Plundering Coalitions*

Weijia Li† Alireza Mohammadi Sepahvand‡

February 6, 2026

Abstract

We develop a model to study coalitions that extract the resources of outsiders. The players in our model are endowed with power and resources. The ruling coalition plunders outsiders, distributes the plundered resources among its members, and guarantees that insiders' resources remain safe. Under natural conditions, we show that a unique ruling coalition exists using both axiomatic and non-cooperative approaches. Our analysis focuses on the resilience of the ruling coalition to shocks affecting the power and resources of both insiders and outsiders, as well as the intensity of plundering. We show that a coalition with a classical hierarchical structure—where power and resources are equal within each "rank" but strictly higher in higher ranks—exhibits greater resilience to external shocks affecting outsiders' power and resources. The only exception arises when plundering intensity is relatively weak, in which case the internal distribution of power and resources does not affect external resilience. Our final results provide insights into how the intensity of plundering impacts the internal and external resilience of ruling coalitions across political environments.

Keywords: Political Economy, Coalition Formation, Institutions, Resilience, Plundering.

1 Introduction

Coalition formation is always challenging (Ray and Vohra (2015a)), and a "plundering coalition" is no exception. For such a coalition, the wealth they can distribute among

*This paper is extracted from the doctoral dissertation of Alireza Mohammadi Sepahvand, completed as part of his Ph.D. studies at Monash University. We thank Arthur Campbell, Yves Zenou, Matthew Jackson, John Moore, Kaveh Majlesi, Massimo Morelli, Paul Raschky, Vai-Lam Mui, Sam Jindani, Solmaz Moslehi, Melika Liporace, Pawel Gola, and seminar and conference participants at Monash University, and the University of Edinburgh for helpful comments and suggestions. All errors are ours.

†Monash University, Australia. Email: Weijia.li@monash.edu.
‡Edinburgh University, UK. Email: mailto:amohamma@ed.ac.uk.

coalition members is plundered from the outsiders. This setup applies to a wide range of important social phenomena, such as an army that plunders the civil society, or an oligarchical government that taxes its citizens (Puga and Trefler (2014); Xu (2018); Sánchez De La Sierra (2020); Henn et al. (2024)). We formally study the problem to form a coalition whose primary objective is to plunder outsiders. To our knowledge, this is the first ever such attempt in coalition formation games. Our model yields a series of novel results. Among others, we study the resilience of a plundering coalition against outsiders, which justifies the key organizational principle of hierarchy for an effective army or a stable oligarchy. By doing so, we also propose a novel methodology to analyze the resilience of an equilibrium coalition against exogenous shocks.

Our model features a society of a finite number of individuals. Each individual has two endowments, power and wealth. The power of a coalition is the summation of all its members' powers. A "winning coalition" can defeat the outsiders with its power.[1] The game starts with an initial winning coalition. A member of the initial coalition may propose to create a new coalition. If all members of the proposed coalition approve and this is a winning coalition, the new coalition is formed and becomes the ruling coalition. Otherwise, another member can make a proposal, and the game continues until either all members of the initial coalition have proposed or a new coalition is formed. If no new coalition is formed and nobody remains from the initial winning coalition to propose, the initial winning coalition becomes the ruling coalition. The emerging ruling coalition will then defeat the outsiders, plunder their wealth, and distribute the plundered wealth among the members of the ruling coalition.

We are primarily interested in the properties of the ruling coalition. The ruling coalition is shaped by the following trade-off. By bringing a new member to the coalition, the new coalition is more powerful against outsiders, therefore being able to plunder more wealth from the defeated outsiders. But a new member is also costly for existing insiders because they cannot plunder the wealth of the new member anymore. We show that a ruling coalition that optimally balances the trade-off exists and is unique, and it corresponds to an axiomatic characterization of the coalition formation game. The ruling coalition has

To prepare our novel analysis of resilience, we prove a necessary and sufficient condition for the ruling coalition in equilibrium. First, the coalition must be better at plundering than any of its sub-coalitions. This motivates us to define a concept of "internal resilience:" a ruling coalition is more "internally resilient" if it is more likely to survive an exogenous perturbation to the power and resources of its own members. Sec-

[1]More rigorously, a coalition is a winning coalition if its power is higher than the β fraction of the total power of society, with $\beta > 1/2$.

ond, the coalition must be better at plundering than any possible alliances between one of its sub-coalitions and any subset of outsiders. This motivates us to define of a concept of "external resilience." Holding the power and resources of its own members constant, a ruling coalition is more "externally resilient" if it is more likely to survive an exogenous perturbation to the power and resources of outsiders. We then focus on the external resilience because it is more challenging to conceptualize and characterize than internal resilience.

To understand the socioeconomic condition of high external resilience, we conduct a thought experiment to make any two coalition members more "homogenous." Specifically, consider an exogenous transfer of power from a stronger member to a weaker member, without flipping their power rank, or a transfer of wealth from a richer member to a poorer member, without flipping their resource rank, or both. This transfer holds the characteristics of the ruling coalition constant, so it is still the unique ruling coalition. But importantly, such a transfer reduces the risk of the more threatening member with stronger power or lower wealth. After the transfer, the ruling coalition becomes more resilient to an alliance between a sub-coalition that includes the more threatening member and any subset of outsiders, where the outsiders are subject to any possible perturbation of their resources and power. At the same time, the ruling coalition is equally resilient to an alliance between a sub-coalition that includes the less threatening member and any subset of outsiders. Therefore, the ruling coalition becomes more externally resilient if two of its members become more homogenous.

It is important to note that the analysis does not imply that a ruling coalition is the most externally resilient if its members are absolute equal. Instead, the analysis implies that more externally resilient than others is a ruling coalition of a classic hierarchical structure. Such a hierarchical coalition consists of well-defined "ranks." Within each "rank," all members are absolutely equal with each other; but higher "ranked" members are both richer and more powerful than lower ranked members. Once such a hierarchy emerges, it is not possible to further improve external resilience through an operation of transfer as above. Our analysis therefore offers a justification for the classical hierarchical structure of many organizations, such as armies and bureaucracy, by their unique capacity in bearing changes to its enemies/subjects. This justification is, as far as we know, novel, in contrast to the conventional emphasis on the advantage of a hierarchical structure in incentive-alignment (Qian (1994); Mookherjee (2013)) or division of labor in (Garicano (2000); Garicano and Rossi-Hansberg (2015)).

Finally, we jointly investigate how internal and external resilience respond to a change in the environment, i.e., a change in the plundering "technology." Consider that, holding the power and wealth of the ruling coalition and society constant, the ruling coalition becomes more capable of extracting wealth from society. This exogenous change raises the cost of keeping a player within the ruling coalition, because the insiders' resources remain safe and are not subject to plundering. As a result, the preference of the members of the ruling coalition for "exclusive" alternatives—less powerful and poorer than the ruling coalition—increases, while their inclination for "inclusive" alternatives—more powerful and richer than the ruling coalition—decreases. The internal threats to the ruling coalition are its sub-coalitions which are exclusive alternatives. Therefore, a stronger plundering process decreases the internal resilience.[2] This contrasts with the naive view that plundering more intensively increases the insiders' attachment to the ruling coalition.

For the external resilience of the ruling coalition, a stronger plundering technology is a double-edged sword. On one hand, exclusive alternatives which involve small segments of society become more threatening to the ruling coalition. On the other hand, inclusive alternatives which encompass broader segments of society become less threatening. Thus, the realization of these alternatives—the specification of shocks—becomes particularly important. If the exclusive alternatives are more likely to emerge, a stronger plundering technology decreases external resilience. Instead, if inclusive alternatives are more likely to appear, a stronger plundering process increases the external resilience. The latter suggests that a ruling coalition that engages in power-light plundering of society benefit more from facing a more powerful and wealthier opposition than a weaker and poorer one.

Lastly, although the direction of change in external resilience generally depends on the realization of powers and resources inside the ruling coalition, we identify a wide range of political environments where this is not the case. That is, the external resilience is robust with respect to changes in internal configuration of powers and resources. In these political environments, the plundering process is "power-intensive;" for instance, it is endowed with better protections of property rights. Precisely, in these contexts, corresponding to any exclusive alternative, there always exists an inclusive alternative that is more threatening to the ruling coalition. This implies that the only factor affecting external resilience is the players' preference for inclusive alternatives. As a result, a stronger plundering technology always increases the external resilience of the ruling coalition, since it renders the inclusive alternatives less beneficial for the players. Thus, in power-intensive plundering environments, there exists a trade-off between external and internal resilience of the ruling coalition with respect to the plundering intensity, regardless of the specifications of internal and external shocks. This offers a novel insight: even imperfect property rights—which do not fully prevent plundering by insiders—could potentially hinder the ruling coalition from achieving both internal and external stability

[2]This generally holds regardless of the specifics of the perturbations.

when plundering technology changes. This contrasts with "power-light plundering" environments, wherein a change in plundering technology could alleviate both internal and external threats to the ruling coalition.

1.1 Relevant Literature

Our paper is relevant to a few strands of literature. The literature on coalition formation largely focus on characterizing the equilibrium coalition (Acemoglu et al. (2008); Ray and Vohra (2015b); Battaglini (2021)), or define stability mainly by incorporating the notion of "farsightedness" (Harsanyi (1974); Ray and Vohra (2015c)). We instead take one step further by analyzing the resilience of the equilibrium coalition against exogenous shocks. By doing so, we make a methodological contribution by proposing a simple framework to analyze the resilience of the equilibrium coalition. This novel focus on resilience also uncovers a lot of new substantive insights.

We bring together the two strands of literature on coalition formation and organizational economics of hierarchy. Existing literature usually focus on how a hierarchy may improve incentive-alignment or division of labor (Qian (1994); Qian et al. (2006); Mookherjee (2013); Garicano (2000); Garicano and Rossi-Hansberg (2015)). We offer a new justification for hierarchy: we show that a hierarchy is uniquely resistant to arbitrary exogenous changes to the characteristics of individuals outside the hierarchy. Our novel justification is relevant to many hierarchies where the characteristics of outsiders are a first order concern, such as armies and fiscal bureaucracies (Besley and Persson (2009); Xu (2018); Sánchez De La Sierra (2020); Henn et al. (2024)).

Our model also makes novel contributions to a few central debates in political economy. First, political economists have uncovered that the interaction between power and wealth is a fundamental thread in political economy (Acemoglu and Robinson (2008); Dal Bó and Dal Bó (2011); Dal Bó et al. (2022); Acemoglu and Robinson (2013)). We contribute to this literature by an in-depth analysis of the power-wealth trade-off through the lens of coalition formation, the first ever attempt to our knowledge. It is through the coalition analysis that we uncover the innovative insight on the unique resilience of a hierarchical organization.

Our analysis also contributes to the burgeoning literature on political economy of non-democracies (Egorov and Sonin (2024)). Specifically, our analysis of internal and external resilience engages with the literature that addresses the trade-offs that authoritarian states resolve while dealing with internal or external threats to their rule. A strand of literature studies the loyalty-competence trade-off, i.e., how autocratic states balance the competence of their officials against their loyalty to prevent internal dissent (Besley and Kudamatsu (2007); Egorov and Sonin (2011); Jia et al. (2015); Zakharov (2016); Bai and

Zhou (2019); Mattingly (2024)). Another strand of literature focuses on external problems such as mass protests, or propaganda (Wintrobe (1990); Wintrobe (2000); Konrad and Skaperdas (2007) Egorov et al. (2009); De Mesquita (2010); Yanagizawa-Drott (2014); Shadmehr (2018)). There are many trade-offs the dictators resolve while tackling external threats, for instance, the one between "informational openness" and "security" (Lorentzen (2013); Gehlbach and Sonin (2014); Lorentzen (2014); Guriev and Treisman (2019); Enikolopov et al. (2020)). Through the novel lens of coalition formation, we contribute to this literature by showing how the internal and external threats are related. In particular, we identify the condition for a trade-off between internal and external resilience driven by the process of coalition formation. Additionally, we provide insights into when this trade-off does not hold, and the characteristics of oppositions that can benefit an autocratic state engaging in intensive plundering.

The remainder of the paper is organized as follows. Section 2 introduces the model. Section 3 presents the preliminary analysis of the coalition formation game. Building on Section 3, we proceed by studying the resilience in Section 4. Section 5 concludes.

2 Environment

There is a set of players $N = \{1, 2, \dots, n\}$. We denote the set of all subsets of N by 2^N. Time is finite and indexed by $t \in \{1, 2, \dots, T\}$. The players are endowed with a pair of power p and resources x, specified by the mappings

$$p(\cdot) : N \to \mathbb{R}_{++},$$
$$x(\cdot) : N \to \mathbb{R}_{++}.$$

We refer to $p_i := p(i)$ and $x_i := x(i)$ as the political power and resources of individual $i \in N$. A non-empty set $I \subseteq N$ is called a coalition. Any player can be a member of only one coalition at any stage of the game. The power and resources of any coalition $I \subseteq N$ are denoted by

$$P_I = \sum_{i \in I} p_i \quad \text{and} \quad X_I = \sum_{i \in I} x_i.$$

Let $P_N := \sum_{i \in N} p_i$ and $X_N := \sum_{i \in N} x_i$. The coalition I is called a winning coalition if $P_I \geq \beta P_N$, where $\beta \in (1/2, 1]$ is a fixed supermajority requirement. Denote the set of all winning coalitions by $\mathcal{W}$. There is a baseline payoff function $U : N \times \mathcal{W} \to \mathbb{R}$ that assigns to any player $i \in N$ the payoff $U_i(I)$ when the winning coalition $I \in \mathcal{W}$ becomes the ruling coalition. We also write $U(i, I) := U_i(I)$.

A ruling coalition of our model is necessarily a winning coalition. As the key novelty of our setup, a ruling coalition can only plunder outsiders, while the resources of its members are safe. This creates a central trade-off for our model. A new member who is brought into the ruling coalition strengthens its capability to plunder outsiders, but the ruling coalition loses the opportunity to plunder this new member. This key trade-off is formally captured by Assumption 1.

Assumption 1. *[Payoffs] For any $i \in N$ and $I \in \mathcal{W}$, $U_i(I) := x_i + w_i(I)$, where $w_i(\cdot)$ satisfies the following properties:*

1. *(Trade-off) If $I \in \mathcal{W} \setminus \{N\}$ and $i \in I$, we have $w_i(I) = G_i(P_I, X_I) > 0$, where $G_i(\cdot,\cdot) : [\beta P_N, P_N) \times [0, X_N) \to \mathbb{R}_{++}$ is a function continuous in P_I and X_I, satisfying the following conditions.*
   - *(a) For all I and $I' \in \mathcal{W}\setminus\{N\}$ with $P_I = P_{I'}$, if $i \in I$ and $i \in I'$, then $G_i(P_I, X_I) > G_i(P_{I'}, X_{I'})$ if and only if $X_I < X_{I'}$.*
   - *(b) For all I and $I' \in \mathcal{W} \setminus \{N\}$ with $X_I = X_{I'}$, if $i \in I$ and $i \in I'$, then $G_i(P_I, X_I) > G_i(P_{I'}, X_{I'})$ if and only if $P_I > P_{I'}$.*
2. *If $I \in \mathcal{W} \setminus \{N\}$ and $i \notin I$, then $w_i(I) < 0$.*
3. *For all $i \in N$, $w_i(N) = 0$.*

Assumption 1 establishes some of the key primitives of the model. In Part 1, the function $G_i(\cdot,\cdot)$ ranks the plundered resources of any individual across non-trivial ruling coalitions of which she is a member.[3] Part 1(a) says that between ruling coalitions with equal power, players prefer the one with fewer internal resources, which permits more external resources for plundering. Meanwhile, between ruling coalitions with equal resources, players prefer the one with larger power (Part 1(b)), as it strengthens the ruling coalition in extracting resources. Both Part 1(a) and Part 1(b) imply that when the ruling coalition is not the grand coalition N, insiders obtain strictly positive payoffs from plundering outsiders. Together with Part 2, this implies that inclusion in the ruling coalition strictly benefits insiders relative to their initial resources, while exclusion strictly harms outsiders relative to their initial resources. Part 3 states that the players' payoff from the plundered resources is zero when the ruling coalition is N, since there are no outsiders to plunder.

Under Assumption 1, a ruling coalition is fully characterized by its power and resources. This keeps the model tractable by eliminating the complexities that arise when the specific combination of players inside the ruling coalition also matters. For the rest

[3]For example, one can view $G_i(P_I, X_I)$ as a combination of a plundering component $F(I) : \mathcal{W} \to \mathbb{R}_{++}$ and a share component $\Pi(i, I) : N \times \mathcal{W} \to [0, 1]$, i.e., $G_i(P_I, X_I) := \Pi(i, I)F(I)$ is the share allocated to individual i within the coalition I from plundered resources $F(I)$.

of this paper, we thus write $G_i(P_I, X_I)$ for player i's plunder gains in coalition I. Assumption 1 immediately yields the following lemma.

Lemma 1. *Under Assumption 1, any player $i \in N$ has strictly increasing and continuous indifference curves over (P, X). The variables P and X are the aggregate power and resources of ruling coalitions that include the player i; $(P, X) \in [\beta P_N, P_N) \times [0, X_N)$.*

Lemma 1 does not imply that indifference curves are identical across individuals. The following assumption imposes common preferences for players, which simplifies notation throughout the paper. We later show that the main results continue to hold under a considerably weaker assumption.

Assumption 2. *For all $I \in \mathcal{W}$ and all $i \in I$, $G_i(I) := g(i)G(I)$, where $g(i) > 0$.*

Under Assumption 2, there are two components in a player's preference over ruling coalitions that include the player: an idiosyncratic component $g(i)$ and a common component $G(I)$, which depends on the aggregate powers and resources of the coalition, (P_I, X_I). This assumption implies that for all players, the indifference curves over the coalitions containing them are the same and determined by the function $G(\cdot)$ (Figure 1). In other words, for any ruling coalitions $I, I' \in \mathcal{W}$ and any $i, j \in I \cap I'$, we have $U_i(I) \geq U_i(I')$ if and only if $U_j(I) \geq U_j(I')$, i.e., the preferences of players over any pair of ruling coalitions including them are identical.

Remark 1. *Appendix B provides a simple microfoundation for Assumption 2. In particular, Assumption 2 holds if insiders' payoffs decompose as $w_i(I) = \Pi_i(I)\, F(I)$, where $\Pi_i(I)$ is an intra-coalition share (e.g., proportional to p_i/P_I) and $F(I)$ is the coalition's total extractable surplus, depending only on aggregate characteristics such as $(P_I, X_N - X_I)$. This structure induces a common ordering over the deviation coalitions relevant for the resilience constraints.*

Furthermore, Assumption 2 is not required for the main results and is imposed mainly to simplify notation. All results can be derived under a weaker condition that requires preference "consistency" only on a restricted domain, i.e., the set of "potential ruling coalitions" Z in Definition 3.1.[4] *This milder assumption is also natural for our focus on resilience. Since we study whether a ruling coalition survives a shock affecting outsiders, it is reasonable to require that its members (weakly) prefer to be in that coalition before the shock; otherwise, external resilience is trivially zero.*

[4]More precisely, the proofs only invoke preference comparisons among coalitions in $\mathcal{Z}$ and among the sub-coalitions and deviation coalitions generated by Z that enter the resilience constraints.

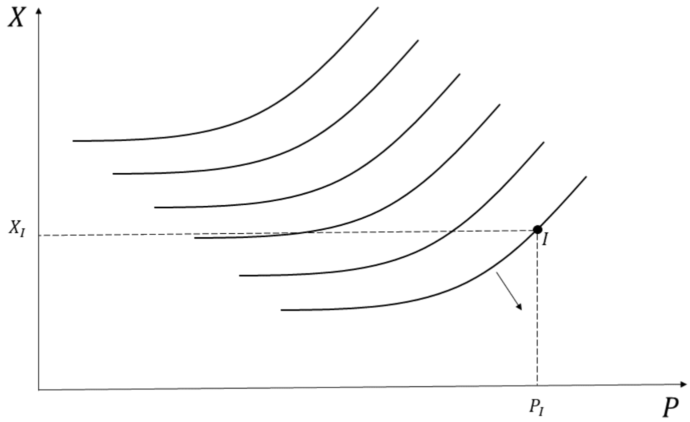


Figure 1: Identical indifference curves under Assumption 2

Definition 1. *Fix a function $G(\cdot)$ that satisfies Assumption 1. For any ruling coalition I, denote the indifference curve through I by $X := H_I(P)$, which is implicitly defined by $G(P, X) = G(P_I, X_I)$.*

Throughout the paper, we assume that the joint power and resources mapping is generic in the sense that for all $I, I' \in \mathcal{W}$, we have $P_I \neq P_{I'}$ or $X_I \neq X_{I'}$.[5] The following assumption helps establish the uniqueness results in the subsequent section.

Assumption 3. *Fix the power and resource mappings. Then, for all $I, I' \in \mathcal{W}$, we have $G(I) \neq G(I')$.*[6]

This assumption implies that players receive strictly different payoffs from different ruling coalitions involving them.

3 Preliminary Analysis of the Coalition Formation Game

This section establishes existence and uniqueness of the coalitional equilibrium, which prepares our analysis of its "resilience," i.e., how the equilibrium responds to exogenous shocks to players' power or resources.

[5] Mathematically, this assumption is without much loss of generality, since the set of vectors $\{(P_I, X_I)\} \in \mathbb{R}_{++}^{2^{|N|+1}}$ that are not generic is the union of finitely many hyperplanes and therefore has Lebesgue measure zero.

[6] This assumption is also made without much loss of generality, as the set of functions from $\mathbb{R}^2$ to $\mathbb{R}$ for which the outputs coincide on a finite set of distinct inputs forms a measure-zero set in the space of all functions from $\mathbb{R}^2$ to $\mathbb{R}$.

3.1 Axiomatic analysis

We begin with an axiomatic analysis. As in Acemoglu et al. (2008) and Acemoglu et al. (2012), it shows that our results are independent of the details of the agenda-setting and voting protocols in the non-cooperative game introduced in Section 3.2. The axiomatic analysis will also help characterize the equilibrium of the non-cooperative game in Section 3.2.

We define a correspondence $\phi : \mathcal{W} \rightrightarrows 2^N$, which identifies the set of ruling coalitions corresponding to each initial winning coalition. We assume that ϕ satisfies the following axioms:

Axiom 1 (Non-triviality)**.** *For any $I \in \mathcal{W}$, $\emptyset \notin \phi(I)$ and $N \notin \phi(I)$.*

Axiom 2 (Super-majority of Power)**.** *For any $I \in \mathcal{W}$ and any $I' \in \phi(I)$, we have $I' \in \mathcal{W}$.*

Axiom 3 (Rationality)**.** *For any $I \in \mathcal{W}$, any $I' \in \phi(I)$, and any $I'' \in \mathcal{W}$,*

$$I'' \notin \phi(I) \iff G(I'') < G(I').$$

These axioms are natural and capture the economic forces that give rise to the pure-strategy SPE of the game in Section 3.2. Axiom 1 requires ϕ to map any initial winning coalition to a non-trivial ruling coalition. Axiom 2 requires any ruling coalition selected by ϕ to be a winning coalition. Axiom 3 imposes payoff-based selection: if $I' \in \phi(I)$, then no winning coalition I'' with strictly lower $G(\cdot)$ can be selected, and conversely any winning coalition with strictly higher $G(\cdot)$ must be selected. Proposition 1 establishes that these axioms pin down a unique mapping under Assumptions 1–2, and that the correspondence is single-valued under Assumptions 1–3.

Proposition 1. *1. (Existence) Under Assumptions 1–2, the unique correspondence that satisfies Axioms 1–3 is*

$$\phi(I) = \arg\max_{W \in \mathcal{W}} G(W).$$

2. (Uniqueness) Under Assumptions 1–3, the correspondence ϕ is single-valued.

Proposition 1 is straightforward. It shows that the ruling coalition is a winning coalition that maximizes plunder, i.e., it maximizes $G(W)$ among all $W \in \mathcal{W}$. Under Assumption 3, this coalition is unique.

3.2 The non-cooperative extensive game

We next define the extensive-form complete-information game

$$\Gamma = \big(N, I_0, p(\cdot), x(\cdot), \{U_i(\cdot)\}_{i\in N}, \beta\big),$$

where N is the set of players, I_0 is the initial winning coalition, $p(\cdot)$ and $x(\cdot)$ are the power and resource mappings, $\{U_i(\cdot)\}_{i\in N}$ are the payoff functions satisfying Assumption 1 and Assumption 2, and $\beta \in (1/2, 1]$ is the supermajority requirement. The game starts with the initial winning coalition $I_0 \in \mathcal{W}$, and the steps are as follows:

1. Nature randomly picks an agenda setter a_q from the initial winning coalition, with $q = 1$, where $q \in \{1, \dots, |I_0|\}$ denotes the round of agenda setting and voting.

2. The agenda setter a_q proposes a coalition $I_q \subseteq N$. If $P_{I_q} < \beta P_N$, then the game proceeds to Step 4. Otherwise, Nature chooses an order of votes and the game proceeds to Step 3.

3. The voting process begins. The coalition I_q forms if and only if the proposal of a_q is accepted by *all* players in I_q. In this case, I_q becomes the ruling coalition, and each player $i \in N$ receives payoff $U_i(I_q) = x_i + w_i(I_q)$. Otherwise, following the first rejection of the proposal, the game proceeds to Step 4.

4. If $q < |I_0|$, Nature randomly picks a *new* agenda setter $a_{q+1} \in I_0 \setminus \{a_1, a_2, \dots, a_q\}$ and the game returns to Step 2. If $q = |I_0|$, then I_0 becomes the ruling coalition and each player $i \in N$ receives payoff $U_i(I_0) = x_i + w_i(I_0)$.

The solution concept is subgame perfect equilibrium (SPE). The extensive-form game specifies players' strategies in any such equilibrium. A pure strategy of any player $i \in I_0$ is a pair of functions $\sigma_i(h) = (v_i(h, \mathcal{P}), \mathcal{P}_i(h))$ specifying her behavior at each decision node h: the function $v_i(h, \mathcal{P})$ specifies player i's vote (either 'Yes' or 'No') in any history h where Nature selects her to vote on a proposal $\mathcal{P}$, and $\mathcal{P}_i(h)$ specifies the coalition that player $i \in I_0$ proposes if selected by Nature as the agenda setter in history h. According to the extensive-form game, if $i \in N \setminus I_0$, player i may only be a voter throughout the game.[7] Thus, the strategy of any $i \in N \setminus I_0$ is the voting function $v_i(h, \mathcal{P})$, which assigns either 'Yes' or 'No' to any proposed ruling coalition $\mathcal{P}$ containing i in any history h where Nature selects her to vote on $\mathcal{P}$.

[7] All results continue to hold if the game is modified so that all players can be both voters and proposers.

We now establish existence and uniqueness of the ruling coalition in the non-cooperative coalition formation game, a preliminary result that prepares our analysis of the equilibrium's resilience to exogenous shocks. We also show that the SPE outcome of the coalition formation game coincides with the ruling coalition characterized by the axiomatic approach in Section 3.1.

Proposition 2. *1. (Existence) Suppose that Assumptions 1–2 hold and that $\phi(I_0)$ satisfies Axioms 1–3. Then, for any $I \in \phi(I_0)$, there exists a pure-strategy SPE σ_I that produces I as the ruling coalition. In this SPE, each player $i \in N$ receives payoff $U_i(I) = x_i + w_i(I)$.*

2. (Uniqueness) Suppose that Assumptions 1–3 hold, that $\phi(I_0)$ satisfies Axioms 1–3, and that $\phi(I_0) = \{I\}$. Then, in any SPE, I is the ruling coalition. In particular, in any SPE, each player $i \in N$ receives payoff $U_i(I) = x_i + w_i(I)$.

The intuition is straightforward given Assumptions 1–3 and the axiomatic characterization in Proposition 1, where $\phi(I_0) = \arg\max_{W \in \mathcal{W}} G(W)$ for any $I_0 \in \mathcal{W}$. Any ruling coalition I identified by the axiomatic analysis (i.e., $I \in \phi(I_0)$) can be supported by an SPE in which every agenda setter from I_0 proposes I, and every voter from I_0 accepts I and rejects any other proposal. Under the supermajority rule $\beta \in (1/2, 1]$, any coalition I' proposed before I must include at least one player from I. Since coalitions form under unanimity, the proposed strategy prevents any such I' from becoming the ruling coalition, ensuring that I forms.[8] Moreover, in the axiomatic analysis, Assumption 3 implies that the correspondence $\phi : \mathcal{W} \to 2^N$ is single-valued, so any SPE yields the same ruling coalition.

Remark 2 (A direct property of the ruling coalition)**.** *We comment on a direct property of the ruling coalition, which is useful for our main analysis on resilience. We define "potential ruling coalitions" as follows.*

Definition 2. *For any power and resource mappings $p(\cdot)$ and $x(\cdot)$, define the set of potential ruling coalitions as*

$$Z := \{I \in \mathcal{W} \mid \nexists I' \in \mathcal{W} \text{ such that } P_{I'} > P_I \text{ and } X_{I'} < X_I\}. \tag{3.1}$$

Figure 2 illustrates the potential ruling coalitions, which are the blue dots in the figure. Intuitively, a winning coalition is a potential ruling coalition if and only if there does not exist another winning coalition with strictly higher power and strictly lower resources.

[8] Off-path equilibrium strategies are reported in Equation 3 in the proof of Proposition 2(1) in the appendix.

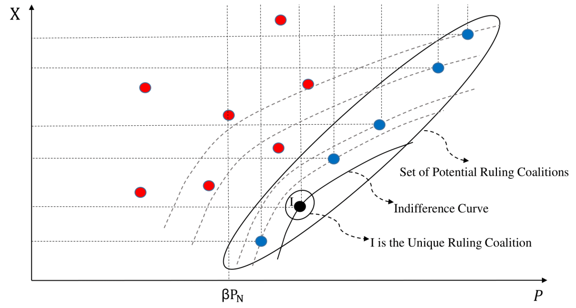


Figure 2: I is the unique ruling coalition of the game.

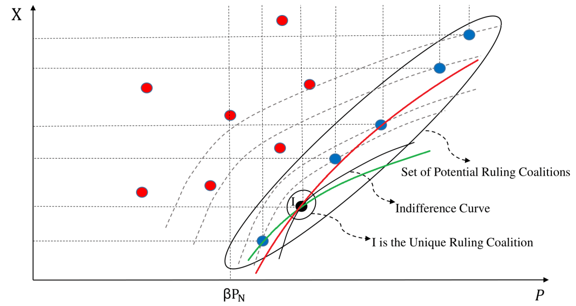


Figure 3: I is the unique ruling coalition under heterogeneous but consistent preferences over Z (each color represents one player's preferences over Z).

It is straightforward that the ruling coalition in Proposition 2 must be a potential ruling coalition. Though simple, this property will be useful in our subsequent analysis.

Remark 3 (Robustness without Assumption 2)**.** *While Assumption 2 pins down a ruling coalition, our resilience analysis does not require it. It is enough that preferences be consistent on the set of potential ruling coalitions Z, so that a ruling coalition is well-defined and we can later study its resilience to external shocks (Figure 3). Moreover, the resilience analysis under Assumption 2 is informative even when preferences differ across members, i.e., when insiders disagree about what constitutes a "good" ruling coalition. As we discuss below, such preference heterogeneity tends to reduce the ruling coalition's external resilience.*

4 Main analysis on coalitional resilience

This section studies the resilience of the ruling coalition, which is the central part of the paper. Our analysis proceeds in three steps. First, Proposition 3 provides a useful characterization of the ruling coalition and allows us to define "internal" and "external" resilience, i.e., robustness to changes in the power and resources of members versus outsiders. Second, Proposition 4 shows that the ruling coalition that is most externally resilient has a hierarchical structure. Third, we show that internal and external resilience can trade off, depending on the "intensity" of plundering. Studying coalitional resilience is also useful for understanding the dynamics of the ruling coalition.

4.1 Internal and external resilience

Section 4.1 first establishes that a ruling coalition must have a relatively higher power-to-resource ratio than alternative winning coalitions. Equally important, Section 4.1

shows that it is sufficient for the ruling coalition to dominate two types of threats: sub-coalitions of the ruling coalition and alternative coalitions that include players outside the ruling coalition. This distinction reflects the central challenges from regime insiders and outsiders (Svolik (2012); Meng (2020); Paine (2021); Egorov and Sonin (2024)), enabling us to distinguish between two notions of resilience.

To proceed, we first define the set of "best sub-coalitions."

Definition 3. *For any $p(\cdot)$ and $x(\cdot)$ and any subset of players I, define the set of best sub-coalitions of I as follows:*

$$\mathcal{A}_I := \{A \subseteq I \,|\, A \neq \emptyset,\ \nexists A' \subseteq I \text{ such that } P_{A'} > P_A \text{ and } X_{A'} < X_A\}\,. \tag{4.1}$$

For any subset of players $I \subseteq N$, $\mathcal{A}_I$ includes the best subsets of I, i.e., those for which there does not exist another subset of I with both higher power and lower resources. Equation 4.1 is analogous to Equation 3.1 in Definition 3.1 for potential ruling coalitions, but restricts attention to sub-coalitions of the coalition in question. We can now prove Proposition 3, which characterizes the two types of threats to the ruling coalition.

Proposition 3. *Fix the game $\Gamma = \left(I_0, p(\cdot), x(\cdot), \{U_i(\cdot)\}_{i\in N}, \beta\right)$ and suppose Assumptions 1–3 hold. Then $\phi(I_0) = \{I\}$ if and only if $I \in \mathcal{W}$ and:*

(i) For all $A^{\text{ins}} \in (\mathcal{A}_I \setminus \{I\}) \cap \mathcal{W}$, $G(I) > G(A^{\text{ins}})$ (i.e., there is no profitable internal secession).

(ii) For all $A^{\text{ext}} \in \mathcal{A}_{N\setminus I}$ and for all $A^{\text{ins}} \in \mathcal{A}_I$ with $A^{\text{ins}} \cup A^{\text{ext}} \in \mathcal{W}$, $G(I) > G(A^{\text{ins}} \cup A^{\text{ext}})$ (i.e., there is no profitable external secession).

Proposition 3 distinguishes between two types of alternatives that the ruling coalition must dominate: its own subsets (Condition (i)) and coalitions that include outsiders (Condition (ii)). It shows that a necessary and sufficient condition for I to defeat all alternative winning coalitions is to dominate (i) all its nontrivial best sub-coalitions and (ii) the combinations of its best sub-coalitions with best sub-coalitions of outsiders. The following example illustrates the proposition.

Example 1. *Suppose that for a ruling coalition I, $\mathcal{A}_I = \left\{A_1^{\text{ins}}, A_2^{\text{ins}}, I\right\}$ where $A_1^{\text{ins}}, A_2^{\text{ins}} \in \mathcal{W}$, and $\mathcal{A}_{N\setminus I} = \{A_1^{\text{ext}}, A_2^{\text{ext}}\}$, where $A_2^{\text{ext}} = N \setminus I$. Condition (i) states that I must be able to win against both A_1^{ins} and A_2^{ins}. Condition (ii) requires I to dominate any coalition of the form $A_j^{\text{ins}} \cup A_k^{\text{ext}}$, where $j, k \in \{1, 2\}$.*

Motivated by Condition (i) of Proposition 3, we now define the key object for our analysis of internal resilience. Using the indifference curve $H_I(\cdot)$ over aggregate power P

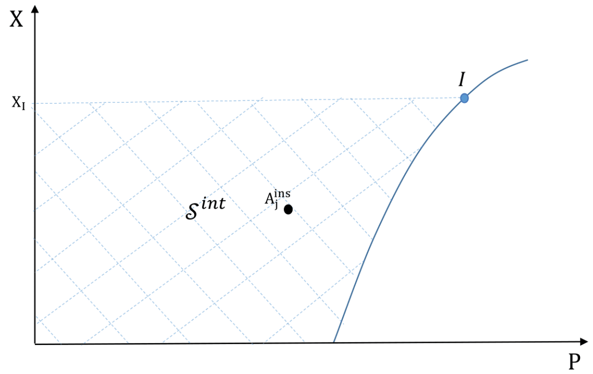


Figure 4: The internal safe area $\mathcal{S}^{\text{int}}$ is the shaded region.

and aggregate resources X defined in Definition 1, Condition (i) can be expressed in (P, X) space. For example, in Example 1, the ruling coalition must satisfy $X_{A_j^{\text{ins}}} > H_I(P_{A_j^{\text{ins}}})$ for $j \in \{1, 2\}$, which identifies a region in the (P, X) plane. This motivates the following definition.

Definition 4. *For the ruling coalition I, the "internal safe area" is*

$$\mathcal{S}_I^{\text{int}} = \left\{(P, X) \in \mathbb{R}_{++}^2 \;\middle|\; X > H_I(P)\right\}. \tag{4.2}$$

A ruling coalition I has the same internal resilience as a ruling coalition J if and only if $\mathcal{S}_I^{\text{int}} = \mathcal{S}_J^{\text{int}}$. A ruling coalition I is strictly (weakly) more internally resilient than a ruling coalition J if and only if $\mathcal{S}_J^{\text{int}} \subsetneq \mathcal{S}_I^{\text{int}}$ (respectively, $\mathcal{S}_J^{\text{int}} \subseteq \mathcal{S}_I^{\text{int}}$).

For any coalition I to be the ruling coalition, all its best sub-coalitions must lie in $\mathcal{S}^{\text{int}}$. This guarantees that all members of I prefer I to any best sub-coalition of I, satisfying Condition (i) of Proposition 3. To simplify notation, we write $\mathcal{S}^{\text{int}}$ when there is no confusion. Figure 4 illustrates the internal safe area $\mathcal{S}^{\text{int}}$ for a coalition I. This set plays a central role in our subsequent analysis of internal resilience. In particular, we will see that a coalition I remains stable if, after an exchange of power and resources within I, all its best sub-coalitions remain inside the internal safe area $\mathcal{S}^{\text{int}}$.

We now turn to "external" threats. For any best sub-coalition of outsiders A^{ext} and any best sub-coalition of insiders A^{ins}, Condition (ii) of Proposition 3 requires $G(I) > G(A^{\text{ins}} \cup A^{\text{ext}})$, which is equivalent to $X_{A^{\text{ins}} \cup A^{\text{ext}}} > H_I(P_{A^{\text{ins}} \cup A^{\text{ext}}})$. Since aggregate power and resources are additive, this condition can be written as $X_{A^{\text{ins}}} + X_{A^{\text{ext}}} > H_I(P_{A^{\text{ins}}} + P_{A^{\text{ext}}})$.

To illustrate the geometric interpretation, consider A_1^{ins} in Example 1. We ask which pairs $(P_{A^{\text{ext}}}, X_{A^{\text{ext}}})$ for a best sub-coalition A^{ext} of outsiders ensure that $A_1^{\text{ins}} \cup A^{\text{ext}}$ is *not* preferred to I. To do so, shift the curve $X = H_I(P)$ by the vector $(-P_{A_1^{\text{ins}}}, -X_{A_1^{\text{ins}}})$ and denote the shifted curve by $X = H_I^1(P)$. Analogously to the internal safe area, we define the "external safe area" for a ruling coalition I.

Definition 5. *Consider a ruling coalition I with its set of best sub-coalitions $\mathcal{A}_I$. For any $A_j^{\text{ins}} \in \mathcal{A}_I$, define*

$$H_I^j(P) := H_I(P + P_{A_j^{\text{ins}}}) - X_{A_j^{\text{ins}}}, \tag{4.3}$$

i.e., the indifference curve shifted by $(-P_{A_j^{\text{ins}}}, -X_{A_j^{\text{ins}}})$.

The area externally safe relative to A_j^{ins} is

$$\mathcal{S}_j^{\text{ext}} \equiv \left\{(P, X) \in \mathbb{R}_{++}^2 \;\middle|\; P < \beta P_N \textit{ and } X > H_I^j(P)\right\}. \tag{4.4}$$

Define the external safe area for the ruling coalition I as

$$\mathcal{S}_I^{\text{ext}} \equiv \bigcap_{A_j^{\text{ins}} \in \mathcal{A}_I} \mathcal{S}_j^{\text{ext}}. \tag{4.5}$$

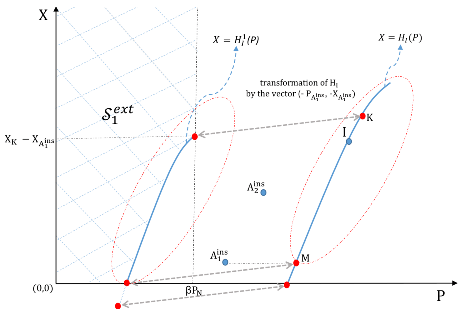
A ruling coalition I has the same external resilience as a ruling coalition J if and only if $\mathcal{S}_I^{\text{ext}} = \mathcal{S}_J^{\text{ext}}$. A ruling coalition I is strictly (weakly) more externally resilient than a ruling coalition J if and only if $\mathcal{S}_J^{\text{ext}} \subsetneq \mathcal{S}_I^{\text{ext}}$ (respectively, $\mathcal{S}_J^{\text{ext}} \subseteq \mathcal{S}_I^{\text{ext}}$).

What is the intuition behind Definition 5 and "external safety"? First, outsiders cannot have $P^{\text{ext}} \geq \beta P_N$. If they did, this outsider group would itself have supermajority power. Since $P_I \geq \beta P_N$ is required for I to be a ruling (winning) coalition, such an outsider group would necessarily be able to form a winning coalition on its own, and I could not remain the ruling coalition. Hence $P^{\text{ext}} < \beta P_N$ is a necessary condition for an outsider group to be "safe" for any ruling coalition I.

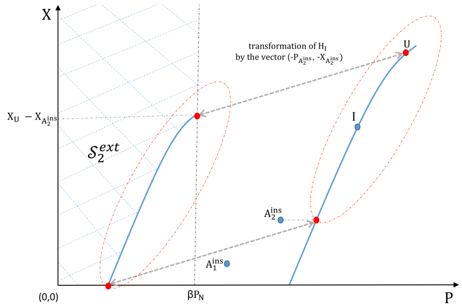
Second, fix a best sub-coalition of insiders $A_j^{\text{ins}} \in \mathcal{A}_I$. For any best sub-coalition of outsiders A^{ext} that lies exactly on the shifted indifference curve $H_I^j(P)$, insiders are indifferent between the current ruling coalition I and the alternative coalition $A_j^{\text{ins}} \cup A^{\text{ext}}$. If A^{ext} lies in the region $\mathcal{S}_j^{\text{ext}}$, then $A_j^{\text{ins}} \cup A^{\text{ext}}$ is strictly worse than I, so outsiders with such (P, X) cannot combine with A_j^{ins} to form a profitable external deviation. In this sense, $\mathcal{S}_j^{\text{ext}}$ is "externally safe" relative to A_j^{ins}. Taking the intersection over all $A_j^{\text{ins}} \in \mathcal{A}_I$ yields the external safe area $\mathcal{S}_I^{\text{ext}}$, i.e., the set of outsider best sub-coalitions that are simultaneously safe against *any* insider best sub-coalition.

Proposition 3 provides an additional insight: it underscores that a ruling coalition must maintain a relatively high power-to-resource ratio compared to relevant alternatives (as reflected in both the internal and external safe areas).[9] For example, the proposition offers a rationale for the voluntary destruction of resources by a ruling coalition when confronting a threatening alternative.

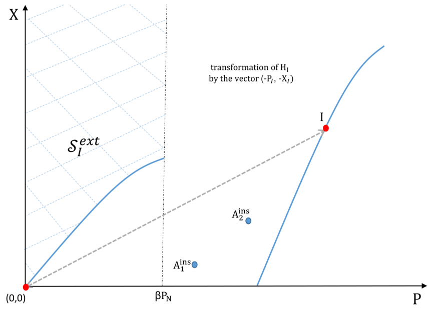
[9] Example 4 in the appendix demonstrates that neither the player with the highest power nor the one with the lowest resources is necessarily included in the ruling coalition.



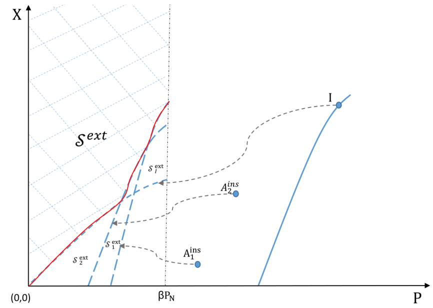
(a) Condition (ii) of Proposition 1 with respect to A_1^{ins}.



(b) Condition (ii) of Proposition 1 with respect to A_2^{ins}..



(c) Condition (ii) of Proposition 1 with respect to I.



(d) Condition (ii) of Proposition 1 with respect to all sub-coalitions of I.

Figure 5: The external safe area—$\mathcal{S}^{ext}$: the blue area above the red curve.

Remark 4 (Heterogeneity of preferences and resilience)**.** *How does preference heterogeneity within the ruling coalition affect external resilience once Assumption 2 is relaxed? Since the external safe area is defined as the region above the envelope of the boundaries induced by insiders' best sub-coalitions, introducing an insider whose preferences differ from others can only add an additional boundary, which can only shift the envelope weakly upward. Hence, the external safe area can only weakly shrink, i.e., external resilience weakly decreases. The same logic applies to the internal safe area.*

Now that we have provided a precise characterization of internal and external safe areas, we are ready to study internal and external resilience of a ruling coalition.

4.2 Which ruling coalitions are more externally resilient?

Consider a ruling coalition I and suppose there are two players $i, j \in I$ with $p_i > p_j$ and $x_i < x_j$. Holding the power and resources of all other players fixed, transfer either (i) a portion of player i's power, with $0 < \Delta p \leq \frac{p_i - p_j}{2}$, or (ii) a portion of player j's resources, with $0 < \Delta x \leq \frac{x_j - x_i}{2}$, to player i (Figure 6). The following proposition establishes that such an equalizing transfer within the ruling coalition (weakly) reduces the risk posed by relatively stronger or poorer members, and therefore (weakly) increases the coalition's external resilience. This result is quite general and does not depend on the precise specification of the plundering function $G(\cdot)$, i.e., on the shape of indifference

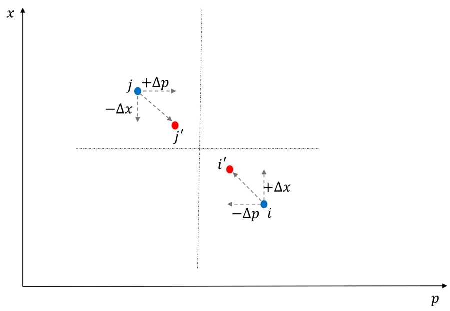


Figure 6: Exchange of powers and resources between player i and player j—from blue to red.

curves.

Proposition 4. *Suppose I is the unique ruling coalition of the game Γ, and there exist $i, j \in I$ with $p_i > p_j$ and $x_i < x_j$. Holding fixed the powers and resources of players in $I \setminus \{i, j\}$, consider the modified coalition $(I \setminus \{i, j\}) \cup \{i', j'\}$ where*

$$p_{i'} = p_i - \Delta p, \quad x_{i'} = x_i + \Delta x, \quad p_{j'} = p_j + \Delta p, \quad x_{j'} = x_j - \Delta x,$$

for any $0 < \Delta p \leq \frac{p_i - p_j}{2}$ and $0 < \Delta x \leq \frac{x_j - x_i}{2}$. Then the external resilience of $(I \setminus \{i, j\}) \cup \{i', j'\}$ is weakly higher than the external resilience of I.

Proposition 4 is the first key result of the paper. The proof is in Appendix A and proceeds in three steps. Step 1 shows that the exchange from $\{i, j\}$ to $\{i', j'\}$ has two effects: (i) some best sub-coalitions of I end up with lower power and higher resources, and hence become less threatening; and (ii) the identity of best sub-coalitions may change, i.e., new best sub-coalitions may emerge and some previously best sub-coalitions may cease to be best. Step 2 shows that effect (i) cannot reduce external resilience. For effect (ii), note that before the exchange there must exist a sub-coalition that was (weakly) more threatening than any emerging best sub-coalition (i.e., it had weakly higher power and weakly lower resources). This is shown by contradiction: otherwise, the emerging best sub-coalition would already have been a best sub-coalition before the exchange. Hence, any newly emerging best sub-coalition cannot be more threatening than a previously best sub-coalition, and therefore cannot reduce external resilience. Moreover, if a previously best sub-coalition ceases to be best after the exchange, it cannot reduce external resilience by Proposition 3. Step 3 argues that these changes do not generate a profitable internal secession. Intuitively, once I can withstand internal secession (Condition (i) of Proposition 3) before the exchange, replacing some best sub-coalitions by a strictly less threatening set after the exchange cannot trigger internal secession.

Corollary 1 (Hierarchy)**.** *As an implication of Proposition 4, iterating the exchanges of power and resources depicted in Figure 6 yields "conditional equality" (or "conditional pro-*

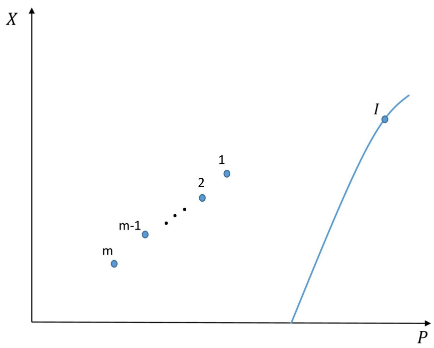


Figure 7: A coalition consisting of m ranks. Each blue dot represents a rank of players with identical power and resources and ranks are totally ordered by power P and resources X.

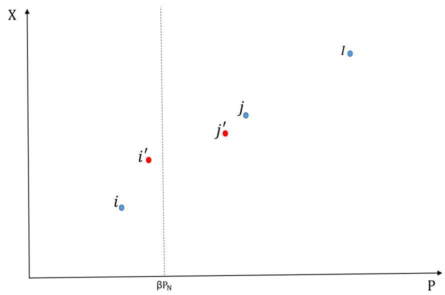


Figure 8: Ambiguity in comparing external resilience of proportional internal distributions.

portionality") within the ruling coalition. Coalition members are partitioned into classes: within each class, members have identical power and resources, and across classes, power and resources are proportional, with the highest class holding the most power and resources, followed by the second class, and so on. At each step, external resilience weakly increases. Consequently, the resulting conditionally proportional allocation has weakly higher external resilience than the initial allocation. In this sense, iterated exchanges produce a "hierarchy" (Figure 7).

It is important to note that comparing the external resilience of different "hierarchies"—i.e., different proportional configurations of power and resources within the ruling coalition—generally requires additional structure on the plundering function $G(\cdot)$. For instance, fix the total power and resources of a two-player ruling coalition I and consider two internal configurations, $\{i, j\}$ and $\{i', j'\}$, such that $P_I = p_i + p_j = p_{i'} + p_{j'}$ and $X_I = x_i + x_j = x_{i'} + x_{j'}$ (Figure 8). Then there is no general argument that ranks which configuration yields higher external resilience without further restrictions on $G(\cdot)$ (e.g., beyond concavity of indifference curves).

Remark 5 (A new perspective on hierarchy)**.** *Corollary 1 offers a perspective on why the most resilient plundering coalitions tend to exhibit hierarchical organization with well-defined ranks. Examples include stable oligarchies, Weberian bureaucracies, and armies.*

Remark 6 (Dynamic implications)**.** *It is natural to ask what happens after outsiders* *are plundered. Plundering lowers outsiders' resources and raises their power-to-resource ratios, potentially making some outsiders more attractive future coalition partners and thereby destabilizing the current ruling coalition. Our static game and resilience analysis can be interpreted as a benchmark for a repeated* stationary-bandit *environment: each season the ruling coalition extracts and consumes the extracted resources, while outsiders rebuild resources through production before the next season. Each season then starts from a recovered distribution $x(\cdot)$ and the same coalition-formation and extraction problem is played again. In such a setting, extraction can persist period after period as long as the ruler does not accumulate resources and preserves outsiders' productive capacity (e.g., does not burn the land or kill the farmers), unless an exogenous shock disrupts it. In this sense, our resilience object provides a simple benchmark for studying how extractive coalitions evolve over time in the presence of shocks.*

Remark 7 (Preference heterogeneity and hierarchy)**.** *Importantly, the logic of the proof of Proposition 4 continues to hold when preferences are heterogeneous within the ruling coalition. If an insider strictly prefers another ruling coalition, then the coalition's resilience is trivially zero, regardless of the external shock. Aside from this degenerate case, Proposition 4 extends under a weaker version of Assumption 2 that allows for preference heterogeneity. The key force in the proof remains unchanged: the exchange shifts certain best sub-coalitions up and left in (P, X)-space, which pushes down the relevant boundary curves. Under Assumption 1, for any insider (regardless of the curvature of her indifference curves), a sub-coalition that moves up/left becomes weakly more attractive after the exchange. Hence, the associated boundary for that insider shifts weakly downward; aggregating across insiders, this weakly expands the external safe area. Therefore, the exchange in Figure 6 enlarges the external safe area in the same direction as in the benchmark case of homogeneous preferences under Assumption 2.*

In Proposition 4, the exchange of power and resources *weakly* increases the external resilience of the ruling coalition. The next sections characterize conditions under which the exchange *strictly* increases external resilience, as well as conditions under which external resilience remains unchanged. In particular, we highlight the role of the shape of indifference curves—convex versus concave—which we interpret as capturing the strength of property-rights protection.

4.3 Power-intensive and power-light plundering

Concave and convex indifference curves capture a fundamental difference in how the marginal value of power varies with a coalition's power. Consider a small increase in a coalition's power. To keep insiders indifferent, by how much must the coalition's resources increase? Recall that insiders dislike resources held *inside* the coalition, since these resources are protected from plunder. Under concave indifference curves, when the coalition starts with a low level of power, it can tolerate a large increase in internal resources while remaining indifferent: the marginal value of power is high, so a small increase in power offsets a large increase in resources. When the coalition starts with a high level of power, it can tolerate only a small increase in resources while remaining indifferent: the marginal value of additional power is low. In other words, under concave indifference curves, the marginal value of power decreases as coalition power increases.

Under concave indifference curves, individuals thus have little appetite for a ruling coalition with very high aggregate power, i.e., for an "inclusive" ruling coalition that brings many players inside.[10] This case corresponds to a power-light plundering environment: additional power has sharply diminishing marginal value, so coalitions prefer to remain relatively exclusive to keep more resources outside and thereby increase the extractable pool.

By contrast, under convex indifference curves, as the power of the ruling coalition increases, the marginal value of additional power also increases. This induces the formation of a more "inclusive" ruling coalition with high aggregate power. For instance, when institutions constrain plundering, inclusive ruling coalitions may have an advantage because greater aggregate power helps overcome these constraints.[11] Accordingly, throughout the rest of the paper, we refer to concave and convex indifference curves as, respectively, power-light and power-intensive plundering environments.

The following proposition shows that, under convex indifference curves or power-intensive plundering environments, the internal configuration of power and resources may not affect external resilience.

Proposition 5 (Power-intensive plundering and invariance of external resilience)**.** *Suppose that preferences over coalitions (P, X) have strictly convex indifference curves. For any ruling coalition I and any $A \in \mathcal{A}_I \setminus \{I\}$, we have*

$$\mathcal{S}_I^{\text{ext}} \subseteq \mathcal{S}_A^{\text{ext}}.$$

Hence,

$$\mathcal{S}^{\text{ext}} := \bigcap_{A \in \mathcal{A}_I} \mathcal{S}_A^{\text{ext}} = \mathcal{S}_I^{\text{ext}}.$$

[10] This preference has an analogy in standard consumer theory: although a consumer prefers more of a good, diminishing marginal utility implies that extremely large amounts of the same good are not valuable at the margin.

[11] Example 3 in the appendix provides a more detailed discussion of power-intensive and power-light plundering environments.

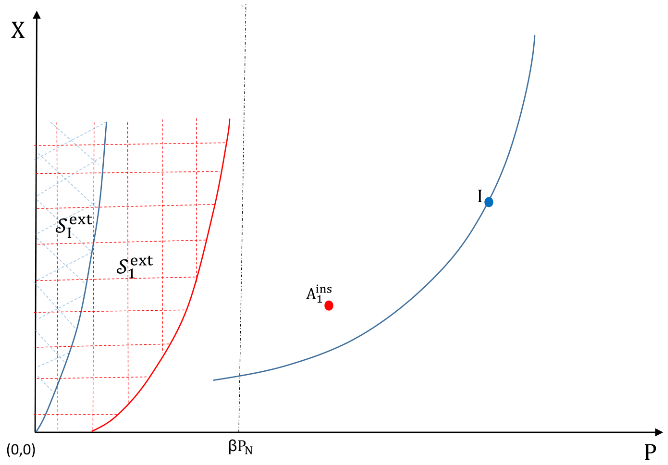


Figure 9: Comparing the external safe area for I and $A_1^{ins} \in \mathcal{A}_I \backslash I$ under convex indifference curves.

Therefore, any exchange of power and resources within I that preserves internal stability leaves the external resilience of I unchanged.

Proposition 5 is proved in Appendix A. Fix a ruling coalition I and $A_1^{\text{ins}} \in \mathcal{A}_I \setminus \{I\}$, a nontrivial best sub-coalition of I (an insider sub-coalition). Under strictly convex indifference curves (power-intensive plundering), insiders strictly prefer more inclusive ruling coalitions. Hence, for any best sub-coalition of outsiders $B \in \mathcal{A}_{N\setminus I}$, insiders in A_1^{ins} prefer $B \cup I$ to $B \cup A_1^{\text{ins}}$. Any profitable and feasible external deviation of the form $B \cup A_1^{\text{ins}}$ is therefore (weakly) dominated by the deviation $B \cup I$. Thus, the set of outsider coalitions that can induce secession with A_1^{ins} is contained in the set that can induce secession with I, which implies $\mathcal{S}_I^{\text{ext}} \subseteq \mathcal{S}_1^{\text{ext}}$, as illustrated in Figure 9. Since I is itself included among the insider best sub-coalitions, we obtain

$$\mathcal{S}^{\text{ext}} := \bigcap_{A_i^{\text{ins}} \in \mathcal{A}_I} \mathcal{S}_i^{\text{ext}} = \mathcal{S}_I^{\text{ext}}.$$

Therefore, as long as internal exchanges of power and resources do not trigger internal secession, they leave $\mathcal{S}^{\text{ext}}$—and hence external resilience—unchanged.

Remark 8. *Proposition 5 suggests that when property rights are better protected—even if plundering is not fully eliminated—a "specialized" coalition that separates political power from economic resources can be highly stable against external shocks.*

Next, we show that when the plundering technology is *sufficiently power-light* (i.e., indifference curves are sufficiently concave), external resilience *strictly* increases as we transition to a hierarchy with well-defined ranks by iterating the exchanges in Figure 6 within the ruling coalition.

To obtain a parametric measure of concavity, we focus on a CES family of plundering functions $\{G_\rho\}_{\rho \neq 0}$. Fix a ruling coalition I with aggregate power and resources (P_I, X_I).

For $\alpha \in (0,1)$ and $\rho \neq 0$, define

$$G_\rho(P, X) := \left[\alpha \left(\frac{P}{P_I}\right)^\rho + (1-\alpha)\left(\frac{\bar{X} - X}{\bar{X} - X_I}\right)^\rho\right]^{1/\rho},$$

where $\bar{X}$ is assumed to be sufficiently large (with $\bar{X} > X_I$) and $(P, X) \in \mathbb{R}_+ \times (0, \bar{X})$. For $\rho \leq 1$, G_ρ is concave in (P, X) (Figure 10), corresponding to increasingly power-light plundering as ρ decreases.[12]

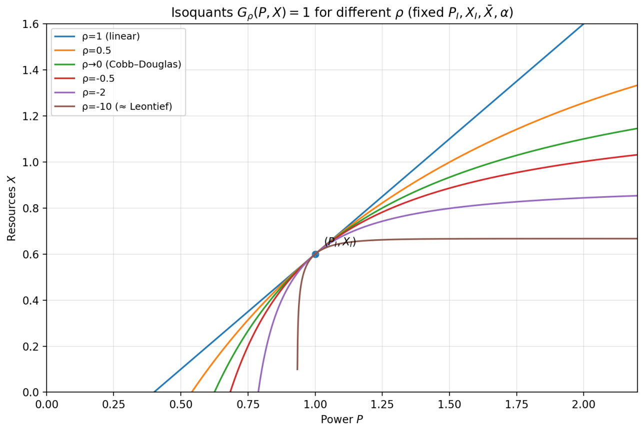


Figure 10: Isoquants of the CES plundering technology $G_\rho(P, X)$ for different values of ρ, normalized to pass through (P_I, X_I). Lower ρ implies a more concave/ Leontief-like shape.

The next proposition shows that for any initial distribution of power and resources within the ruling coalition, there exists an indifference curve passing through I that is sufficiently concave—i.e., a sufficiently small ρ in the CES family $\{G_\rho\}$—such that iterating the exchange in Figure 6 until the coalition becomes hierarchical *strictly* increases external resilience.[13]

Proposition 6 (Power-light plundering and strict gains from hierarchy)**.** *Suppose that I is a ruling coalition and preferences over coalitions (P, X) are given by $G_\rho(\cdot)$ for some $\rho < 1$. Further suppose that I' is the allocation obtained from I by iterating the bilateral exchanges in Figure 6 until the coalition becomes a hierarchy with well-defined ranks. Then there exists $\bar{\rho} < 1$ such that if $\rho \leq \bar{\rho}$,*

$$\mathcal{S}_I^{\text{ext}} \subsetneq \mathcal{S}_{I'}^{\text{ext}},$$

so reaching a hierarchical allocation under power-light plundering strictly increases external resilience.

[12] The elasticity of substitution between power P and the resource-loss slack $\bar{X} - X$ is $\sigma = \frac{1}{1-\rho}$. Thus, lower ρ implies lower σ, and the Leontief limit obtains as $\rho \to -\infty$.

[13] The logic is not specific to the CES class. We use this family only because it provides a transparent notion of "sufficient concavity" (via ρ) and allows for a rigorous and tractable argument.

The proof is in Appendix A. Consider the sequence of bilateral exchanges within I described in Figure 6, and suppose these exchanges are repeated until there remain two insiders $i, j \in I$ such that $p_i > p_j$ and $x_i < x_j$ and the exchange is still applicable. If the external safe area has already expanded at some earlier step, the result follows because Proposition 4 implies that the final exchange cannot shrink the external safe area. Thus, assume that external resilience has remained constant up to this point.

Denote the sub-coalition $A_1^{\text{ins}} := I \setminus \{j\}$ and define the shifted indifference curve through this sub-coalition as

$$H_I^1(P) := H_I\big(P + P_{A_1^{\text{ins}}}\big) - X_{A_1^{\text{ins}}}.$$

It is straightforward to show that A_1^{ins} is a best insider sub-coalition of I. After the exchange between i and j, A_1^{ins} moves left and up in (P, X)-space, e.g., to A_2^{ins} as in Figure 12. Under Assumption 1, for *any* indifference curve, a move from A_1^{ins} to A_2^{ins} strictly enlarges the corresponding external safe area, i.e.,

$$\mathcal{S}_1^{\text{ext}} \subsetneq \mathcal{S}_2^{\text{ext}}.$$

[14] Moreover, no sub-coalition that includes j but excludes i can be a best insider sub-coalition after the exchange. Otherwise, replacing j with i would yield a sub-coalition with weakly higher power and weakly lower resources (with at least one strict inequality), which is a contradiction.

Therefore, the external safe area can expand strictly at this step only if $H_I^1(P)$ is binding in the construction of $\mathcal{S}_I^{\text{ext}}$, i.e., only if $H_I^1(P)$ uniquely determines the upper envelope

$$\max_{A_i^{\text{ins}} \in \mathcal{A}_I} H_I^i(P)$$

for some $P \in (0, \beta P_N)$. The proof shows that for any given initial allocation of power and resources within the ruling coalition, there exists a sufficiently concave indifference curve H_I (equivalently, a sufficiently small ρ) such that the boundary $H_I^1(P)$ is *uniquely* binding at some $P \in (0, \beta P_N)$; that is, $H_I^1(P)$ uniquely attains $\max_{A_i^{\text{ins}} \in \mathcal{A}_I} H_I^i(P)$ at that P.

To shed further light on the geometry, consider the extreme case $\rho \to -\infty$, where the plundering technology becomes Leontief and the indifference curve in (P, X)-space is an inverse-L with a kink at (P_I, X_I) (Figure 11). In this limit, for any initial allocation of power and resources within I and any associated collection of best sub-coalitions $\mathcal{A}_I$, the boundary induced by each best sub-coalition (e.g., A and B in Figure 11) uniquely

[14]This is shown in the proof of Proposition 4.

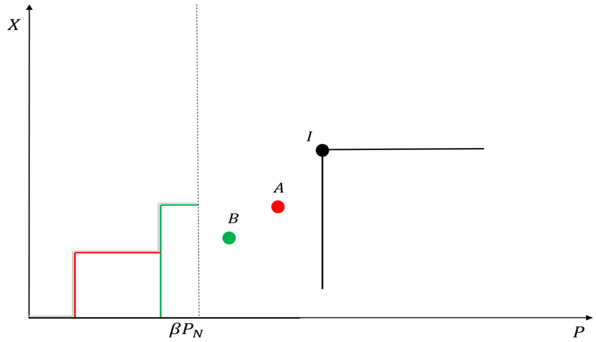


Figure 11: Under sufficiently concave (Leontief-like in limit) indifference curves, the boundary corresponding to a best sub-coalition is uniquely binding on the upper envelope that defines the external safe area.

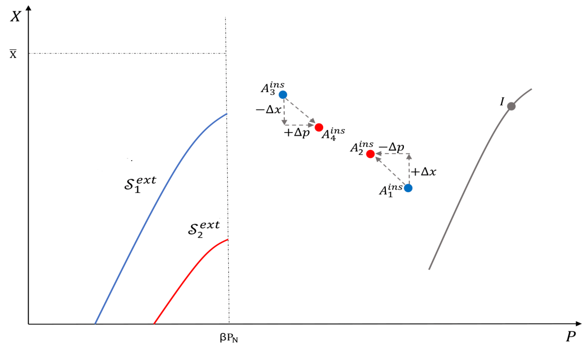


Figure 12: Under concave indifference curves, the exchange in Figure 6 expands the external safe area for best sub-coalitions that contain i (the member with higher power but lower resources than j).

determines the upper envelope

$$\max_{A_i^{\text{ins}} \in \mathcal{A}_I} H_I^i(P)$$

on a nonempty interval of $P \in (0, \beta P_N)$. For instance, as Figure 11 illustrates, the boundary of the external safe area of I (the grey envelope) is piecewise determined in an ordered way: the right segment (green) is pinned down by the least-powerful best sub-coalition B, the middle segment (red) by the next-most powerful best sub-coalition A, and the left segment (black) by I itself (viewed as a best sub-coalition of itself). Therefore, when ρ is sufficiently low—i.e., indifference curves are sufficiently concave—the exchange between i and j strictly expands the external safe area, and iterating exchanges until the coalition becomes hierarchical strictly increases external resilience under power-light plundering.

Remark 9 (Weak property rights and the emergence of hierarchy)**.** *Our analysis so far highlights a central insight: the* resilience advantage *of hierarchy is stronger when property rights are weak than when they are strong. Proposition 4 shows that moving the ruling coalition toward a hierarchical allocation never decreases external resilience, under* any *plundering function $G(\cdot,\cdot)$ satisfying Assumptions 1–3. We then contrast power-intensive and power-light plundering environments, interpreting convex versus concave indifference curves as capturing the strength of constraints on extraction (e.g., property-rights protection). As a hierarchy emerges, external resilience is unchanged under strong property-rights protection (power-intensive plundering; Proposition 5), but it* strictly *increases when property rights are sufficiently weak (power-light plundering; Proposition 6). Taken together, these results suggest that when constraints on plundering are weak, organizing the ruling coalition as a hierarchy can deliver a strict resilience gain against outsider threats, whereas this benefit is absent when property rights are well protected.*

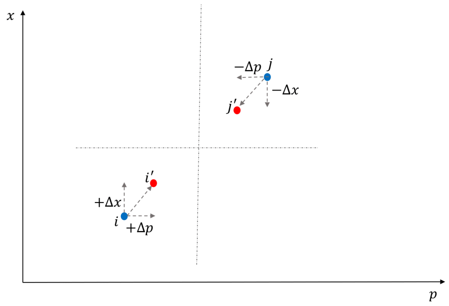


Figure 13: Exchange of power and resources between player i and player j—from blue to red.

4.3.1 Coalitions of absolutely equal members

It is important to note that the analysis above does not imply that a ruling coalition is, in general, most externally resilient when its members are absolutely equal. Without further restrictions on the plundering function and on the joint distribution of power and resources within the ruling coalition, no general ranking is possible. That said, an absolutely equal coalition does maximize external resilience in a particular case: when the most powerful member in the ruling coalition also has the lowest resources, the second most powerful has the second lowest resources, and so on. In that case, it is straightforward to implement a sequence of exchanges of the type in Figure 6 that converges to absolute equality.[15] The following examples illustrate a more general case in which the exchange of power and resources depicted in Figure 13 within the ruling coalition also increases external resilience. Together with Proposition 4, this generates a broad range of single-class coalitions.

Example 2. *Consider a ruling coalition $I = \{i, j\}$ with $p_i < p_j$ and $x_i < x_j$. Let $H_I(\cdot)$ denote the (concave) indifference curve through I. Suppose player i is more threatening than j, i.e., $\mathcal{S}_i^{\text{ext}} \subset \mathcal{S}_j^{\text{ext}}$. If we perform the exchange in Figure 13 between i and j and obtain $\mathcal{S}_i^{\text{ext}} \subset \mathcal{S}_{i'}^{\text{ext}}$ (the red region covers the blue region in Figure 14), then the external safe area of the ruling coalition expands under this exchange.*

Intuitively, this occurs because player i is much more threatening than player j given $H_I(\cdot)$. After the exchange, taking resources from player j may make her more threatening (moving from j to j'), but this effect is dominated by the reduction in the threat posed by player i when she becomes better endowed (moving from i to i'). This force is strongest under power-light plundering (high concavity), where low-resource insiders are especially threatening because they have high power-to-resource ratios.

Remark 10. *Although we do not provide a general ranking here, Example 1 suggests an important intuition: when property-rights protection is extremely weak and there exists a*

[15]More precisely, start by exchanging between the two most powerful players until they become equal in both power and resources; then continue with the third most powerful player until the top three become equal; and so forth.

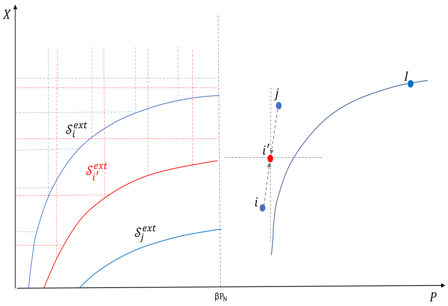


Figure 14: The increases in the external resilience due to an exchange of powers and resources as in Figure 12.

poor but sufficiently powerful group within the ruling coalition, a single-class coalition can be particularly externally resilient. This perspective may help rationalize the egalitarian thrust of some communist revolutions—but specifically in settings where the revolutionary base is both economically disadvantaged and sufficiently powerful (e.g., urban proletarians in some historical contexts).

4.4 Trade-off between internal and external resilience

We now link external and internal resilience by characterizing when a trade-off may arise between them.

4.4.1 Power-intensive plundering

We begin by showing that when plundering is power-intensive, a trade-off can arise between external and internal resilience as the plundering environment changes. Specifically, shifting toward more power-intensive plundering lowers external resilience but raises internal resilience. To formalize this comparative-static notion, we introduce a more general definition of the "intensity of plundering." If the marginal rate of substitution between power and resources is strictly higher, then the marginal value of adding power to the ruling coalition is higher relative to the marginal cost of retaining internal resources that cannot be plundered. This corresponds to a more power-intensive (i.e., less power-light) plundering environment.

Definition 6. *The indifference curve $H_I(\cdot)$ represents a more power-light plundering environment than $H_I'(\cdot)$ if and only if for all $P \in [\beta P_N, P_N]$,*

$$\frac{d}{dP} H_I(P) \; < \; \frac{d}{dP} H_I'(P),$$

denoted by $H \succ H'$.

Proposition 7. *Fix a ruling coalition I. Suppose the indifference curves $H_I(\cdot)$ and $H'_I(\cdot)$ are convex, and that internal and external shocks are independently distributed. Then $H \succ H'$ if and only if the internal resilience of I under H is lower than under H' and its external resilience under H is higher than under H'.*

In a power-intensive plundering environment, the most threatening external deviations are those that include outsiders *and* all members of the ruling coalition, rather than those formed by outsiders together with a non-trivial insider subset. The reason is that power-intensive plundering (convex indifference curves) makes aggregate power increasingly valuable at the margin. Holding the outsider block fixed, expanding the insider component from a proper subset $A^{\text{ins}} \subsetneq I$ to the full coalition I raises aggregate power, and under convexity this gain in power more than compensates for the fact that bringing in additional insiders also brings in protected internal resources that cannot be plundered. Hence, for any outsider best sub-coalition, deviations of the form $A^{\text{ext}} \cup I$ are weakly preferred to deviations of the form $A^{\text{ext}} \cup A^{\text{ins}}$, implying that the external-safe constraint is effectively pinned down by deviations that involve I itself.

This has a direct implication for resilience. As the environment becomes more power-intensive, the propensity toward inclusiveness strengthens, so it becomes easier for outsiders to construct a profitable deviation precisely of the most threatening form $A^{\text{ext}} \cup I$. The external safe area therefore shrinks, and external resilience falls. At the same time, the same force raises internal resilience: because aggregate power is increasingly valuable, insiders are less tempted by breakaway sub-coalitions with lower power, so internal secession is less attractive. Thus, power-intensive plundering induces a trade-off: external resilience decreases while internal resilience increases (Figure 15).

Remark 11. *This implies that when property rights are relatively well protected—so plundering is more power-intensive—shifting toward lower plundering intensity can reduce the likelihood of an insider coup d'etat while increasing the risk of a popular uprising, and vice versa. In this sense, the ruling coalition faces a trade-off between stabilizing power-sharing among insiders and maintaining authoritarian control against outsiders.*

4.4.2 Power-light plundering

As in the convex case, it is straightforward to show that moving toward less power-light plundering increases internal resilience when indifference curves are concave (Figure 16). However, how a change in plundering intensity affects external resilience is generally ambiguous in this region. Example 3 shows that a shift toward more power-intensive (i.e., less power-light) plundering can yield either higher or lower external resilience, depending on the nature of external perturbations.

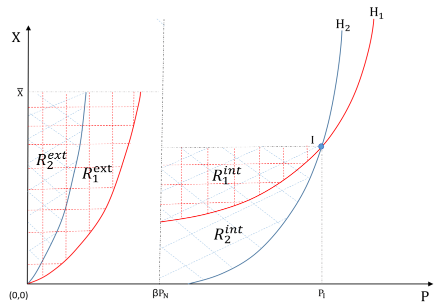


Figure 15: Trade-off between internal and external resilience under convex indifference curves.

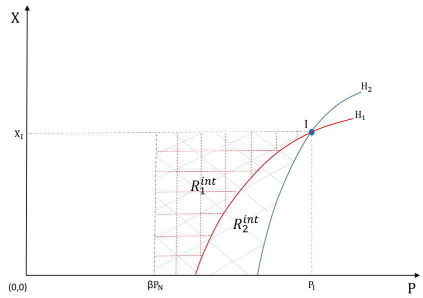


Figure 16: The increasing effect of a weaker plundering on internal resilience under concave indifference curves.

Example 3. *Consider a ruling coalition I and a plundering technology represented by the indifference curve H_1 passing through I. Suppose the plundering environment shifts toward more power-intensive plundering, moving from H_1 to H_2. Suppose further that there is a best insider sub-coalition $A_1^{\text{ins}} \in \mathcal{A}_I$ that is realized with probability one. The effect of the shift from H_1 to H_2 on external resilience is ambiguous (Figure 17): depending on the configuration of external perturbations, the external safe area associated with A_1^{ins} may expand or shrink.*

For instance, consider two distributions over external shocks, $Pr_1^{\text{ext}}(\cdot,\cdot)$ and $Pr_2^{\text{ext}}(\cdot,\cdot)$, where Pr_1^{ext} assigns relatively more probability mass to the outsider best sub-coalition A_2^{ext} than to A_1^{ext} (compared to Pr_2^{ext}), as illustrated in Figure 17. Since internal and external shocks are independent, under Pr_1^{ext} the coalition $A_2^{\text{ext}} \cup A_1^{\text{ins}}$ is more likely to arise than $A_1^{\text{ext}} \cup A_1^{\text{ins}}$. Suppose that moving from H_1 to H_2 makes $A_1^{\text{ext}} \cup A_1^{\text{ins}}$ newly preferred to I (while it is not preferred under H_1). Then external resilience falls under Pr_1^{ext} relative to Pr_2^{ext}, because the newly dangerous deviation receives more probability weight.

Alternatively, suppose Pr_2^{ext} assigns relatively more probability mass to A_1^{ext} than to A_2^{ext} (compared to Pr_1^{ext}), so that $A_1^{\text{ext}} \cup A_1^{\text{ins}}$ is more likely to arise than $A_2^{\text{ext}} \cup A_1^{\text{ins}}$ under $Pr_2^{\text{ext}}(\cdot,\cdot)$. Suppose further that moving from H_1 to H_2 makes $A_2^{\text{ext}} \cup A_1^{\text{ins}}$ no longer preferred to I (while it is *preferred under H_1). Then external resilience falls under Pr_2^{ext} relative to Pr_1^{ext}, because the deviation that becomes newly safe receives more probability weight under Pr_2^{ext}.*

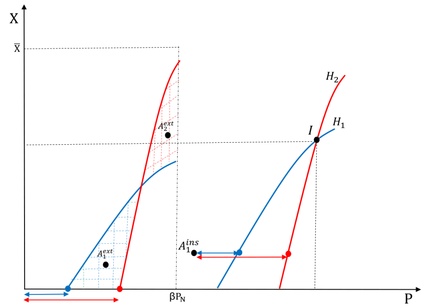


Figure 17: The ambiguous effect of a weaker plundering on external resilience.

Thus, the effect of shifting from H_1 to H_2 on external resilience depends on how the change in the indifference curve reclassifies which outsider coalitions are dangerous, and on where the external-shock distribution places probability mass. In particular, the trade-off with internal resilience need not hold: it is possible that moving from H_1 to H_2 increases both internal and external resilience of the ruling coalition.

Although a general argument cannot be made, the above examples suggest that when external shocks favor coalitions with low power and resources, weaker plundering might lead to both higher external and internal resilience. This occurs because weaker plundering increases the tendency to form an inclusive ruling coalition. External sub-coalitions with low power and resources cannot typically form inclusive coalitions. Therefore, if the distribution of external shocks gives a higher probability to these external coalitions, external resilience would increase. This implies that if lower classes with low political power are more likely to emerge in society, a weaker plundering process might reduce both the threat of a *coup d'etat* and a popular uprising.

Remark 12. *The above discussion suggests that even imperfect property rights, which do not entirely prevent exploitation by insiders, can impede the ruling coalition from attaining both internal and external stability when plundering technology changes. Moreover, a ruling coalition that engages in power-light plundering may prefer to confront a more powerful and richer opposition compared to a weaker but poorer one.*

5 Concluding remarks

This paper develops a theory of plundering coalitions. We conclude by commenting on further applications and directions for future work. Our theory shows that the ruling coalition has a relatively higher power-to-resources ratio than its alternatives. This result formalizes the concept of *Asabiyyah*, usually translated as "social cohesion," which is a central concept for understanding political dynamics in the Middle East (Khaldun (1967); Kuran (1996); Alatas (2014)). Specifically, the great historian and sociologist

Ibn Khaldun argued that nomadic tribes had a much higher level of "social cohesion" than urban civilizations, and that this strong social cohesion facilitated the conquest of urban civilizations by nomadic tribes. Our model microfounds the higher social cohesion of nomadic tribes through their relative poverty compared to urban civilizations, which generates a high power-to-resources ratio. It is therefore easier for nomadic tribes to form a coalition to plunder cities, which is a repeated pattern in the pre-modern world. Similar logic can apply to communist revolutions (Morishima (1974); Roemer (1980); Roemer (1981); Brewer (2002)), where increasing inequality widens the power-to-resources gap, thereby incentivizing the proletariat to rebel against the capitalists. Importantly, our analysis may provide a clue to understanding the oligarchic tendencies of these plundering coalitions. Our model may explain why successful nomadic conquerors and communist parties, even when starting as movements of radical equality, eventually evolved into strictly hierarchical structures.

In future research, our framework may be useful as a methodological approach for studying the resilience of coalitions to exogenous changes in players' characteristics and in the environment governing coalition formation. Moreover, although we study how coalitions respond to changes in power, resources, and plundering technology, these objects are exogenous in our model. Endogenizing them could be informative and suggests several extensions. A natural extension is to allow players to invest in power prior to coalition formation, which would clarify how the initial distribution of resources shapes power investment incentives and, ultimately, the ruling coalition. Another extension is to study environments in which the ruling coalition receives an exogenous flow of resources in addition to exploiting outsiders.

Another extension is to endogenize the plundering technology—interpreted as property-rights protection—in a dynamic version of our framework in which ruling coalitions can invest in institutions over time. This would contribute to the literature on the emergence and evolution of property-rights protection (Andolfatto (2002); Hafer (2006); Diermeier et al. (2017)) from the perspective of resilience. Moreover, institutions are persistent in many settings, so early institutional choices can have long-lasting effects on subsequent political and economic outcomes (Persson (2002); Michalopoulos and Papaioannou (2013); Lowes et al. (2017)). Finally, incorporating networks into the coalition-formation process is another promising direction. For example, König et al. (2017) studies how a network of military alliances affects conflict intensity. Extending our model along these lines could clarify how players' connections shape political alliances and their resilience.

Appendix A Proofs

Proof of Lemma 1. Fix $u \in \mathbb{R}$. Player i's indifference curve at utility level u is the level set

$$I(u) := \{(P, X) \in \mathbb{R}^2 : G_i(P, X) = u\}.$$

To see that $I(u)$ is closed, take any sequence $\{(P_k, X_k)\}_{k\geq 1} \subset I(u)$ with $(P_k, X_k) \to (P, X)$. By continuity of G_i,

$$G_i(P, X) = \lim_{k\to\infty} G_i(P_k, X_k) = \lim_{k\to\infty} u = u,$$

so $(P, X) \in I(u)$.

Next, we show that indifference curves are strictly increasing. Let $P' > P$ and $X' < X$. By Assumption 1(1a), $G_i(P', X') > G_i(P, X')$, and by Assumption 1(1b), $G_i(P, X') > G_i(P, X)$. Hence $G_i(P', X') > G_i(P, X)$.

Suppose, toward a contradiction, that $I(u)$ is not strictly increasing. Then there exist $(P_1, X_1), (P_2, X_2) \in I(u)$ with $P_2 > P_1$ and $X_2 \leq X_1$. If $X_2 = X_1$, then Assumption 1(1a) implies $G_i(P_2, X_2) = G_i(P_2, X_1) > G_i(P_1, X_1)$, contradicting $G_i(P_2, X_2) = G_i(P_1, X_1) = u$. If $X_2 < X_1$, then by the monotonicity implication above, $G_i(P_2, X_2) > G_i(P_1, X_1)$, again a contradiction. Therefore, along any indifference curve, an increase in power must be accompanied by an increase in resources, so the curve is strictly increasing. □

Proof of Proposition 1. Define, for each $I \in \mathcal{W}$,

$$\phi(I) := \arg\max_{W\in\mathcal{W}} G(W).$$

Since $\mathcal{W}$ is finite and non-empty, $\arg\max_{W\in\mathcal{W}} G(W)$ is well-defined and non-empty, so $\phi(I) \neq \emptyset$, establishing the first part of Axiom 1. Moreover, by construction $\phi(I) \subseteq \mathcal{W}$, so Axiom 2 holds.

If $I' \in \phi(I)$, then $I' \in \arg\max_{W\in\mathcal{W}} G(W)$, so for any $I'' \notin \phi(I)$ we have $G(I'') < G(I')$. Conversely, if $G(I'') < G(I')$, then $I'' \notin \arg\max_{W\in\mathcal{W}} G(W)$, hence $I'' \notin \phi(I)$. This is exactly Axiom 3. Finally, Assumption 1 implies $G(N) = 0$, while for any $W \in \mathcal{W} \setminus \{N\}$ we have $G(W) > 0$. Hence N cannot be a maximizer of G over $\mathcal{W}$, so $N \notin \phi(I)$ and the second part of Axiom 1 holds. This proves existence.

For uniqueness, suppose there is another mapping ϕ' satisfying Axioms 1–3. Fix $I \in \mathcal{W}$ and take $I'' \in \phi'(I)$. If $I'' \notin \arg\max_{W\in\mathcal{W}} G(W)$, let $I' \in \arg\max_{W\in\mathcal{W}} G(W)$ (non-empty by the argument above). Then $G(I'') < G(I')$, so by Axiom 3 for ϕ' we must have $I'' \notin \phi'(I)$, a contradiction. Hence $\phi'(I) \subseteq \arg\max_{W\in\mathcal{W}} G(W)$, i.e. $\phi'(I) \subseteq \phi(I)$.

Conversely, suppose $I' \in \phi(I)$ but $I' \notin \phi'(I)$. By Axiom 1, there exists $I'' \in \phi'(I)$.

Since $I'' \in \phi'(I)$ and $I' \notin \phi'(I)$, Axiom 3 implies $G(I'') > G(I')$, which contradicts $I' \in \phi(I) = \arg\max_{W \in \mathcal{W}} G(W)$. Therefore $\phi(I) \subseteq \phi'(I)$, and thus $\phi = \phi'$.

For the second statement, Assumption 3 implies that G takes distinct values on distinct coalitions in $\mathcal{W}$, so $\arg\max_{W \in \mathcal{W}} G(W)$ is a singleton. Hence ϕ is single-valued. This completes the proof. □

Proof of Proposition 2(1). Fix $I_0 \in \mathcal{W}$ and assume $\beta > \frac{1}{2}$. Let $C := \phi(I_0) = \arg\max_{W \in \mathcal{W}} G(W)$ (Proposition 1). Fix any $I \in C$.

Claim 1 (overlap under super-majority). If $\beta > \frac{1}{2}$, then for any $W, W' \in \mathcal{W}$ we have $W \cap W' \neq \emptyset$.

Proof. If $W \cap W' = \emptyset$, then $P_{W \cup W'} = P_W + P_{W'} \geq 2\beta P_N > P_N$, contradicting $P_{W \cup W'} \leq P_N$. ■

Step 1: primitives for continuation values. For any $W \subseteq N$ and $i \in N$, define the *net* payoff component

$$\mathcal{G}_i(W) := \begin{cases} g(i)\, G(W) & \text{if } i \in W, \\ 0 & \text{if } i \notin W, \end{cases}$$

so that $U_i(W) = x_i + \mathcal{G}_i(W)$.

For each agenda-setter $a \in I_0$, fix an arbitrary selection

$$\psi(a) \in \arg\max_{W \in \mathcal{W}:\ a \in W} G(W).$$

This is well-defined because $I_0 \in \mathcal{W}$ and $a \in I_0$.

At any history h, let $A^-(h)$ be the set of agenda-setters already used (including the current one if h is a voting node on the current proposal), and let

$$R(h) := I_0 \setminus A^-(h)$$

be the set of remaining agenda-setters. For any $i \in N$ and any $R \subseteq I_0$, define

$$m_i(R) := \begin{cases} \max_{a \in R}\ \mathcal{G}_i(\psi(a)) & \text{if } R \neq \emptyset, \\ \mathcal{G}_i(I_0) & \text{if } R = \emptyset. \end{cases}$$

Thus $m_i(R)$ is the highest net payoff i can obtain in continuation equilibrium after rejecting the current proposal, given that only agenda-setters in R remain (and the terminal fallback is I_0 when $R = \emptyset$).

Step 2: define a strategy profile σ^I.

Agenda setting. At any agenda-setting history h with agenda-setter $a = a(h) \in I_0$,

$$\mathcal{P}^a(h) = \begin{cases} I & \text{if } a \in I, \\ \psi(a) & \text{if } a \notin I. \end{cases}$$

Voting. Consider any voting history h on a proposal $\mathcal{P}$ with voter $v \in \mathcal{P}$, and let $R = R(h)$ be the remaining agenda-setters *after* the current agenda-setter (i.e., the agenda-setter of $\mathcal{P}$ is already in $A^-(h)$).

- If $\mathcal{P} = I$, then v votes YES.
- If $\mathcal{P} \neq I$ and $R \cap I \neq \emptyset$, then any $v \in I \cap \mathcal{P}$ votes NO (and voters not in I vote arbitrarily, say YES).
- If $\mathcal{P} \neq I$ and $R \cap I = \emptyset$, then v votes YES iff $\mathcal{G}_v(\mathcal{P}) \geq m_v(R)$, and NO otherwise.

Step 3: outcome on the equilibrium path. By Claim 1, $I \cap I_0 \neq \emptyset$ because both I and I_0 are in $\mathcal{W}$. Hence with positive probability Nature selects an agenda-setter in $I \cap I_0$; under σ^I such a proposer offers I, all members of I vote YES on I, and the game ends with ruling coalition I. Moreover, if an agenda-setter $a \notin I$ is selected before that happens, then $\mathcal{P} = \psi(a) \in \mathcal{W}$ (it must reach voting), and by Claim 1 we have $\psi(a) \cap I \neq \emptyset$. While $R \cap I \neq \emptyset$, some voter in $I \cap \psi(a)$ vetoes, so $\psi(a)$ cannot form. Therefore, before all agenda-setters in $I \cap I_0$ are exhausted, the only coalition that can possibly form is I.

Step 4: sequential rationality.

(a) Subgames with $R \cap I = \emptyset$. In such a subgame, no future agenda-setter belongs to I. Voting rules are exactly the standard acceptance rule relative to the continuation value $m_v(R)$, and each remaining agenda-setter $a \in R \subseteq I_0$ proposes $\psi(a)$, which maximizes $G(\cdot)$ (and hence maximizes her payoff) among winning coalitions containing a. Backward induction on the finite number of remaining agenda-setters implies that, in every subgame with $R \cap I = \emptyset$, the continuation strategies form an SPE.

(b) Subgames with $R \cap I \neq \emptyset$. Fix any such subgame and consider any voting history on a proposal $\mathcal{P} \neq I$. Because $\mathcal{P}$ reaches voting it must be in $\mathcal{W}$, so Claim 1 implies $\mathcal{P} \cap I \neq \emptyset$. Let $v \in \mathcal{P} \cap I$. Since $R \cap I \neq \emptyset$, v votes NO under σ^I, so $\mathcal{P}$ is rejected. Hence, as long as $R \cap I \neq \emptyset$, no proposal different from I can ever form.

Now consider a member $v \in I$ at a voting node on $\mathcal{P} \neq I$ with $R \cap I \neq \emptyset$. If v deviates to YES, the proposal still cannot form unless *all* insiders in $\mathcal{P} \cap I$ also vote YES; but at least one insider votes NO by strategy, so v is not pivotal and cannot gain. At a voting node on $\mathcal{P} = I$, voting YES yields $U_v(I)$ immediately. Deviating to NO delays the game; since $R \cap I \neq \emptyset$, some future agenda-setter in $I \cap I_0$ will propose I again and it will be accepted, so deviation cannot improve the payoff.

Finally, consider an agenda-setting node with $R \cap I \neq \emptyset$ and proposer $a \in I_0$. If $a \in I$, proposing I yields $U_a(I)$ immediately. Any deviation to $\mathcal{P} \neq I$ is rejected by an insider veto (as argued above), and the continuation still leads to I, so deviation cannot improve a's payoff. If $a \notin I$, then any proposal that reaches voting must lie in $\mathcal{W}$ and hence overlaps I by Claim 1, so it is vetoed while $R \cap I \neq \emptyset$; proposing $\psi(a)$ is therefore (weakly) optimal since no deviation can change the continuation outcome I.

Combining (a) and (b), σ^I is sequentially rational in every subgame; hence it is an SPE. By Step 3, this SPE produces I as the ruling coalition. This proves Proposition 2.1.

□

Proof of Proposition 2(2). Assume $\beta \in (\frac{1}{2}, 1]$ and $\phi(I_0) = \{I\}$. By Proposition 1 and Assumption 3, I is the unique maximizer of G over $\mathcal{W}$, i.e.,

$$G(I) > G(W) \quad \text{for all } W \in \mathcal{W} \setminus \{I\}. \tag{A.1}$$

For any $i \in I$ and any $W \in \mathcal{W}$ with $i \in W$, we have $U_i(W) = x_i + w_i(W)$ and (by the payoff structure) $w_i(W) = g(i)\, G(W)$, so (A.1) implies

$$U_i(I) > U_i(W) \quad \text{for all } i \in I \text{ and all } W \in \mathcal{W} \setminus \{I\} \text{ with } i \in W. \tag{A.2}$$

First, if $\beta > \frac{1}{2}$, then any two winning coalitions intersect: indeed, if $W, W' \in \mathcal{W}$ and $W \cap W' = \emptyset$, then $P_{W \cup W'} = P_W + P_{W'} \geq 2\beta P_N > P_N$, a contradiction.

For any history h, let $R(h) \subseteq I_0$ be the set of remaining agenda-setters at h (those not yet removed after having their proposal rejected), and define $R_I(h) := R(h) \cap I$ and $k(h) := |R_I(h)|$. We prove by induction on $k(h)$ that in any subgame starting at h with $k(h) \geq 1$, every SPE yields ruling coalition I.

If $k(h) = 1$, let a be the unique remaining agenda-setter in $R_I(h)$ and consider any SPE of the subgame at h. If a proposes I, then when members of I are called to vote on I, voting YES is weakly optimal: a unilateral NO rejects I and moves to a continuation in which no agenda-setter from I remains, so the eventual ruling coalition (if any forms) must be some $J \neq I$; for any voter $i \in I$, whenever $i \in J$ we have $U_i(I) > U_i(J)$ by (A.2), and whenever $i \notin J$ we have $U_i(J) = 0 < U_i(I)$. Hence all members of I vote YES and I is accepted. If instead a proposes some $W \neq I$, then either W is accepted, in which case a strictly prefers deviating to proposing I by (A.2), or W is rejected, in which case deviating to proposing I yields immediate acceptance and payoff $U_a(I)$. Thus, in any

SPE, a proposes I and I is accepted, so the ruling coalition is I.

Now fix $k \geq 2$ and assume the statement holds for all smaller values. Consider a subgame starting at some history h with $k(h) = k$, and fix any SPE of this subgame. We claim that no winning coalition $W \neq I$ can be accepted along the induced play. Suppose, for a contradiction, that some $W \in \mathcal{W}$ with $W \neq I$ is accepted at some history $\tilde{h}$. By the overlap property, $W \cap I \neq \emptyset$. Let $i \in W \cap I$ be the first member of $W \cap I$ (under the realized Nature order) who is called to vote on W along this history. Since W is accepted, at the moment i is called, voting NO rejects W. If i deviates to NO, the game continues with the current proposer removed from $R(\cdot)$, so the continuation starts at a history h' with $k(h') \geq k - 1 \geq 1$. By the induction hypothesis, every SPE of the continuation subgame yields ruling coalition I. Thus the deviation yields outcome I, while not deviating yields $W \neq I$. Since $i \in I \cap W$, (A.2) gives $U_i(I) > U_i(W)$, so the deviation is profitable, a contradiction. Therefore no winning $W \neq I$ can be accepted.

It follows that along any SPE play from h, either proposals are rejected until some agenda-setter in I proposes I, or I is proposed immediately. When I is proposed, the same voting argument as above implies it is accepted. Hence the ruling coalition is I in any SPE of the subgame at h.

Finally, since $I_0 \in \mathcal{W}$ and $I \in \mathcal{W}$, the overlap property implies $I_0 \cap I \neq \emptyset$, so at the initial history h_0 we have $k(h_0) = |I_0 \cap I| \geq 1$. Applying the induction result at h_0 yields that in any SPE of the full game the ruling coalition is I. Payoffs are then $U_i(I) = x_i + w_i(I)$ for all $i \in N$ by definition. □

Proof of Proposition 3. Fix $\Gamma = (I_0, p(\cdot), x(\cdot), \{U_i(\cdot)\}, \beta)$ and suppose Assumptions 1–3 hold. Recall from Proposition 1 that

$$\phi(I_0) = \arg\max_{W \in \mathcal{W}} G(W),$$

and by Assumption 3 the maximizer is unique whenever it exists. Also since $\beta > \frac{1}{2}$, any two winning coalitions intersect.

Only if. Suppose $\phi(I_0) = \{I\}$. Then $I \in \mathcal{W}$ and for every $W \in \mathcal{W} \setminus \{I\}$ we have $G(I) > G(W)$. In particular, for every $A^{ins} \in (\mathcal{A}_I \setminus \{I\}) \cap \mathcal{W}$ we have $G(I) > G(A^{ins})$, which is (i). Also, for every $A^{ext} \in \mathcal{A}_{N \setminus I}$ and every $A^{ins} \in \mathcal{A}_I$ such that $A^{ins} \cup A^{ext} \in \mathcal{W}$, we have $G(I) > G(A^{ins} \cup A^{ext})$, which is (ii).

If. Now suppose $I \in \mathcal{W}$ and conditions (i)–(ii) hold. We show that $G(I) > G(W)$ for every $W \in \mathcal{W} \setminus \{I\}$, implying $\phi(I_0) = \{I\}$.

Take any $W \in \mathcal{W}$ with $W \neq I$. There are two cases.

Case 1: $W \subseteq I$**.** If $W \in \mathcal{A}_I \setminus \{I\}$, then $W \in (\mathcal{A}_I \setminus \{I\}) \cap \mathcal{W}$ and condition (i) gives

$G(I) > G(W)$.

If $W \notin \mathcal{A}_I$, then by Definition 3 there exists $A^{ins} \in \mathcal{A}_I$ such that $P_{A^{ins}} > P_W$ and $X_{A^{ins}} < X_W$. Since $W \in \mathcal{W}$ and $P_{A^{ins}} > P_W \geq \beta P_N$, we have $A^{ins} \in \mathcal{W}$. By Assumption 1, $(P_{A^{ins}}, X_{A^{ins}})$ strictly dominates (P_W, X_W), so $G(A^{ins}) > G(W)$. If $A^{ins} = I$, then $G(I) > G(W)$ directly. If $A^{ins} \neq I$, then $A^{ins} \in (\mathcal{A}_I \setminus \{I\}) \cap \mathcal{W}$ and (i) implies $G(I) > G(A^{ins}) > G(W)$.

Case 2: $W \not\subseteq I$. Write $W = W^{ins} \cup W^{ext}$ where $W^{ins} := W \cap I$ and $W^{ext} := W \setminus I$. Because W and I are both winning and $\beta > \frac{1}{2}$, the overlap fact implies $W^{ins} \neq \emptyset$, and by assumption $W^{ext} \neq \emptyset$.

Choose $A^{ins} \in \mathcal{A}_I$ such that either $A^{ins} = W^{ins}$ if $W^{ins} \in \mathcal{A}_I$, or else $P_{A^{ins}} > P_{W^{ins}}$ and $X_{A^{ins}} < X_{W^{ins}}$ (guaranteed by Definition 3). Similarly, choose $A^{ext} \in \mathcal{A}_{N\setminus I}$ such that either $A^{ext} = W^{ext}$ if $W^{ext} \in \mathcal{A}_{N\setminus I}$, or else $P_{A^{ext}} > P_{W^{ext}}$ and $X_{A^{ext}} < X_{W^{ext}}$.

Let $\widetilde{W} := A^{ins} \cup A^{ext}$. By construction,

$$P_{\widetilde{W}} \geq P_W \geq \beta P_N \qquad \text{and} \qquad X_{\widetilde{W}} \leq X_W,$$

with at least one inequality strict (since $W \neq I$ and at least one of the two parts is strictly improved unless both parts are already best and equal). Hence $\widetilde{W} \in \mathcal{W}$ and Assumption 1 implies $G(\widetilde{W}) \geq G(W)$, with strict inequality whenever at least one part was strictly improved. In any case, condition (ii) applies to (A^{ins}, A^{ext}) (since $\widetilde{W} \in \mathcal{W}$) and yields $G(I) > G(\widetilde{W}) \geq G(W)$, hence $G(I) > G(W)$.

Since $G(I) > G(W)$ for all $W \in \mathcal{W} \setminus \{I\}$, I is the unique maximizer of G on $\mathcal{W}$, i.e. $\phi(I_0) = \{I\}$. This completes the proof. □

Proof of Proposition 4. Fix a ruling coalition I and two members $i, j \in I$ with $p_i > p_j$ and $x_i < x_j$. Let the post-exchange characteristics be

$$p_{i'} = p_i - \Delta p, \quad x_{i'} = x_i + \Delta x, \qquad p_{j'} = p_j + \Delta p, \quad x_{j'} = x_j - \Delta x,$$

where $0 < \Delta p \leq (p_i - p_j)/2$ and $0 < \Delta x \leq (x_j - x_i)/2$. Hence $p_{i'} \geq p_{j'}$ and $x_{i'} \leq x_{j'}$, and the aggregate characteristics of the ruling coalition are unchanged:

$$P_I = P_{I'} \quad \text{and} \quad X_I = X_{I'}.$$

Let $H_I(\cdot)$ denote the indifference curve through (P_I, X_I).

Step 1 (best insider coalitions and what moves). Before the exchange, no best insider sub-coalition can contain j but not i. Indeed, if $A \subseteq I$ contains j and excludes i, then replacing j by i yields a coalition with strictly higher power and strictly lower internal resources, contradicting Definition 3. After the exchange, the same conclusion holds whenever $(p_{i'}, x_{i'}) \neq (p_{j'}, x_{j'})$; on the knife-edge case $(p_{i'}, x_{i'}) = (p_{j'}, x_{j'})$, any best set can be chosen so that no best coalition contains j' without i' (since swapping produces the same (P, X)).

Therefore, both before and after the exchange, every best insider sub-coalition is of one of the following types:

1. contains i but not j;
2. contains both i and j;
3. contains neither i nor j.

Only type-(1) coalitions move. Specifically, if A is type-(1), then after the exchange it becomes

$$(P'_A, X'_A) = (P_A - \Delta p,\ X_A + \Delta x),$$

while type-(2) and type-(3) coalitions do not change.

Step 2 (weakening an insider coalition expands its external safe area). For any insider coalition $A \subseteq I$, define its shifted boundary

$$H_A(P) := H_I(P + P_A) - X_A,$$

and the associated external safe area

$$\mathcal{S}_A^{\text{ext}} := \big\{(P, X) \in \mathbb{R}^2_{++} :\ P < \beta P_N,\ X > H_A(P)\big\}.$$

(Equivalently, $(P, X) = (P_{A^{ext}}, X_{A^{ext}}) \in \mathcal{S}_A^{\text{ext}}$ iff $G(I) > G(A \cup A^{ext})$.)

Suppose A is replaced by another insider coalition $\widetilde{A}$ with $P_{\widetilde{A}} \leq P_A$ and $X_{\widetilde{A}} \geq X_A$ (at least one inequality weakly strict). Since H_I is strictly increasing in P (Lemma 1), for every P,

$$H_{\widetilde{A}}(P) = H_I(P + P_{\widetilde{A}}) - X_{\widetilde{A}} \leq H_I(P + P_A) - X_A = H_A(P).$$

Hence $\mathcal{S}_A^{\text{ext}} \subseteq \mathcal{S}_{\widetilde{A}}^{\text{ext}}$ (with strict inclusion if at least one inequality is strict). In particular, every type-(1) best insider coalition becomes weaker and more resource-heavy after the exchange, so its associated external safe area weakly expands; type-(2) and type-(3) safe areas are unchanged.

Step 3 (intersection, new best coalitions, and internal stability). External resilience is the intersection of the external safe areas across best insider coalitions:

$$\mathcal{S}_I^{\text{ext}} = \bigcap_{A \in \mathcal{A}_I} \mathcal{S}_A^{\text{ext}}, \qquad \mathcal{S}_{I'}^{\text{ext}} = \bigcap_{A \in \mathcal{A}_{I'}} \mathcal{S}_A^{\text{ext}}.$$

Consider the change from $\mathcal{A}_I$ to $\mathcal{A}_{I'}$.

(a) Coalitions that remain best. For every $A \in \mathcal{A}_I \cap \mathcal{A}_{I'}$, Step 2 implies $\mathcal{S}_A^{\text{ext}}$ weakly expands if A is type-(1) and is unchanged otherwise.

(b) Coalitions that cease to be best. If some $A \in \mathcal{A}_I$ is no longer best after the exchange, then its safe area is removed from the intersection, which can only weakly enlarge the intersection.

(c) Newly best coalitions. Let $C \in \mathcal{A}_{I'} \setminus \mathcal{A}_I$ be a newly best insider coalition. We claim C cannot be type-(1). Indeed, every type-(1) coalition shifts by the same vector $(-\Delta p, +\Delta x)$, so dominance relations among type-(1) coalitions are preserved; moreover, relative to any coalition that does not move, a type-(1) coalition becomes weakly *less* powerful and weakly *more* resource-heavy. Hence a type-(1) coalition cannot become newly undominated. Therefore C is type-(2) or type-(3), so it does not move.

Since C was not best before the exchange, there exists at least one *pre-exchange* best coalition $A \in \mathcal{A}_I$ that strictly dominates it:

$$P_A > P_C, \qquad X_A < X_C.$$

Necessarily such an A must be type-(1) (otherwise the dominance would persist after the exchange and C could not become best). By Step 2 applied to (A, C) we have $\mathcal{S}_A^{\text{ext}} \subseteq \mathcal{S}_C^{\text{ext}}$. Since $\mathcal{S}_I^{\text{ext}} \subseteq \mathcal{S}_A^{\text{ext}}$, it follows that $\mathcal{S}_I^{\text{ext}} \subseteq \mathcal{S}_C^{\text{ext}}$. Thus adding C to the intersection cannot exclude any point that was externally safe before.

Putting (a)–(c) together, we obtain

$$\mathcal{S}_I^{\text{ext}} \subseteq \mathcal{S}_{I'}^{\text{ext}},$$

so external resilience weakly increases.

Internal stability. Because (P_I, X_I) is unchanged, the internal safe region $\{(P, X) : X > H_I(P)\}$ is unchanged. Any insider coalition that moves does so left/up, making $X > H_I(P)$ easier to satisfy; coalitions that do not move are unaffected. Hence no new profitable internal secession is created, so internal stability is preserved.

Therefore, the exchange weakly increases the external resilience of the ruling coalition, proving the proposition. □

Proof of Proposition 5. Fix a ruling coalition I and let $H_I(\cdot)$ be the indifference curve through (P_I, X_I), so $X_I = H_I(P_I)$. Take any non-trivial best insider sub-coalition $A^{\text{ins}} \in \mathcal{A}_I \setminus \{I\}$ and write $(P_A, X_A) := (P_{A^{\text{ins}}}, X_{A^{\text{ins}}})$.

Since I is a ruling coalition, internal stability implies $G(I) > G(A^{\text{ins}})$, equivalently

$$X_A > H_I(P_A). \tag{A.3}$$

For any $K \subseteq I$, define the translated boundary and external safe region

$$H_I^K(P) := H_I(P + P_K) - X_K, \qquad \mathcal{S}_K^{\text{ext}} := \{(P, X) \in \mathbb{R}^2_{++} : \; P < \beta P_N, \; X > H_I^K(P)\}.$$

In particular, $H_I^I(P) = H_I(P + P_I) - X_I$ and $H_I^A(P) = H_I(P + P_A) - X_A$.

We claim that for every $P \geq 0$,

$$H_I^I(P) > H_I^A(P), \tag{A.4}$$

which implies $\mathcal{S}_I^{\text{ext}} \subseteq \mathcal{S}_{A^{\text{ins}}}^{\text{ext}}$. To prove (A.4), note that $P_I > P_A$ and set $\Delta := P_I - P_A > 0$. Then

$$H_I^I(P) - H_I^A(P) = \big[H_I(P + P_A + \Delta) - H_I(P + P_A)\big] - (X_I - X_A).$$

Because H_I is strictly convex, the increment $H_I(t + \Delta) - H_I(t)$ is weakly increasing in t. With $t = P + P_A \geq P_A$,

$$H_I(P + P_A + \Delta) - H_I(P + P_A) \;\geq\; H_I(P_A + \Delta) - H_I(P_A) = H_I(P_I) - H_I(P_A).$$

Using $X_I = H_I(P_I)$ and (A.3), we have $H_I(P_I) - H_I(P_A) > X_I - X_A$, hence $H_I^I(P) - H_I^A(P) > 0$ for all $P \geq 0$, proving (A.4).

Therefore, for every $A^{\text{ins}} \in \mathcal{A}_I \setminus \{I\}$, $\mathcal{S}_I^{\text{ext}} \subseteq \mathcal{S}_{A^{\text{ins}}}^{\text{ext}}$. Since $I \in \mathcal{A}_I$,

$$\mathcal{S}^{\text{ext}} = \bigcap_{A^{\text{ins}} \in \mathcal{A}_I} \mathcal{S}_{A^{\text{ins}}}^{\text{ext}} = \mathcal{S}_I^{\text{ext}}.$$

Finally, any internal exchange that preserves (P_I, X_I) leaves H_I and thus $\mathcal{S}_I^{\text{ext}}$ unchanged, so external resilience is invariant to such internal reallocations. □

Proof of Proposition 6. Fix a ruling coalition I and consider the finite sequence of within-I exchanges in Figure 6 that keeps (P_I, X_I) fixed and terminates at a hierarchical allocation $I' := I^T$. Let I^t denote the allocation after t exchanges. By Proposition 4, each exchange weakly enlarges the external safe area, so

$$\mathcal{S}_I^{\text{ext}} \subseteq \mathcal{S}_{I'}^{\text{ext}}.$$

If the inclusion is strict at some intermediate step, we are done. Hence suppose $\mathcal{S}^{\text{ext}}_{I^t} = \mathcal{S}^{\text{ext}}_{I}$ for all $t < T$.

Immediately before the last exchange, there is (by construction) a remaining misordered pair $i, j \in I$ with $p_i^{T-1} > p_j^{T-1}$ and $x_i^{T-1} < x_j^{T-1}$. Consider any best insider sub-coalition $C \in \mathcal{A}_{I^{T-1}}$. As in the proof of Proposition 4, no best insider sub-coalition can contain j but exclude i: if $j \in C$ and $i \notin C$, then replacing j by i strictly increases power and strictly decreases internal resources, contradicting the definition of $\mathcal{A}_{I^{T-1}}$. Therefore, the only best insider sub-coalitions whose (P, X)-location changes in the last exchange are those that contain i and exclude j.

Fix any $A \in \mathcal{A}_{I^{T-1}}$ with $i \in A$ and $j \notin A$, and let A' be its post-exchange counterpart. By the definition of the exchange, p_i decreases and x_i increases, while $j \notin A$, so

$$P_{A'} < P_A \quad \text{and} \quad X_{A'} > X_A.$$

Recall the translated boundary $H_C(P) := H_I(P + P_C) - X_C$ and the associated external safe region $\mathcal{S}^{\text{ext}}_C := \{(P, X) \in \mathbb{R}^2_{++} : \ P < \beta P_N, \ X > H_C(P)\}$. Since H_I is strictly increasing, for every $P < \beta P_N$ we have

$$H_{A'}(P) - H_A(P) = \big[H_I(P + P_{A'}) - H_I(P + P_A)\big] - (X_{A'} - X_A) < 0,$$

hence $H_{A'}(P) < H_A(P)$ for all $P < \beta P_N$, and therefore

$$\mathcal{S}^{\text{ext}}_A \subsetneq \mathcal{S}^{\text{ext}}_{A'}. \tag{A.5}$$

The external safe area of the ruling coalition is $\mathcal{S}^{\text{ext}}_{I^{T-1}} = \bigcap_{C \in \mathcal{A}_{I^{T-1}}} \mathcal{S}^{\text{ext}}_C$. In the last exchange, (i) every best sub-coalition that does not contain i is unchanged, hence its $\mathcal{S}^{\text{ext}}_C$ is unchanged; (ii) every best sub-coalition that contains i and excludes j expands strictly as in (A.5); (iii) any change in the best-subcoalition set can only remove dominated constraints (which weakly enlarges the intersection), and cannot create a new best sub-coalition that contains j but excludes i (by the dominance argument above). Consequently, the last exchange yields a strict expansion of the intersection whenever at least one affected boundary H_A is binding somewhere in the upper envelope $\max_{C \in \mathcal{A}_{I^{T-1}}} H_C(P)$ on $(0, \beta P_N)$.

It remains to show that for sufficiently concave CES preferences (i.e. ρ sufficiently negative), such binding occurs. As $\rho \to -\infty$, G_ρ converges pointwise to the Leontief aggregator $G_{-\infty}$, and the indifference curve through I becomes kinked; each translated boundary $H_C(\cdot)$ inherits a single kink. Since $\mathcal{A}_{I^{T-1}}$ is finite, the upper envelope of these kinked boundaries is piecewise and is attained by a single boundary except at finitely many kink/intersection points. In particular, some affected best sub-coalition $A \in \mathcal{A}_{I^{T-1}}$ with $i \in A$ and $j \notin A$ attains the envelope at some $P^* \in (0, \beta P_N)$ in the Leontief limit (cf. Figure 11). By continuity of H_I and hence $H_C(\cdot)$ in ρ on compact P-sets, there exists $\bar{\rho} < 1$ such that for all $\rho \leq \bar{\rho}$, the same coalition A remains binding at some $P^* \in (0, \beta P_N)$. For such ρ, the last exchange strictly lowers the envelope at P^* (by (A.5)), and thus strictly enlarges $\mathcal{S}^{\text{ext}}_{I^{T-1}}$, implying

$$\mathcal{S}^{\text{ext}}_{I} \subsetneq \mathcal{S}^{\text{ext}}_{I'}.$$

□

Appendix B Examples

The following example illustrates a typical function $G_i(\cdot)$ satisfying Assumptions 1-3 and clarifies the distinction between environments that generate inclusive versus exclusive ruling coalitions.

Example 4. *For any $i \in I \in \mathcal{W} \setminus \{N\}$, consider*

$$w_i(I) = G_i(I) := \left(\frac{p_i}{P_I}\right)\left(\frac{P_I}{P_N}\right)^{\alpha+1}\left(\frac{X_N}{X_I}\right), \tag{B.1}$$

where $\alpha > 0$. The term $\frac{p_i}{P_I}$ is i's share of plundered resources, proportional to her relative power in the ruling coalition, and the plunder function $\left(\frac{P_I}{P_N}\right)^{\alpha+1}\left(\frac{X_N}{X_I}\right)$ ranks ruling coalitions by the resources they extract. One can verify that $\alpha > 0$ is required for parts (i)-(ii) of Assumption 1; if $\alpha < 0$, these are violated.

Normalize $P_N = X_N = 1$. Then

$$G_i(I) = p_i P_I^{\alpha} \frac{1}{X_I}.$$

Fix a payoff level $\overline{G}_i$ and let I be a ruling coalition containing i. The indifference curve of i through I is the locus of (P, X) with $G_i(P, X) = \overline{G}_i$:

$$X = C_i(I) P^{\alpha}, \tag{B.2}$$

where $C_i(I) := \frac{p_i}{\overline{G}_i}$ and $P \in [\beta, 1]$. Along such an indifference curve the marginal rate of substitution between power and resources is

$$MRS_{PX} = -\alpha \frac{X}{P}.$$

The parameter α governs the relative valuation of power versus resources. A higher α makes indifference curves steeper: for given (P, X), the marginal value of power relative to internal resources is higher, so players are more willing to sacrifice resources to gain power. Since $P \in [\beta, 1]$, a higher α reduces P_I^α (for $P_I < 1$) and thus dampens the increase in plunder from further increases in power. This corresponds to a relatively weak *plundering technology: very powerful coalitions obtain relatively modest incremental gains from additional power, and insiders are more willing to form large, inclusive ruling coalitions.*

Conversely, lower values of α flatten indifference curves and raise P_I^α for $P_I < 1$, making payoffs more sensitive to power and less constrained by internal resources. This corresponds to a relatively intensive *plundering technology: small, powerful coalitions can extract much more from outsiders, and insiders are more reluctant to dilute power, so ruling coalitions tend to be more exclusive.*

The next example shows that, without further restrictions on the joint distribution of power and resources and the plundering function, there is no general characterization of the ruling coalition's composition. This follows, first, from Proposition 2(1), which implies that any coalition in the set of potential ruling coalitions may be the ruling coalition for some range of indifference curves; and second, from the fact that the set of potential ruling coalitions itself cannot be sharply characterized without additional structure on $(p_i, x_i)_{i \in N}$. In particular, there is no guarantee that the ruling coalition contains the most powerful player, the player with the fewest resources, or the player with the highest power-to-resource ratio (Figure 18).[16] Example 5 illustrates this point.

Example 5. *Suppose $N = \{1, 2, 3, 4\}$, with*

$$p_1 = 6,\ x_1 = 5, \quad p_2 = 3,\ x_2 = 3, \quad p_3 = 2,\ x_3 = 6, \quad p_4 = 5,\ x_4 = 4,$$

and let $\beta = \frac{1}{2} + \epsilon$. Then the set of winning coalitions is

$$\mathcal{W} = \{\{1, 2\}, \{1, 4\}, \{2, 4\}, \{1, 2, 4\}\}.$$

As shown in Figure 18, different indifference curves select different ruling coalitions: when the indifference curve is H_1, the ruling coalition is $\{2, 4\}$, which excludes the player with the highest power (player 1); when it is H_3, the ruling coalition is $\{1, 4\}$, which excludes the player with the lowest resources; and when it is H_2, the ruling coalition excludes the player with the highest power-to-resource ratio (player 4). Thus, absent additional structure, the ruling coalition need not contain the most powerful, or the poorest player.

[16] Example 4 in the appendix shows that a sharper characterization is possible when powers, resources,

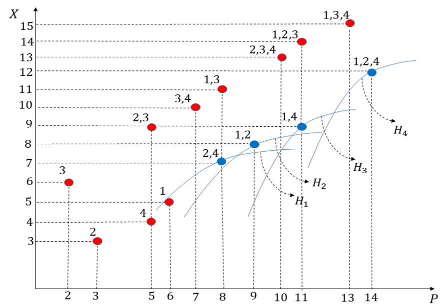


Figure 18: Ruling coalition under different indifference curves.

References

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

——— (2012): "Dynamics and stability of constitutions, coalitions, and clubs," *American Economic Review*, 102 (4), 1446–1476.

Acemoglu, Daron and James A Robinson (2008): "Persistence of power, elites, and institutions," .

——— (2013): "Economics versus politics: Pitfalls of policy advice," *Journal of Economic perspectives*, 27 (2), 173–192.

Alatas, Syed Farid (2014): *Applying Ibn Khaldūn: The recovery of a lost tradition in sociology*, Routledge.

Andolfatto, David (2002): "A theory of inalienable property rights," *Journal of Political Economy*, 110 (2), 382–393.

Bai, Ying and Titi Zhou (2019): ""Mao's last revolution"': a dictator's loyalty-competence tradeoff," *Public Choice*, 180 (3), 469–500.

Battaglini, Marco (2021): "Coalition formation in legislative bargaining," *Journal of Political Economy*, 129 (11), 3206–3258.

Besley, Timothy and Torsten Persson (2009): "The origins of state capacity: Property rights, taxation, and politics," *American economic review*, 99 (4), 1218–1244.

Besley, Timothy J and Masayuki Kudamatsu (2007): "Making autocracy work," *LSE STICERD Research Paper No. DEDPS48.*

Brewer, Tony (2002): *Marxist theories of imperialism: A critical survey*, Routledge.

or both are equally distributed in society.

Dal Bó, Ernesto and Pedro Dal Bó (2011): "Workers, warriors, and criminals: social conflict in general equilibrium," *Journal of the European Economic Association*, 9 (4), 646–677.

Dal Bó, Ernesto, Pablo Hernández-Lagos, and Sebastián Mazzuca (2022): "The paradox of civilization: Preinstitutional sources of security and prosperity," *American Political Science Review*, 116 (1), 213–230.

De Mesquita, Ethan Bueno (2010): "Regime change and revolutionary entrepreneurs," *American Political Science Review*, 104 (3), 446–466.

Diermeier, Daniel, Georgy Egorov, and Konstantin Sonin (2017): "Political economy of redistribution," *Econometrica*, 85 (3), 851–870.

Egorov, Georgy, Sergei Guriev, and Konstantin Sonin (2009): "Why resource-poor dictators allow freer media: A theory and evidence from panel data," *American Political Science Review*, 103 (4), 645–668.

Egorov, Georgy and Konstantin Sonin (2011): "Dictators and their viziers: Endogenizing the loyalty–competence trade-off," *Journal of the European Economic Association*, 9 (5), 903–930.

——— (2024): "The political economics of non-democracy," *Journal of Economic Literature*, 62 (2), 594–636.

Enikolopov, Ruben, Alexey Makarin, and Maria Petrova (2020): "Social media and protest participation: Evidence from Russia," *Econometrica*, 88 (4), 1479–1514.

Garicano, Luis (2000): "Hierarchies and the Organization of Knowledge in Production," *Journal of political economy*, 108 (5), 874–904.

Garicano, Luis and Esteban Rossi-Hansberg (2015): "Knowledge-based hierarchies: Using organizations to understand the economy," *economics*, 7 (1), 1–30.

Gehlbach, Scott and Konstantin Sonin (2014): "Government control of the media," *Journal of public Economics*, 118, 163–171.

Guriev, Sergei and Daniel Treisman (2019): "Informational autocrats," *Journal of economic perspectives*, 33 (4), 100–127.

Hafer, Catherine (2006): "On the origins of property rights: Conflict and productionin the state of nature," *The Review of Economic Studies*, 73 (1), 119–143.

Harsanyi, John C. (1974): "An equilibrium-point interpretation of stable sets and a proposed alternative definition," *Management Science*, 20 (11), 1472–1495.

Henn, Soeren J, Christian Mastaki Mugaruka, Miguel Ortiz, Raúl Sánchez de la Sierra, and David Qihang Wu (2024): "Monopoly of Taxation without a Monopoly of Violence: The Weak State's Trade-Offs from Taxation," *Review of Economic Studies*, rdae036.

Jia, Ruixue, Masayuki Kudamatsu, and David Seim (2015): "Political selection in China: The complementary roles of connections and performance," *Journal of the European Economic Association*, 13 (4), 631–668.

Khaldun, Ibn (1967): "The Muqaddimah, an Introduction to History, tr. by Franz Rosenthal, 3 vols," .

König, Michael D, Dominic Rohner, Mathias Thoenig, and Fabrizio Zilibotti (2017): "Networks in conflict: Theory and evidence from the great war of africa," *Econometrica*, 85 (4), 1093–1132.

Konrad, Kai A and Stergios Skaperdas (2007): "Succession rules and leadership rents," *Journal of Conflict Resolution*, 51 (4), 622–645.

Kuran, Timur (1996): "The discontents of Islamic economic morality," *The American Economic Review*, 86 (2), 438–442.

Lorentzen, Peter (2014): "China's strategic censorship," *American Journal of political science*, 58 (2), 402–414.

Lorentzen, Peter L (2013): "Regularizing rioting: Permitting public protest in an authoritarian regime," *Quarterly Journal of Political Science*, 8 (2), 127–158.

Lowes, Sara, Nathan Nunn, James A Robinson, and Jonathan L Weigel (2017): "The evolution of culture and institutions: Evidence from the Kuba Kingdom," *Econometrica*, 85 (4), 1065–1091.

Mattingly, Daniel C (2024): "How the party commands the gun: The foreign–domestic threat dilemma in China," *American Journal of Political Science*, 68 (1), 227–242.

Meng, Anne (2020): *Constraining dictatorship*, Cambridge University Press Cambridge.

Michalopoulos, Stelios and Elias Papaioannou (2013): "Pre-colonial ethnic institutions and contemporary African development," *Econometrica*, 81 (1), 113–152.

Mookherjee, Dilip (2013): "Incentives in hierarchies," *The Handbook of Organizational Economics*, 764–798, edited by Robert Gibbons and John Roberts.

Morishima, Michio (1974): "Marx in the light of modern economic theory," *Econometrica: Journal of the Econometric Society*, 611–632.

Paine, Jack (2021): "The Dictator's Power-Sharing Dilemma: Countering Dual Outsider Threats," *American Journal of Political Science*, 65 (2), 510–527.

Persson, Torsten (2002): "Do political institutions shape economic policy?" *Econometrica*, 70 (3), 883–905.

Puga, Diego and Daniel Trefler (2014): "International trade and institutional change: Medieval Venice's response to globalization," *The Quarterly Journal of Economics*, 129 (2), 753–821.

Qian, Yingyi (1994): "Incentives and loss of control in an optimal hierarchy," *The Review of Economic Studies*, 61 (3), 527–544.

Qian, Yingyi, Gerard Roland, and Chenggang Xu (2006): "Coordination and experimentation in M-form and U-form organizations," *Journal of Political Economy*, 114 (2), 366–402.

Ray, Debraj and Rajiv Vohra (2015a): "Coalition Formation," in *Handbook of Game Theory with Economic Applications*, ed. by H. Peyton Young and Shmuel Zamir, North Holland, 240–322.

——— (2015b): "Coalition formation," *Handbook of game theory with economic applications*, 4, 239–326.

——— (2015c): "The farsighted stable set," *Econometrica*, 83 (3), 977–1011.

Roemer, John E (1980): "A general equilibrium approach to Marxian economics," *Econometrica: Journal of the Econometric Society*, 505–530.

——— (1981): *Analytical foundations of Marxian economic theory*, Cambridge University Press.

Sánchez De La Sierra, Raúl (2020): "On the origins of the state: Stationary bandits and taxation in eastern congo," *Journal of Political Economy*, 128 (1), 000–000.

Shadmehr, Mehdi (2018): "Multiplicity and Uniqueness in Generalized Regime Change Games," *Available at SSRN 3121241*.

Svolik, Milan W (2012): *The politics of authoritarian rule*, Cambridge University Press.

Wintrobe, Ronald (1990): "The tinpot and the totalitarian: An economic theory of dictatorship," *American political science review*, 84 (3), 849–872.

——— (2000): *The political economy of dictatorship*, Cambridge University Press.

Xu, Guo (2018): "The costs of patronage: Evidence from the british empire," *American Economic Review*, 108 (11), 3170–3198.

Yanagizawa-Drott, David (2014): "Propaganda and conflict: Evidence from the Rwandan genocide," *The Quarterly Journal of Economics*, 129 (4), 1947–1994.

Zakharov, Alexei V (2016): "The loyalty-competence trade-off in dictatorships and outside options for subordinates," *The Journal of Politics*, 78 (2), 457–466.