly to build up this capacity. We specify this in a very simple form here, with the cost of investments increasing linearly as last period's capacity falls below the threshold γ_x. This increasing returns aspect plays an important role in our analysis as we emphasize below.

The cost of investment for the elite during a period of length Δ is similarly given as $\Delta \cdot \tilde{C}_s(i_t^s, s_{t-\Delta})$ where

$$\tilde{C}_s(i_t^s, s_{t-\Delta}) = \begin{cases} c_s(i_t^s) & \text{if } s_{t-\Delta} > \gamma_s, \\ c_s(i_t^s) + (\gamma_s - s_{t-\Delta}) i_t^s & \text{if } s_{t-\Delta} \leq \gamma_s. \end{cases}$$

In these expressions, it will often be more convenient to eliminate investment levels and directly work with the two stock variables, x_t and s_t, especially when we take Δ to be small and transition to continuous time. In preparation for this transition, let us substitute out the investment levels and observe that the cost function for society and state can be written as:

$$C_x(x_t, x_{t-\Delta}) = c_x\left(\frac{x_t - x_{t-\Delta}}{\Delta} + \delta\right) + \max\left\{\gamma_x - x_{t-\Delta}, 0\right\}\left(\frac{x_t - x_{t-\Delta}}{\Delta} + \delta\right),$$

and

$$C_s(s_t, s_{t-\Delta}) = c_s\left(\frac{s_t - s_{t-\Delta}}{\Delta} + \delta\right) + \max\left\{\gamma_s - s_{t-\Delta}, 0\right\}\left(\frac{s_t - s_{t-\Delta}}{\Delta} + \delta\right),$$

where the increasing returns to scale nature of the cost function is now captured by the max term.[7]

[6] Assuming that depreciation is independent of the current level of the capacity of the state or society is for convenience only. In addition, we can easily allow the two stock variables to have different depreciation rates, but do not do so in order to prevent the notation from becoming more cumbersome.

[7] Note that when we consider the limit $\Delta \to 0$, we obtain

$$\begin{aligned} C_x(\dot{x}_t) &= c_x(\dot{x}_t + \delta) + \max\left\{\gamma_x - x_t, 0\right\}(\dot{x}_t + \delta), \\ C_s(\dot{s}_t) &= c_s(\dot{s}_t + \delta) + \max\left\{\gamma_s - s_t, 0\right\}(\dot{s}_t + \delta). \end{aligned}$$

During the lifetime of each generation, a polity with state capacity s_t and societal capacity x_t produces income/output given by

$$f(x_t, s_t), \tag{3}$$

where f, the production function, is assumed to be nondecreasing and differentiable.[8] We adopt a reduced-form approach in modeling the contribution of state and society capacities, and assume that they contribute to aggregate income as well as impacting the conflict between them. In reality, the ability of society to coordinate and the infrastructural power of the state increase the productivity of producers, which then affects the income to be divided between different parties.[9] The dependence of the total output of the economy on state capacity captures the various efficiency-enhancing roles of the state. In addition, we allow for output to depend on the capacity of civil society as well, because its greater cooperation and coordination also improves economic efficiency (Ostrom, 1990).

We next discuss how the aggregate income is distributed between the elite (controlling the state) and society. At date t, if the elite and society fight and one side wins and capture all of the income of the economy, and the other side receives zero. Winning probabilities are functions of relative capacities. In particular, the elite will win if

$$s_t \geq x_t + \sigma_t, \tag{4}$$

where σ_t is drawn from the distribution H independently of all past events. We denote the density of the distribution function H by h. The existence of the random term σ_t captures the fact that various stochastic factors impact the outcome of any conflict.[10]

This specification of a stochastic contest function, and a symmetry assumption which we will shortly impose, implies that when the capacities of society and state are given, respectively, by x and s, the probability that the elite will win the conflict is $H(s-x)$, and the probability that the society will do so is $1 - H(s-x) = H(x-s)$, a property we will use frequently below.[11]

2.2 Simplifying Assumptions

We next introduce three assumptions. The first one is a simplifying assumption, which we impose initially and then relax subsequently:

Assumption 1 $f(x, s) = 1$ for all $x \in [0, 1]$ and $s \in [0, 1]$.

[8] The fact that (3) refers to output during the lifetime of each generation means that each generation will produce this quantity regardless of $\Delta > 0$.

[9] See the Appendix for a micro-foundation. As we show more explicitly in footnote 13, this feature is important to ensure that the incentives for investment do not vanish when we consider short-lived players as in the next section and $\Delta \to 0$. (When we return to long-lived, forward-looking players, incentives for investment will not vanish and similar results apply as $\Delta \to 0$ even if (3) is multiplied with the period of length Δ.

[10] Scholars in international relations have provided a great deal of evidence on the applicability of these contest functions, for example Reiter (1996) and Buhaug (2010).

[11] In the Appendix, we also show that the most important qualitative features implied by this formulation of conflict between the elite and society are shared by other formulation of the contest between these parties.

This assumption makes it transparent that the multiple equilibria and their dynamics — our main focus — are driven by the dynamic contest between the elite and society, not because of changes in the value of the prize in this contest. It will be relaxed in the Appendox.

The next two assumptions are imposed throughout.

Assumption 2 1. c_x and c_s are continuously differentiable, strictly increasing and weakly convex over $\mathbb{R}_+$, and satisfy $\lim_{x\to\infty} c'_x(x) = \infty$ and $\lim_{x\to\infty} c'_s(s) = \infty$.

2.
$$c'_s(\delta) \neq c'_x(\delta).$$

3.
$$\frac{|c''_s(\delta) - c''_x(\delta)|}{\min\{c''_x(\delta), c''_s(\delta)\}} < \frac{1}{\sup_z |h'(z)|}.$$

4.
$$c'_s(0) + \gamma_s \geq c'_x(\delta) \text{ and } c'_x(0) + \gamma_x > c'_s(\delta).$$

Part 1 of Assumption 2 is standard. Part 2 is imposed for simplicity and rules out the non-generic case where the marginal cost of investment at δ is exactly equal for the two parties. Part 3 is also imposed for technical convenience, and is quite weak. For example, if the gap between $c''_x(\delta)$ and $c''_s(\delta)$ is small, this condition is automatically satisfied. We will flag its role when we come to our analysis, but anticipating that discussion, it makes it much easier for us to establish the instability of some substantively uninteresting equilibria. Part 4 ensures that the marginal cost of each player in the increasing returns region (when $x < \gamma_x$ or $s < \gamma_s$) when making zero investment is greater than the marginal cost of the other player outside this region when evaluated at δ — the marginal cost on the right-hand side is evaluated at δ since, as our above transformation showed, the level of investment necessary for maintaining any positive equilibrium level of capacity is δ. We will flag the role of this assumption when we come to our formal analysis.

Assumption 3 1. h exists everywhere, and is differentiable, single-peaked and symmetric around zero.

2. For each $z \in \{x, s\}$,
$$c'_z(0) > h(1).$$

3. For each $z \in \{x, s\}$,
$$\min\{h(0) - \gamma_z; h(\gamma_z)\} > c'_z(\delta).$$

Part 1 contains the second key assumption for our analysis — single peakedness and symmetry of h around 0 (differentiability is standard). This assumption not only simplifies our analysis as it ensures that $h(x - s) = h(s - x)$ and $h'(x - s) = -h'(s - x)$, but also implies that incentives for

investment are strongest when x and s are close to each other. We highlight the role of this feature below as well.[12]

Part 2 imposes that when a player has the maximum gap between itself and the other player, it has no further incentives to invest. Part 3, on the other hand, ensures that at or near the point where capacities are equal, there are sufficient incentives to increase capacity. Both of these assumptions restrict attention to the part of the parameter space of greater interest to us.

3 Equilibrium with Short-Lived Players

We now present our main results about the dynamics of the capacity of state and society, focusing on the non-overlapping generations setup, where at each point in time, each side of the conflict is represented by a single short-lived agent who will be replaced by a new agent from the same side next period.

3.1 Preliminaries

Suppose that the above-described polity is populated by non-overlapping generations of agents — on the one side representing the elite and on the other, society.

With these assumptions, at each time t, society maximizes

$$H(x_t - s_t) - \Delta \cdot C_x(x_t, x_{t-\Delta})$$

by choosing x_t (or equivalently i_t^x), taking $x_{t-\Delta}$ as given. Simultaneously, the elite maximize

$$H(s_t - x_t) - \Delta \cdot C_s(s_t, s_{t-\Delta})$$

by choosing s_t, taking $s_{t-\Delta}$ as given. A dynamic (Nash) equilibrium with short-lived players is given by a sequence $\{x^*_{k\Delta}, s^*_{k\Delta}\}_{k=0}^{\infty}$ such that $x^*_{k\Delta}$ is a best response to $s^*_{k\Delta}$ given $x^*_{(k-1)\Delta}$, and likewise $s^*_{k\Delta}$ is a best response to $x^*_{k\Delta}$ given $s^*_{(k-1)\Delta}$.

The investment decisions of both elites and society are then determined by their respective first-order conditions (with complementary slackness). In particular, at time t, we have:[13]

$$\begin{array}{ll} h(x_t - s_t) \leq c'_x(\frac{x_t - x_{t-\Delta}}{\Delta} + \delta) + \max\{0; \gamma_x - x_{t-\Delta}\} & \text{if } \frac{x_t - x_{t-\Delta}}{\Delta} = -\delta \text{ or } x_t = 0, \\ h(x_t - s_t) \geq c'_x(\frac{x_t - x_{t-\Delta}}{\Delta} + \delta) + \max\{0; \gamma_x - x_{t-\Delta}\} & \text{if } x_t = 1, \\ h(x_t - s_t) = c'_x(\frac{x_t - x_{t-\Delta}}{\Delta} + \delta) + \max\{0; \gamma_x - x_{t-\Delta}\} & \text{otherwise,} \end{array}$$

[12] The result that incentives for investment are strongest when the two sides are evenly matched is more general than the specification used here. For example, suppose that we have a contest function where the probability that the elite wins is $\frac{k(s)}{k(s)+k(x)+\eta}$ and the probability that society wins is $\frac{k(x)}{k(s)+k(x)+\eta}$, where $k(\cdot)$ is an increasing, differentiable function, and $\eta \geq 0$ is a constant. In this case, the marginal return to increasing investment for the elite is $\frac{k'(s_t)(k(x_t)+\eta)}{(k(s)+k(x)+\eta)^2}$, and the expression for society is also similar. It can be verified that, when $\eta = 0$, the cross-partial derivative of this expression is positive when $s_t > x_t$, and negative when $s_t < x_t$. When $\eta > 0$, the same result holds provided that s_t is sufficiently larger than x_t.

[13] Following up on footnote 8, we can more clearly see the role that Δ in front of the cost function plays here: without this term (or equivalently if the return was also multiplied by Δ), the marginal cost of investment would be multiplied by $1/\Delta$, and thus as $\Delta \to 0$, investments would converge to zero. This is because short-lived players that are not forward-looking do not take the impact of their instantaneous investments on future stocks (and have infinitesimal impact on the current stock).

and

$$\begin{array}{ll} h(s_t - x_t) \leq c_s'(\frac{s_t - s_{t-\Delta}}{\Delta} + \delta) + \max\{0; \gamma_s - s_{t-\Delta}\} & \text{if } \frac{s_t - s_{t-\Delta}}{\Delta} = -\delta \text{ or } s_t = 0, \\ h(s_t - x_t) \geq c_s'(\frac{s_t - s_{t-\Delta}}{\Delta} + \delta) + \max\{0; \gamma_s - s_{t-\Delta}\} & \text{if } s_t = 1, \\ h(s_t - x_t) = c_s'(\frac{s_t - s_{t-\Delta}}{\Delta} + \delta) + \max\{0; \gamma_s - s_{t-\Delta}\} & \text{otherwise.} \end{array}$$

The first line of either expression applies when the relevant player has chosen zero investment so that its stock variable shrinks as fast as it can (at the rate δ), or is already at its lower bound $x_t = 0$ or $s_t = 0$. In this case, we have the additional cost of investment on the right-hand side, and also the optimality condition is given by a weak inequality, since at this lower boundary, the marginal benefit could be strictly less than the marginal cost of investment. The second line, on the other hand, applies when the stock variable takes its maximum value, 1, and in this case the marginal benefit could be strictly greater than the marginal cost of investment. Away from these boundaries, the third line applies and requires that the marginal benefit equal the marginal cost. Note also that the marginal benefit for society is the same as the marginal benefit for the elite — since $h(s_t - x_t) = h(x_t - s_t)$. On the other hand, we also have from Assumption 3 that changes in the marginal benefits of the two players are the converses of each other — that is, $h'(s_t - x_t) = -h'(x_t - s_t)$.

Now letting $\Delta \to 0$, we obtain the following continuous-time first-order optimality (and thus equilibrium) conditions

$$\begin{array}{ll} h(x_t - s_t) \leq c_x'(\dot{x}_t + \delta) + \max\{0; \gamma_x - x_t\} & \text{if } \dot{x}_t = -\delta \text{ or } x_t = 0, \\ h(x_t - s_t) \geq c_x'(\dot{x}_t + \delta) + \max\{0; \gamma_x - x_t\} & \text{if } x_t = 1, \\ h(x_t - s_t) = c_x'(\dot{x}_t + \delta) + \max\{0; \gamma_x - x_t\} & \text{otherwise,} \end{array} \tag{5}$$

and

$$\begin{array}{ll} h(s_t - x_t) \leq c_s'(\dot{s}_t + \delta) + \max\{0; \gamma_s - s_t\} & \text{if } \dot{s}_t = -\delta \text{ or } s_t = 0, \\ h(s_t - x_t) \geq c_s'(\dot{s}_t + \delta) + \max\{0; \gamma_s - s_t\} & \text{if } s_t = 1, \\ h(s_t - x_t) = c_s'(\dot{s}_t + \delta) + \max\{0; \gamma_s - s_t\} & \text{otherwise.} \end{array} \tag{6}$$

In what follows, we work directly with these continuous-time first-order optimality conditions. Moreover, it is straightforward to see that in continuous time, away from the boundaries of $[0,1]^2$ these first-order optimality conditions will hold as equality, and thus the dynamics of state and society capacity can be represented by the following two differential equations:

$$\begin{aligned} \dot{x} &= \max\{(c_x')^{-1}(h(x-s) - \max\{\gamma_x - x, 0\}); 0\} - \delta \\ \dot{s} &= \max\{(c_s')^{-1}(h(s-x) - \max\{\gamma_s - s, 0\}); 0\} - \delta. \end{aligned} \tag{7}$$

The roles of the two key assumptions highlighted above — the single-peakedness of h and the increasing returns aspect of the cost function — are evident from (7). First, when x and s are close to each other, $h(x-s)$ is large, and thus both of these variables will tend to grow further. Conversely, when x and s are far apart, $h(x-s)$ is small, and investment by both parties is discouraged. This observation captures the key force that will lead to the emergence of different dynamics of elite-society relations and different types of states in our setup.[14] Secondly, the presence of the max term implies that once the capacity of a party falls below a critical threshold (γ_x or γ_s), there is an additional force pushing towards further reduction in this capacity.

[14] In the Appendix we show that this same property holds with other formulations of the contest function.

3.2 Dynamics of the Capacity of Society and the State

Our main result in this section is summarized in the next proposition.

Proposition 1 *Suppose Assumptions 1, 2 and 3 hold. Then there are three (locally) asymptotically stable (Nash) equilibria:*

1. $x^* = s^* = 1$.
2. $x^* = 0$ *and* $s^* \in (\gamma_s, 1)$.
3. $x^* \in (\gamma_x, 1)$ *and* $s^* = 0$.

This proposition shows that there exist three relevant (asymptotically stable) equilibria, one corresponding to an inclusive state, one corresponding to a despotic state and one to a weak state. The intuition, as already anticipated, is that when we are in the neighborhood of the steady state $x^* = s^* = 1$, $h(x - s)$ is large, encouraging both parties to move further towards $x^* = s^* = 1$. In contrast, in the neighborhood of $x^* = 0$ or $s^* = 0$, $h(x - s)$ is small, and neither party has as strong incentives to invest, and in fact, one of them ends up with zero capacity.[15]

The equilibria presented in Proposition 1 and their local dynamics are exactly the same as those shown in Figure 1 in the Introduction. This can also be seen from the numerically-constructed Figure 3 presented here.[16]

3.3 Conditional Comparative Statics

One of the most important results of our framework is that comparative statics —which show how the equilibrium changes when perturbed by exogenous factors—are *conditional*, and in particular depend on the prevailing balance between elites and society (where we are in the diagram in Figure 1). The easiest way of seeing this is to consider an increase in s_0 to $s_0 + \bar{s}$ (this may be driven, for example, by changes in military technology or international factors, which necessitate greater building of state capacity for military reasons, as in Tilly's (1990) discussion). As Figure 2 in the Introduction illustrates, such a change can leave a polity in the same region as before, in which case the equilibrium trajectory will be shifted uniformly up, but the long-run outcome will remain unchanged (this will correspond to the Montenegrin case we discuss in Section 5). Alternatively, this increase can shift us from, say, Region III to Region II, in which case not only the equilibrium trajectory but also the long-run outcome will change, and in fact it will involve greater state capacity (corresponding to the Swiss case below). However, depending on the exact value of (x_0, s_0), the same increase of $\bar{s}$ in state capacity could also shift us from Region II to Region I, in which case the impact on the long-run

[15] Under Assumption 1, there is no social benefit in reaching the equilibrium $x^* = s^* = 1$, since the capacities of the state and society do not contribute to the size of aggregate income. This will be relaxed below when we consider the general environment in which x and s contribute to income.

[16] In particular, we take $f(x, s) = 0.6$, choose H to be a raised cosine distribution over $[-1, 1]$ with mean $\mu = 0$, and set $c_x(i) = 3.25 \times i^2$ (for $i \in [0, 10]$) and $c_s(i) = 3.25 \times i^2$ (for $i \in [0, 15]$), with $\gamma_x = 0.35$, $\gamma_s = 0.35$, and $\delta = 0.1$.

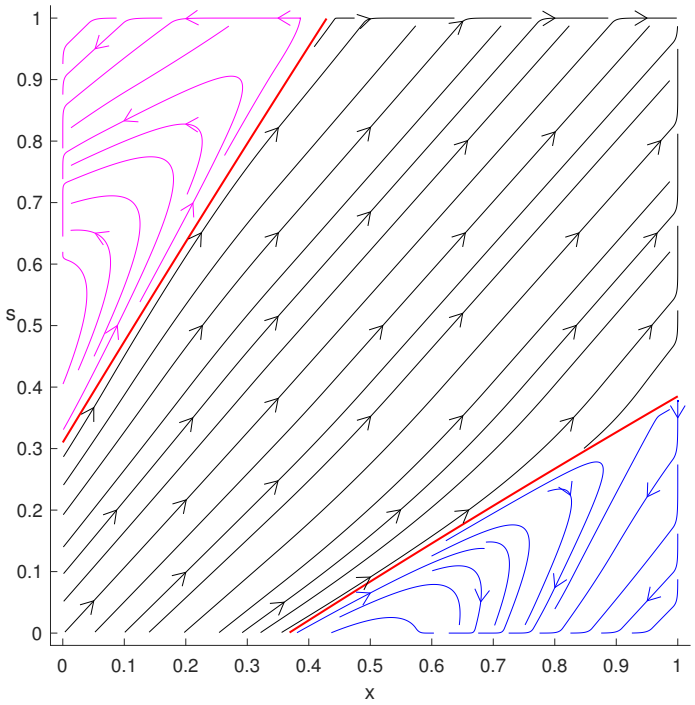

Figure 3: The direction of change of the capacities of state and society in a simulated example.

state capacity will be negative instead of positive as in the Prussian case also discussed below). This discussion thus establishes:

Proposition 2 *The effects of changes in the initial conditions* (x_0, s_0) *on equilibrium dynamics and the long-run outcome for state capacity are conditional in the sense that these depend on which region we move out of and into.*

A more rigorous discussion of conditional comparative statics is provided in the Appendix, where we generalize the model so that the f function depends on state and societal capacity, and we consider variations in the importance of these capacities for income. In that case, even for given initial conditions, dynamics can change fundamentally as the boundaries between regions shift, and once again exactly where we are in terms of initial conditions will determine how elites and society respond to the same change in structural factors.

4 Equilibrium with Forward-Looking Players

In this section, we briefly outline how the results generalize to a setting with long-lived, forward-looking players.

The technology of investment and conflict are the same as in the previous section. The only difference is that now both society and elite are long-lived and forward-looking. To maximize the parallel with the model with short-lived players, we assume that both players again correspond to sequences of non-overlapping generations, but each generation has an exponentially-distributed life-

time or equivalently, a Poisson end date with parameter, $1-\beta$, where $\beta = e^{-\rho\Delta}$. We assume that this random end date is the only source of discounting. Clearly, this specification guarantees that as the period length Δ shrinks, discounting between periods will also decline (and the discount factor will approach 1). Again to maximize the parallel with our myopic model, we assume that in expectation, there is one instance of conflict between the two players during the lifetime of each generation. Since with this Poisson specification, the expected lifetime of his generation is $1/(1-\beta)$, this implies that a conflict arrives at the rate $1-\beta$.[17]

The main result of this section is contained in the next proposition.

Proposition 3 *Suppose Assumptions 1, 2 and 3 hold. Then there exist discount rates $\bar{\rho} \geq \underline{\rho} > 0$ such that for all $\rho > \bar{\rho}$, there are three (locally) asymptotically stable equilibria:*

1. $x^* = s^* = 1$.
2. $x^* = 0$ *and* $s^* \in (\gamma_s, 1)$.
3. $x^* \in (\gamma_x, 1)$ *and* $s^* = 0$.

Moreover, for all $\rho < \underline{\rho}$, there exists a unique globally stable equilibrium $x^* = s^* = 1$.

This result thus shows that the main insights from our analysis apply provided that players, though forward-looking, are sufficiently impatient.[18] We note that this result is not a simple consequence of the fact that as we consider larger and larger values of ρ, players are becoming closer to myopic. It necessitates establishing properties of the the relevant value functions and their derivatives in the limit. Details as well as numerical illustrations are provided in the Appendix.[19]

5 European State Divergence in Historical Perspective

In the Introduction, we discussed the divergent paths of political development in Switzerland, Prussia and Montenegro. In this section, we elaborate on these historical experiences, illustrating how our theory is useful for interpreting these divergent trajectories. This section has two other related aims. First, we illustrate the notion of conditional comparative statics using the salient example of inter-state warfare. Finally, we provide evidence that state capacity is greatest when there is more balanced contestation between state and society, rather than under the auspices of a despotic state.

[17] An alternative specification of the model with long-lived players which leads to identical equations, but eschews the parallel with the myopic model, is to assume that both players are infinitely lived and discount the future at the rate $\beta = e^{-\rho\Delta}$ and there is a conflict during each interval of length Δ. Recall from footnote 13 that in this case there will be no investment when $\Delta \to 0$ with short-lived players (because they do not take into account the benefit from increasing future conflict capabilities), but incentives for investment do not disappear with long-lived, forward-looking players even as $\Delta \to 0$ (because they *do* take into account the benefit from increasing future capacities).

[18] The dependence of the equilibrium set on the degree of impatience, or alternatively how forward looking the players are, is standard in dynamic models. In our case positing a sufficient degree of impatience seems plausible. For example, elites lines may die out, or they may be killed in warfare or their state eliminated in geopolitical competition. These are all interpretations which lead to impatience.

[19] When ρ is between $\underline{\rho}$ and $\bar{\rho}$, we may have a situation in which one of the two corner equilibria disappears while the other one still exists.

5.1 Switzerland

Switzerland was historically on the periphery of the Holy Roman Empire. The Empire still had an emperor, but it had fragmented into many small and relatively independent polities. The component Swiss polities, the cantons, had their own systems of assemblies. These were legacies of two forces, first of the broader pattern of assembly politics inherited from German tribes, which had been institutionalized initially by the Merovingians, and second, from the political consequences of the communal economic organization of the medieval period (Wickham, 2017). Swiss cantons did not just have assemblies; they also featured feudal, often monastic, elites.

The Swiss confederation started in 1291 with men from the cantons of Uri, Schwyz and Unterwalden taking oaths in Rütli, a meadow above Lake Uri, and signing the *Bundesbrief* (Federal Charter). The charter was concerned with public order and lawlessness and committed the three cantons to come to each other's aid and provided a framework for resolving disputes.[20] The disputes were often with the feudal elites. In Schwyz, for example, cantonal authorities were in conflict with Einsiedeln Abbey over access to lands, which the Abbey claimed. In the conflict Schwyz was excommunicated by the Bishop of Constance, and in retaliation the Abbey was attacked and looted.

We see here two very concrete ways in which the power of society to act collectively was manifested. One was through the direct democracy of the assemblies; "The symbol of Swiss freedom has thus been the *Landesgemeinden* — the cantonal assemblies through which direct participation was assured, and on-going self-government guaranteed" (Barber, 1974, p. 11). The other was the process of oath taking, which started in 1291. Oaths were taken, and continually re-affirmed, to coordinate and commit adult men to common goals and help solve the collective action problem. This was such a prevalent feature of Switzerland that when the Swiss Confederacy first appeared, it was known to contemporaries as the *Eidgenossenschaft*, the "oath comradeship".

Uri, Schwyz and Unterwalden were in principle subservient not just to local elites but also to the Habsburg Duke of Austria. The charter stated: "We have further unanimously vowed and established that we in these valleys shall accept no judge who has gained his office for money or for any other price and who is not our resident or native", clearly undercutting the authority of Habsburg judges. After Einsiedeln was sacked, Uri and Unterwalden were also excommunicated. The Habsburgs attempted to use force to help local elites gain control, but were defeated in 1315 at the battle of Morgarten. The three cantons could invoke the *Bundesbrief* to field a collective army. More pacts (and oaths) followed and what came to be known as the Old Swiss Confederacy spread. The Confederacy was continually threatened, for example by Duke Leopold II of Austria, whose army was decisively defeated at the battle of Sempach in 1386 by the combined Swiss forces. The Treaty of Basle in 1499 established the Confederacy's de facto autonomy.

The Swiss state emerged therefore out of a constant conflict between communes (and the cantons) and local elites backed up by the higher political institutions of the Holy Roman Empire. After

[20] https://www.admin.ch/gov/en/start/federal-council/history-of-the-federal-council/federal-charter-of-1291.html

1291 it was the cantons that accumulated power and took on a collective identity. After 1415 an assembly made up of delegates from all the cantons in the Confederacy began to meet regularly thus weakening the power of elites and building republican institutions. The independent Swiss state was in the making. Fribourg and Solothurn were admitted in 1481, Basel and Schaffhausen in 1501, and Appenzell in 1513. The independence of the this state from the Holy Roman Empire was finally recognized by the Treaty of Westphalia in 1648 which ended the Thirty Years War though the consolidation of a modern central state continued for another 200 years.

5.2 Prussia

Prussia was never part of the Holy Roman Empire, but in 1618 it merged through marriage with Brandenburg, which was. The ruling family of Brandenburg, the Hohenzollerns, became the ruling family of Brandenburg-Prussia, with the ruler known as the Elector. During the Thirty Years War the newly-united territories were devastated. Brandenburg might have lost as much as half of its population (Blanning, 2016, p. 12).

In 1640 Frederick William I came to the throne as the new Elector. Known as the Great Elector, he ruled for 48 years. He charted a new course for Brandenburg-Prussia based on the despotic path of state building. One of his main aims was to build a much more effective military. To achieve this, Frederick William needed tax revenues. Taxes had to be negotiated with various representative bodies, such as the Estates of Kurmark in Brandenburg. He started out by trying to get permanent grants of taxation which would free him from the need to endlessly negotiate them. In 1653 he negotiated the so-called Recess that gave him 530,000 thalers over a period of six years. Crucially, he, rather than the Estates, got to collect the taxes which allowed him to start buiding a bureacracy. In exchange he gave the nobility, which made up a chamber in the Estates, tax-exempt status. He went on to extract similar concessions from the Prussian Estates. Frederick William then over-rode the authority of the Estates and started to tax without their agreement. In 1655 he initiated the Kriegskommissariat (the "war commissary") which covered both tax collection and military organization which he also stripped Estates of their militias which responded "with bitter protests" (Clark, 2006, p. 56). By 1659 the Estates retreated to local issues. They attempted to combine their forces, but their resistance was broken by "coercion and force" and "Leading Estates activists were intimidated or arrested" (Clark, 2006, p. 57, and see Carsten, 1951).

With the Estates sidelined despotic state building commenced. In 1733 the basis of military recruitment was reorganized. The country was divided the territory into cantons of 5,000 households with a regiment assigned to each for recruitment. At least a quarter of the male population was included in the rolls dramatically increasing the potential size of the army. Frederick the Great, assuming power in 1740, further expanded the tax base and strengthened the Prussian military machine. The French philosopher Voltaire (supposedly) summarized the Prussian situation as: "Other states possess an army; Prussia is an army which possesses a state."

Our main point here is that this very stark difference from Switzerland evolved out of initial

circumstances that were far more similar than different. Swiss local autonomy and democracy were not entirely unique; there were similar models all over Germany (Brady, 1985) and the representative institutions of the Estates had considerable powers in many parts of the territories. Yet in Prussia, such institutions weakened in the conflict with the Hohenzollerns. This did allow for the accumulation of state capacity, including an expanded fiscal system and army. Yet, as we will see, the Prussian despotic state eventually ended up with less capacity than the Swiss Republic.

5.3 Montenegro

Montenegro was made up of kinship groups, referred to as clans, and lacked the elements of centralization that the Swiss (and Prussians) had inherited from the Carolingians and the Holy Roman Empire. Such kin groups had a great ability to coordinate and they persistently opposed the creation of a state since "Continued attempts to impose centralized government were in conflict with tribal loyalty" (Simić, 1967, p. 87). Prior to 1852 Montenegro was a theocracy, but where the ruling Bishop, the Vladika, could exercise no coercive authority over the clans that dominated the society and "It was only when their central leader attempted to institutionalize forcible means of controlling feuds that the tribesmen stood firm in their right to follow their ancient traditions. This was because they perceived in such interference a threat to their basic political autonomy" (Boehm, 1986, p. 186). Here Boehm is referring to the attempts of Vladika Njegoš to develop a state in Montenegro in the 1840s. Djilas describes the situation as "It was a clash between two principles — the state and the clan. The former stood for order and a nation, and against chaos and treason; the latter stood for clan freedoms and against the arbitrary actions of an impersonal central authority - the Senate, the Guard, the captains" (Djilas, 1966, p. 107). Djilas records that Njegoš' reforms were immediately confronted by the revolt of the Piperi and Crmnica clans motivated by the fact that "The imposition of government and a state was putting an end to the independence and internal freedom of the clans" (Djilas, 1966, p. 115).

Njegoš was succeeded by his nephew Danilo who made himself the first secular Prince of Montenegro in 1851, but his efforts to centralize authority also ran into fierce opposition. An attempt to raise taxes in 1853 led the clans to revolt with the Piperi, Kuči and Bjelopavlići, declaring themselves an independent state. Danilo's attempt failed, and a member of the Bjelopavlići assassinated him in 1860.

5.4 The Nature of the Comparison

Methodologically, this section uses Mill's (1872) "most similar" design for choosing case studies. Switzerland, Prussia and Montenegro diverged in terms of the capacities of their states, and we relate this primarily to differences in initial conditions, in particular how much capacity their state had in the contest with society. Apart from these differences, there were many similarities.

Consider Switzerland and Prussia. We have emphasized the similarities in terms of culture, ethnicity, religion and of the institutions of the Holy Roman Empire. The key difference is that the

Elector of Brandenburg-Prussia had more control over his territories than the Habsburg Dukes had over the Swiss cantons. The Elector was able to override the power of the Estates, while the local Swiss nobility could not stop the cantons signing an agreement that, for example, forbid Austrian judges.

Turning to Switzerland and Montenegro, they both had been peripherally part of the Roman Empire and shared the same type of mountainous ecology and an economy based on herding. We also argued that there were similarities in social structures, in particular clans. The important distinction is that the state was stronger in Switzerland where there was at least the legacy of Carolingian centralization which never existed in Montenegro.

These small differences around 1600 ended up creating very different dynamics of state-building and societal capacity in the early modern period because, in terms of our model, they put the different polities into three different basins of attraction.

5.5 Conditional Comparative Statics in Action

Our case study illustrates our central notion of conditional comparative statics. We focus on the impact of warfare, as discussed in Proposition 2.

For the Swiss, the threat of warfare, in particular the persistent threat from the Habsburgs to reinstate the over-lordship of the feudal local elites, seems to have been an important incentive for the otherwise individualistic cantons and cities to unite into a larger confederation and to accede to more centralized authority and decision-making. This seems to have overcome what might otherwise have been strong centrifugal tendencies. Thus the centralization that started in 1291 was likely encouraged by the pressure of warfare. As Figure 2 in the Introduction highlighted, it is plausible that this impulse towards greater state capacity pushed the Swiss cantons into the basin of attraction of the inclusive state.

The outcome in Prussia was very different. Though Brandenberg had many of the same structural features of Switzerland and Prussia was run by its ruling house, the territories to the east did not have the same history of communes and quasi-democratic politics like the Swiss cantons. Society was more easily dominated. Nevertheless, at the start of the early modern period it is plausible to believe that Prussia was in, or at the very least in the vicinity of, the basin of attraction of the inclusive state. In this light, the impact of Thirty Year War can be thought of as forcing Prussia out of the basin of attraction of the inclusive state and into that of the despotic path. This is how Frederick the Great himself viewed the situation: “So long as God gives me breath, I shall assert my rule like a despot ” (quoted in Blanning, 2015, p. 127)

Finally, war did not make the state in Montenegro and the incentives it created were not powerful enough to move the country out of the basin of attraction of the weak state. The impact of continuous warfare with the Ottomans did induce the clans to try to coordinate more and create more central structures (see Durham, 1928). As we have seen however, this impulse was not sufficiently powerful to create a state.

5.6 The Capacity of the State

One of the most surprising implications of our analysis is that it is not the despotic state which has more capacity, but the inclusive one. This implies an ordering of Switzerland then Prussia and then Montenegro in terms of their levels of state capacity, exactly as in Figure 2.

This can be seen when we focus on the early modern period. We focus on a central metric of state capacity in the literature — the size of the fiscal state.[21] The earliest available evidence suggests that despite the development of the absolutist Prussian state, the tax take relative to national income was higher in Switzerland. This is in line with the predictions of our model. Data in Aidt and Jensen (2009) suggest that central government tax revenues compared to national income were around 1.8 % in the period 1860-1880 in Switzerland rising to 2.1 % in 1881-1914. But the central government was the smallest part of Swiss government tax revenues and expenditures. Cantons had raised income taxes since 1840, while the federal government did not do so until the 1930s. Though we have not been able to identify historical estimates of cantonal taxes, the OECD provides a comprehensive breakdown since 1965.[22] In 2018 cantonal and municipal tax revenues on individuals were five times the levels of federal income taxes while the local corporate taxes were about the same. In 1965 these differences were far larger, while local income and corporate taxes were over 11 and four times as large as federal taxes, respectively. Since in 1860, for example, there were local income taxes but no federal income tax the number of 1.8 % of GDP is clearly a large under-estimate of the actual extent of taxation in Switzerland (Bullock (1924) provides an extensive discussion of the numerous taxes levied by Swiss cantons in the 19th century). Spoerer's (2010) estimates for Prussia are that tax revenues were 5% of national income in 1860. We can therefore conservatively estimate that in 1860 tax revenues relative to national income were twice as high in Switzerland, about 10%, as in Prussia, and thus Switzerland appears to have achieved greater fiscal capacity with an inclusive state.

6 Conclusion

There is a great deal of diversity in the nature of states and their capacity around the world today. But societies, not just states, differ enormously as well. Some are highly mobilized and organized collectively, with high levels of 'social capital', and institutions that facilitate collective actions. In contrast, in others non-elites are weak and incapable of contesting for power against elites and the state. The capabilities of states and societies go together.

[21] There is a clear ranking of other aspects of state capacity in these polities also. This is particularly so with respect to dispute resolution. Prussia retained large feudal and pre-bureaucratic elements in the state, as Rosenberg (1958) documented (see also Ziblatt (2009), and local courts were controlled by feudal elites (See Clark (2009, p. 160), Carsten (1954) and Cerman (2012). This led to an endemic lack of cooperation with society and an inability to implement many policies (as documented bu Raeff (1983), see pp. 45-46, 51). This is in severe contrast to Switzerland which was founded on a demand for the objective resolution of disputes. At the commune level, magistrates were elected and Swiss society then fought a long, and ultimately successful battle, against precisely the type of local despotism that Prussian peasants had to put up with (see Schläppi (2009) for a detailed relevant study of Berne). In Montenegro disputes were mediated by the clans and the feud. This clearly corresponds to much less capacity than the Swiss case, where an institutionalized system of law and justice developed.

[22] https://stats.oecd.org/Index.aspx?DataSetCode=REVCHE

In this paper we have developed a new theory for studying the variation in state capacity and state-society relations, arguing that states endogenously acquire capacity in a dynamic contest between elites and society. At the heart of our model is the notion that elites that control states must contest with society for control over political power and the distribution of resources. If the state accumulates capacity, then this helps elites win this contest. But in response society can also accumulate capacity, and this contestation from society in turn encourages the elite to build further capacity. In our model, this logic leads to three distinct equilibria with very different constellations of state and societal capacity. In one equilibrium, which we called despotic, the state acquires far more capacity than society, in a sense dominating it. In the reverse situation, where society accumulates more capacity than the state, we have a weak state. Finally, a rough balance of power between state and society leads to the emergence of an inclusive state. Our model clarifies how the competition between elites and society in this case is the engine behind the emergence of the greatest state capacity. Elites in despotic states, because they can easily dominate society, have less reason to accumulate as much power and capacity.

Bibliography

Abramson, Scott F. (2017) "The Economic Origins of the Territorial State," *International Organization*, 71, Winter 2017, 97–130.

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

Acemoglu, Daron and James A. Robinson (2019) *The Narrow Corridor: States, Societies and the Fate of Liberty*, New York: Penguin.

Aidt, Toke S. and Peter S. Jensen (2009) "Tax structure, size of government, and the extension of the voting franchise in Western Europe, 1860–1938 " *International Tax and Public Finance*, 16, 362–394.

Amsden, Alice (1989) *Asia's Next Giant: South Korea and Late Industrialization*, New York: Oxford University Press.

Anderson, Perry (1974) *Lineages of the Absolutist State*, London: Verso.

Arjona, Ana (2016) *Rebelocracy*, New York: Cambridge University Press.

Barber, Benjamin R. (1974) *The Death of Communal Liberty: A History of Freedom in a Swiss Mountain Canton*, Princeton: Princeton University Press.

Berwick, Elissa and Fotini Christia (2011) "State Capacity Redux: Integrating Classical and Experimental Contributions to an Enduring Debate, " *Annual Review of Political Science*, 21, 71-91.

Blanning, Timothy (2016) *Frederick the Great: King of Prussia*, New York: Random House.

Blaydes, Lisa (2017) "State Building in the Middle East," *Annual Review of Political Science*,

20, 487-504.

Boehm, Christopher (1986) *Blood Revenge: The Enactment and Management of Conflict in Montenegro and Other Tribal Societies*, Philadelphia: University of Pennsylvania Press.

Brady, Thomas A. (1985) *Turning Swiss: Cities and Empire 1450-1550*, New York: Cambridge University Press.

Brewer, John (1990) *The Sinews of Power: War, Money and the English State, 1688-1783*, Cambridge: Harvard University Press.

Buhaug, Halvard (2010) "Dude, Where's My Conflict?: LSG, Relative Strength, and the Location of Civil War," *Conflict Management and Peace Science*, 27, 2, 107-128.

Bullock, Charles J. (1924) *Selected Readings in Public Finance*, Boston: Ginn and Company.

Cammett, Melani (2014) *Compassionate Communalism: Welfare and Sectarianism in Lebanon*, Ithaca: Cornell University Press.

Cammett, Melani and Lauren M. McLean eds. (2014) *The Politics of Non-State Social Welfare*, Ithaca: Cornell University Press.

Carsten, F.L. (1954) "The Resistance of Cleves and Mark to the Despotic Policy of the Great Elector, " *English Historical Review*, 66, 219-241.

Carsten, F.L. (1954) *Origins of Prussia*, Oxford: Oxford University Press.

Centeno, Miguel A. (1997) "Blood and Debt: War and Taxation in Nineteenth-Century Latin America," *American Journal of Sociology* , 102, 6, 1565-1605.

Centeno, Miguel A., Atul Kohli, Deborah J. Yashar and Dinsha Mistree (2017) *States in the Developing World*, Princeton: Princeton University Press.

Cerman, Markus (2012) *Villagers and Lords in Eastern Europe, 1300-1800*, New York: Palgrave Macmillan.

Clark, Christopher (2009) *Iron Kingdom: The Rise and Downfall of Prussia, 1600-1947*, Cambridge: Bellknap Press.

Dharmapala, Dhammika, Joel Slemrod and John D. Wilson (2011) "Tax Policy and the Missing Middle: Optimal Tax Remittances with Firm-Level Administrative Costs" *Journal of Public Economics*, 95, 9-10, 1036–1047.

Dincecco, Mark, James Fenske and Massimiliano G. Onorato (2019) "Is Africa Different? Historical Conflict and State Development," *Economic History of Developing Regions*, 34, 2, 209-50.

Dincecco, Mark and Yuhua Wang (2018) "Violent Conflict and Political Development over the Long Run: China versus Europe," *Annual Review of Political Science*, 21, 341-58.

Djilas, Milovan (1966) *Njegoš*, New York: Harcourt, Brace and World, Inc.

Durham, M. Edith (1928) *Some Tribal Origins, Laws and Customs of the Balkans*, London: George Allen and Unwin.

Ertman, Thomas (1997) *The Birth of Leviathan: Building States and Regimes in Medieval and Early Modern Europe*, Cambridge: Cambridge University Press.

Fauvelle-Aymar, Christine (1999) "The Political and Tax Capacity of Government in Developing Countries," *Kyklos*, 52(3), 391–413.

textbfEvans, Peter B. (1995) *Embedded Autonomy: States and Industrial Transformation*, Princeton: Princeton University Press.

Fearon, James D. and David D. Laitin (2003) "Ethnicity, Insurgency, and Civil War, " *American Political Science Review*, 97(1), 75–90.

Fukuyama, Francis (2011) *The Origins of Political Order: From Prehuman Times to the French Revolution*, New York: Farrar, Straus and Giroux.

Geddes, Barbara (1994) *Politician's Dilemma: Building State Capacity in Latin America*, Berkeley: University of California Press.

Grzymala-Busse, Anna (2007) *Rebuilding Leviathan Party Competition and State Exploitation in Post-Communist Democracies*, New York: Cambridge University Press.

Grzymala-Busse, Anna (2020) "Beyond War and Contracts: The Medieval and Religious Roots of the European State," *Annual Reviews of Political Science*, 23, 19-36.

Harris, Christopher and John Vickers (1985) "Perfect Equilibrium in a Model of a Race," *Review of Economic Studies* 52 (2), 193-209.

Hechter, Michael and William Brustein (1980) "Regional Modes of Production and Patterns of State Formation in Western Europe," *American Journal of Sociology*, 85(5), 1061-1094.

Herbst, Jeffrey I. (2000) *States and Power in Africa*, Princeton: Princeton University Press.

Hirshleifer, Jack (1989) "Conflict and Rent-Seeking Success Functions: Ratio Versus Difference Models of Relative Success," *Public Choice*, 63, 101–112.

Hoffman, Philip T. (2015) "What Do States Do? Politics and Economic History," *Journal of Economic History*, 75 (2), 303-332.

Hui, Victoria Tin-bor (2005) *War and State Formation in Ancient China and Early Modern Europe*, New York: Cambridge University Press.

Huntington, Samuel (1968) *Political Order in Changing Societies*, New Haven: Yale University Press.

Levi, Margaret (1989) *Of Rule and Revenue*, Berkeley: University of California Press.

Mann, Michael (1993) *The Sources of Social Power: Volume 2, The Rise of Classes and Nation-States, 1760-1914*, New York: Cambridge University Press.

Marwell, Gerald and Pamela Oliver (1993) *The Critical Mass in Collective Action : A Micro-social Theory*, New York, Cambridge University Press.

Mill, John Stuart (1872) *System of logic*, 8th ed. London: Longmans. Green.

Migdal, Joel (1988) *Strong Societies and Weak States: State-Society Relations and State Capabilities in the Third World*, Princeton: Princeton University Press.

Ostrom, Elinor (1990) *Governing the Commons*, New York: Cambridge University Press.

Pearson, Paul T. (2000) "Increasing Returns, Path Dependence, and the Study of Politics," *American Political Science Review*, 94, 2, 251-267.

Powell, Robert (1999) *In the Shadow of Power*, Princeton: Princeton University Press.

Raeff, Marc (1983) *The Well-Ordered Police State*, New Haven: Yale University Press.

Reiter, Dan (1996) *Crucible of Beliefs*, Ithaca: Cornell University Press.

Rosenberg, Hans (1958) *Bureaucracy, Aristocracy and Autocracy*, Cambridge: The Beacon Press.

Rueschemeyer, Dietrich, Evelyne Huber Stephens and John D. Stephens (1992) *Capitalist Development and Democracy*, New York: Cambridge University Press.

Saylor, Ryan (2014) *State Building in Boom Times: Commodities and Coalitions in Latin America and Africa*, New York: Oxford University Press.

Schläppi, Daniel (2009) "Corporate Property, Collective Resources and Statebuilding in Older Swiss History," in Wim Blockmans, André Holenstein, Jon Mathieu and Daniel Schläppi eds. (2009) *Empowering Interactions: Political Cultures and the Emergence of the State in Europe 1300–1900*, Burlington: Routledge.

Scott, James C. (2010) *The Art of Not Being Governed*, New Haven: Yale University Press.

Simić, Andrei (1967) "The Blood Feud in Montenegro," University of California at Berkeley, Kroeber Anthropological Society Special Publications 1.

Skaperdas, Stergios (1992) "Cooperation, Conflict, and Power in the Absence of Property Rights," *American Economic Review*, 82, 4, 720-739.

Skocpol, Theda (1979) *States and Social Revolutions*, New York: Cambridge University Press.

Slater, Daniel (2010) *Ordering Power: Contentious Politics and Authoritarian Leviathans in Southeast Asia*, New York: Cambridge University Press.

Spoerer, Mark (2010) "The Evolution of Public Finances in Nineteenth-Century Germany " in José Luís Cardoso and Pedro Lains eds. *Paying for the Liberal State. The Rise of Public Finance in Nineteenth Century Europe*, New York: Cambridge University Press.

Steele, Abbey (2017) *Democracy and Displacement in Colombia's Civil War*, Ithaca: Cornell University Press.

Swidler, Ann (1986) "Culture in Action: Symbols and Strategies," *American Sociological Review*, 51, 2, 273-286.

Steinberg, Jonathan (2016) *Why Switzerland?* , New York: Cambridge University Press.

Taylor, Brian D. and Roxana Botea (2008) "Tilly Tally: War-Making and State-Making in the Contemporary Third World," *International Studies Review*, 10, 1, 27-56.

Thies, Cameron G. (2005) "War, rivalry, and state building in Latin America," *American Journal of Political Science*, 49 (3), 451-465.

Tilly, Charles (1990) *Coercion, Capital and European States*, Oxford: Basil Blackwell.

Tilly, Charles (1995) *Popular Contention in Great Britain, 1758 to 1834*, London: Paradigm Publishers.

Tilly, Charles (2007) *Democracy*, New York: Cambridge University Press.

Tullock, Gordon (1980) "Efficient rent seeking," in James M. Buchanan, Robert D. Tollison and Gordon Tullock eds., *Towards a Theory of the Rent Seeking Society*, College Station: Texas A&M University Press.

Viterna, Jocelyn (2013) *Women in War: The Micro-Processes of Mobilization in El Salvador*, New York: Oxford University Press.

Wade, Robert H. (1990) *Governing the Market: Economic Theory and the Role of Government in East Asian Industrialization*, Princeton: Princeton University Press.

Wickham, Christopher (2017) *Medieval Europe*, New Haven: Yale University Press.

Wood, Elizabeth J. (2003) *Insurgent Collective Action and Civil War in El Salvador*, New York: Cambridge University Press.

Ziblatt, Daniel (2009) "Shaping Democratic Practice and the Causes of Electoral Fraud," *American Political Science Review*, 103 (1): 1-21.

Online Appendix (Not for Publication)

In this Appendix, we first provide the proofs of all of the results stated in the text. We then present some further discussion on global dynamics, and then provide a generalization of the model where there are direct transitions from despotic to week and from weak to despotic states. The Appendix also includes additional numerical results and a discussion of the microfoundations of the assumptions used in the text.

Proof of Proposition 1

We start with a series of lemmas on the equilibria of this model, and their stability properties. Before presenting these results, we remark that, mathematically, there can be three types of equilibria: (i) those in which the party in question (say society) chooses a positive level capacity, and thus we will have $x_t^* = x^* \in (0,1)$, so that the marginal cost of investment is simply $c_x'(\delta) + \max\{\gamma_x - x^*, 0\}$, which is equal to the benefit from this capacity; (ii) those in which we have zero capacity, in which case the marginal cost of investment, $c_x'(0) + \gamma_x$, is greater than or equal to the benefit from building further capacity; (iii) those in which the party in question has capacity equal to 1, in which case marginal cost of investment, $c_x'(\delta)$, is less than or equal to the benefit from building additional capacity.

Lemma 1 *There exists a (locally) asymptotically stable equilibrium with $x^* = s^* = 1$.*

Proof of Lemma 1. At $x^* = s^* = 1$, the marginal cost of investment for player $z \in \{x, s\}$ is $c_z'(\delta)$, while the marginal benefit starting from this point is $h(0)$, so Assumption 3 ensures that the marginal benefit strictly exceeds the marginal cost, and neither player has an incentive to reduce its investment. Furthermore, because 1 is the maximum level of investment, neither party has the ability to increase it.

We turn next to asymptotic stability of this equilibrium. First note that from (7), the laws of motion of x and s in the neighborhood of $(x^* = 1, s^* = 1)$ are given by

$$\begin{aligned} c_x'(\dot{x} + \delta) &= h(x - s) \text{ if } x < 1 \text{ and } \dot{x} = 0 \text{ if } x = 1 \\ c_s'(\dot{s} + \delta) &= h(s - x) \text{ if } s < 1 \text{ and } \dot{s} = 0 \text{ if } s = 1, \end{aligned} \tag{A1}$$

where we are exploiting the fact that once we are away from the equilibrium, there cannot be an immediate jump and thus the first-order conditions have to hold in view of Assumption 2. We have also used the information that we are in the neighborhood of the equilibrium $(1,1)$ in writing the system for $x > \gamma_x$ and $s > \gamma_s$. Now to establish asymptotic stability, we will show that

$$L(x,s) = \frac{1}{2}(1-x)^2 + \frac{1}{2}(1-s)^2$$

is a Lyapunov function in the neighborhood of the equilibrium $(1,1)$. Indeed, $L(x,s)$ is continuous and differentiable, and has a unique minimum at $(1,1)$. We next verify that in a sufficiently small

neighborhood of $(1,1)$, $L(x,s)$ is decreasing along solution trajectories of the dynamical system given by (A1). Since L is differentiable, for $x \in (\gamma_x, 1)$ and $s \in (\gamma_s, 1)$, we can write

$$\frac{dL(x,s)}{dt} = -(1-x)\dot{x} - (1-s)\dot{s}.$$

First note that since $h(x-s) > c'_x(\delta)$ and $h(s-x) > c'_s(\delta)$ for x and s in a sufficiently small neighborhood of $(1,1)$, we have both $\dot{x} > 0$ and $\dot{s} > 0$. This implies that, in this range, both terms in $\frac{dL(x,s)}{dt}$ are negative, and thus $\frac{dL(x,s)}{dt} < 0$. Moreover, the same conclusion applies when $x = 1$ (respectively when $s = 1$), with the only modification that $\frac{dL(x,s)}{dt}$ no longer includes the $\dot{s}$ (respectively the $\dot{x}$) term, but still continues to be strictly negative, even on the boundary of $[0,1]^2$. Then the asymptotic stability of $(1,1)$ follows from LaSalle's Theorem (which takes care of the fact that our equilibrium is on the boundary of the domain of the dynamical system in question, see, e.g., Walter, 1998). ■

This Lemma shows that under our maintained assumptions, both parties investing at their maximum capacity is a stable equilibrium. Intuitively, this proposition exploits the fact that when the two players are "neck and neck," they both have strong incentives to invest. If instead we had, say, x much larger than s, then from part 1 of Assumption 3, both $h(x-s)$ and $h(s-x)$ would be smaller than $h(0)$, reducing the investment incentives of both parties. The stronger investment incentives around $x^* = s^* = 1$ are key for maintaining this combination as an (asymptotically stable) equilibrium — combined with part 2 of Assumption 3, which ensures that these strong incentives are sufficient to guarantee a corner solution. If the inequality in part 2 of Assumption 3 did not hold, $x^* = s^* = 1$ could not be an equilibrium, and in this case, the only possible equilibria would be those identified in Lemma 2 below.

The local stability of this equilibrium is then established by constructing a Lyapunov function. The use of this method is necessitated by the fact that $x^* = s^* = 1$ is at the corner of the feasible set, $[0,1]^2$, and thus dynamics around it cannot be characterized by using linearization methods.

Our next result identifies two additional locally asymptotically stable equilibria

Lemma 2 *There exist two additional (locally) asymptotically stable equilibria:*

1. *one with $x^* = 0$ and $s^* \in (\gamma_s, 1)$, and*
2. *one with $s^* = 0$ and $x^* \in (\gamma_x, 1)$.*

Proof of Lemma 2. We start with the first statement. Suppose first that $x^* = 0$. Then from (6) an interior equilibrium level of investment requires

$$h(s) = c'_s(\delta) + \max\{0; \gamma_s - s\}.$$

Note that Assumption 3 implies that at $s = 1$, $h(1) < c'_s(\delta)$, and at $s = \gamma_s$, $h(\gamma_s) > c'_s(\delta)$, thus by the intermediate value theorem, there exists s^* between γ_s and 1 satisfying

$$h(s^*) = c'_s(\delta). \tag{A2}$$

Moreover, because h is single peaked and symmetric around 0, $h(s)$ is decreasing in $s \geq \gamma_s$, and thus only a unique s^* satisfying this relationship exists.

We next verify that $x^* = 0$ is indeed consistent with the optimization of society. This follows immediately since

$$h(-s^*) = h(s^*) = c_s'(\delta) < c_x'(0) + \gamma_x,$$

where the first equality follows from the symmetry of h, the second one simply replicates (A2), and the inequality follows from Assumption 2, and establishes that $x^* = 0$ is optimal for society.

The local stability is again established using a Lyapunov argument as in the proof of Lemma 1. Now in the neighborhood of the equilibrium $(x = 0, s = s^*)$, the dynamical system in (7) can be written as

$$\begin{aligned} c_x'(\dot{x} + \delta) &= h(x - s) + \gamma_x - x \text{ if } x > 0 \text{ and } \dot{x} = 0 \text{ if } x = 0, \text{ and} \\ c_s'(\dot{s} + \delta) &= h(s - x), \end{aligned}$$

where we are now using the fact that we are in the neighborhood of $(0, s^*)$ so that $x < \gamma_x$ and $s > \gamma_s$. The dynamical system in (7) in this case can be written as

$$\begin{aligned} \dot{x} &= (c_x')^{-1}(h(x - s) + \gamma_x - x) - \delta \\ \dot{s} &= (c_s')^{-1}(h(s - x)) - \delta. \end{aligned} \tag{A3}$$

We now choose the Lyapunov function

$$L(x, s) = \frac{1}{2}x^2 + \frac{1}{2}(s - s^*)^2,$$

which is again continuous and differentiable, and has a unique minimum at $(0, s^*)$. We next verify that in the neighborhood of $(0, s^*)$, $L(x, s)$ is decreasing along solution trajectories of the dynamical system given by (A3). Specifically, since L is differentiable, for $x \in (0, \gamma_x)$ and $s \in (\gamma_s, 1)$, we can write

$$\frac{dL(x, s)}{dt} = x\dot{x} + (s - s^*)\dot{s}.$$

First note that as $h(-s^*) < c_x'(\delta) + \gamma_x$, for x and s in the neighborhood of $(0, s^*)$,

$$\dot{x} = (c_x')^{-1}(h(x - s) + \gamma_x - x) - \delta < 0. \tag{A4}$$

Then, using a first-order Taylor expansion of (A3) in this neighborhood, we obtain

$$(s - s^*)\dot{s} = \frac{1}{c_s''(\delta)}h'(s^*)(s - s^*)(s - x - s^*) + o(\cdot), \tag{A5}$$

where $o(\cdot)$ denotes second-order terms in x and $s - s^*$.

The desired result follows from the following arguments: (i) for $x \in (0, \gamma_x)$ and $s \in (\gamma_s, 1)$, $|x\dot{x}| > |(s - s^*)\dot{s}|$, regardless of the sign of $(s - s^*)\dot{s}$, as $x \to 0$ and $s \to 0$, $(s - s^*)(s - x - s^*)/x \to 0$, because in the neighborhood of the equilibrium $(0, s^*)$, $\dot{s}$ is of the order $s - s^*$, while $h(-s^*) < c_x'(\delta) + \gamma_x$,

ensuring that $\dot{x} < 0$). Therefore, in the range where $x \in (0, \gamma_x)$ and $s \in (0, \gamma_s)$, $\frac{dL(x,s)}{dt} < 0$. (ii) when $x = 0$, (A5) implies that $(s - s^*)\dot{s} < 0$ in view of the fact that $h'(s^*) < 0$, and thus we have $\frac{dL(x,s)}{dt} < 0$. (iii) when $s = s^*$, (A4) ensures that $\dot{x} < 0$, so that we again have $\frac{dL(x,s)}{dt} < 0$. Then in all three cases, the asymptotic stability of $(0, s^*)$ follows from LaSalle's Theorem (e.g., Walter, 1998).

The proof of the existence, uniqueness and asymptotic stability of the equilibrium with $s^* = 0$ and $x^* \in (\gamma_x, 1)$ is analogous, and is omitted. ■

These two additional equilibria have a very different flavor than the equilibrium in Lemma 1. Now both parties have a lower level of capacity, and one of them is in fact at zero. The intuition is again related to the incentives for investment in capacity: when one party is at zero capacity, $h(\cdot)$ is small for both players, which encourages the first player to build a state with low capacity, and discourages the other player from building further capacity.

Assumptions 2 and 3 play an important role in this lemma as well. Without the boundary conditions in Assumption 3, there could be other equilibria with some of them including investments below γ_x and γ_s. Though these equilibria would be locally unstable (with the same argument as in Lemma 4 below), it would also become harder to ensure that there exists a locally stable equilibrium, making us prefer these assumptions.

The next lemma rules out several types of equilibria.

Lemma 3 *There is no equilibrium with (i) $x^* = s^* = 0$; or (ii) $x^* = 0$ and $s^* \in (0, \gamma_s)$, or $s^* = 0$ and $x^* \in (0, \gamma_x)$; or (iii) $x^* \in (\gamma_x, 1)$ and $s^* \in (\gamma_s, 1)$.*

Proof of Lemma 3. Claim (i) follows immediately, since from part 3 of Assumption 3, we have $h(0) - \gamma_s > c_s'(0)$, so that when $x^* = 0$, the elite will deviate from $s = 0$. Claim (ii) follows directly from the proof of Lemma 2. Finally, for claim (iii), note that an equilibrium with $x^* \in (\gamma_x, 1)$ and $s^* \in (\gamma_s, 1)$ would necessitate

$$\begin{aligned} h(s^* - x^*) &= c_s'(\delta) \\ h(x^* - s^*) &= c_x'(\delta), \end{aligned} \tag{A6}$$

but then from the symmetry of the h function around zero, we have that $h(s^* - x^*) = h(x^* - s^*)$, so that

$$c_s'(\delta) = h(s^* - x^*) = c_x'(\delta),$$

which contradicts part 2 of Assumption 2. ■

There are other types of equilibria that could exist, but the next lemma shows that when they do, they will all be asymptotically unstable.

Lemma 4 *All other (possible) equilibria are asymptotically unstable.*

Proof of Lemma 4. We will prove this lemma by considering three types of equilibria, which exhaust all possibilities.

Type 1: $x^* \in (0, \gamma_x)$ and $s^* \in (0, \gamma_s)$.

The optimality conditions in such a equilibrium are

$$\begin{aligned} h(s^* - x^*) &= c_s'(\delta) + \gamma_s - s^* \\ h(x^* - s^*) &= c_x'(\delta) + \gamma_x - x^*. \end{aligned}$$

The dynamical system (7) now becomes

$$\begin{aligned} \dot{x} &= (c_x')^{-1}(h(x^* - s^*) + \gamma_x - x^*) - \delta \\ \dot{s} &= (c_s')^{-1}(h(s^* - x^*) + \gamma_s - s^*) - \delta. \end{aligned}$$

Since the equilibrium levels of state and civil society capacity are defined by equality conditions in this case, local dynamics can be determined from the linearized system, with characteristic matrix given by

$$\begin{pmatrix} \frac{1}{c_s''(\delta)}[h'(s^* - x^*) + 1] & -\frac{1}{c_s''(\delta)}h'(s^* - x^*) \\ -\frac{1}{c_x''(\delta)}[h'(x^* - s^*) & \frac{1}{c_x''(\delta)}[h'(x^* - s^*) + 1] \end{pmatrix}.$$

Using the fact that from Assumption 3, $h'(s^* - x^*) = -h'(x^* - s^*)$, the determinant of this matrix can be computed as $\frac{1}{c_s''(\delta)c_x''(\delta)} > 0$. Moreover, from part 2 of Assumption 2, we can show that the trace of this matrix is

$$\frac{1}{c_s''(\delta)}[h'(s^* - x^*) + 1] + \frac{1}{c_x''(\delta)}[h'(x^* - s^*) + 1].$$

Once again using Assumption 3, this expression is positive provided that

$$h'(s^* - x^*)(c_s''(\delta) - c_x''(\delta)) \leq c_x''(\delta) + c_s''(\delta). \tag{A7}$$

Assumption 2 ensures that

$$\left|c_s''(\delta) - c_x''(\delta)\right| \leq \frac{c_x''(\delta)}{|h'(s^* - x^*)|},$$

which is a sufficient condition for (A7), establishing that both eigenvalues are positive, and we have asymptotic instability.

Type 2: $x^* \in (\gamma_x, 1)$ and $s^* \in (0, \gamma_s)$, or $x^* \in (0, \gamma_x)$ and $s^* \in (\gamma_s, 1)$. Consider the first of these,

$$\begin{aligned} h(s^* - x^*) &= c_s'(\delta) + \gamma_s - s^* \\ h(x^* - s^*) &= c_x'(\delta). \end{aligned}$$

Now once again, local dynamics can be determined from the linearized system, with characteristic matrix

$$\begin{pmatrix} \frac{1}{c_s''(\delta)}[h'(s^* - x^*) + 1] & -\frac{1}{c_s''(\delta)}h'(s^* - x^*) \\ -\frac{1}{c_x''(\delta)}[h'(x^* - s^*) & \frac{1}{c_x''(\delta)}h'(x^* - s^*) \end{pmatrix}.$$

The trace of this matrix is

$$\frac{1}{c_s''(\delta)}[h'(s^* - x^*) + 1] + \frac{1}{c_x''(\delta)}h'(x^* - s^*),$$

which is positive provided that

$$h'(s^* - x^*)(c''_s(\delta) - c''_x(\delta)) \leq c''_x(\delta).$$

The same argument as in the proof of Type 1 shows that this condition follows from Assumption 2, implying that at least one of the eigenvalues is positive and thus establishing asymptotic instability. The argument for the case where $x^* \in (0, \gamma_x)$ and $s^* \in (\gamma_s, 1)$ is analogous.

Type 3: $s^* = 1$ and $x^* < 1$ or $x^* = 1$ and $s^* < 1$.

We prove the first case (the proof for the second is analogous). Such a equilibrium exists only if

$$\begin{aligned} h(1 - x^*) &\geq c'_s(\delta) \\ h(x^* - 1) &= c'_x(\delta) + \max\{\gamma_x - x^*, 0\}. \end{aligned}$$

Exploiting these conditions, we will show that such an equilibrium cannot be asymptotically stable. To do this, let us distinguish between $x^* > \gamma_x$ and $x^* \leq \gamma_x$. Consider the first one of these. Then consider a perturbation that keeps s^* constant and reduces x^* to $x^* - \varepsilon_x$ for $\varepsilon_x > 0$ small (since it is sufficient to show asymptotic instability for a specific set of perturbations). Then, we have

$$\dot{x} = -\frac{1}{c''_x(\delta)} h'(x^* - 1) - \delta < 0.$$

The sign follows because $h'(x^* - 1) > 0$ from Assumption 3, and implies that x^* decreases away from the equilibrium in question, establishing asymptotic instability. Consider finally the second possibility, with the same perturbation which yields

$$\dot{x} = -\frac{1}{c''_x(\delta)} [h'(x^* - 1) + 1] - \delta < 0,$$

which is also locally asymptotically unstable. This completes the proof of the lemma. ■

Proposition 1 then follows straightforwardly by combining these lemmas. Figure 4 provides a visual representation.

Global Dynamics

We next partially characterize the global dynamics. In particular, we will determine three regions, as shown in Figure 5, separating the phase diagram into basins of attraction of the three asymptotically stable equilibria characterized in the previous subsection. For example, starting from Region I, the dynamics converge to the equilibrium with $x^* = 0$ and $s^* \in (\gamma_s, 1)$; from Region II, convergence is to the equilibrium with $x^* = s^* = 1$; and from Region III, convergence will be to the equilibrium with $x^* \in (\gamma_x, 1)$ and $s^* = 0$. Unfortunately, it is not possible to determine the boundaries of these regions analytically, but we will be able to characterize subsets thereof explicitly.

Consider first Region II, which is the basin of attraction of the equilibrium $x^* = s^* = 1$. Recall that the dynamical system for the behavior of the capacities of society and state take the form given in (7) above. We proceed by first noting that any subset S of $[0, 1]^2$ for which there exists a Lyapunov

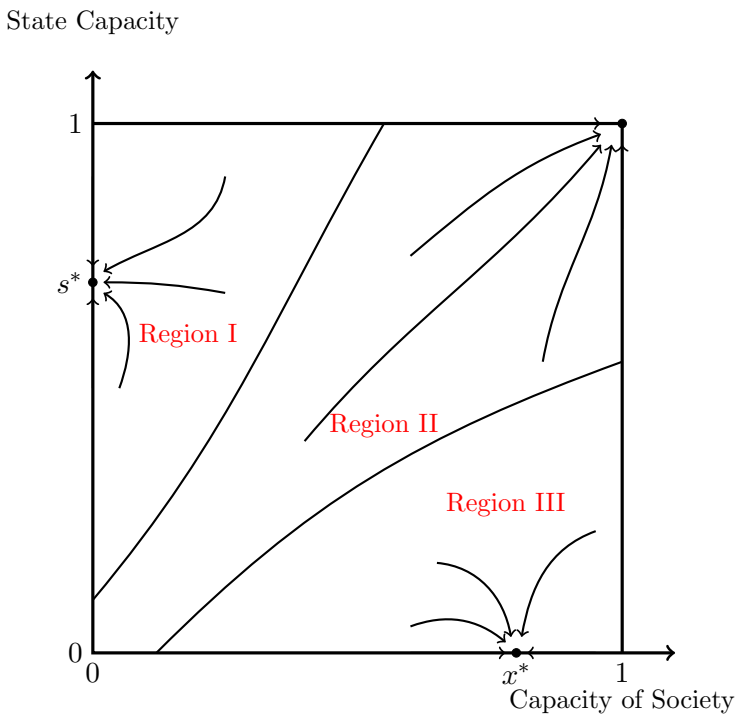

Figure 4: Stable Equilibria and their local dynamics.

function $L(x,s)$ such that (i) $S = \{(x,s) : L(x,s) \leq K\}$ for some $K > 0$; (ii) $L(x,s) \geq 0$ for all $(x,s) \in S$, with equality only if $x = s = 1$; and (iii) $\partial L(x,s)/\partial t \leq 0$ for all $(x,s) \in S$, with equality only if $x = s = 1$, is part of the basin of attraction of this equilibrium.

Let us first construct a subset of the parameters (x,s) such that $\dot{x} \geq 0$ and $\dot{s} \geq 0$, with one of them holding as strict inequality. Let us define $\bar{x}$ such that $c'_x(\delta) = h(\bar{x} - 1)$. Clearly, from Assumption 2 $c'_s(\delta) < h(1 - \bar{x})$. This defines $\mathcal{R}''_{II} = \{(x,s) : x \geq \max\{\gamma_x, \bar{x}\}$ and $s \geq \max\{\gamma_s, \bar{x}\}\}$. This region can be further extended by noting that any combination of (x,s) such that $(c'_x)^{-1}(h(x - s) - \max\{\gamma_x - x, 0\}) - \delta \geq 0$ and $(c'_s)^{-1}(h(s - x) - \max\{\gamma_s - s, 0\}) - \delta \geq 0$ also satisfies $\dot{x} \geq 0$ and $\dot{s} \geq 0$. Let us define $\bar{s}(x)$ such that $h(\bar{s}(x) - x) - \max\{\gamma_s - \bar{s}(x), 0\} - c'_s(\delta) = 0$. Similarly, define $\bar{x}(s)$ such that $h(\bar{x}(s) - s) - \max\{\gamma_x - x, 0\} - c'_x(\delta) = 0$. Both $\bar{s}(x)$ and $\bar{x}(s)$ are upward sloping, and in fact correspond to lines with slope 1 when $s \geq \gamma_s$ and $x \geq \gamma_x$, respectively. Then starting within $\mathcal{R}'_{II} = \{(x,s) : s \leq \bar{s}(x)$ and $x \leq \bar{x}(s)\}$, we also have $\dot{x} \geq 0$ and $\dot{s} \geq 0$ (and in fact, $\mathcal{R}''_{II} \subset \mathcal{R}'_{II}$). This region, as well as $\mathcal{R}''_{II}$, is depicted in Figure 5. The shape of the region is intuitive.

Now consider the family of functions, $L(x,s \mid l_x, l_s) = \frac{l_x}{2}(1-x)^2 + \frac{l_s}{2}(1-s)^2$, indexed by $l_x > 0$ and $l_s > 0$. Clearly, for any member of this family, we have that for all $(x,s) \in \mathcal{R}'_{II} \backslash (1,1)$,

$$\frac{\partial L(x,s \mid l_x, l_s)}{\partial t} = l_x(1-x)\dot{x} - l_s(1-s)\dot{s} < 0.$$

So if we in addition define the subset $\mathcal{R}_{II}$ of $\mathcal{R}'_{II}$ where $L(x,s \mid l_x, l_s) \leq K$, then $\mathcal{R}_{II}$ satisfies the above conditions and by construction is part of the basin of attraction of the equilibrium $(1,1)$.

Now consider the problem of choosing K, l_x and l_s such that we achieve the largest set $\mathcal{R}_{II} =$

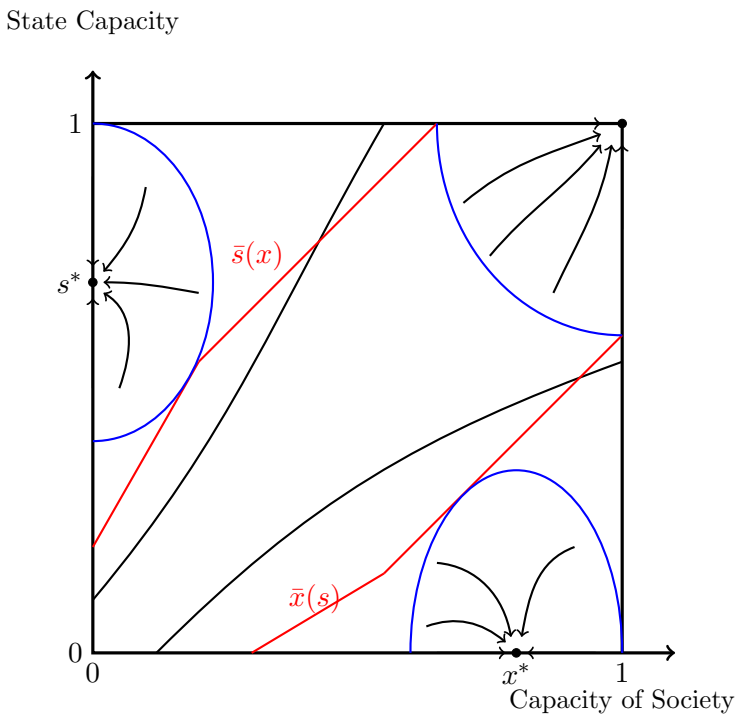

Figure 5: Global Dynamics.

$\{(x, s) : L(x, s \mid l_x, l_s) \leq K\}$ contained in $\mathcal{R}'_{II}$. Mathematically, let $\mathcal{A}(\mathcal{R}_{II})$ be the area of set $\mathcal{R}_{II}$. Then the problem is to choose

$$\max_{K, l_x, l_s > 0} \mathcal{A}(\mathcal{R}_{II}).$$

Figure 5 shows the construction of region $\mathcal{R}_{II}$ in this manner, which is by construction part of the basin of attraction of the equilibrium $(1, 1)$.

Subsets of the basins of attraction of the other equilibria can be constructed analogously and are shown in Figure 5.

Forward-Looking Model

We first provide the basics of the argument for Proposition 3 and then present some additional details and numerical illustration.

Proof of Proposition 3

With the specification introduced above, we can straightforwardly represent the maximization problem of each player as a solution to a recursive, dynamic programming problem, written as

$$V_x(x_{t-\Delta}, s_{t-\Delta}, \beta; \Delta) = \max_{x_t \in [0,1]} \left\{(1-\beta)H(x_t - s_t) - \Delta \cdot C_x\left(x_t, x_{t-\Delta}\right) + \beta V_x(x_t, s'^*(x_{t-\Delta}, s_{t-\Delta}, \beta; \Delta), \beta; \Delta)\right\}, \tag{A8}$$

and

$$V_s(x_{t-\Delta}, s_{t-\Delta}, \beta; \Delta) = \max_{s_t \in [0,1]} \left\{ (1-\beta) H(s_t - x_t) - \Delta \cdot C_s\left(s_t, s_{t-\Delta}\right) + \beta V_s(x'^*(x_{t-\Delta}, s_{t-\Delta}, \beta; \Delta), s_t, \beta; \Delta) \right\}. \tag{A9}$$

Several things are important to note. First, as anticipated in the previous section, we multiply the flow costs with Δ, but not the benefits, since these capture life-time benefits from conflict, and we have conditioned on Δ in writing the value functions for emphasis. Second, notice that we have already imposed the boundary conditions, $x_t \in [0,1]$ and $s_t \in [0,1]$, in the maximization problems. Third, $x'^*(x, s, \beta; \Delta)$ and $s'^*(x, s, \beta; \Delta)$ are the policy functions, which give the next period's values of the state variables as a function of this period's values (and are explicitly conditioned on $\Delta > 0$).

A dynamic equilibrium in this setup as given by a pair of policy functions, $x'^*(x, s, \beta; \Delta)$ and $s'^*(x, s, \beta; \Delta)$ which give the next period's values of the state variables as a function of this period's values (for $\Delta > 0$), and each solves the corresponding value function taking the policy function of the other party is given. Once these policy functions are determined, the dynamics of civil society and state strength can be obtained by iterating over these functions.

Since these are standard Bellman equations, the following result is immediate (throughout this proof we take $(x,\ s,\ \beta) \in [0,1]^3$).

Lemma 5 *For any $\Delta > 0$, $V_x(x, s, \beta; \Delta)$ and $V_s(x, s, \beta; \Delta)$ exist and are continuously differentiable in x, s and Δ. Moreover, $x'^*(x, s, \beta; \Delta)$ and $s'^*(x, s, \beta; \Delta)$ and are continuous in x, s and Δ.*

In particular, from (A8) and (A9), as $\beta \to 0$, $V_x(x, s, \beta; \Delta) \to V_x(x, s, \beta = 0; \Delta)$ and $V_s(x, s, \beta; \Delta) \to V_s(x, s, \beta = 0; \Delta)$. But since $x'^*(x, s, \beta; \Delta)$ and $s'^*(x, s, \beta; \Delta)$ are maximizers of the continuous (and bounded) functions, (A8) and (A9), we can apply Berge's maximum theorem to conclude that $x'^*(x, s, \beta; \Delta)$ and $s'^*(x, s, \beta; \Delta)$ are also continuous, particularly in β, and thus $x'^*(x, s, \beta; \Delta) \to x'^*(x, s, \beta = 0; \Delta)$ and $s'^*(x, s, \beta; \Delta) \to s'^*(x, s, \beta = 0; \Delta)$, and thus for β sufficiently close to 0, we have that $x'^*(x, s, \beta; \Delta)$ and $s'^*(x, s, \beta; \Delta)$ are approximately the same as their myopic values. Therefore, there exists $\bar{\beta} > 0$, such that for all $\beta < \bar{\beta}$, a steady state of the dynamical system given by $x'^*(x, s, \beta; \Delta)$ and $s'^*(x, s, \beta; \Delta)$ exists and is locally stable if and only if it is a locally stable steady state of the myopic model.

This argument establishes that the forward-looking, discrete-time dynamics when the discount factor is sufficiently close to 0 will have the same locally stable steady states as the myopic, discrete-time dynamics. In the previous section, we approximated the discrete-time dynamics with their continuous-time limit, and it is also convenient to do the same here, and to maximize the parallel, this is how we have stated the proposition.

We can also observe that when the discount factor $\beta \to 1$, the two steady states other than $(1, 1)$ disappear. The argument is simple: take the steady state with $x = 0$, where the society's flow return is zero. If civil society invests at a high level for a finite number of periods, this will ensure that $x \geq \gamma_x$, eliminating the region of higher costs of investment for civil society, and thus taking x to 1

(which gives the society a positive flow return). When β is arbitrarily close to 1, the costs of investing at a high level for a finite number of periods are negligible, and hence such a deviation is profitable for civil society. This argument, again from continuity, ensures that there exists $\underline{\beta}^x < 1$ such that for $\beta > \underline{\beta}^x$, $x = 0$ is not consistent with a steady state. With the parallel argument, we also have that there exists $\underline{\beta}^s < 1$ such that for $\beta > \underline{\beta}^s$, $s = 0$ cannot be part of a steady state. Then, for $\beta > \underline{\beta} = \max\{\underline{\beta}^x, \underline{\beta}^s\}$ only $(1, 1)$ remains as an asymptotically stable steady state.

The next subsection discusses the continuous-time limit and also derives the continuous-time Hamilton-Jacobi-Bellman (HJB) equations, which can be used to characterize the equilibrium more generally. We then come back to completing the proof of Proposition 3.

Continuous-Time Approximation

For characterizing the equilibrium for any value of the players' impatience, we can once again use the continuous-time approximation by taking the limit $\Delta \to 0$, which shrinks the period length (and correspondingly adjusts the discount factor $\beta = e^{-\rho\Delta}$, so that the discount rate remains constant at ρ). In this limit, conditions on β translate into conditions on ρ. More specifically, we have:

Lemma 6 *As $\Delta \to 0$, the value functions $V_x(x, s, \beta; \Delta)$ and $V_s(x, s, \beta; \Delta)$ converge to their continuous-time limits $V_x(x, s)$ and $V_s(x, s)$, and the policy functions $x'^*(x, s, \beta; \Delta)$ and $s'^*(x, s, \beta; \Delta)$ converge to their continuous-time limits $x'^*(x, s)$ and $s'^*(x, s)$.*[23]

Proof of Lemma 6. This follows given the continuous differentiability of $V_x(x, s, \beta; \Delta)$ and $V_s(x, s, \beta; \Delta)$ and of $x'^*(x, s, \beta; \Delta)$ and $s'^*(x, s, \beta; \Delta)$ for all $\Delta > 0$. ■

The continuous-time Hamilton-Jacobi-Bellman (HJB) equations can be obtained as follows. First rearrange (A8) evaluated at the optimal choices and divide both sides by Δ to obtain

$$\begin{aligned} & \frac{1-\beta}{\Delta} V_x(x_{t-\Delta}, s_{t-\Delta}, \beta; \Delta) \\ = \; & \max_{x_t \geq 0} \left[\frac{1-\beta}{\Delta} H(x_t - s_t) - C_x\left(x_t, x_{t-\Delta}\right) + \beta \frac{V_x(x_t, s^*_\Delta(x_{t-\Delta}, s_{t-\Delta}), \beta; \Delta) - V_x(x_{t-\Delta}, s_{t-\Delta}, \beta; \Delta)}{\Delta} \right]. \end{aligned}$$

Now note that as $\Delta \to 0$, $(1 - \beta) \to 0$ and $(1 - \beta)/\Delta \to \rho$. Moreover the last term in the previous expression tends to the total derivative of the value function with respect to time. Therefore, the continuous-time HJB equation for civil society is

$$\rho V_x(x, s) = \rho H(x - s) + \max_{\dot{x} \geq -\delta} \left\{ -C_x(x, \dot{x}) + \frac{\partial V_x(x, s)}{\partial x} \dot{x} \right\} + \frac{\partial V_x(x, s)}{\partial s} \dot{s}^*(x, s),$$

where we have used the notation $C_x(x, \dot{x})$ to denote the continuous-time cost function as a function of the change in the conflict capacity of civil society, while $\dot{x}^*(x, s)$ and $\dot{s}^*(x, s)$ designate the continuous-time policy functions, conveniently written in terms of the time derivative of the conflict capacities of the two parties. We have also imposed that $\dot{x}$ cannot be less than $-\delta$.

[23] We also drop the conditioning on the discrete-time discount factor β in writing the continuous-time value and policy functions and do not add conditioning on its continuous-time equivalent, the discount rate ρ to simplify the notation.

Applying the same argument to (A9) and denoting the continues-time cost function for the state by $C_s(s, \dot{s})$, we also obtain

$$\rho V_s(x,s) = \rho H(s-x) + \max_{\dot{s} \geq -\delta} \left\{ -C_s(s,\dot{s}) + \frac{\partial V_s(x,s)}{\partial s}\dot{s} \right\} + \frac{\partial V_s(x,s)}{\partial x}\dot{x}^*(x,s),$$

The first-order optimality conditions for civil society are given by

$$\begin{aligned} \frac{\partial C_x(x,\dot{x})}{\partial \dot{x}} &= \frac{\partial V_x(x,s)}{\partial x} && \text{if} \quad -\delta < \dot{x}(x,s), \text{ and } x \in (0,1), \\ \frac{\partial C_x(x,\dot{x})}{\partial \dot{x}} &\leq \frac{\partial V_x(x,s)}{\partial x} && \text{if} \quad x = 1, \\ \frac{\partial C_x(x,\dot{x})}{\partial \dot{x}} &\geq \frac{\partial V_x(x,s)}{\partial x} && \text{if} \quad \dot{x}(x,s) = -\delta \text{ or } x = 0. \end{aligned} \tag{A10}$$

In the first case, when we have an interior solution, we can also write

$$\dot{x} = \begin{cases} (c_x')^{-1}\left(\frac{\partial V_x(x,s)}{\partial x} - \gamma_x + x\right) & \text{if } x \leq \gamma_x \\ (c_x')^{-1}\left(\frac{\partial V_x(x,s)}{\partial x}\right) & \text{if } x > \gamma_x \end{cases}. \tag{A11}$$

The first-order conditions for the state are also similar, and for an interior solution, they yield

$$\dot{s} = \begin{cases} (c_s')^{-1}\left(\frac{\partial V_s(x,s)}{\partial s} - \gamma_s + s\right) & \text{if } s \leq \gamma_s \\ (c_s')^{-1}\left(\frac{\partial V_s(x,s)}{\partial s}\right) & \text{if } s > \gamma_s \end{cases}. \tag{A12}$$

Numerical Characterization

We next provide a numerical characterization of the dynamics in the forward-looking model. As in the text, we take $f(x,s) = 0.6$, and choose H to be a raised cosine distribution over $[-1,1]$ with mean $\mu = 0$, which is single-peaked and symmetric consistent with Assumption 3.[24] The cost functions of the state and civil society, once again as in the text, are

$$c_x(i) = 3.25 \times i^2 \text{ (for } i \in [0,10]) \text{ and } c_s(i) = 3.25 \times i^2 \text{ (for } i \in [0,15]),$$

and outside of these ranges, the cost functions become vertical, placing a bound on investment levels.[25] In addition, we set $\gamma_x = 0.35$, $\gamma_s = 0.35$, and $\delta = 0.1$. The critical threshold for ρ is computed as $\bar{\rho} = 100$, and for discount rates above this value, the vector field is identical to the one we obtain for the same parameter values in the static model (thus confirming that for high discount rates the equilibrium dynamics of the model with forward-looking agents coincide with the equilibrium of the model with myopic agents as claimed in Proposition 1). Figure 6, presented here, can be contrasted with Figure 3 in the text, which also applies in this dynamic model when $\rho \geq \bar{\rho}$. On the other hand, Figure 6 shows the implied vector field when ρ is smaller than $\bar{\rho}$, illustrating the very different dynamics with smaller discount rates.

[24] Assumption 1 imposed that $f(x,s) = 1$ rather than setting it equal to a constant, say ϕ_0, in order to reduce the number of parameters. We consider a more general surplus function in Assumption 1′ below. Setting $\phi_0 = 0.6$ enables us to construct an example with more equally-sized regions.

[25] This bound plays no role in the numerical results reported here, but facilitates convergence when we consider the dynamic model with the same parameterization and low discount rates in the next section.

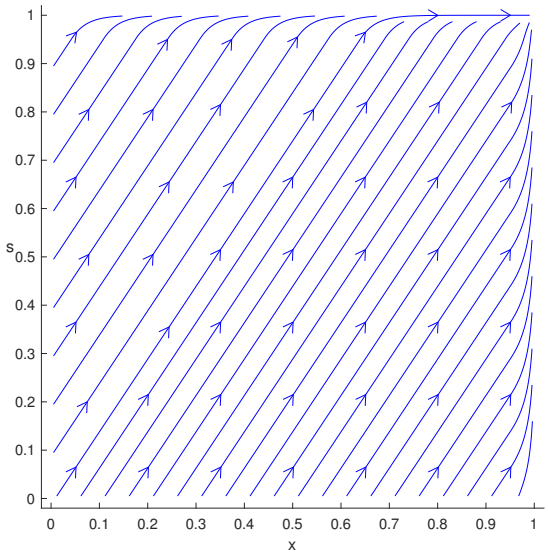

Figure 6: The direction of change of the power of state and society in a simulated example with $\rho = 30$.

Dynamics under Concavity

We now show numerically that the same results as those provided above generalize even when f is concave. Figure 7 depicts the dynamics of state and civil society when we consider the concave surplus function,

$$f(x, s) = .6 + 0.1x^{0.8} + 0.1s^{0.8}.$$

We can see that in this case, the dynamics are very similar to the ones studied in the section where the surplus function is linear.

General Characterization

In this part of the Appendix, we relax Assumption 1. Since we have established the equivalence of the myopic and forward-looking models when the discount rate is sufficiently large in the latter (which is a result that does not depend in any way on Assumption 1), here we focus on a model with forward-looking players. We also simplify the analysis throughout by assuming that f is linear as specified in the next assumption, which replaces Assumption 1.

Assumption 1′ $f(x, s) = \phi_0 + \phi_x x + \phi_s s$, where $\phi_0 > 0$, $\phi_x > 0$ and $\phi_s > 0$.

Our other two assumptions also require some minor modifications, which are provided next.

Assumption 2′ 1. c_x and c_s are continuously differentiable, strictly increasing and weakly convex, and satisfy $\lim_{x\to\infty} c'_x(x) = \infty$ and $\lim_{x\to\infty} c'_s(s) = \infty$.

2.

$$c'_s(\delta) \neq c'_x(\delta).$$

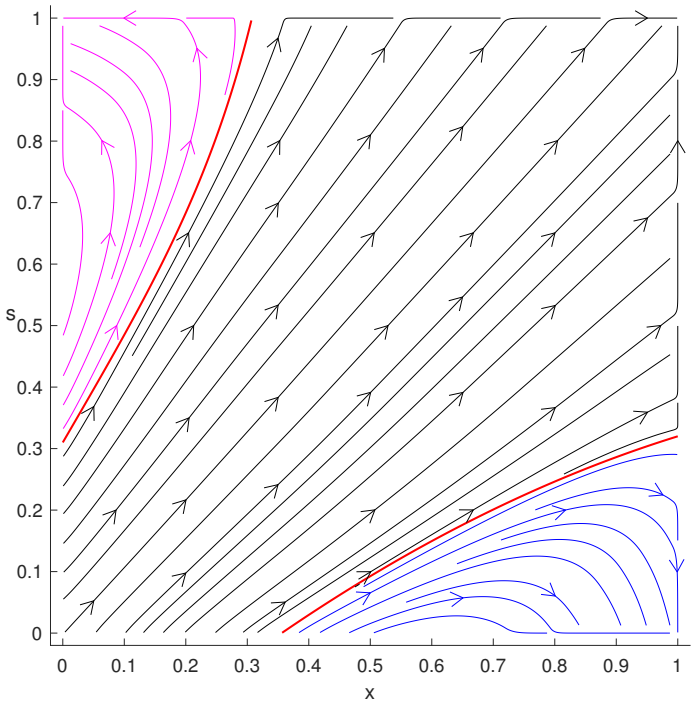

Figure 7: Dynamics when the aggregate output function, $f(x,s)$, is concave.

3.
$$\frac{|c_s''(\delta) - c_x''(\delta)|}{\min\{c_x''(\delta), c_s''(\delta)\}} < \inf_z \frac{2h(z)(\phi_s + \phi_x)}{|h'(z)|\,(\phi_0 + \phi_s + \phi_x)}.$$

4.
$$c_s'(0) + \gamma_s \geq c_x'(\delta) \text{ and } c_x'(0) + \gamma_x > c_s'(\delta).$$

The minor modifications in parts 3 and 4 are in view of the fact that marginal benefits of investment are different between the state and civil society. For same reason, we also modify Assumption 3 as follows.

Assumption 3′ 1. h exists everywhere, and is differentiable, single-peaked and symmetric around zero.

2. For each $z \in \{x, s\}$,
$$c_z'(0) > h(1)(\phi_0 + \phi_z) + H(1)\phi_z.$$

3. For each $z \in \{x, s\}$,
$$\min\{h(0)\phi_0 + H(0)\phi_z - \gamma_z; h(\gamma_z)(\phi_0 + \phi_z\gamma_z) + H(\gamma_z)\phi_z\} > c_z'(\delta).$$

Under these assumptions, the first-order optimality conditions with short-live players (in continuous time) are modified in the following straightforward fashion:

$$\begin{array}{ll}
h(x_t - s_t)(\phi_0 + \phi_x x + \phi_s s) + H(x_t - s_t)\phi_x \leq c_x'(\dot{x}_t + \delta) + \max\{0; \gamma_x - x_t\} & \text{if } \dot{x}_t = -\delta \text{ or } x_t = 0, \\
h(x_t - s_t)(\phi_0 + \phi_x x + \phi_s s) + H(x_t - s_t)\phi_x \geq c_x'(\dot{x}_t + \delta) + \max\{0; \gamma_x - x_t\} & \text{if } x_t = 1, \\
h(x_t - s_t)(\phi_0 + \phi_x x + \phi_s s) + H(x_t - s_t)\phi_x = c_x'(\dot{x}_t + \delta) + \max\{0; \gamma_x - x_t\} & \text{otherwise.}
\end{array}$$

and

$$\begin{array}{ll} h(s_t - x_t)(\phi_0 + \phi_x x + \phi_s s) + H(s_t - x_t)\phi_s \leq c_s'(\dot{s}_t + \delta) + \max\{0; \gamma_s - s_t\} & \text{if } \dot{s}_t = -\delta \text{ or } s_t = 0, \\ h(s_t - x_t)(\phi_0 + \phi_x x + \phi_s s) + H(s_t - x_t)\phi_s \geq c_s'(\dot{s}_t + \delta) + \max\{0; \gamma_s - s_t\} & \text{if } s_t = 1, \\ h(s_t - x_t)(\phi_0 + \phi_x x + \phi_s s) + H(s_t - x_t)\phi_s = c_s'(\dot{s}_t + \delta) + \max\{0; \gamma_s - s_t\} & \text{otherwise,} \end{array}$$

Main Result

We have the following straightforward result.

Proposition 4 *Suppose that Assumptions 1′, 2′ and 3′ hold. Then Propositions 1 and 3 apply.*

Proof of Proposition 4. The proof of this proposition follows directly from the proofs of Propositions 1 and 3, with only minor changes to Lemma 4, which we provide next, ruling out the stability of three different types of steady states. We again treat each type separately.

Type 1: $x \in (0, \gamma_x)$ and $s \in (0, \gamma_s)$.

The optimality conditions in such a steady state are

$$\begin{aligned} h(s - x)(\phi_0 + \phi_x x + \phi_s s) + H(s - x)\phi_s &= c_s'(\delta) + \gamma_s - s \\ h(x - s)(\phi_0 + \phi_x x + \phi_s s) + H(x - s)\phi_x &= c_x'(\delta) + \gamma_x - x. \end{aligned}$$

Local dynamics are in turn given by

$$\begin{aligned} h(s - x)(\phi_0 + \phi_x x + \phi_s s) + H(s - x)\phi_s &= c_s'(\dot{s} + \delta) + \gamma_s - s \\ h(x - s)(\phi_0 + \phi_x x + \phi_s s) + H(x - s)\phi_x &= c_x'(\dot{x} + \delta) + \gamma_x - x. \end{aligned}$$

Since the steady-state levels of state and civil society strength are defined by equality conditions in this case, local dynamics can be determined from the linearized system, with characteristic matrix given by

$$\begin{pmatrix} \frac{1}{c_s''(\delta)}[h'(\cdot)(\phi_0 + \phi_x x + \phi_s s) + 2h(\cdot)\phi_s + 1] & \frac{1}{c_s''(\delta)}[-h'(\cdot)(\phi_0 + \phi_x x + \phi_s s) + h(\cdot)(\phi_x - \phi_s)] \\ \frac{1}{c_x''(\delta)}[h'(\cdot)(\phi_0 + \phi_x x + \phi_s s) + h(\cdot)(\phi_s - \phi_x)] & \frac{1}{c_x''(\delta)}[-h'(\cdot)(\phi_0 + \phi_x x + \phi_s s) + 2h(\cdot)\phi_x + 1] \end{pmatrix},$$

where we wrote $h(\cdot)$ or $h'(\cdot)$ instead of $h(s-x)$ and $h'(s-x)$ in order to save space (and we will adopt this shorthand whenever we write matrices or long expressions below). From part 2 of Assumption 3′, we can show that the trace of this matrix is positive. In particular, the trace is given by

$$\frac{1}{c_s''(\delta)}[h'(\cdot)(\phi_0 + \phi_x x + \phi_s s) + 2h(\cdot)\phi_s + 1] + \frac{1}{c_x''(\delta)}[-h'(\cdot)(\phi_0 + \phi_x x + \phi_s s) + 2h(\cdot)\phi_x + 1].$$

Using Assumption 3′, this expression is positive if

$$h'(s - x)(c_s''(\delta) - c_x''(\delta))(\phi_0 + \phi_x x + \phi_s s) \leq (c_x''(\delta) + c_s''(\delta))(1 + 2h(s - x)(\phi_s + \phi_x)). \tag{A13}$$

Assumption 2′ ensures that

$$\left|c_s''(\delta) - c_x''(\delta)\right| \leq \frac{c_x''(\delta)(1 + 2h(s - x)(\phi_s + \phi_x))}{|h'(s - x)|\,(\phi_0 + \phi_x + \phi_s)},$$

which is a sufficient condition for (A13), establishing that at least one of the eigenvalues is positive, and we have asymptotic instability.

Type 2: $x \in (\gamma_x, 1)$ and $s \in (0, \gamma_s)$, or $x \in (0, \gamma_x)$ and $s \in (\gamma_s, 1)$. Consider the first of these,

$$\begin{aligned} h(s-x)(\phi_0 + \phi_x x + \phi_s s) + H(s-x)\phi_s &= c_s'(\delta) + \gamma_s - s \\ h(x-s)(\phi_0 + \phi_x x + \phi_s s) + H(x-s)\phi_x &= c_x'(\delta). \end{aligned}$$

Now, once again, local dynamics can be determined from the linearized system, with characteristic matrix

$$\begin{pmatrix} \frac{1}{c_s''(\delta)}[h'(\cdot)(\phi_0 + \phi_x x + \phi_s s) + 2h(\cdot)\phi_s + 1] & \frac{1}{c_s''(\delta)}[-h'(\cdot)(\phi_0 + \phi_x x + \phi_s s) + h(\cdot)(\phi_x - \phi_s)] \\ \frac{1}{c_x''(\delta)}[-h'(\cdot)(\phi_0 + \phi_x x + \phi_s s) + h(\cdot)(\phi_s - \phi_x)] & \frac{1}{c_x''(\delta)}[h'(\cdot)(\phi_0 + \phi_x x + \phi_s s) + 2h(\cdot)\phi_x] \end{pmatrix}.$$

The trace of this matrix can now be computed as

$$\begin{aligned} &\frac{1}{c_s''(\delta)}[h'(s-x)(\phi_0 + \phi_x x + \phi_s s) + 2h\phi_s + 1] \\ &+\frac{1}{c_x''(\delta)}[h'(x-s)(\phi_0 + \phi_x x + \phi_s s) + 2h(x-s)\phi_x]. \end{aligned}$$

which is positive if

$$h'(s-x)(c_s''(\delta) - c_x''(\delta))(\phi_0 + \phi_x x + \phi_s s) \leq (c_x''(\delta) + c_s''(\delta))(2h(s-x)(\phi_x + \phi_s)) + c_x''(\delta).$$

The same argument as in the proof of Type 1 establishes that this condition follows from Assumption 2', and thus at least one of the eigenvalues is positive and the steady state in question is asymptotically unstable. The argument for the case where $x \in (0, \gamma_x)$ and $s \in (\gamma_s, 1)$ is analogous.

Type 3: $x = 1$ and $s < 1$ or $s = 1$ and $x < 1$.

Let us prove the first case. Such a steady state would require

$$\begin{aligned} h(1-s)(\phi_0 + \phi_x + \phi_s s) + H(1-s)\phi_x &\geq c_x'(\delta) \\ h(s-1)(\phi_0 + \phi_x + \phi_s s) + H(s-1)\phi_s &= c_s'(\delta) + \max\{0, \gamma_s - s\}. \end{aligned}$$

We distinguish between $s \leq \gamma_s$ and $s > \gamma_s$. Consider the first one of these. Consider a perturbation to $s + \varepsilon_s$ for $\varepsilon_s > 0$ (it is sufficient to consider perturbations that maintain x constant). Then the local dynamics of s are given by:

$$\dot{s} = \frac{1}{c_s''(\delta)}[h'(s-1)(\phi_0 + \phi_x + \phi_s s) + 2h(s-1)\phi_s + 1]\varepsilon_s.$$

From Assumption 3', $h'(s-1) > 0$, the conflict capacity of the state locally diverges from this steady state, establishing asymptotic instability. Consider next the second possibility. In this case, for $s + \varepsilon_s$, we have

$$\dot{s} = \frac{1}{c_s''(\delta)}[h'(s-1)(\phi_0 + \phi_x + \phi_s s) + 2h(s-1)\phi_s]\varepsilon_s,$$

which is also locally asymptotically unstable. The other case is proved identically. ■

Comparative Statics

In this subsection, we discuss how changes in parameters affect the steady states and the dynamics of equilibrium. We focus on the effects of changes in the parameters ϕ_x, ϕ_s, γ_x and γ_s as well as the cost functions c_x and c_s. The effects of changes in initial conditions are identical to those already discussed in the text.

Assumption 3′ guarantees that $x^* = 1$ and $s^* = 1$ is a steady state. There are also at least two interior steady states. These steady states are one of two types. The first type is given by $x^* = 0$ and any s^* that satisfies the following equation:

$$h(s)(\phi_0 + \phi_s s) + H(s)\phi_s = c_s'(\delta).$$

The second type is given by $s^* = 0$ and any x^* that satisfies the following equation

$$h(x)(\phi_0 + \phi_x x) + H(x)\phi_x = c_x'(\delta).$$

Assumption 3′ guarantees that at least one steady state of each type exists. We impose the following assumption to make sure that only one steady state of each type exist:

Assumption 4 $h(y)(\phi_0 + \phi_z y) + H(y)\phi_z$ is a decreasing function of $y \geq 0$ for $z \in \{s, x\}$.

This assumption is fairly mild. The following two conditions would be sufficient to guarantee it: (i) ϕ_z is small, in which case the fact that, from Assumption 3′, $h(y)$ is decreasing for $y \geq 0$ ensures that this assumption is also satisfied, or that (ii) the elasticity of the h function is greater than 1/2, in which case for any value of ϕ_0, Assumption 4 is satisfied.

Let us focus on the comparative statics of the steady state with $x^* = 0$ and $s^* \in (\gamma_s, 1)$. The other case is identical. s^* solves the following equation:

$$h(s^*)(\phi_0 + \phi_s s^*) + H(s^*)\phi_s = c_s'(\delta). \tag{A14}$$

The parameter ϕ_x does not directly appear in this equation. Therefore, $\partial s^*/\partial \phi_x = 0$. Next, implicitly differentiating with respect to ϕ_0 and rearranging, we obtain

$$\frac{\partial s^*}{\partial \phi_0} = \frac{-h(s^*)}{h'(s^*)(\phi_0 + \phi_s s^*) + 2h(s^*)\phi_s} > 0,$$

where the inequality follows from Assumption 4. Implicitly differentiating equation (A14) with respect to ϕ_s, we analogously get

$$\frac{\partial s^*}{\partial \phi_s} = \frac{-h(s^*)s - H(s^*)}{h'(s^*)(\phi_0 + \phi_s s^*) + 2h(s^*)\phi_s} > 0,$$

where again the inequality is a consequence of Assumption 4. Turning next to comparative statics with respect to the cost function, it is straightforward to observe thatγ_s, γ_s, and $c_x(\cdot)$ do not affect the solution of equation (A14). But the marginal cost of increasing capacity affects the location of

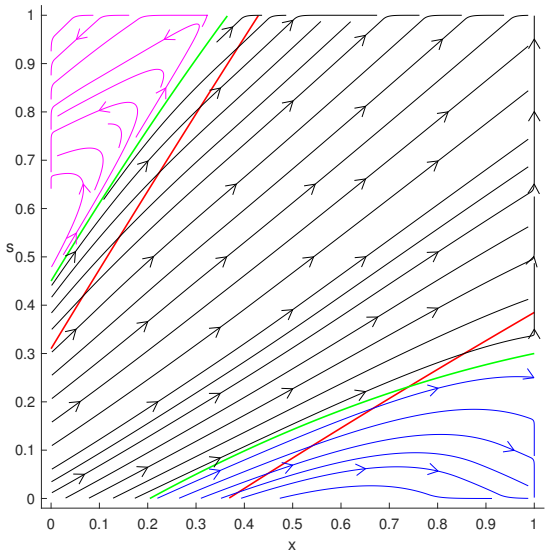

Figure 8: Changes in steady states and dynamics in response to an increase in ϕ_x. The red curves depict the boundaries between the basins of attraction of the different steady states when $\phi_x = 0$ and the green curves show the same boundaries when $\phi_x = 0.1$.

the steady state. To quantify this effect, let us implicitly differentiate equation (A14) with respect to $c_s'(\delta)$ and rearrange to obtain:

$$\frac{\partial s^*}{\partial c_s'(\delta)} = \frac{1}{h'(s^*)(\phi_0 + \phi_s s^*) + 2h(s^*)\phi_s} < 0.$$

Even though there are unambiguous comparative statics of changes from changes in the output and cost functions on s^* and x^*, it has to be borne in mind that these are the values of state and civil society capacity in a given steady state. The more important conclusion continues to be the one already highlighted in Proposition 2, that comparative statics in this model are *conditional*. Proposition 2 emphasized this for initial conditions, but they are no less true when we consider changes in the output or cost functions. For instance, an increase in the marginal benefit of the capacity of civil society on output, ϕ_x, increases x^* as we have just shown. However, such a change also shifts the boundaries of the basins of attraction of the different steady states as depicted in Figure 8. As a result, an economy that was previously in Region II — the basin of attraction of the steady state $(1, 1)$ — can now shift to the basin of attraction of the corners steady state $(0, x^*)$ in Region III. Consequently, the long-run state capacity may end up decreasing rather than increasing following an increase in ϕ_x. This reiterates the conclusions of Proposition 2.

Numerical Results

Figure 8 illustrates how the steady states and the basins of attraction change when we increase ϕ_s, making the capacity of the state more important for overall output. To draw this figure, we use exactly the same parameterization as in the simulation reported in Figure 6, which corresponds to the case in which $\phi_x = \phi_s = 0$ in terms of the model of this section. We then show how the steady

states and dynamics are affected when we increase ϕ_x to .1. Particularly noteworthy are the shifts in the boundaries between the regions, which show that the same type of conditional comparative statics in response to shifts in initial conditions now apply when we consider changes in parameters such as the sensitivity of aggregate surplus to the capacity of the state.

Direct Transitions between Region I and Region III

Figure 3 demonstrates how in our main model, the state space is divided into three regions, and Region II always lies between Regions I and III. However, throughout much of pre-modern history, we have many examples of societies approximating our Regions I and III, but relatively fewer examples of Region II. Perhaps more challengingly for our model, we observe several transitions from Region I directly into Region III, which would not be possible in our baseline model, since Region II is in-between and should be traversed. Here we present a simple modification of the model where Region II shrinks, and creates a subset of the state space (with low levels of state and civil society strength) where Regions I and III are adjacent. The basic idea is to modify the model such that the economies of scale in the cost of investment function becomes dependent on relative strengths.

Suppose that the cost functions for the two players take the form

$$C_x(x_t, x_{t-\Delta}) = c\left(\frac{x_t - x_{t-\Delta}}{\Delta} + \delta\right) + \left[\max\left\{\gamma - x_{t-\Delta}, 0\right\} - \max\left\{\gamma - s_{t-\Delta}, 0\right\}\right]\left(\frac{x_t - x_{t-\Delta}}{\Delta} + \delta\right),$$

and

$$C_s(s_t, s_{t-\Delta}) = c\left(\frac{s_t - s_{t-\Delta}}{\Delta} + \delta\right) + \left[\max\left\{\gamma - s_{t-\Delta}, 0\right\} - \max\left\{\gamma - x_{t-\Delta}, 0\right\}\right]\left(\frac{s_t - s_{t-\Delta}}{\Delta} + \delta\right),$$

where we have made two changes relative to our baseline model. First, we have made c and γ the same for the two players, which is just for simplicity's sake. Second and more important, we have changed the formulation of economies of scale in conflict, so that it is the relative strength of the two players that matters. In particular, when both x and s are less than γ, the second term in the cost function becomes simply a function of the gap between x and s. Clearly this leaves the dynamics when $x_t > \gamma$ and $s_t > \gamma$ unchanged. Consider the case in which $x_t < \gamma$ and $s_t < \gamma$. The differential equations for the strength of society and state can now be written as

$$\begin{aligned} \dot{x} &= (c')^{-1}(h(x-s) + x - s) - \delta \\ \dot{s} &= (c')^{-1}(h(s-x) + s - x) - \delta. \end{aligned}$$

Therefore, defining a new variable $z = x - s$, we have

$$\dot{z} = (c')^{-1}(h(z) + z) - (c')^{-1}(h(z) - z).$$

Or approximating this around $z = 0$, we have

$$\dot{z} = \frac{2z}{c''(\delta)}.$$

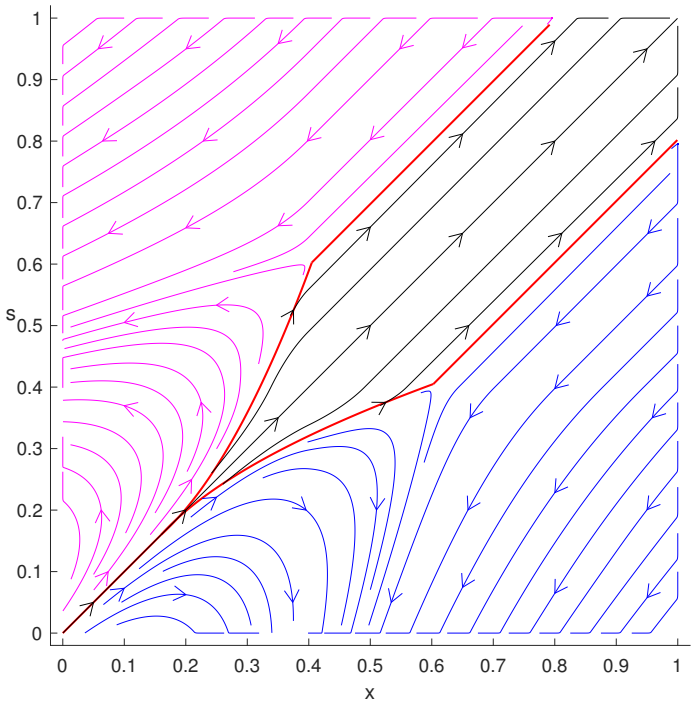

Figure 9: Dynamics with the relative formulation of increasing returns to scale in the cost function for investing in capacity for the state and society. We see that shrinking of Region II and the possibility of direct transitions between Regions I and III.

Thus regardless of whether $x \gtrless s$, the gap between these two variables will grow, with either x or s increasing. Moreover, with x and s sufficiently small, this implies that we converge to one of these two variables being zero. Therefore, we can conclude that there exists a neighborhood of (0, 0), such that starting in this neighborhood, Region II is absent, and the economy will go to either of the two steady states in Regions I or III. This is depicted in Figure 9, where we use exactly the same parameterization as in Figure 3, except that we use the cost functions in this section and also set $c_x(i) = c_s(i) = 9 \times i^2$, and let $\gamma_x = \gamma_s = 0.4$. This pattern implies that starting with low values of state and civil society strength, a society that starts with a weak state could transition directly into one on the path to a despotic state. However, when we consider societies with sufficiently developed states and civil societies, transitions from despotic or weak states could take us towards an inclusive state.

Microfoundations for Economic and Political Decisions

The model presented so far is reduced-form in many dimensions. One of those is the nature of the actions taken by "society". In this section, we briefly outline a model of conflict and production, which maps into the reduced-form setup described so far. Suppose that society consists of a state (ruler) and a number of small producers, each with the production function

$$F(g_t, k_{it}),$$

where g_t is a measure of public good provision (such as infrastructure, bureaucratic services or law enforcement) at time t, and k_{it} designates the capital investment of producer i. To simplify notation in this subsection, we suppress time subscripts.

The cost of public good investment depends on the state's "infrastructural power", which is denoted by s_t. We write this cost as $\Gamma_g(g \mid s)$. This dependence captures the fact that investing in public good provision will be much more difficult for the state when it is not otherwise powerful. There is also a separate cost of increasing the infrastructural power of the state as specified in the text. In addition, this infrastructural power of the state will also determine the state's relationship with society.

The producers, on the other hand, individually choose their capital level, but also jointly choose the extent to which they coordinate their political (and perhaps also economic) actions, which we denote by x. A higher degree of coordination among the producers might (but need not) impact their costs of investing in capital, which we write as $\Gamma_k(k \mid x)$, and this dependence might reflect the fact that a greater degree of coordination among the producers enables them to help each other or develop greater trust in production relations or internalize some externalities. More importantly, such coordination impacts how they can deal with the state's demands, and in the context of our model also stands for social norms that society develops for managing political hierarchy as our historical cases also emphasize. We assume that the cost of investing in x is as specified earlier in this section.

Note that the assumptions that only s and x, and not g and k, build on their non-depreciated stock is for simplicity, and facilitate the comparison with our reduced-form model.

The political game takes the following form: first, the state and civil society simultaneously choose their investments, g and k. Then, the state announces a tax rate τ on the output of the producers. If the producers accept this tax rate, it is collected and the remainder is kept by the producers. If they refuse to recognize this tax rate, there will be a conflict between state and society, the outcome of which will be determined by s and x in a manner similar to the conflict in the text. In particular, the state will win this conflict if (4) above holds, and if so, it can extract the entire output of producers, while if the inequality is reversed, society wins, and the state will not be able to collect any taxes.

The equilibrium can be solved by backward induction within the period, starting from the tax decision of the state. Given the conflict technology we have just specified, it is clear that if the tax rate τ is greater than the likelihood of the state winning the conflict, $H(s-x)$, then there will be a conflict. We may thus focus, without loss of any generality, on the case in which $\tau = H(s-x)$. Then the state's maximization problem can be written as

$$H(s-x)F(g,k) - \Gamma_g(g \mid s) - \tilde{C}_s(s, s_{-\Delta}),$$

where $\tilde{C}_s$ is a cost function for the power of the state similar to the one specified in the text, $s_{-\Delta}$ denotes last period's state strength, and k is the common physical capital investment level of all agents. The solution to this problem for g can be summarized as

$$g = g^*(x, k, s).$$

Note that even though $s_{-\Delta}$ influences s, it does not directly impact the choice of g.

Similarly, recalling that $1 - H(s - x) = H(x - s)$, the maximization problem of citizens can be written as

$$H(x - s)F(g, k) - \Gamma_k(k \mid x) - \tilde{C}_x(x, x_{-\Delta}),$$

with solution

$$k = k^*(x, g, s).$$

Solving this equation together with the equation for g, we can eliminate dependence on the economic decision of the other party, and obtain an equilibrium, expressed as $g = g^{**}(x, s)$, and $k = k^{**}(x, s)$. Substituting these into the payoff functions, we obtain a simplified maximization problem for both players similar to the one described above. In particular, the relevant equations become:

$$H(s - x)f(x, s) - C_s(s, s_{-\Delta} \mid x),$$

and

$$H(x - s)f(x, s) - C_x(x, x_{-\Delta} \mid s),$$

where

$$f(x, s) = F(g^{**}(x, s), k^{**}(x, s)),$$

$$C_s(s, s_{-\Delta} \mid x) = \Gamma_g(g^{**}(x, s) \mid s) + \tilde{C}_s(s, s_{-\Delta})$$

and

$$C_x(x, x_{-\Delta} \mid s) = \Gamma_k(k^{**}(x, s) \mid s) + \tilde{C}_x(x, x_{-\Delta}).$$

The only complication relative to the model presented so far is that because the cost functions depend on the equilibrium action choices of the other player, there may be non-uniqueness issues, and thus the relevant statements now will have to be conditional on a particular equilibrium selection.

Additional Reference:

Walter, Wolfgang (1998) *Ordinary Differential Equations*, New York: Springer-Verlag.