Online Appendix (Not for Publication)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

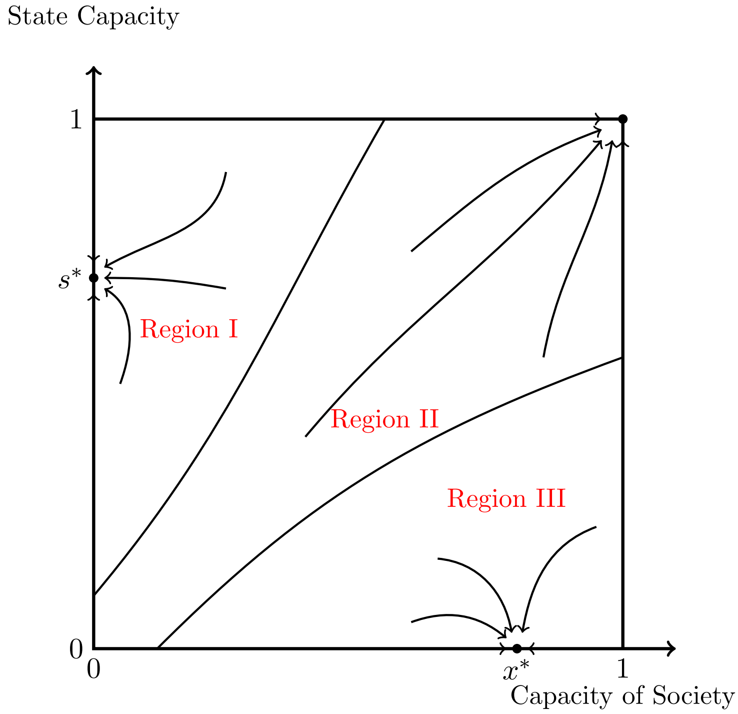

Figure 4: Stable Equilibria and their local dynamics.

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 $\mathcal{R}_{II} =$

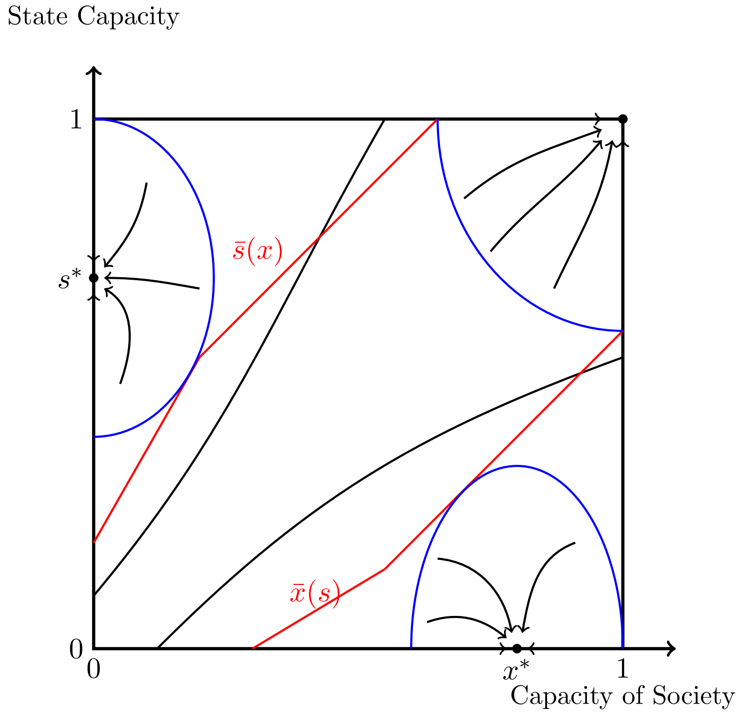

Figure 5: Global Dynamics.

$\{(x, s) : L(x, s \mid l_x, l_s) \leq K\}$ 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

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

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.

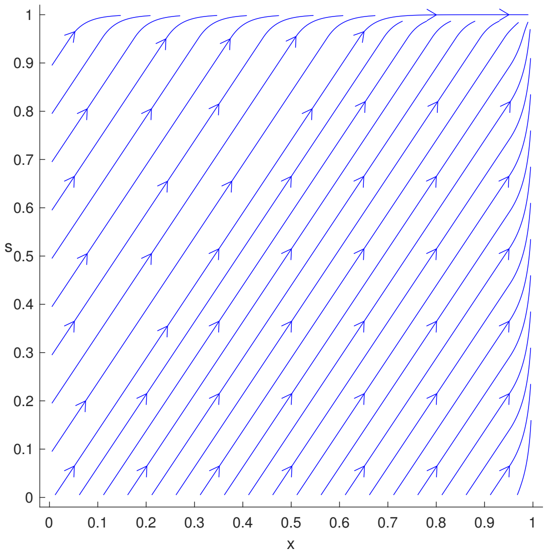

Figure 6: The direction of change of the power of state and society in a simulated example with $\rho = 30$.

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.

$$c_s'(\delta) \neq c_x'(\delta).$$

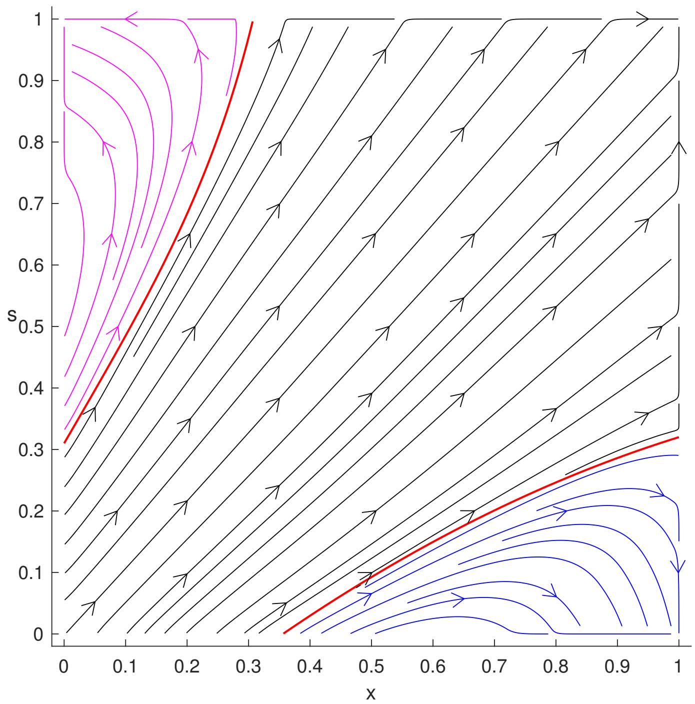

Figure 7: Dynamics when the aggregate output function, $f(x, s)$, is concave.

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:

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

and

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

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

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

Assumption 2′ ensures that

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

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

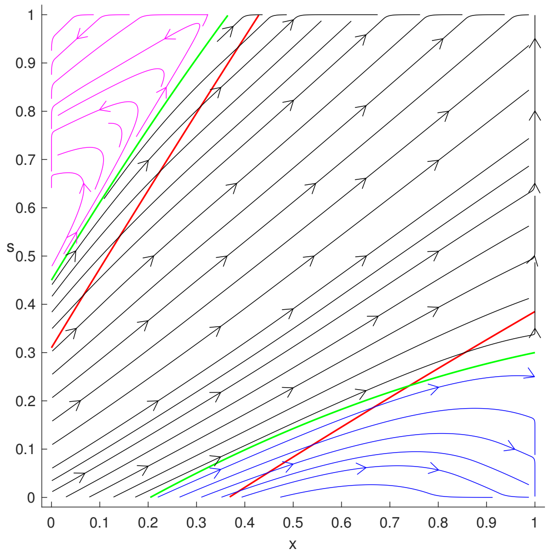

Figure 8: Changes in steady states and dynamics in response to an increase in ϕ_x. The red curves depict the boundaries between the basins of attraction of the different steady states when $\phi_x = 0$ and the green curves show the same boundaries when $\phi_x = 0.1$.

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

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

and

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

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

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

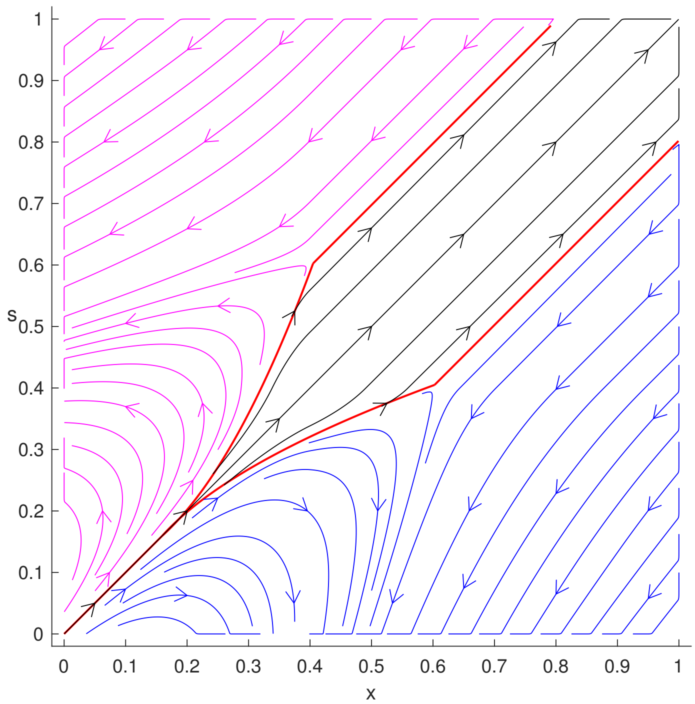

Figure 9: Dynamics with the relative formulation of increasing returns to scale in the cost function for investing in capacity for the state and society. We see that shrinking of Region II and the possibility of direct transitions between Regions I and III.

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.