Elite Persistence, Power Struggles, and Coalition Dynamics*

Xinyu Fan†

Abstract

We study why a ruling minority (the elite) inevitably emerges from power struggles. Power struggles are modeled as iterative coalition formations where players use their power to form alliances and eliminate others without formal commitment. We show that when players can strategically relinquish power, i.e., burn power, it allows a vested interested group to commit to a smaller ruling circle, but also enables outsiders to preemptively cede power to deter regime changes. This ensures the survival of the weak at the cost of power centralization. The elite persist after power perturbations if their commitment to the ruling circle remains intact. The persistence is softened when the weak is productive, when there are external threats, but, surprisingly, not when there are external arbitrators.

JEL Code: D72, D74, O43, P41

Keywords: Power Struggles, Power Relinquishment, Coalition Formation, Elite Persistence

*The author is deeply indebted to Simon Board, as well as Michael Chwe, Moritz Meyer-ter-Vehn, Marek Pycia, Tomasz Sadzik and Daniel Treisman. The author is also grateful to Omer Ali, Nick Doran, Hugo Hopenhayn, Jay Lu, Zhuoran Lu, Ichiro Obara, Lingwei Wu, Chenggang Xu, Feng Yang, Leslie Young, and the seminar participants at UCLA, National University of Singapore, Hong Kong Baptist University, Cheung Kong Graduate School of Business, Peking University, and the Ronald Coase Institute Workshop for Institutional Analysis.

†Cheung Kong Graduate School of Business, Beijing, China. Email: xyfan@ckgsb.edu.cn.

"*Even when the discontent of the masses culminates in a successful attempt to deprive the bourgeoisie of power, this is ... effected only in appearance; always and necessarily there springs from the masses a new organized minority which raises itself to the rank of a governing class.* "

Robert Michels, *Political Parties: A sociological study of the oligarchical tendencies of modern democracy,* 1915, p.390.

1 Introduction

Why do revolutions that aim to install social equality quickly reproduce the same type of hierarchy that the revolutionaries sought to destroy? In George Orwell's terms, it turns out that "some animals are more equal than others". The Bolshevik Revolution, rather than establishing the rule of the proletariat, quickly degenerated into an oligarchy led by Lenin and later the 30-year dictatorship of Stalin (Pipes 2011). Likewise, the French revolutionaries, after overthrowing their king, soon produced an emperor (De Tocqueville 2008). Meanwhile, elites tend to persist after grasping power, as with the al Assads in Syria or the Kims in North Korea. The durability of unequal power distribution has often been attributed to culture - the "serf mentality" of Russians (Havel and Wilson 1985) or Confucian values in Asia (Zhang et al. 2005) - or to the rapacious energy of dominant individuals (Hobbes 1651). A breakdown of hierarchies, as suggested in the seminal paper of Acemoglu, Egorov, and Sonin (2008, hereafter AES), can occur when the alliance of the weak collectively expels the strong in a power struggle and ensures no further fights within such alliance. That is, to sustain a regime when no agreements are credible, power itself should become the *de facto* commitment device. However, it remains a puzzle why the centralization of power is often restored soon after the old regime falls, or in Michels' (1915) words, why "always and necessarily there springs from the masses a new organized minority which raises itself to the rank of a governing class".[1]

The main contribution of this paper is that it connects power adjustment and the inevitable trend of centralization in power struggles. The paper explains how the ability to *relinquish* power serves as an additional commitment device in power struggles that leads to the concentration of power in an elite group, and how this elite group persists during perturbations. Power struggles are modeled as an iterative coalition formation process wherein a group of players uses its power to form alliances, eliminate others, and split fixed resources according to to relative power if it survives the struggles. Power struggles follow the law of the jungle: players cannot commit to protecting themselves from others because there is no one to enforce such

[1]In AES, players have fixed power levels and thus are unable to to respond to elimination threats by adjusting to a lower-ranked position, like Nikolai Bukharin of the Soviet Union after 1929 (Gregory 2013); or by pledging allegiance to others, like the Senat Conservateur to Napoleon Bonaparte after the coup of 18 Brumaire (Doyle 1989).

agreements. This means your friend today can be your enemy tomorrow. However, players have the option to strategically relinquish power, i.e., burn power, before eliminating some others. By doing so, they credibly invite new alliances or buy off key members to avoid future conflicts. In equilibrium, the weak outsiders cede power (and consequently, rents) to a few strong insiders who are capable of ruling in an adjusted smaller circle, until the insiders are sufficiently compensated to preserve the status quo. This ensures the survival of the weak at the cost of increased inequality. Thus, the elite class emerges. We characterize the equilibrium power structures and show that regardless of the immediate direction of small power perturbations, power necessarily ends up concentrated in a few elite members. Thus, the elites persist during ups and downs.

We highlight this inevitable emergence of elites in an *institution-free* environment. We believe it is fundamentally important to understand the origin of elite persistence, without assumptions of external enforcement. The intuition is as follows. Because there is no credible commitment in power struggles, any alliance that intends to eliminate others must ensure that it can survive future power struggles from within. Therefore, a stable regime change often involves power restructuring among the regime changers prior to (the completion of) a revolution or a coup. In particular, a strong player sometimes relinquishes power to credibly assure his or her allies that he or she is not a future threat, in the effort to form a smaller rent-splitting circle. This, on the other hand, gives the excluded players an opportunity to buy off the strong by ceding power preemptively: the strong players receive sufficient benefits from the ceded power to become reluctant to support further regime changes. That is, the flexibility of power relinquishment serves as a commitment device by threatening to form a smaller ruling circle. In response, the players outside the ruling circle (outsiders) must give away their power until some of the players (elites) in the smaller ruling circle prefer the status quo. As a result, power becomes more concentrated, and a hierarchy is established in equilibrium between the outsiders and the elites. After the elites establish their prestigious position, any perturbations that make them stronger naturally centralize the power structure. Meanwhile, any (small) perturbations that weaken their positions re-establish the incentives to eliminate outsiders, which forces the outsiders to further cede power to restore the prestige of the elites. In the end, the hierarchy persists, regardless of the initial direction of the power shift. In more recently popular terms, the enduring inner circle that possesses regime changing and preserving abilities resembles the *deep state* (Lofgren 2016, O'Neil 2017).

To be specific, the model consists of a coalition of agents who split a fixed rent of one according to their respective, endowed levels of power. For example, in a 4-agent coalition with power structure $(2, 4, 5, 10)$, the first agent (with power level of 2) claims rents of $\frac{2}{2+4+5+10} = \frac{2}{21}$, while the remaining agents claim $\frac{4}{21}$, $\frac{5}{21}$ and $\frac{10}{21}$, respectively. If the agents are unsatisfied with the current rent distribution, they can engage in a power struggle to attempt to change the structure. Power struggles are modeled as an iterative process of

two meetings. The first meeting is called a *power burning* meeting, where each agent sequentially proposes a power relinquishment scheme to a subcoalition of agents (an alliance) to prepare for the coming struggle. We assume that agents only adjust power downwards (e.g., disarm an army, resign from office, migrate to a foreign country, etc.), and refer to this process as *power burning*. A burning scheme is enacted only with unanimous consent from the included agents. Otherwise, no power is burnt, and it becomes the next agent's turn to propose. Of course, an individual can burn his own power at will. The burning meeting continues until no agent or subcoalition wishes to further relinquish power.

The second meeting is an *elimination meeting*. Each agent sequentially proposes a new and potentially smaller ruling coalition. The proposal passes if the coalition commands strict majority power and only with the unanimous consent of all agents within the coalition. Otherwise, the proposal is rejected, and it is the next agent's turn to propose. If all proposals are rejected, the game ends, and the rents are split between agents based on the current power structure. If a new coalition is formed, all excluded agents are eliminated (e.g., retired, exiled, or executed). The power struggle cycle continues in the new coalition: first the burning meetings and then the elimination meetings. An equilibrium is reached when no further burning or eliminations take place. Agents collect payoffs in equilibrium according to their status: a surviving agent splits the rents proportional to his or her power; An eliminated agent receives a large negative payoff. The structure of the model emphasizes the far sight in power struggles: one has to foresee the end game when deploying today's strategy. Other players can quickly alternate between being a friend or an enemy during the process.

We show that a subgame perfect Nash equilibrium (SPE) always exists for the power struggle game, and the equilibrium structure is unique given the initial power structure and the sequence of moves. Next, we study the formation of an equilibrium structure to show how hierarchies resurge in equilibrium and how the elites maintain their prestige during power perturbations.

We illustrate the main economic forces using the following example. In power struggles, agents seek to form smaller alliances to obtain higher rents while simultaneously guarding against future eliminations. Recall the example of $(2, 4, 5, 10)$, which resembles a hierarchy with an (initial) elite agent holding power of 10, two medium agents holding power of 4 and 5, and a weak agent holding power of 2. Based on the rules of power struggles, an alliance of the non-elites, $(2, 4, 5)$ is strong enough to overthrow the elite $(2+4+5 > 10)$ in a revolution and split the rent within the alliance.

Power struggles do not stop when the old hierarchy falls. After the old elite is eliminated in the new regime of $(2, 4, 5)$, a new alliance between the stronger agents with power 4 and 5 is powerful enough to eliminate the weak agent. However, the coalition $(4, 5)$ is not stable from within because the stronger agent

with power 5 can further expel the agent with power 4 and become a dictator. More importantly, there is no way for the stronger agent to credibly commit not to do so. Anticipating future power struggles, $(4, 5)$ is not a profitable deviation for the agent with power 4. The current structure $(2, 4, 5)$ is thus stable. This is the idea of self-enforcing coalitions in AES, where power itself becomes the *de facto* commitment device. The self-enforcing coalitions explain the stability of a power structure, but not the direction it may evolve.

However, when agents are allowed to adjust power, the *levels* of power becomes an additional commitment device. First, during the fall of the old hierarchy, the most powerful agent with power 10, facing elimination, can cede power completely to survive, $(2, 4, 5, 10) \to (2, 4, 5, 0)$. Next, in the new regime, $(2, 4, 5, 0)$, the agent with power 5 could credibly commit to its alliance by burning power down to 4, thus inducing a post-burning ruling alliance $(4, 4)$. Such an alliance is not only rent-improving for both participants but also safe from future power struggles because neither individual is powerful enough to expel the other. We call such alliances *stable improving* coalitions. And the existence of such coalitions hinges precisely on the assumption of power relinquishment.

Anticipating the formation of a stable improving alliance, the excluded agent must cede power to survive. It continues to do so until some insider of the alliance becomes indifferent between the status quo and the alternative. In the example, the weakest agent (with power 2) burns power to 1, $(2, 4, 5, 0) \to (1, 4, 5, 0)$ so that the strong insider (with power 5) is rent-indifferent between the current regime $\left(\frac{5}{1+4+5}\right)$ and the stable improving alternative $\left(\frac{4}{4+4}\right)$. The new power structure $(1, 4, 5, 0)$ is the equilibrium structure, wherein the weak cedes power to survive. Lastly, the weak agent in equilibrium enjoys greater rents than in the original power structure $\left(\frac{1}{10} > \frac{2}{21}\right)$, which confirms its incentive to join in the regime change in the first place: it is not because of the interim peace but due to the rents obtained after all future struggles are resolved.

We characterize the equilibrium structure for small-sized coalitions in Section 3, and characterize the boundaries for large-sized equilibrium structures in Section 4. When coalition sizes are small, a power structure reaches equilibrium in power struggles only when there does not exist a strictly stable improving alternative. Furthermore, the shape of the regime necessarily converges to one of the three forms: a dictatorship where all power is concentrated in one individual; an oligarchy where half of the power is concentrated in one individual; or a polyarchy where all individuals have identical power levels. We then show that any power perturbations to the equilibrium structure, regardless of direction, lead to power centralization in the already powerful. This resembles the Matthew effect (Merton et al. 1968), where the rich become richer, and the poor become poorer.[2] In the literature of organizational structures, this is also referred to as the iron law of oligarchy.(Michels 1915)[3] When the coalition sizes become larger, more flexible power structures can

[2]The original idea, as the name suggests, comes from *Matthew 13-12*: Whoever has will be given more, and they will have an abundance. Whoever does not have, even what they have will be taken from them.

[3]The iron law is supported by considerable observations in reality. Magaloni and Kricheli (2010) demonstrate that one-party

be supported in equilibrium. In particular, it is possible that a stable improving alternative exists but fails to be implemented because the transitional path is blocked by early-moving powerful agents that guard the regime. We then categorize the shapes that equilibrium structures can take and provide conditions under which the iron law of oligarchy holds in larger coalitions.

To see the iron law in the model, recall the equilibrium structure in the previous example $(1, 4, 5)$.[4] Suppose a negative perturbation hits the elite, $(1, 4, 5) \to (1, 4, 5 - \epsilon)$. After the perturbation, the stable improving coalition, $(4, 4)$, remains feasible, and moreover, strictly profitable ($\frac{1}{2} > \frac{5-\epsilon}{1+4+5-\epsilon}$). To block this alliance, the weak again cedes power until the new structure becomes $(1 - \epsilon, 4, 5 - \epsilon)$, where the elite restores its rents and thus is reluctant to launch any regime changes. Likewise, suppose a positive perturbation empowers the weak, $(1, 4, 5) \to (1 + \epsilon, 4, 5)$. The alliance $(4, 4)$ again becomes profitable, which forces the weak to cede back the increased power. The stable structure recovers to $(1, 4, 5)$, and the weak remains weak.

On the other hand, suppose a positive perturbation hits the elite, $(1, 4, 5 + \epsilon)$, or a negative perturbation hits the weak $(1 - \epsilon, 4, 5)$. Then the elite has unchecked power to eliminate the rest of the coalition ($5 + \epsilon > 1 + 4$) and become a dictator. To avoid elimination, the two other agents cede power completely to the elite by burning down to 0, and the equilibrium structure ends up as $(0, 0, 5 + \epsilon)$ and $(0, 0, 5)$, respectively. Combining the two cases, a perturbation, regardless of of direction, has a bias towards centralization, and a small perturbation may collapse the power structure to a dictatorship.

Although the model is abstract, the insights apply to many real life scenarios. For example, the founding of the Song Dynasty in imperial China captures all of the previous forces. In 960 AD, at the Bridge of Chen, a military general named Zhao Kuangyin was persuaded by his lieutenants to take the throne and became the new emperor. His lieutenants put a yellow dragon robe, which represented imperial power, on General Zhao. The incident resulted in a peaceful coup as the sitting emperor, 9-year-old Chai Zongxun agreed to cede power. Chai stepped down and lived a peaceful life thereafter (Paludan 1998). Known as the Coup of the Dragon Robe, this encapsulates the change $(2, 4, 5, 10) \to (2, 4, 5, 0)$. However, the power struggles did not end after General Zhao became Emperor Zhao. Two years later, Emperor Zhao encouraged some of his old lieutenants to retire in a grand banquet. The lieutenants agreed with "the tears of joy and regret (of not volunteering to do so earlier)", ceding their military command back to the regime (Tao 2004, Vol. 1). The regime became more centralized, which captures the changes $(2, 4, 5) \to (1, 4, 5)$. After grasping military control, Emperor Zhao further took advantage of a small misstep in local jurisdiction to centralize

rule has become increasingly rampant in modern times. Svolik (2012) and Kendall-Taylor and Frantz (2016) highlight the persistence of successors' power after the deaths of previous rulers.

[4]Here, we ignore the agent with zero power, as he does not participate in the rent distribution.

both administrative and jurisdiction control and soon seized all power for himself. This last episode captures $(1-\epsilon, 4, 5) \to (0, 0, 5)$ (Tao 2004, Vol. 14).

Though the persistence of elites seems inevitable, we consider three scenarios in Section 5, to explore whether it may be softened. The first scenario focuses on the tradeoff between economic development and extraction: Production increases the size of the pie, which benefits all; meanwhile, the power of the producer also grows, which may threaten existing rule. Consequently, when initiating a power struggle is a policy choice for leadership, a far-sighted ruler may adopt a "sow and reap" strategy: soften the squeeze on the weak and aim for economic development for some time, then initiate power struggles before the potential opponents grow too powerful, and resume economic development afterwards. An outcome following such a strategy is the cycling between decentralization (reform) and centralization (extraction). Meanwhile, it predicts that an authoritarian ruler, having consolidated power, would prioritize economic development for some time, before re-centralization. The early years of the Bashar al-Assad regime in Syria and the recent developments under the Kim Jong-un regime in North Korea are consistent with this prediction.

The second and third scenarios concern domestic conflict resolution in the presence of international influence. Intuitively, imminent foreign threats (Nazis on the border of the Soviet Union, Mongolian riders at the gate, etc.) soften domestic conflicts because a united domestic power stands a better chance against the threat. However, when external forces are no longer a threat, domestic coalitions resume their power struggles. As a result, conflicts can be inefficiently prolonged if the involved parties concern the purges that follow the resolution of the conflict. We examine the power struggles in post-war military leaderships in the Soviet Union as a case study. On the other hand, an agent facing expulsion threats can make himself indispensable in the power balance by inviting a strong foreign arbitrator to settle the dispute. As an illustration, we point to the regime changes in Caribbean colonies, where the increasingly powerful citizenry induced the elite to cede power back to the British Crown (Ashdown 1979; Carvalho and Dippel 2016). These two scenarios suggest that foreign threats may soften elite persistence, by making the to-be-expelled indispensable, while foreign arbitration may consolidate elite persistence, by making the elite indispensable.

We conclude the paper in Section 6, by discussing the important assumptions made in the model and the future directions of research. In particular, we highlight the impact of elimination costs (e.g., war casualties) in power struggles. Surprisingly, when elimination costs are positive, centralized regimes are sometimes more likely to be formed. Implications for international relations are discussed.

1.1 Literature

Our discussion of power struggles is related to a number of studies. We echo the models of power struggles (Acemoglu, Egorov, and Sonin 2006, 2008, 2012; Jandoc and Juarez 2017; Guimaraes and Sheedy 2017) in that participants' concern goes beyond the immediate impact of power structure changes to include the aftermath, i.e., the struggles after the struggle. Building on AES, this is the first paper that allows agents to adjust their power to impose elimination threats or respond to such threats. Moreover, we identify the novel role of power relinquishment as a commitment device. While power levels guarantee the stability of coalitions, power relinquishment further shapes the tendency towards power centralization.

Our model is also among the first to combine the role of power in rent distribution (Baron and Ferejohn 1989) with the role of power in endogenously determining coalition composition (Pycia 2012; Ray and Vohra 2015). The seminal work by Baron and Ferejohn (1989) argues that the resource distribution in a legislature depends on the exogenous participation constraints of the involved parties. Further, one must accommodate the necessary alliances to rule. De Mesquita (2005) extend the insight to political survival and argue that the ruler must accommodate the interests of his smallest winning coalition. However, it remains unclear how the participation constraints and the size of the winning coalitions are determined. Our paper provides a microfoundation of how the winning coalitions are endogenously formed in power struggles, thus linking the distribution role of power to its coalition formation role.

In the literature of elite persistence in organizations, especially in non-democracies, Acemoglu and Robinson (2006, 2008) and Robinson (2012) emphasize that office-holding elites implement policies and investments that favor themselves in the future. These insights are also confirmed. in empirical studies of democracies in the context of the United States Congress (Dal Bó, Dal Bó, and Snyder 2009). Their models focus on the competition between an existing elite and, essentially, a representative citizen. This paper adds to the discussion by showing how elites emerge and persist from a generic coalitional structure. To the best of our knowledge, this is the first paper to show that the persistence of elites can be established and sustained through power struggles and far-sighted coalition formation in a zero-sum game, without the opportunity for elites to accumulate power through investments (e.g., Robinson 2008).

Our model aims to provide a framework of regime change without explicit assumptions about the institutional details. In the classical literature about regime change, a menu of established regime choices are usually prerequisites: Acemoglu and Robinson (2005, 2013) on democracy vs. dictatorship; Fearon (2011) on ballot vs. bullet, etc. Discussions based on established regimes provide clear-cut comparisons but implicitly assume away the origin of institutions, i.e., where does this menu come from? However, there are growing attempts to discuss regime changes separate from institutional details (e.g., Aghion and Bolton

2003; Aghion, Alesina, and Trebbi 2004; Gehlbach, Sonin, and Svolik 2016). Following the latter approach, this paper contains no presumed forms of institutions, thus adding to the literature by showing how familiar institutional designs may emerge as equilibrium outcomes.

2 Model

2.1 Setup

Players, power, coalitions. We denote the set of finite *agents* by N_o. Agents are individuals or representatives of groups.[5] A *coalition*, $N = \{1, 2, ..., n\}$, is a group of agents, $N \subset N_o$. Each agent in the coalition is endowed with some power $0 \leq \gamma_i < \infty$. γ_i is discrete and exists on small grids $\mathbb{G} = \{0, \epsilon, 2\epsilon, ...\}$.[6] Agents are labeled from weak to strong in the *power vector*, $\gamma_N = (\gamma_1, \gamma_2, ..., \gamma_n)$, $\gamma_1 \leq \gamma_2 \leq ... \leq \gamma_n$. The power of a coalition N is simply the sum of its members' powers, $\sum_{i \in N} \gamma_i$. Within coalition N, a subcoalition $C \subset N$, where $\gamma_{i \in C} = \gamma_{i \in N}$ for all $i \in C$, is *winning* if $\sum_{i \in C} \gamma_i > \sum_{j \in N \setminus C} \gamma_j$. That is, a subcoalition is winning if it commands strict majority power. Furthermore, we denote $N_+ \subset N$ as the subcoalition in which players have strictly positive power levels.

Players move sequentially in the game. An *agenda* $s \in S$ specifies the sequence of moves and is fixed in a game.[7] We refer to a particular agenda as *top-down* when the (initially) strongest agent always moves first, followed by the second strongest, all the way down the hierarchy to the least powerful agent. An agenda is imposed on individuals but not on power structures.

Strategies. Agents take action in two types of meetings: iteratively. The game starts with stage-j coalition N_j, where in the first stage $N_1 = N$ and proceeds in the following steps:

1. [**Burning meeting**]: The burning meeting takes place in rounds, and the history is visible to all agents. In Round 1, each agent, i, sequentially makes a burning proposal, $0 \leq \rho_i^{j,1}\left(\gamma_{B_{i,1}}^{j,0} | h_{1,s_i}\right) \leq \gamma_{B_{i,1}}^{j,0}$, to agents in subcoalition $B_{i,1} \subset N$. h_{1,s_i} denotes the history up to Round 1 according to agenda s before i burns. Then, agents in $B_{i,1}$ cast their votes sequentially according to s. Each agent votes yes or no. The proposal is passed if and only if all agents included in the proposed coalition vote yes. If passed, the proposal is implemented and power is burnt accordingly; otherwise, no power relinquishments take

[5] For the latter reason, we refer to an agent using the non-gendered "it".

[6] We assume power is discrete to cut the game into a finite horizon for the existence of equilibrium. In real life, power is often discrete, such as the size of an army or the population of one's followers. We also assume that the sizes of grids are uniform. It is useful to think of grids as very small, such that the power relinquishment can almost be continuous, and $\frac{1}{\epsilon} \in \mathbb{N}$, so that all integers are available in the grids.

[7] We assume a fixed agenda because turbulent power struggles often take place in a short period of time, where the sequence of moves of is hard to change due to geographic limits, etc.

place, and we proceed to the next agent. Of course, if an agent chooses to form a one-man coalition of himself, then he can freely burn his power as he likes. A round finishes after all agents have made their proposals, and then proceeds to the next round. In Round 2, each agent continues to make burning proposals $0 \leq \rho_i^{j,2}\left(\gamma_{B_{i,2}}^{j,1}|h_{2,s_i}\right) \leq \gamma_{B_{i,2}}^{j,1}$, following s. The process continues until Round $t+1$ when no burning takes place: that is, for all $i \in N_j$, either $\rho_i^{j,t+1}\left(\gamma_{B_{i,t}}^{j,t}|h_{t,s_i}\right) = \gamma_{B_{i,t}}^{j,t}$, $t \geq 1$, or no proposal is passed. A burning meeting lasts for at least two rounds but ends in finite rounds because power is discrete. When the burning meeting ends, the game proceeds to Step 2.

2. [**Elimination meeting**]: The elimination meeting also takes place in rounds. Following the same agenda s, skipping any eliminated agents or agents with zero power, an agent, i, makes a proposal of a winning subcoalition $P_i^j \subset N_j$ as the new ruling coalition. Agents then cast their votes sequentially according to s. Each agent in P_i^j votes yes or no. The proposal is passed if and only if all agents in P_i^j vote yes. The game proceeds to Step 3 below if the proposal is accepted; otherwise, it becomes the next agent s_{i+1}'s according to s turn to propose. If no individual proposal is accepted in a round, the status quo proposal $P^j = N_j$ is accepted automatically. As a result, an elimination meeting always ends in one round.

3. Agents not in the proposal are eliminated. If the size of the winning coalition shrinks, i.e., $|P_i^j| < |N_j|$, we go to stage $j+1$, where $N_{j+1} = P_i^j$, and start from Step 1 again. If $|P_i^j| = |N_j|$, we go to Step 4.

4. Payoffs are assigned.

Payoffs. Staying in the coalition generates office-holding rents. The total rents to share are fixed and normalized to one. Agents receive rents proportional to their relative power if they survive the power struggles. An agent who survives in coalition N with power structure γ receives payoffs $U_i(N,\gamma) = \frac{\gamma_i}{\sum_{i\in N}\gamma_i}$. An eliminated agent receives $U_i = -n$, which captures the high cost of losing office.

During the elimination meetings, if multiple non-proposing agents have the same power levels, we assume they have equal probabilities of being included in an elimination proposal. Given the large and negative payoff of being eliminated, an agent does not gamble for probabilistic inclusion in the stable coalition:[8]

Assumption 1. Non-proposing agents with the same power levels have equal probabilities of being included in an elimination proposal.

We apply the following tie-breaking assumptions when an agent receives the same rents or has the same power levels as another agent in the elimination meetings:

[8]For example, suppose there is a coalition with power structure $(3,3,3)$ and the agenda is top-down. Any two agents can form a winning coalition and eliminate the remaining coalition. When agent 1 chooses its alliance, it is indifferent between choosing agent 2 or 3, thus assigning probability $\frac{1}{2}$ to including either one. Both agents 2 and 3 face negative expected utility $\frac{1}{2}\left(\frac{1}{2}-n\right) < 0$. Under the top-down agenda, the late-mover of the two, i.e., agent 3, has a better response by burning down to 0 and guaranteeing a payoff of 0.

Assumption 2. Agents always prefer a coalition where they burn less power while receiving the same rents.

Assumption 3. Agents always prefer a coalition formed with fewer meetings when the rents received and the power burnt are the same.

Agents stop attending meetings after they burn down to zero power. Assumption 2 captures the costly nature of power burning, for instance, the dismissal costs of a disarmament. Assumption 3 captures the meeting costs. These costs are usually small in scale relative to the rents, so we only consider them when breaking the ties. The lexicographic tie-breaking assumptions help to identify a unique stable structure in equilibrium.

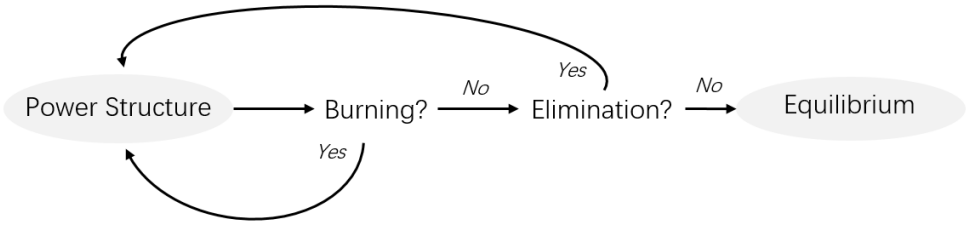

Figure 1. Timing of the game

2.2 Existence and uniqueness of equilibrium

The solution concept is a subperfect Nash equilibrium (SPE). A coalition N^* with a power structure γ^* is in equilibrium if it survives the iterations of power burning and elimination: $(N^*, \gamma^*) \in \phi(N, \gamma|s)$, where ϕ is the correspondence generated by the power struggle game, N is the initial coalition, γ is the initial power structure, and s is the agenda. Power struggles do not last forever because the power levels are discrete and thus the room for adjustment is limited. Consequently, power structures stabilize in finite numbers of meetings. Furthermore, with the tie-breaking assumptions, we are able to select the unique equilibrium that aims to minimize power burning and the number of meetings.

Theorem 1. *There exists an SPE of the power struggle game. For a given agenda, s, and under Assumptions 1–3, the game induces a unique equilibrium power structure for initial coalition N with power structure γ, that is, $(N^*, \gamma^*) = \phi(N, \gamma|s)$.*

Proof. See Appendix. □

Theorem 1 shows that a unique outcome is reached after power struggles if the initial power structure and the agenda are fixed: the direction of coalition dynamics is clear. Denote the initial power structure of

a coalition as the *status quo*. By Theorem 1, if multiple equilibrium structures are reached starting from the same status quo, it is either due to agenda differences or to violations of at least one of the three assumptions. By contrast, multiple status quos may converge to the same equilibrium structure.[9]

Given the significant costs of being eliminated, the losing agents in power struggles always cede power completely to the rest of the coalition upon expulsion and forego the right to claim rents. As a result, no elimination takes place on the equilibrium path.[10]

Corollary 1. *No elimination occurs on the equilibrium path. In any equilibrium,* $(N^*, \gamma^*) = \phi(N, \gamma|s)$, $N^* = N$.

Furthermore, agents who do not have elimination concerns may also give up power to trigger favorable regime changes, thus enriching the potential stable structures. Surprisingly, such richness sharpens the predictions we can make regarding equilibrium structures, which, as shown in the following sections, often have a tendency to centralize.

Hereafter, we suppress notations when there is no confusion about the coalition size or the agenda referred to and denote the equilibrium structure as $\gamma^* = \phi(\gamma)$. Additionally, we refer to to "power struggles taking place" as the case where some power is burnt during the formation of equilibrium, that is, $\gamma^* \neq \gamma$.

3 Elite Persistence in Small-sized Coalitions

The insight of the argument is most easily understood in a 3-agent coalition, which is the minimal organization size that allows both competition for higher rents and alliances for survival.

3.1 Three-Agent Coalitions

Consider a generic 3-agent coalition wherein agents have different power levels. We identify the agents as weak (w), medium (m) and elite (e) based on their power levels. They form a coalition with initial power structure[11] $\gamma = (\gamma_w = a, \gamma_m = b, \gamma_e = c)$, where $0 < a < b < c$. In the absence of a power struggle, the *elite* receives rent of $\frac{c}{a+b+c}$; the *medium* receives $\frac{b}{a+b+c}$; and the *weak* receives $\frac{a}{a+b+c}$.

During power struggles, less powerful agents tend to form alliances to defeat the strong while watching out for future power struggles within the alliance. Such an alliance may not always be feasible: when $a + b < c$, the collective force of the weak and the medium cannot compete with the strength of the elite. Facing the threat of elimination, they must cede power to the elite completely by burning down to 0, $(a, b, c) \to (0, 0, c)$.

[9] In the online appendix, we discuss the impact of different agendas on equilibrium structures and the convergence of status quos in detail.

[10] This is in contrast with AES, where agents can be eliminated due to a lack of opportunities to adjust power to survive.

[11] In the following part of the paper, we use "power structure" and "regime" interchangeably.

As a result, the equilibrium structure becomes $(\gamma_w = 0, \gamma_m = 0, \gamma_e = c)$. An extreme hierarchy is formed, which we naturally refer to as a dictatorship.

When $a + b > c$, no single agent is powerful enough to dictate, but any two-agent alliance is powerful enough to eliminate the third. A myopic weak agent believes that an alliance with the medium agent could collectively expel the elite and lead to higher payoffs in the new regime $(\gamma_w = a, \gamma_m = b)$. However, a far-sighted weak agent understands that without credible commitment, the medium agent becomes dominant after ousting the elite and will proceed to eliminate the weak. As a result, the weak agent expects zero payoff. In anticipation, the alliance of (a, b) is not rent-improving. Likewise, no other two-agent alliance is stable at the current power levels, which makes the original structure (a, b, c) stable in power struggles. This structure is the same as the idea of self-enforcing coalitions in AES.

To solve the commitment problem, agents strategically burn power to prepare for the upcoming struggles. Each agent understands that a two-agent alliance is stable only if the agents are of equal power: each agent receives rent of $\frac{1}{2}$, which is rent-improving for both participants. Meanwhile, the post-burning structure must be powerful enough to oust their opponents. For example, when $a < \frac{c}{2}$, an alliance between the weak and the medium, $\left(\gamma_w^{'}, \gamma_m^{'}\right) = (a, a)$ fails to eliminate the elite. As a comparison, an alliance where the elite burns down to match power with the medium, $\left(\gamma_m^{'}, \gamma_e^{'}\right) = (b, b)$, is powerful enough to triumph in the power struggles.[12]

Anticipating the formation of (b, b) and its consequent elimination, the weak agent adjusts power to block the alliance. It does so by burning power until the elite agent becomes rent-indifferent between the current coalition and the alternative $(\frac{1}{2})$. To be precise, the weak burns to $c - b$, which adjusts the regime to $\gamma^{'} = \left(\gamma_w^{'} = c - b, \gamma_m^{'} = b, \gamma_e^{'} = c\right)$. The structure is in equilibrium because no agent wishes to further burn power or eliminate other agents: the flexibility in power relinquishment provides credible threats leading the elite to burn power, which further leads to the weak agent burning power to survive.

For any agenda, we can follow the practice to look for rent-improving alliances that are powerful enough to win the current power struggles, and are stable to survive into the future. If such an alliance exists, excluded agents must cede power for survival. For 3-agent coalitions, we characterize the equilibrium structures as follows:

Proposition 1. *Starting from a generic power structure* $\gamma = (\gamma_w = a, \gamma_m = b, \gamma_e = c)$, $0 < a < b < c$, *the equilibrium structure converges to one of three forms:*

1. *a dictatorship:* $\gamma^* = (0, 0, c)$ *when* $a + b < c$;
2. *a duopoly:* $\gamma^* = (a, a, 0)$ *when* $a + b \geq c$, $a > \frac{c}{2}$, $s = (m, w, e)$ *or* $s = (w, m, e)$;

[12] There are other options for stable and rent-improving alliances, such as $\left(\gamma_m^{'}, \gamma_e^{'}\right) = (b - \epsilon, b - \epsilon)$ for small ϵ. However, the tie-breaking assumption in Section 2.1 selects the outcome with the least power burnt, i.e., (b, b).

3. an oligarchy: $\gamma^* = (c-b, b, c)$ *otherwise.*

Proof. See Appendix. □

As the proposition shows, a society with unbalanced power necessarily falls into a dictatorship without external forces to enforce rules or commitments. The other social classes fail to collectively compete with the elite and must rally behind the powerful agent for survival.[13] By contrast, when the initial power distribution is relatively equal, the equilibrium structure often ends results in oligarchy, where the rents received by the weak are irrelevant to the initial power distribution. Instead, the structure is ironed down by the power difference between the medium and the elite agents. Power ironing suggests that a variety of initial structures lead to the same "top-heavy" equilibrium, which explains why only a few common forms of organizational structure are observed across geographic locations, cultures, and initial power endowments.[14] We highlight the contrast with a classic paradigm of regime change where the small collectively overthrow the strong and rule without concern of future power changes. The weak agent in the model chooses to cede power instead of allying with the medium agent not because of a lack of ability to overthrow the elite, but because of survival concerns in the future. In this sense, power burning widens participation at the cost of power centralization.

3.2 Characterization

The solution of the 3-person coalition sheds light on the characterization of more general coalition structures. In general, agents seek alliances to eliminate others and split resources within a smaller circle. To prevent future struggles, the power within the alliance should be restructured before elimination. Similar to Proposition 1, we follow an induction approach. Suppose we have solved all $(n-1)$-agent equilibrium structures under all agendas ($n \geq 4$); we now turn to solve the equilibrium for an n-agent coalition under a specific agenda s by examining whether a subcoalition wishes to eliminate the remaining agents. As in 3-agent coalitions, a sub-group of agents in an n-agent coalition chooses to form a smaller coalition among themselves if they are able to adjust into a power structure that is strong enough to win in the current power struggles, is stable in all future struggles and raises rents for participating agents. Formally, we define power structures satisfying such properties to be *stable improving*:

[13] This outcome resembles the *established autocrats* in Svolik (2012), when the autocrats have no comparable competitors. The complete ceding of power to the ruling elite is also consistent with Acemoglu and Robinson's (2017) recent study, which shows that regimes turn despotic when the power of the state is significantly stronger than the power of society.

[14] The balance between the oligarch elite and the alliance of the remaining agents resembles the *contested autocracies* in Svolik (2012), where the autocrat's power is delicately balanced with that of other players. A natural question is whether and how the balance of power may break down and degenerate into unchecked power, i.e., the transition from contested to established autocracies. We address this question in Section 3.2 below.

Definition 1. *$C \subset N$ with γ_C is stable improving given N, γ, s if there exists $\gamma_C^{'} \leq \gamma_C$, s.t. $\sum \gamma_{i\in C}^{'} > \sum \gamma_{j\in N\setminus C}$, and $U_i\left(C, \gamma_C^*\right) \geq U_i\left(N, \gamma\right)$ for all $i \in C$, where $\gamma_C^* = \phi\left(\gamma_C^{'} | s\right)$.*

In particular, (C, γ_C) is strictly stable improving, if $U_i\left(C, \gamma_C^*\right) > U_i\left(N, \gamma\right)$ for all $i \in C$.

In the definition of stable improving coalitions, the adjustable structure $\gamma_C^{'}$ might not coincide with the equilibrium structure γ_C^*. That is, agents may aim for a temporarily unstable deviation if the subsequent stable deviation is desirable. This logic echoes the idea of far-sighted stability in Chwe (1994). It is crucial when some interim power structure, $\gamma_C^{'}$, is winning while γ_C^* is not ($\sum \gamma_{i\in C}^{'} > \sum \gamma_{i\in N\setminus C} \geq \sum \gamma_{i\in C}^*$), that agents in C are willing to first deviate to $\gamma_C^{'}$ to eliminate agents in $N\setminus C$ and then further adjust to γ_C^*.[15] Therefore, the nonexistence of strictly stable improving coalitions is stricter than the nonexistence of immediately profitable deviations because it requires the deviation to be in equilibrium. However, it is weaker than the concept of a coalition-proof equilibrium in Bernheim, Peleg, and Whinston (1987) because it allows for deviations after the deviation.

For small-sized coalitions, we show that there is a one-to-one relationship between strictly stable improving coalitions and equilibrium power structures.

Theorem 2. *For $|N| \leq 7$, a coalition is in equilibrium if and only if there does not exist a subcoalition C that is strictly stable improving. Moreover, equilibrium power structures always take one of the following three forms:*

1. *a dictatorship: $\gamma_i^* > 0$, and $\gamma_j^* = 0$ for all $j \neq i$;*
2. *an oligarchy: $\gamma_i^* = \sum_{j\neq i} \gamma_j$, for some $i \in N$;*
3. *a polyarchy: $N_+^* \in \left\{2^k\right\}$, $k \in \mathbb{N}_+$, and $\gamma_i^* = \gamma_j^*$, for all $i, j \in N_+^*$*

Proof. See Appendix. □

There are two takeaways from Theorem 2. First, in small-sized coalitions, stability is guaranteed by that all subcoalitions with regime-changing capabilities have at least one agent that prefers the status quo. The key reason driving this result is that the degree of freedom in terms of power restructuring is limited when coalition sizes are small. As shown later, the theorem fails to hold for larger coalitions because of a relatively equal structure, with a large number of players enabling a powerful non-elite group to check and balance potential deviations. Second, for small-sized coalitions, equilibrium structures converge to three familiar institutional designs. In most cases, there is either an unequivocal dictator or a strongman there

[15] An example is $\gamma = (3, 4, 6, 6, 6)$ when the agenda is bottom-up, i.e., from the weakest agent to the strongest. A stable improving subcoalition C has a power structure $\gamma^{'} = \left(\gamma_1^{'} = 3, \gamma_2^{'} = 4, \gamma_3^{'} = 6\right)$, and $\gamma_C^* = \left(\gamma_1^* = 2, \gamma_2^* = 4, \gamma_3^* = 6\right)$. Anticipating this, agents 4 and 5 must burn power to avoid elimination, and the equilibrium is $(2, 4, 6, 0, 0, 0)$. In equilibrium, the first agent burns power $(3 \to 2)$ but enjoys higher rents $\left(\frac{2}{12} > \frac{3}{25}\right)$.

is by several less powerful agents who provide checks and balances. This shows that the daily observations of dictatorships and authoritarianisms are necessary equilibrium outcomes instead of random dice rolls. There are occasionally polyarchies formed in equilibrium. In fact, it is easy to show that the three forms of institution design of any size are all equilibrium structures.

Corollary 2. *Dictatorships, oligarchies and polyarchies of all sizes are equilibrium power structures.*

However, as we show in the next sections, the reverse is not true: equilibrium structures can take more forms than the three intuitive designs described above. Moreover, polyarchies are fragile and have a tendency to fall into oligarchies or dictatorships in face of exogenous shocks, while power in oligarchies tends to centralize given perturbations. Therefore, equilibrium outcomes for small-sized coalitions after power struggles not only converge to certain shapes but also converge in a certain direction: centralization.

3.3 Elite persistence

In this section, we analyze how small-sized equilibrium structures respond to power perturbations. Power perturbations are small positive or negative shocks to individual agents[16]. We show that regardless of the equilibrium structure and the direction of the perturbation, power in the post-perturbation structure always becomes weakly more concentrated in the already powerful.

Using a 3-agent coalition as an example, when the equilibrium structure is a dictatorship and the dictator is perturbed, either $(0, 0, c + \epsilon)$ or $(0, 0, c - \epsilon)$ remains an equilibrium and a dictatorship. When the equilibrium is a polyarchy, any perturbation gives the slightly stronger group and advantage in power struggles. As a result, the balance between almost equal players is highly fragile. For instance, a perturbation on a duopoly leads the slightly stronger agent to dictate, i.e., $(a + \epsilon, a, 0) \to (a + \epsilon, 0, 0)$ or $(a - \epsilon, a, 0) \to (0, a, 0)$.

When the equilibrium structure is an oligarchy, power is balanced between the elite circle and the non-elites. We separate the perturbations into those that increase equality within the coalition and those that reduce it. For perturbations that increase equality, suppose a small positive shock $\epsilon > 0$ hits the weak, creating a new power structure $(c - b + \epsilon, b, c)$, or a small negative shock hits the elite, creating a new power structure $(c - b, b, c - \epsilon)$. In both cases, the elite agent receives lower rents immediately after the shock. Thus, the stable and rent-improving alliance $\left(\gamma_m^{'}, \gamma_e^{'}\right) = (b, b)$ becomes more attractive to the elite. Applying Proposition 1, the post-perturbation equilibrium structures are $(c - b + \epsilon, b, c) \to (c - b, b, c)$ and $(c - b, b, c - \epsilon) \to (c - b - \epsilon, b, c - \epsilon)$, respectively. In both cases, the elite agent restores rents of $\frac{1}{2}$. Next, for perturbations that reduce equality, such as $(c - b - \epsilon, b, c)$ or $(c - b, b, c + \epsilon)$, by Proposition 1, the structures

[16]Here we emphasize the impact of unexpected shocks to power structure, because otherwise the players will make buffering room in the power structure.

directly collapse into dictatorships: $(c-b-\epsilon, b, c) \to (0, 0, c)$ and $(c-b, b, c+\epsilon) \to (0, 0, c+\epsilon)$. That is, small changes in favor of equality are usually ineffective in oligarchies as long as the vested interest group retains regime-changing abilities. However, small changes in favor of inequality might lead to a drastic descent into dictatorship because an unchecked power is created. In either case, the prestige of the elite persists, if not strengthens. It is easy to show that the result generalizes to dictatorships, oligarchies and polyarchies of all sizes. Denote the set of all dictatorships, oligarchies and polyarchies as $\mathbb{D}$, and we obtain the following theorem.

Theorem 3. *For equilibrium structure $\gamma_N^* \in \mathbb{D}$ of any size, power is weakly more centralized after perturbation on any individual agents.*

That is, there exists some agent $i \in N$, where $U_i(\gamma_N^*) \geq U_j(\gamma_N^*)$ for all $j \neq i$, and there exists a threshold, $\underline{\eta} > 0$, such that for any perturbation $\eta \in \mathbb{G}$, $\eta \in (-\underline{\eta}, 0) \cup (0, \underline{\eta})$, on any $j \in N$, after the perturbation, $U_i\left(\phi\left[\left(\gamma_j^* + \eta, \gamma_{-j}^*\right)\right]\right) \geq U_i(\gamma^*)$, and $U_i\left(\phi\left[\left(\gamma_j^* + \eta, \gamma_{-j}^*\right)\right]\right) \geq U_j\left(\phi\left[\left(\gamma_j^* + \eta, \gamma_{-j}^*\right)\right]\right)$.

Proof. See Appendix. □

Theorem 3 shows that the most powerful agent before perturbation remains so after perturbation, regardless of the direction of the power shifts. Furthermore, the agent has weakly higher relative power and thus weakly more rents. Combining Theorems 2 and 3, elite persistence for small-sized coalitions are inevitable after power perturbations.

Corollary 3. *For any equilibrium structure with sizes $|N| \leq 7$, power is weakly more centralized after perturbation on any individual agents.*

While it is not surprising that an elite becomes more powerful after positive perturbations, in reality, instances where a weakened elite regains power are not uncommon. In authoritarian regimes, leadership is often transmitted from experienced old rulers to less powerful successors; for example, from Kim Jung-il to Kim Jung-un of North Korea and from Hafez to Bashar al-Assad in Syria. Both sons were less experienced in politics and were not raised as successors to the throne for the majority of their youth. However, both managed to seize power while the old lieutenants ceded theirs. Similar persistence of elite families and identities are documented in other countries.[17] In the next section, we discuss, in more detail, why and how the persistence can be sustained in a general coalition with larger sizes.

[17] See Stone (1990) and Paige (1999) for Central American countries; Dalton (1965) for Liberia; Woodward (1955) and Wiener (1982) for the US South after the civil war.

4 Elite Persistence in Large-sized Coalitions

This section introduces the conditions under which the elite persistence can be extended to larger-sized coalitions. We then provide the real-life interpretations of such conditions.

4.1 Equilibrium characterization

While a power structure remains stable when any subcoalition fails to form a strictly stable improving alternative when coalition sizes increase, there are other ways to sustain stability. In particular, the structure is stable when there exists a stable improving alternative, but the formation of such an alternative is blocked by the mimicking behavior of a prior-moving agent according to the agenda. Denote $N_{\succ_s i}$ as the set of agents who move before agent i in agenda s. The existence of stable improving subcoalitions means some agents are able and willing to form their own alliances and expel the others. However, their restructuring attempt may trigger respective adjustment from other agents in the coalition, which then leaves the original subcoalition undesirable. In fact, the equilibrium structure can be bounded by the nonexistence of strictly stable improving subcoalitions and the nonexistence of an early-moving blocking agent within such subcoalitions.

Theorem 4. *For $|N| \geq 8$, a coalition is in equilibrium if there does not exist a subcoalition C that is strictly stable improving and only if there does not exist a subcoalition C that is strictly stable improving and $\gamma'_{i\in C} > \max\{\gamma_j \mid j \in N_{\succ_s i} \backslash C\}$ for all $i \in C$.*

Proof. See Appendix. □

A summary of the proof follows. If a coalition (N, γ) starts in equilibrium with a stable improving subcoalition, C, then agents in such subcoalition should burn power down to γ'_C to raise their rents. They are able to do so if there is no powerful early-mover who can block the burning process by mimicking ($\gamma'_{i\in C} > \max\{\gamma_j \mid j \in N_{\succ_s i} \backslash C\}$). Therefore, the original power structure cannot be in equilibrium. On the other hand, if a power structure does not have a stable improving subcoalition, then for every adjustable winning subcoalition, C, such that $\gamma'_C \leq \gamma_C$ and $\sum \gamma'_{i\in C} > \sum \gamma_{i\in N\backslash C}$, there is some agent who does not enjoy more rents in the deviation than in the current power structure. Consequently, the deviation proposal fails to pass, and the original structure is in equilibrium. However, the theorem does not eliminate the possibility of an equilibrium structure where rent-improving regime change is blocked by a powerful agent who has first-mover advantage and enjoys more rents in the status quo.

Theorem 4 establishes the boundaries of any equilibrium structure after power struggles. More importantly, it categorizes the two ways to support an equilibrium structure in general. The first is the nonexistence of strictly stable improving coalitions, as in the small-sized coalitions: every coalition with regime-changing

capabilities has an agent that prefers the status quo, as it frees the agents from the concern of "my friend today becomes my foe tomorrow"[18]. The second way to support an equilibrium structure is to have one (or more) powerful early-moving agent, who prefers staying in the current power structure and is able to block a strictly stable improving coalition by threatening to mimic their burning choices.

4.2 Elite persistence

In a small-sized coalition, the elite persist because the inner circle with regime-changing abilities remains intact during equality-enhancing perturbations. Moreover, the power struggles triggered by perturbations only end when the outsiders cede just enough power to accommodate the insiders. In this section, we show that such intuitions extend to a general environment where power burning actually takes place during the formation of the equilibrium.

When an agent burns power for survival, it either faces inevitable elimination and thus has to completely cede power and give away any claims to rent, or it adjusts to block a stable improving subcoalition.[19] In general, for a coalition N with an initial structure γ and equilibrium structure γ^*, recall that N_+ is the coalition of agents with strictly positive power, we have the following lemma.

Lemma 1. *When power is partially burnt during power struggles, some agent in the equilibrium structure receives the same rents in the current structure as in a stable improving alternative. That is, if* $0 < \gamma_i^* < \gamma_i$, *for some* $i \in N$, *there exists* $C \subsetneqq N_+$ *with* $\gamma_C^{'}$, *where* $\gamma_C^{'} \leq \gamma_C^*$ *and* $\gamma_C^{'} > \gamma_{N\backslash C}^*$, *such that for some* $j \in C$, $U_j\left(N, \gamma^*\right) = U_j\left(C, \phi\left(\gamma_C^{'}\right)\right)$

Proof. See Appendix. □

Lemma 1 captures a key economic force within the model. Agents facing the threat of elimination give away power until the deviating agents are indifferent between the rents from the current structure and those from the stable improving subcoalition. Denote the set of agent(s) whose deviation constraint binds as G, $G = \left\{x \in C | U_x\left(N, \gamma^*\right) = U_x\left(C, \phi\left(\gamma_C^{'}\right)\right)\right\}$. We call the agents in G "guardians" because they guard the stable structure from deviations. Denote the complement set of C in N as $A = N\backslash C$. We call agents in A "accommodators" because they give away power to accommodate the guardians. Applying this terminology to the 3-agent example $\gamma^* = (\gamma_w^* = 1, \gamma_m^* = 4, \gamma_e^* = 5)$, The stable improving coalition with power structure

[18]The concept resembles that of AES but carries two major differences: 1. AES does not allow deviations after a deviation because power levels are fixed; 2. The algorithm in AES is independent from the agenda, while in our algorithm, different agendas lead to different equilibrium structures.

[19]An example of the first case is the power structure $(1, 2, 6, 8)$: the agent with power 1 can avoid elimination only by burning power to 0. An example of the second case is the agent with power 2 in $(2, 4, 5)$, who burns power partially to block the stable improving coalition of $(4, 4)$, which creates the equilibrium structure $(1, 4, 5)$. For the block to succeed, some agents in the original stable improving coalition should be indifferent between the rents received in the deviation and the post-burning structure.

$\gamma_C^{'} = \left(\gamma_m^{'} = 4, \gamma_e^{'} = 4\right)$ is blocked by the accommodator, w, who partially burns $\gamma_w^* = 1 < \gamma_w = 2$, to ensure that guardian e enjoys the same rents in the equilibrium $\left(\frac{5}{1+4+5}\right)$ as in the deviation $\left(\frac{4}{4+4}\right)$.[20]

As with small-sized coalitions, we are interested in whether shocks that weaken the strong (the guardians) and shocks that empowers the weak (the accommodators) actually lead to increased equality in equilibrium. In both cases, we focus on equilibrium structures where partial power is burnt during the formation of equilibrium: that is, the set of guardians is non-empty.

Suppose a small negative shock hits one of the guardians. There are two possible outcomes. First, if there are multiple guardians and one of them is weakened, then the remaining guardians now enjoy even higher rents than in the stable improving alternative. Consequently, the deviation remains unattractive to them, and no power struggles take place. Second, suppose only one guardian is hit by a small, negative shock. The weakened guardian receives fewer rents, thus allowing for a *strictly* stable improvement. As a result, the accommodators respond by burning their own power to appease the guardian. Formally, we have the following theorem.

Theorem 5. *If $0 < \gamma_i^* < \gamma_i$ for some $i \in N$, and G is a singleton, then there exists a threshold, $\underline{\eta} > 0$, such that for any negative perturbation $\eta \in \mathbb{G}$, $0 < \eta < \underline{\eta}$, on $x \in G$, $U_x(\gamma^*) = U_x\left(\phi\left[\left(\gamma_x^* - \eta, \gamma_{-x}^*\right)\right]\right)$.*

Proof. See Appendix. □

Theorem 5 illustrates the persistence of the elite after they are weakened. The elite recovers from negative shocks in terms of the rents received because the "outside option" of the elite remains unchanged, and the accommodators bear the burden of the negative shock.[21] Such an idea, known as the iron law of oligarchy, is well documented in the studies of power structures. A common scenario for a weakened elite group is the death of its leader. Kendall-Taylor and Frantz (2016) studied all 79 dictatorships between 1946 and 2012 and concluded, "...a leader's death in office almost never leads to significant near-term liberalization. Likewise, only rarely does it spell the end of the regime or precipitate instability in the form of coups or protests. On the contrary, authoritarian regimes have proven to be remarkably resilient when a leader dies." In the meantime, the trend of organizations towards oligarchy continued to spread throughout the first decade of the twenty-first century. According to Magaloni and Kricheli (2010), one-party regimes have now become the most common type of authoritarian rule, constituting 57 percent of the authoritarian regimes from 1950 to 2006 and 33 percent of the total regimes worldwide.

[20] In Michels' (1915) words, the guardians exist because " (the) society cannot exist without a dominant or political class, and that the ruling class, while its elements are subject to frequent partial renewal, nevertheless constitutes the only factor of sufficiently durable efficacy in the history of human development."

[21] In Michels' (1915) words, "The government, or, ... the state, cannot be anything other than the organization of a minority. It is the aim of this minority to impose upon the rest of society a legal order that is the outcome of the exigencies of dominion and of the exploitation of the mass."

Meanwhile, if a positive shock hits the guardian, it always receives at least as many rents as before because the guardian always has the option to burn up the positive shock and return to the previous equilibrium structure. Therefore, an immediate corollary follows: regardless of the direction of the perturbations, the power of the strong always becomes weakly more centralized.

Corollary 4. *If $0 < \gamma_i^* < \gamma_i$ for some $i \in N$, and G is a singleton, then there exists $\underline{\eta} > 0$, such that $U_x(\gamma^*) \leq U_x\left(\phi\left[\left(\gamma_x^* + \eta, \gamma_{-x}^*\right)\right]\right)$, for any $\eta \in \mathbb{G}$, $\eta \in \left(-\underline{\eta}, \underline{\eta}\right)$ applied on $x \in G$.*

Next, we consider the other side of the same coin: the shock that increases the power of an accommodator. When an accommodator is empowered, the guardian is necessarily weakened, which again triggers a strictly improving subcoalition.

Theorem 6. *If $0 < \gamma_i^* < \gamma_i$ for some $i \in N$, then there exists $\overline{\eta} > 0$ such that $U_x(\gamma^*) = U_x\left(\phi\left[\left(\gamma_j^* + \eta, \gamma_{-j}^*\right)\right]\right)$, for any $\eta \in \mathbb{G}$, $0 < \eta < \overline{\eta}$ and $\gamma_j^* < \gamma_j$, $j \in A$.*

Proof. See Appendix. □

Theorem 6 suggests that initiatives that aim to empower the weak may not increase social equality because the stable improving coalition always needs to be accommodated. In the extreme it can be totally ineffective, as in the 3-agent example. This theorem provides a unified explanation for the consolidation of an autocratic rule when facing challenges in favor of the non-elites. Such challenges may be trade-related, as in Guatemala and Mexico in the 19th century; the abolition of slavery as in West Africa (Robinson 2012); or the abolition of Mita in Peru (Dell 2008). Therefore, one should not take the rise of the non-elites for granted when a favorable junction arrives. In fact, as long as the elite circle remains intact, empowerment of the weak seldom emerges.[22]

5 Applications: what softens elite persistence?

Though the persistence of elites seems inevitable, there are several intuitive forces that may soften the tendency. In this section, we explore three of them. First, we introduce the possibility of production and discuss the boundary of power struggles when the to-be-expelled happens to be productive. Next, we consider the impact of a foreign threat on domestic power structures. Last, we explore the possibility of inviting a foreign arbitrator to settle in domestic disputes. Throughout the section, we highlight that initiating power

[22]Sometimes social empowerment does induce dramatic changes in the political landscape. Such empowerment usually involves a significant increase in power or a simultaneous change in both power *and* agenda. In the Arab Spring, the Tunisian government under Ben Ali focused on foreign direct investment, open trade, and open media, which not only empowered the masses but also changed domestic political agendas. The success of the revolution reveals the importance of coordinating social empowerment with agenda-setting abilities.

struggles is a *choice* for some particular agent in the coalition.[23] And the to-be-expelled agent has to make itself indispensable in the power structure, in order not to be purged. Some of the insights in the following scenarios may be delivered using separate models. However, we intend to answer the questions in far-sighted power struggles to provide a unified perspective.

5.1 Economic development

Power struggles aim to change how the pie is split, whereas production aims to enlarge the pie. A natural tradeoff exists if the choices are mutually exclusive. From a social welfare perspective, production is always preferred to power struggles. Consider an environment where the ruler and ordinary citizens/producers share the pie according to to their relative power, as in the model. The ruler, as the most powerful player in the population, does not produce, while the citizens produce to enlarge to the pie and gain more power during the process. Without power struggle concerns, the resource distribution shifts in favor of the producers with the passage of time, and the economy grows. However, when power struggles are possible, a ruler prioritizes economic development only after he is free from survival concerns (Gehlbach and Keefer 2011).[24] Meanwhile, because economic growth fosters stronger citizens, the ruler will not advocate for too much growth and allow the citizens to become too powerful. Consequently, a ruler follows, dynamically, a "sow and reap" strategy: it should aim for economic development for a period, then initiate power struggles before the potential opponents become too powerful, and then resume economic development after power consolidation. This strategy could explain the variations, and sometimes the cyclicality of development policies seen in some regimes across time. Furthermore, it sheds lights on the logic behind periodic campaigns initiated by the ruling elites aiming to restore social order. The analysis explains when a regime turns from "extractive" to "inclusive", in Acemoglu and Robinson's (2013) terminology, which may not come from a fundamental shift in institutions, but from the careful calculations of the ruler. In the online Appendix, we provide a simple extension of the 3-agent model to illustrate that indeed, production may soften elite persistence, but only for a period of time.

5.2 Foreign threats

When there is a foreign threat, domestic power should be strong enough collectively to avoid elimination. Since power burning distorts agents' strength, a significant foreign threat softens domestic power struggles.

[23] In reality, initiation of power struggles can indeed be a policy choice. The national development of China was guided by the principle of "taking class struggles as central tasks" starting in 1949, especially during the Cultural Revolution, before the ruling party decided to "concentrate on modernization" in 1977. Vietnam, by contrast, turned to "renovation" in 1986 with the goal of creating a "socialist-oriented market economy" after 15 years of leftist policies.

[24] For more recent events, North Korea announced its plan to boost economic development after Kim Jong-Un consolidated power.

However, the fragile domestic peace may easily break down after the foreign threat disappears, which provides extra incentives for the potential losing players in domestic struggles to prolong the conflict with the foreign threat. In the online Appendix, an extension based on the 3-agent coalition is included to formally address this question. Again, foreign threats may soften elite persistence, at the cost of prolonged conflicts. And the softening effect disappears once the threats are resolved.

We point to the career of General Georgy Zhukov to illustrate our point. Zhukov was one of the most important military commanders of the Soviet Union in World War II. He led and won multiple battles, including the Battle of Berlin, which put an end to the War in Europe. During the war, Zhukov held several prestigious positions in the army and in the political system because of his outstanding military talents. He served as Chief of the General Staff and then Deputy Commander-in-Chief. However, after the War, Zhukov was viewed as a potential threat to Stalin's leadership because of his popularity in the army. Stalin replaced Zhukov with Vasily Sokolovsky as Commander-in-Chief of the Soviet Ground Forces, and Zhukov was assigned to command the Odessa Military District, away from Moscow and lacking in strategic significance and troops. Zhukov's military career illustrates how power struggles can be softened when the threat is imminent and resume after the threat is exterminated. The Soviet Union is not a lone example. In China, during World War II, the Kuomintang and the Communist Party of China, who were previously in a civil war, united forces against Japan. The two parties resumed their civil war after defeating the Japanese. Likewise, in the 7th century, the Spanish kingdoms ceased their conflict to fight against the Moors.

5.3 Foreign arbitration

During domestic power struggles, a foreign arbitrator may be summoned to the aid of certain domestic player(s), which then leads to the restructuring of the domestic power structure and the redistribution of rents. A powerful foreign arbitrator may ease domestic conflicts, but it should not be so dominant as to dictate. Consequently, local elites may cede to a superior foreign power when their domestic rule is threatened. Surprisingly, the prestige of local elites may still persist among local communities after inviting a foreign power to join in domestic affairs. This is because by inviting a strong foreign arbitrator to settle the dispute, local elites can make themselves indispensable in the power balance, thus freeing themselves from struggle concerns. Therefore, a seemingly equality-promoting request for foreign arbitration may end up consolidating prestige of local elites. We include an extension based on 3-agent coalitions in the online appendix and refer to two cases in international relations below to illustrate our point.

The evidence from the 19th century Caribbean islands and modern-day Sri Lanka is highly consistent with the analysis. In the 19th century, there were fourteen British sugar colonies in the Caribbean. The

sugar cane plantation owners were the dominant social class, and slaves were the major source of labor. In 1833, the British parliament passed *An Act for the Abolition of Slavery*, which emancipated the slaves. Consequently, the poor were significantly empowered and began to play a larger role in local governance. According to Carvalho and Dippel (2016), "Starting with Montserrat in 1861 and ending with Grenada in 1876, all but one of the powerful assemblies voluntarily dissolved themselves, and invited the Crown to write a new constitution for them with a legislature appointed by the governor". Ashdown (1979) argues that "the colonies gave up their elected assemblies voluntarily, for in most cases the white, privileged classes preferred direct imperial government to the government of the colored classes who were slowly obtaining greater representation in the legislative councils." As a result, the elites in the Caribbean Islands invited the British to slow domestic regime change and managed to maintain their relative prestige.

Relatedly, if an ongoing domestic conflict cannot be resolved by domestic powers alone, a local force may invite external allies to eliminate the threats for them. The civil war in Sri Lanka lasted for more than 25 years and was ended with the external help from China, who "provided lethal military equipment and post-war diplomatic support"[25]). In exchange for their aid, Sri Lanka signed a $1.1 billion deal with China for the control and development of the southern deep-sea port of Hambantota, which has potential military uses.[26] For more examples of the power balancing role of external intervention, see Levine et al. (2016).

6 Discussions

In this section, we discuss the critical assumptions made in the paper and whether relaxing them changes the model's insights. To begin, we assume elimination is costless, which is an approximation of the case where the rents received by staying in the coalition significantly outweigh the potential costs of elimination. Adding a positive elimination cost does not change the resurgence of hierarchies nor the persistence of elites given perturbations. However, drastic regime changes from oligarchies to dictatorships take place less frequently when there are positive elimination costs. An example of such a scenario is included in the online appendix.

Another assumption we make in the model is that, although commitment is not possible during power relinquishment and power struggles, we assume that the outcome of power burning is implemented when a proposal reaches unanimous agreement from the proposed subcoalition. Alternatively, if agents can only make individual burning choices, and a burning scheme requires simultaneous power adjustment of multiple players, Theorem 2 fails to hold, and the set of stable structures enlarges. The intuition is that a stable improving coalition may not be implemented if other undesirable deviations are triggered during the transition path.

[25]See: https://www.brookings.edu/blog/order-from-chaos/2015/05/01/india-v-china-in-sri-lanka-lessons-for-rising-powers/
[26]See: http://www.bbc.com/news/world-asia-40761732

Mapping into the reality, regime changes can be blocked by deviations on the transition path, including those with a desirable destination.

A key difference between our model and AES is that we allow for a fully flexible power relinquishment while AES assumes fixed power levels.[27] The real world lies somewhere between the two. One way to reconcile the two models is to introduce a lower bound of power burning, making power adjustable only above a lower bound. The lower bound may be determined by bloodlines, religious beliefs, ethnicity (Montalvo and Reynal-Querol 2005), institutional legacies (Evans 2012), etc. For example, if one has a royal name or is of royal blood, it is often not sufficient to cede power and claim loyalty to the current ruler. The leader will still sometimes choose to eliminate the bloodline to tie up loose ends. This practice endures from the sibling killing tradition in the Ottoman Empire, to the assassination of Kim Jung-Nam in February 2017, and to the recent arrests of Prince Alwaleed bin Talal and Prince Mutaib bin Abdullah of Saudi Arabia.

7 Concluding Remarks

The goal of this paper is to study the persistence of elites in an institution-free environment. It does so by constructing a model of power struggles, where players use their power to eliminate others and split resources in a coalition. The modeling innovation is that it allows players to relinquish power when they are about to be expelled, and the paper shows that such flexibility can be a commitment device in power struggles, which then leads to the centralization of power. A power structure is in equilibrium when the opponents of the current structure fail to adjust to a stable, rent-improving alternative. We further show that power often ends up more concentrated in a few elite members when equilibrium structures are perturbed. We use the model to explain why, in many cases, revolutions that aim to install social equality quickly reproduce the same type of hierarchy that the revolutionaries sought to destroy. We highlight the emergence of familiar institutional forms out of an institution-free environment, and the important role that power struggles play in institution building. The elite persistence can be softened when the outsiders to be expelled happen to be productive, under which case the elite adopt a "reap and sow" strategy to alternate centralization and decentralization. The persistence can also be softened when the grand coalition faces external threats, by making some to-be-expelled players valuable in fighting the threats - although prolonged conflicts may be the consequences of players avoiding future struggles. Counter-intuitively, a foreign arbitrator may consolidate domestic elite persistence, by making the local elites indispensable in the power balance.

[27] It is true that sometimes power struggles involve executing your enemies such as in the Great Purge in the Soviet Union. However, in many other cases, the struggles end with the relevant parties ceding power but keeping their lives. Even in the Soviet Union, the power transitions since Nikita Khrushchev, despite the struggles, have been largely without bloodshed. The same applies to the transition after the Cultural Revolution in China, the transition from Chiang Ching-kuo to Lee Teng-hui in Taiwan, the transition from Thein Sein to Aung San Suu Kyi in Myanmar, and the power consolidation of Bashar al Alssad in Syria.

As final remarks, we discuss several other directions we would like to explore in the future. First, the current analysis focuses on non-targeted transfers (burning).[28] An immediate generalization would be to allow targeted transfers, where a member can transfer power to another individual. Second, the coalition in the current framework is assumed to be fully connected. An interesting practice is to explore community-based power struggles, where local power struggles can be a means to access or defy a larger coalition (Dziubinski, Goyal, and Minarsch 2017). Third, the agents in the current model are purely rent maximizing. One further line of research would be to add commitment types or idealists that aim to prevent conflicts. Related questions include the number of commitment types that are needed to ensure a peaceful environment and whether long-lasting peace is possible if people believe that some commitment types exist in the coalition. Lastly, one could extend power to multiple dimensions, such as allowing an agent to hold some economic power, as well as some political power. Economic power can be invested into building political power and vice versa. The dynamics of the interaction may shed light on the transition from one institution to another.

[28] Burning $\Delta = \gamma_i - \rho_i(\gamma_i)$ pieces of power for agent i is equivalent to ceding power $\frac{\Delta}{n-1}$ uniformly to every other agent in the coalition. Burning power is a special case of non-targeted power transfer. We adopt power burning as the only form of power transfer to emphasize the difficulty of such process. Unlike physical assets, power is often hard to transfer from one person to another, such as an army switching allegiance to another commander. Other examples of burning power include resigning or exiting from office (Steve Bannon), migrating to a foreign country (Leon Trotsky), abandoning access to weapons (chemical, nuclear and other weapons of mass destruction), etc.

A Appendix

A.1 Omitted proofs

Proof of Theorem 1. **Existence.** First, observe that the game has a finite horizon. The burning meetings last for a finite number of rounds because power is discrete, so there are at most $\frac{\prod_{i=1}^{n}\gamma_i}{\epsilon^n}$ rounds of burning meetings. The elimination meetings last for a finite number of rounds because there are finite proposals and finite voters, and if no proposal is accepted, the status quo is enforced. By Zermelo's Theorem, a finite horizon game of perfect information admits an SPE.

Uniqueness of equilibrium structure. Given agenda s, suppose there exists another equilibrium $\left(N^{'},\gamma^{'}\right) \neq (N^*,\gamma^*)$. There are two cases. First, $N^* \neq N^{'}$. Without loss of generality assume that $N^* \subsetneqq N^{'}$. Then, some agent i exists such that $i \in N^{'}$ but $i \notin N^*$. Therefore, in the first equilibrium $U_i = -n$, which cannot happen in equilibrium because agent i can simply burn power down to zero in the previous burning meeting such that $U_i = 0$, which would be an improvement. Therefore, (N^*,γ^*) cannot be in equilibrium. By the same logic, we know that $N^* = N$ because otherwise any agent in $N\backslash N^*$ has a better response by burning down to zero.

Consequently, elimination does not take place on the equilibrium path. Now, suppose $N^* = N^{'}$ but $\gamma^* \neq \gamma^{'}$. Denote the first time (round) that an agent burns power differently on the equilibrium path as t, the member as j, the strategy as $s^*_{j|h_t}$, $s^{'}_{j|h_t}$. Therefore, the member is indifferent between two burning choices, $U_j\left(s^*_{j|h_t}\right) = U_j\left(s^{'}_{j|h_t}\right)$, which implies that $\frac{\gamma^*_j}{\sum\gamma^*_i} = \frac{\gamma^{'}_j}{\sum\gamma^{'}_i}$. Furthermore, the tie-breaking rule suggests that $\gamma^*_j = \gamma^{'}_j$. Therefore, j cannot burn power differently, nor can any other member, $\gamma^* = \gamma^{'}$, which is a contradiction.

Thus far, we have shown that an equilibrium structure exists and that it is unique for a given N,γ,s. There might exist multiple SPEs because in the voting stage, if a previous member included in the proposal votes no, and the next member is indifferent between yes and no, the proposal is rejected either way. However, the constitution of the equilibrium structure remains the same. □

Proof of Proposition 1. When the elite agent is very powerful, it naturally dictates. When the elite agent is not sufficiently powerful, the key is whether the weak and the medium agent could form a powerful alliance against him.

When $a \leq \frac{c}{2}$, an alliance between the weak and the medium is not sufficiently powerful. Therefore, the unique, stable rent-improving coalition remains $\left(\gamma^{'}_m = b, \gamma^{'}_e = b\right)$. The resulting equilibrium is $(c-b,b,c)$.

Now we consider the cases under $a > \frac{c}{2}$. When $s = (m,w,e)$ or (w,m,e), the weak and the medium can form a subcoalition $\left(\gamma^{'}_w = a, \gamma^{'}_m = a\right)$ before the elite moves. Facing (a,a,c), the elite, as the last mover, has

no choice but to burn down to zero because gambling for probabilistic inclusion is never optimal. Therefore, the equilibrium structure is $(a, a, 0)$. In any other agenda where the elite is not the last mover, the elite can always mimic the medium agent's burning strategy and secure rent of $\frac{1}{2}$. As a result, the weak agent should always accommodate the elite by burning down to make up the power difference between the elite and the medium agent. Therefore, the equilibrium structure is $(c - b, b, c)$. □

Proof of Theorem 2. We start to follow an induction approach to characterize the equilibrium structure for coalitions with size $|N| \leq 9$. First, the theorem holds for any 3-agent coalition, as suggested in Proposition 1. It is easy to check that all three forms are indeed equilibrium structures. Denote the set of structures that fall into either one of the three forms as Γ.

For coalitions of all sizes, when there exists some i such that $\gamma_i > \sum_{j \neq i} \gamma_j$, then the equilibrium structure naturally converges to a dictatorship. In the following discussion, we thus focus on the cases where $\gamma_i \leq \sum_{j \neq i} \gamma_j$. Without loss of generality, we reorder the agents from weakest to strongest, as in the original power structure.

When $|N| = 4$, suppose there is an equilibrium structure, $\gamma^{'} \notin \Gamma$. The most powerful agent, which we denote γ_4, without loss of generality enjoys rents strictly less than $\frac{1}{2}$, because otherwise the structure will become a dictatorship. Consequently, the two most powerful players, which we denote γ_3 and γ_4, have incentives to form a rent-improving alliance $\left(\gamma_3^{'} = \gamma_3, \gamma_4^{'} = \gamma_3\right)$. Therefore, the original structure cannot be in equilibrium.

When $|N| = 5$, similarly, suppose there exists an equilibrium structure $\gamma^{'} \notin \Gamma$; then the most powerful agent has rents $U_5 < \frac{1}{2}$. If $2\gamma_4 > \gamma_1 + \gamma_2 + \gamma_3$, then γ_5 can threaten to burn power to match γ_4, resulting in oligarchy. If $2\gamma_4 \leq \gamma_1 + \gamma_2 + \gamma_3$, it must be true that $2\gamma_5 \geq \gamma_1 + \gamma_2$. When $2\gamma_5 > \gamma_1 + \gamma_2$, there exists an oligarchy with $\left(\gamma_5^{'} = \gamma_5, \gamma_4^{'}, \gamma_3^{'}\right)$ where $\gamma_4^{'} + \gamma_3^{'} = \gamma_5$, which is strictly rent-improving, therefore the original structure is not in equilibrium. When $2\gamma_5 = \gamma_1 + \gamma_2$, the original structure is equal-powered; thus, it is easy to show that a polyarchy of 4 agents is a profitable deviation.

When $|N| = 6$, suppose $U_6 < \frac{1}{4}$; then a profitable deviation is again a 4-agent polyarchy. Therefore, $U_6 \geq \frac{1}{4}$. Furthermore, $U_5 < \frac{1}{4}$, otherwise an oligarchy is formed. Consequently, $U_4 < \frac{1}{4}$, and $U_1 + U_2 + U_3 \leq \frac{1}{2}$. It cannot be the case that $U_1 + U_2 + U_3 = \frac{1}{2}$, in which case it is an equal-powered coalition, which contradicts $U_6 \geq \frac{1}{4}$. Therefore, $U_1 + U_2 + U_3 < \frac{1}{2}$, and $2U_6 \geq \frac{1}{2}$. This is a profitable deviation to a 3-agent oligarchy centered around agent 6.

When $|N| = 7$, similarly, we know that $U_7 \geq \frac{1}{4}$. Furthermore, $U_6 < \frac{1}{4}$, $U_5 < \frac{1}{4}$. Suppose $U_1 + U_2 + U_3 + U_4 \geq \frac{1}{2}$, then $U_5 < \frac{1}{6}$. Consequently, $U_4 < \frac{1}{6}$, then it is profitable to deviate to a 4-agent oligarchy centered around agent 7. Suppose $U_1 + U_2 + U_3 + U_4 < \frac{1}{2}$, then it is immediately clear that a 3-agent oligarchy

centered around agent 7 is profitable.

When $|N| = 8$, the theorem no longer holds because a relatively equal structure now enables a powerful non-elite group to have checks and balances over potential deviations. As a counter-example, consider the following: coalition structure: $(5.2, 5.1, 5, 5, 5, 5, 5, 5)$. □

Proof of Theorem 3. For dictatorships, pick the dictator as agent i.

For oligarchies, pick the agent that enjoys half of the rent in the original equilibrium as agent i, and the rest follows from the analysis for 3-agent coalitions.

For polyarchies of size 2^k, all agents are equally powerful in the original equilibrium. For $k = 1$, any perturbation collapses the regime into a dictatorship. For $k = 2$, positive perturbation turns the regime into an oligarchy centered around the empowered agent (which we pick as agent i); negative perturbation turns the regime into a polyarchy of size 2, with the two earliest-moving non-weakened agents. In this case, we pick either one of them as agent i.

For $k \geq 3$, if the perturbation is positive, pick the empowered agent as agent i, who now becomes the uniquely most powerful agent. If the perturbation is negative, pick the earliest-moving non-weakened agent as agent i because the post-perturbation regime is a polyarchy of size 2^{k-1} in which i is included. □

Proof of Theorem 4. For a general coalition N,

"If" condition.

Suppose there does not exist a stable improving subcoalition, that is, for any winning subcoalition $C \subset N$ such that $\gamma_C^{'} > \gamma_{N \setminus C}$, where $\gamma_i^{'} \leq \gamma_i$, and $\gamma_C^* = \phi\left(\gamma_C^{'}\right)$, there exists some agent $i \in C$ that $U_i\left(C, \gamma^*\right) \leq U_i\left(N, \gamma\right)$. If $U_i\left(C, \gamma^*\right) < U_i\left(N, \gamma\right)$, then, agent i rejects the proposal when C is proposed in the elimination meeting. If $U_i\left(C, \gamma^*\right) = U_i\left(N, \gamma\right)$, should an equilibrium with inner circle C be formed, the agents have to participate in at least 3 rounds of burning meetings, which are dominated by maintaining the status quo with the tie-breaking assumptions. Therefore, agent i rejects the proposal in the elimination meeting as well. Anticipating the rejection, the other members in C have no incentives to adjust power in the burning meeting; thus, the status quo coalition remains unchanged, and the original structure is in equilibrium.

"Only if" condition.

Suppose a coalition N with γ is in equilibrium, and there is a subcoalition C with adjusted structure $\gamma_C^{'}$ that is stable improving. Pick any agent in C who makes burning proposals in the burning meeting: if he or she proposes $\gamma_C^{'}$, it is a better response for all agents involved in C to agree on the proposal and secure a rent-improving payoff. He or she is able to do so because $\gamma_{i \in C}^{'} > \max\left\{\gamma_j \mid j \in N_{\succ_s i} \setminus C\right\}$ and therefore,

no prior-moving agent can mimic the move. Thus, the original structure is not an equilibrium, which is a contradiction. □

Proof of Lemma 1. Because some power is burnt, then for any adjustable subcoalition $C \subset N$ such that $\gamma_C^{'} > \gamma_{N\setminus C}^{*}$ and $\gamma_C^{'} \leq \gamma_C^{*}$, there exists some agent $j \in C$ that $U_j\left(N, \gamma^{*}\right) \geq U_j\left(C, \phi\left(\gamma_C^{'}\right)\right)$.

If $U_j\left(N, \gamma^{*}\right) = U_j\left(C, \phi\left(\gamma_C^{'}\right)\right)$ for some j, the statement is proved. If $U_j\left(N, \gamma^{*}\right) > U_j\left(C, \phi\left(\gamma_C^{'}\right)\right)$ for all j and all such subcoalitions C, pick agent i as the last agent who burns partial power, $0 < \gamma_i^{*} < \gamma_i$, and slightly increase his power in his last burning choice. Then, the inequalities for all j's still hold (because i is the last agent burning partial power. A complete burner does not change the result; otherwise, he or she does not burn completely) and agent i enjoys more rent. Thus, this is a better response for i. Agent i can continue to do so until for some $j^{'}$,$U_{j^{'}}\left(N, \gamma^{*}\right) = U_{j^{'}}\left(C, \phi\left(\gamma_C^{'}\right)\right)$. □

Proof of Theorem 5. By Lemma 1, $U_x\left(N, \gamma^{*}\right) = U_x\left(C\phi\left(\gamma_C^{'}\right)\right)$. Notice that $\gamma_{x\in C}^{'} < \gamma_{x\in N}^{*}$, otherwise $\frac{\gamma_x^{*}}{\sum \gamma_{i\in N}^{*}} = \frac{\gamma_x^{'}}{\sum \gamma_{i\in C}^{'}}$; $C = N_{+}$, thus, is not a proper subcoalition. Consequently, $\gamma_C^{'}$ is adjustable after the small shocks $0 < \eta < \underline{\eta_0} = \gamma_{x\in N}^{*} - \gamma_{x\in C}^{'}$.

There are two cases to consider.

In the first case, $\left(C, \gamma_C^{'}\right)$ is the only deviation coalition whose deviation constraint binds. That is, for any winning subcoalition $\widetilde{C} \neq C$ such that $\sum \gamma_{i\in\widetilde{C}}^{'} > \sum \gamma_{i\in N\setminus\widetilde{C}}^{*}$, $\gamma_{\widetilde{C}}^{'} \leq \gamma_{\widetilde{C}}^{*}$, we have $U_j\left(N, \gamma^{*}\right) > U_j\left(\widetilde{C}, \phi\left(\gamma_{\widetilde{C}}^{'}\right)\right)$ for some $j \in \widetilde{C}$. Then, for a small enough grid ν, after a small negative shock hits x, we still have $U_j\left(N, \gamma^{*,\eta}\right) > U_j\left(\widetilde{C}, \phi\left(\gamma_{\widetilde{C}}^{'}\right)\right)$, where $\gamma^{*,\eta} = \left(\gamma_x^{*} - \eta, \gamma_{-x}^{*}\right)$. This means that $\left(C, \gamma_C^{'\eta}\right)$ is the only stable improving subcoalition. Thus, agents in $N\setminus C$ have to burn additional power to accommodate x, until $U_x\left(N, \gamma^{*}\right) = U_x\left(C, \phi\left(\gamma_C^{'}\right)\right)$. In particular, denote the set of $\widetilde{C}$ as $\widetilde{\mathcal{C}}$, denote $\eta_{\widetilde{C}}$ as the largest shock that makes $U_{j\in N}\left(\gamma^{*,\eta}\right) \geq U_{j\in\widetilde{C}}\left(\phi\left(\gamma_{\widetilde{C}}^{'}\right)\right)$ for some $j \in \widetilde{C}$, then we can calculate $\underline{\eta_1} = \min\left\{\eta_{\widetilde{C}}\right\}_{\widetilde{C}\in\widetilde{\mathcal{C}}}$.

The second case is if there are other deviation coalitions whose deviation constraint binds. There are two scenarios: 1. the binding agent is not x. Then, after the shock, the deviation constraint is slack. We return to the first case; 2. the binding agent is x, then we calculate the upper bound of the shock size by $\underline{\eta_2} = \min\left\{\gamma_{x\in N}^{*} - \gamma_{x\in\widehat{C}}^{'}\right\}_{\widehat{C}\in\widehat{\mathcal{C}}}$, where $\widehat{\mathcal{C}}$ is the set of deviation coalitions with binding agent x.

Lastly, we denote $\underline{\eta} = \min\left\{\underline{\eta_0}, \underline{\eta_1}, \underline{\eta_2}\right\}$. □

Proof of Theorem 6. We once again try to show that $\left(C, \gamma_C^{'\eta}\right)$ remains the unique stable improving subcoalition after the shock. The proof is similar to that of Theorem 5. A small positive shock from the agents in A makes $\left(C, \gamma_C^{'}\right)$ feasible but keeps the incentive constraints of other winning coalitions slack. Denote the threshold derived from such a process as $\overline{\eta_1}$. The next step is to check whether an empowered agent j brings additional stable improving coalitions to consider. Pick $\overline{\eta_2} < \gamma_j - \gamma_j^{*}$; suppose there is such a new

stable improving coalition $\widehat{C}$, where $j \in \widehat{C}$, $\widehat{\gamma}_j < \gamma_j$, then we know that for agent j, γ_j^* is a better response than $\widehat{\gamma}_j$. Therefore, $\left(C, \gamma_C^{'\eta}\right)$ remains the unique stable improving coalition to consider. Lastly we denote $\overline{\eta} = \min\{\overline{\eta_1}, \overline{\eta_2}\}$. □

A.2 Remarks on agenda

When agents have fixed power levels, the outcomes of power struggles are independent of the agenda (see AES). However, as Proposition 1 suggests, when burning power is possible, the formation of the equilibrium does depend on the agenda:[29] the sequence of moves determines which alternate alliance can be formed first. As Olson (2000, p. xii) writes, the opportunity to consolidate authority "...occurs at times that Polish reformer Leszek Balcerowicz refers to as *moments of extraordinary politics*." Furthermore, the same initiated power structure could lead to drastically different equilibrium structures. For instance, starting from the same initial power structure (a, b, c) when $a > \frac{c}{2}$, two alliances are stable and rent-improving: $\left(\gamma_w^{'}, \gamma_m^{'}\right) = (a, a)$ and $\left(\gamma_m^{'}, \gamma_e^{'}\right) = (b, b)$. When the elite moves last, i.e., $s = (m, w, e)$ or (w, m, e), the weak and the medium agent preempt to form the alliance of (a, a), which forces the elite agent to respond by ceding power. The equilibrium structure results in $(a, a, 0)$. However, in all other agendas, for example, $s = (w, e, m)$, after the medium agent burns power to a, the elite can react by mimicking the first mover and also burn down to a. This leads to a new structure, (a, a, a). The tie-breaking assumption suggests that the last mover should yield, thus blocking the alliance of $\left(\gamma_w^{'}, \gamma_m^{'}\right) = (a, a)$ in the first place. This illustrates the survival concerns of the elite. When the initial power structure is highly skewed, the elite is always free of survival concerns. However, when the initial power structure is relatively equal, the elite could continue to expropriate the weak and strengthen the rule with advantageous timing but may lose power otherwise. We show that the probability of successful expropriations has important implications in how the elite determines the economic development policies in Section 5.1.

The first-mover advantages in mobilizing allies during regime changes are consistent with reality. In the attempted coup d'etat against Recep Erdoğan of Turkey in July 2016, the failed efforts of the rebellion to block social media and the swift reaction of Erdoğan[30] jointly determined the failure of the coup. As a comparison, the 1969 Libyan coup succeeded because the coup soldiers cut off the connection between King

[29] This is consistent with the literature of agenda setting. McCombs and Shaw (1972), followed by Scheufele (1999), Scheufele and Tewksbury (2007), and recently, Helfer (2016) discuss the role of media in determining whether the relevant parties are aware of the potential changes of other members or the potential policies that are put into motion. In the context of Schatz (2009) shows that the authoritarian elites make effective use of agenda-setting tools to strengthen their rule.

[30] One hour after the coup started, rebellion soldiers occupied Taksim Square in central Istanbul and were inside the buildings of the state broadcaster, the Turkish Radio and Television Corporation (TRT), in Ankara, trying to control the media. They also tried to block social media such as Twitter, Facebook and YouTube. At the time, Erdoğan was on holiday outside Turkey. Two hours after the coup, Erdoğan performed a Facetime interview with CNN Turk, in which he called upon his supporters to fight against the coup soldiers. Erdoğan returned to Istanbul five hours after the coup and later wiped out the coup soldiers with the support of Turkish military.

Idris and his potential supporters, which made King Idris the last mover in the power struggle.

A.3 An extension of production and economic development

Power struggles aim to change how the pie is split, whereas production aims to enlarge the pie. A natural tradeoff exists if the choices are mutually exclusive. From a social welfare perspective, production is always preferred to power struggles. Intuitively, a ruler prioritizes economic development only after he is free from survival concerns (Gehlbach and Keefer 2011).[31] However, we show that power struggles may turn in the opposite direction: a ruler free from survival concerns may focus on expropriations, while a ruler with survival concerns may focus on economic development.

Players. Consider a society consisting of three representative agents: the elite agent (e) who rules; the magistrate (m) who helps to rule; and the worker (w) who produces. Denote the initial power structure $\gamma_0 = \left(\gamma_w^0 = k + h, \gamma_m^0 = 1 - k, \gamma_e^0 = 1\right)$, where $k > 0$, $h \geq 0$. By Proposition 1, the structure is in equilibrium when $h = 0$, thus h measures how decentralized power is from the equilibrium. We also assume that $k + h < 1 - k$, such that the magistrate holds more power than the worker.

Strategies and Timing. The game begins with the initial power structure γ_0 and agenda s. The elite chooses whether to initiate power struggles. If power struggles are initiated, the game is played as in Section 2, and agents receive rents according to to the equilibrium structure. If the ruler chooses not to initiate power struggles, then the worker produces. Production increases the worker's power by $\beta > 0$. Production plus rents are distributed according to $\gamma_0^{'} = (k + h + \beta, 1 - k, 1)$.

Payoffs. There is an existing pie of size π to share. Production enlarges the pie by $\alpha\gamma_w^0$, where $\alpha > 0$ is a measure of productivity.

An intuitive approach suggests a direct comparison between the increased rents from expropriation (power struggles) and the increased size of the pie due to production. However, as Proposition 1 shows, the elite may lose power in the struggles, which makes expropriation risky. We therefore distinguish elites by their likelihood of survival in power struggles and provide their optimal decisions in the following proposition:

Proposition 2. *Rulers with survival concerns prioritize economic growth over expropriation more often than those without survival concerns. That is:*

When $s \neq (m, w, e)$ and $s \neq (w, m, e)$, the ruler focuses on production if and only if

$$\frac{1}{2}\pi\beta + \left(\frac{1}{2}\pi - \alpha\right) \leq k < \frac{1 - h}{2}$$

[31] For more recent events, North Korea announced its plan to boost economic development, after Kim Jong-Un consolidated power.

When $s = (m, w, e)$ or $s = (w, m, e)$, the ruler focuses on production if and only if

$$\min\left\{\frac{1}{2} - h,\ \frac{1}{2}\pi\beta + \left(\frac{1}{2}\pi - \alpha\right)\right\} \leq k < \frac{1-h}{2}$$

Proof. Without concerns of losing power, the ruler conducts a direct comparison between enlarging the pie and getting a larger share of the existing pie: $\frac{1}{2+h+\beta}\left[\alpha\left(k+h\right)+\pi\right] > \frac{1}{2}\pi$, which gives the left-hand side of the first inequality, $k > \frac{1}{2}\pi\beta + \left(\frac{1}{2}\pi - \alpha\right)h$. The right-hand side of the first inequality is given by the assumption that $k + h < 1 - k$.

When the ruler has survival concerns, he decentralizes whenever the worker is strong enough to overthrow him in the agenda $s = (m, w, e)$ or $s = (w, m, e)$, therefore he decentralizes whenever $k + h > \frac{1}{2}$. When $k + h \leq \frac{1}{2}$,, the elite ruler has no survival concerns and behaves the same as in the first case. Combining the two, we obtain the expression on the left-hand side of the second inequality. □

It is intuitive that power struggles are more likely to be reduced by economic development when productivity (α) is higher, when the existing pie is smaller and when the empowerment from production is insignificant (β is low). As a result, high value-added agents are less likely to be expropriated, whereas a regime abundant in natural resources often suffers from power struggles.

A key implication of the proposition is that elite rulers facing survival concern, favor economic development in a larger set of initial power structures. They are forced to prioritize economic development when expropriation triggers a regime change. In the meantime, the proposition suggests that if the ruler allows the productive player(s)to accumulates its power, it will eventually be forced to decentralize and may lose power if the agenda-setting rights are shifted to the more powerful players in the game. An immediate takeaway is that the ruler should keep a close watch on the power accumulation of the productive. It is safe to continue decentralization as long as the rule is secured, but when reaching the threshold, the ruler should initiate purges to restore an oligarchic structure.

A.4 An extension of power struggles with a foreign threat

Players. Considers a domestic country with three warlords, w, m and e. The current power structure is $\gamma_0 = (\gamma_w = k + h, \gamma_m = 1 - k, \gamma_e = 1)$. The agenda is top-down. A foreign country with a power level g, $2 < g < 2 + h$, attacks the domestic country with probability $0 \leq \mu \leq 1$. The war is won by domestic forces if and only if they are collectively more powerful than the enemy $\gamma_p + \gamma_m + \gamma_e \geq g$.

Strategies and Timing. The game proceeds as follows: 1. Domestic agents conduct power struggles as in Section 2; 2. The foreign country attacks; 3. The resource is distributed to the winner(s).

Payoffs. The winners of the war share the resources according to to their power levels, and the loser gets zero.

It is easy to see that a foreign threat softens domestic power struggles because $g > 2 > 2 - 2k$, thus warlords m and e are no longer interested in seeking the power burning alliance $(1-k, 1-k)$ because such an alliance fails to defeat the invader. Consequently, the domestic agents set aside power struggles and join hands when foreign forces are likely to invade.

Proposition 3. *Power struggles are softened when the external threat is high. In particular, power struggles cease when $\mu > \underline{\mu}$, where $\underline{\mu} = \frac{h}{2+h}$.*

Proof. The cut-off is determined by $\left(1-\underline{\mu}\right)\frac{1}{2} = \frac{1}{2+h}$, which gives $\underline{\mu} = \frac{h}{2+h}$. □

There are two implications of this proposition. First, the cut-off is independent of the level of external threat g because if the invasion should take place, the outcome is always binary. Therefore, when the threat is credible ($2 < g < 2+h$), domestic conflict resolution depends on the probability of invasion (μ) instead of on the strength of the invader.[32] Second, the cut-off increases in h because when the invasion does not take place, a resourceful warlord further incentivizes more extraction.

A.5 An extension of power struggles with a foreign arbitrator

Players. Consider a similar domestic country with three political parties representing the workers, the middle class, and the elite, with an initial domestic structure $\gamma_0 = (\gamma_w = k+h, \gamma_m = 1-k, \gamma_e = 1)$. The agenda is (m, w, e). The elite has connections with a foreign ally (a), who holds power $1 < d < 2$. For simplicity, the foreign arbitrator is assumed to move after the domestic players, if it is invited to participate in domestic affairs. Thus, the revised agenda is (m, w, e, a).

Strategies and Timing. 1. The elite decides whether to invite a foreign arbitrator; 2. Power struggles take place as in Section 2, either in a 3-agent or a 4-agent coalition; 3. Payoffs are assigned.

Payoffs. Agents split rents of 1 according to their power levels after the power struggles.

By Proposition 1, when $k + h > \frac{1}{2}$, the worker and the middle class form an alliance to overthrow the elite in a 3-agent coalition. Anticipating the elimination, the elite has an incentive to invite the foreign arbitrator, which transforms the power structure to $(d+k-2, 1-k, 1, d)$. This is the equilibrium of the 4-agent coalition after power struggles. The elite receives $\frac{1}{2d} > 0$, which is preferred to ceding power completely, as they would in the 3-agent coalition.

In general, we have the following result for any coalition with domestic conflict, $\phi(\gamma) \neq \gamma$.

[32] We leave an alternative model, with probabilistic victory, for future research.

Proposition 4. *Any domestic conflict can be stopped by a foreign power. That is, for any coalition N with domestic conflicts such that $\phi(\gamma|s) \neq \gamma$, there exists an external agent with a power level γ_f such that $\phi((\gamma, \gamma_f)|s) = (\gamma, \gamma_f)$.*

Proof. Take $\gamma_f = \sum \gamma_{i \in N}$, by Theorem 2, $\phi((\gamma, \gamma_f)) = (\gamma, \gamma_f)$ □

Notice that despite the disappearance of conflict, it might not always be in the interest for the elite to do so.

A.6 An example of costly elimination

Recall a 3-agent power structure, $(2, 4, 5)$. Suppose every elimination costs $\kappa > 0$ to each agent involved, the equilibrium structure of a 3-agent coalition $(2, 4, 5)$ becomes $\left(1 + \frac{20\kappa}{1-2\kappa}, 4, 5\right)$ because the stable improving alternative of $(4, 4)$ only guarantees rent of $\frac{1}{2} - \kappa$ instead of $\frac{1}{2}$. As a result, if the elite agent is hit by a small positive perturbation, $\left(1 + \frac{20\kappa}{1-2\kappa}, 4, 5 + \epsilon\right)$, the perturbed structure remains in equilibrium, since $1 + \frac{20\kappa}{1-2\kappa} + 4 > 5 + \epsilon$. That is, positive elimination costs offer equilibrium structures some slackness in how they can perturbed. Of course, the costs of elimination may depend on the power difference of the involved agents and may be asymmetric even for agents with the same power, for instance, geographical reasons. Caselli, Morelli, and Rohner (2015) showed that conflicts are more likely when the resources of interest are located near the border of two countries because it is less costly for the attacking side to occupy the resources. Meanwhile, such slackness breaks down when perturbations accumulate to a tipping point and an imbalanced power is formed, e.g., $5 + \sum_t \epsilon_t > 1 + \frac{20\kappa}{1-2\kappa} + 4$, and we thus expect to see a sudden change in regime - in this very example, the sudden fall to dictatorship after several periods of oligarchies.

References

Acemoglu, Daron, and James A. Robinson. 2005. *Economic Origins of Dictatorship and Democracy.* Cambridge: Cambridge University Press.

Acemoglu, Daron, and James A Robinson. 2006. "De Facto Political Power and Institutional Persistence." *The American Economic Review*, 96(2): 325–30.

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

Acemoglu, Daron, and James A Robinson. 2013. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty.* New York: Crown Business.

Acemoglu, Daron, and James A Robinson. 2017. "The Emergence of Weak, Despotic and Inclusive States." `http://dx.doi.org/10.3386/w23657`.

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin. 2006. "Coalition Formation in Political Games." Cambridge: National Bureau of Economic Research.

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin. 2008. "Coalition Formation in Non-democracies." *The Review of Economic Studies*, 75(4): 987–1009.

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin. 2012. "Dynamics and Stability of Constitutions, Coalitions, and Clubs." *The American Economic Review*, 102(4): 1446–76.

Aghion, Philippe, Alberto Alesina, and Francesco Trebbi. 2004. "Endogenous Political Institutions." *The Quarterly Journal of Economics*, 119(2): 565–611.

Aghion, Philippe, and Patrick Bolton. 2003. "Incomplete Social Contracts." *Journal of the European Economic Association*, 1(1): 38–67.

Ashdown, Peter. 1979. *Caribbean History in Maps.* [Port of Spain], Trinidad: Longman Caribbean.

Bernheim, B. Douglas, Bezalel Peleg, and Michael D. Whinston. 1987. "Coalition-Proof Nash Equilibria I. Concepts." *Journal of Economic Theory*, 42(1): 1–12.

Caselli, Francesco, Massimo Morelli, and Dominic Rohner. 2015. "The Geography of Interstate Resource Wars." *The Quarterly Journal of Economics*, 130(1): 267–315.

Chwe, Michael Suk-Young. 1994. "Farsighted Coalitional Stability." *Journal of Economic Theory*, 63(2): 299–325.

Dal Bó, Ernesto, Pedro Dal Bó, and Jason Snyder. 2009. "Political Dynasties." *The Review of Economic Studies*, 76(1): 115–42.

Dalton, George. 1965. "History, Politics, and Economic Development in Liberia." *The Journal of Economic History*, 25(4): 569–91.

Dell, Melissa. 2008. "The Mining Mita: Explaining institutional persistence." Manuscript, Department of Economics, MIT.

De Mesquita, Bruce Bueno. 2005. *The Logic of Political Survival.* Cambridge: MIT press.

De Tocqueville, Alexis. 2008. *Ancien Regime and the French Revolution.* London, UK: Penguin.

Doyle, William. 1989. *The Oxford History of the French Revolution.* Oxford: Oxford Paperbacks.

Dziubinski, Marcin, Sanjeev Goyal, and David EN Minarsch. 2017. "The Strategy of Conquest." Department of Applied Economics.

Evans, Peter B. 2012. *Embedded Autonomy: States and Industrial Transformation.* Princeton: Princeton University Press.

Fearon, James D. 2011. "Self-Enforcing Democracy." *The Quarterly Journal of Economics*, 126(4): 1661–708.

Gehlbach, Scott, and Philip Keefer. 2011. "Investment Without Democracy: Ruling-Party Institutionalization and Credible Commitment in Autocracies." *Journal of Comparative Economics*, 39(2): 123–39.

Gehlbach, Scott, Konstantin Sonin, and Milan W Svolik. 2016. "Formal Models of Nondemocratic Politics." *Annual Review of Political Science*, 19: 565–84.

Gregory, Paul R. 2013. *Politics, Murder, and Love in Stalin's Kremlin: The Story of Nikolai Bukharin and Anna Larina.* Stanford: Hoover Press.

Guimaraes, Bernardo, and Kevin D. Sheedy. 2017. "Guarding the Guardians." *The Economic Journal.* 127(606): 2441–77.

Havel, Vaclav, and Paul Wilson. 1985. "The Power of the Powerless." *International Journal of Politics*, 15(3/4): 23–96.

Helfer, Luzia. 2016. "Media Effects on Politicians: An Individual-Level Political Agenda-Setting Experiment." *The International Journal of Press/Politics*, 21(2): 233–52.

Hobbes, Thomas. 2006. *Leviathan.* New York: A&C Black.

Jandoc, Karl, and Ruben Juarez. 2017. "Self-Enforcing Coalitions with Power Accumulation." *International Journal of Game Theory*, 46(2): 327–55.

Kendall-Taylor, Andrea, and Erica Frantz. 2016. "When Dictators Die." *Journal of Democracy*, 27(4): 159–71.

Levine, David K., Salvatore Modica, et al. 2016. "An Evolutionary Model of Intervention and Peace." UCLA Department of Economics.

Lofgren, M. 2016. *The Deep State: The Fall of the Constitution and the Rise of a Shadow Government.* Penguin.

Magaloni, Beatriz, and Ruth Kricheli. 2010. "Political Order and One-Party Rule." *Annual Review of Political Science*, 13: 123–43.

McCombs, Maxwell E., and Donald L. Shaw. 1972. "The Agenda-Setting Function of Mass Media." *Public Opinion Quarterly*, 36(2): 176–87.

Merton, Robert K., et al. 1968. "The Matthew Effect in Science." *Science*, 159(3810): 56–63.

Michels, Robert. 1915. *Political Parties: A Sociological Study of the Oligarchical Tendencies of Modern Democracy.* Hearst's International Library Company.

Olson, Mancur. 2000. *Power and Prosperity: Outgrowing Communist and Capitalist Dictatorships: Outgrowing Communist and Capitalist Dictatorships.* No. 338.9 Ol84p Ej. 1 025107. Basic books.

O'Neil, P. H. 2017. "The Deep State: An Emerging Concept in Comparative Politics." *Available at SSRN 2313375.*

Paige, Jeffrey M. 1999. *Coffee and Power: Revolution and the Rise of Democracy in Central America.* Cambridge: Harvard University Press.

Paludan, Ann. 1998. *Chronicle of the Chinese Emperors: The Reign-by-Reign Record of the Rulers of Imperial China.* London: Thames and Hudson.

Pipes, Richard. 2011. *Russia Under the Bolshevik Regime.* New York: Vintage.

Pycia, Marek. 2012. "Stability and Preference Alignment in Matching and Coalition Formation." *Econometrica*, 80(1): 323–62.

Ray, Debraj, and Rajiv Vohra. 2015. "The Farsighted Stable Set." *Econometrica*, 83(3): 977–1011.

Robinson, James A. 2012. "Elites and Institutional Persistence." *The Role of Elites in Economic Development*, 29–52. http://dx.doi.org/10.1093/acprof:oso/9780199659036.003.0003

Schatz, Edward. 2009. "The Soft Authoritarian Tool Kit: Agenda-Setting Power in Kazakhstan and Kyrgyzstan." *Comparative Politics*, 41(2): 203–22.

Scheufele, Dietram A. 1999. "Framing as a Theory of Media Effects." *Journal of Communication*, 49(1): 103–22.

Scheufele, Dietram A., and David Tewksbury. 2007. "Framing, Agenda Setting, and Priming: The Evolution of Three Media Effects Models." *Journal of Communication*, 57(1): 9–20.

Stone, Samuel Z. 1990. *The Heritage of the Conquistadors: Ruling Classes in Central America from the Conquest to the Sandinistas.* Lincoln: University of Nebraska.

Svolik, Milan W. 2012. *The Politics of Authoritarian Rule.* Cambridge: Cambridge University Press.

Tao, Li. 2004. "Xu zizhi tongjian changbian." *Collected data for a continuation of the Comprehensive mirror for aid in government.*

Wiener, Jonathan M. 1982. "Social Origins of the New South: Alabama, 1860–1885." *Science and Society*, 46(1): 97–100.

Woodward, Comer Vann. 1955. *The Strange Career of Jim Crow.* USA: Oxford University Press.

Zhang, Yan Bing, Mei-Chen Lin, Akihiko Nonaka, and Khisu Beom. 2005. "Harmony, Hierarchy and Conservatism: A Cross-Cultural Comparison of Confucian Values in China, Korea, Japan, and Taiwan." *Communication Research Reports*, 22(2): 107–15.