Plundering Coalitions*

Alireza Mohammadi Sepahvand† Weijia Li ‡

November 22, 2024

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 resource 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 a higher "rank," exhibits a greater resilience to external shocks affecting the outsider's power and resources. The only exception is when plundering intensity is "relatively weak," where the internal distribution of power and resources does not impact external resilience. Our final results provide insights into how the intensity of plundering impacts the internal and external resilience of the ruling coalition in various political environments.

Keywords: Coalition Formation, Plundering, Resilience, Hierarchies, Political Economy

---

*We are We thank Arthur Campbell, Yves Zenou, Paul Raschky, Vai-Lam Mui, Mathew Jackson, Sam Jindani, Massimo Morreli, Sebastiano Della Lena, Kaveh Majlesi, Solmaz Moslehi, Melika Liporace, Thomas BrÃ€uninger, Mehdi Khodakarami, Nurlan Lalayev and all participants of various conferences, seminars and workshops for their valuable comments and insights. All errors are ours.

†Monash University, Australia. Email: Alireza.mohammadisepahvand@monash.edu.

‡Monash University, Australia, E-mail: Weijia.li@monash.edu.

1 Introduction

Coalition formation is always challenging (Ray and Vohra (2015a)), and a "plundering coalition" is no exception. For such a coalition, the wealth it distributes among coalition members is plundered from outsiders. This setup applies to a wide range of important social phenomena, such as an army that plunders 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. Our model yields a series of novel results, both substantive and methodological, by focusing on the resilience of the equilibrium plundering coalition against exogenous shocks.

Specifically, our model features a society of a finite number of individuals. Each individual has two endowments, namely power and wealth. The power of a coalition is the summation of all its members' powers. A "winning coalition" can defeat outsiders with its power.[1] The game starts with an initial winning coalition. A member of the initial coalition may propose the creation of 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 outsiders, plunder their wealth, and distribute the plundered wealth among its members.

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 in the coalition, the new coalition is more powerful against outsiders, therefore being able to plunder more wealth from them. But a new member is also costly for existing insiders because they can no longer plunder the wealth of the new member. 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.

To prepare our novel analysis of resilience, we prove a necessary and sufficient condition

[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 $\beta > 1/2$.

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. Second, 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. We first focus on 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 to a weaker member, without flipping their power rank, or a transfer of wealth from a richer member to a poorer one, 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 classical 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 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); Halac et al. (2021); Halac et al. (2024)) or division of labor (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 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 internal resilience.[2] This contrasts with the naive view that plundering more intensively increases insiders' attachment to the ruling coalition.

For external resilience of the ruling coalition, a stronger plundering technology is a double-edged sword. On the one hand, exclusive alternatives that involve small segments of society become more threatening to the ruling coalition. On the other hand, inclusive alternatives that 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 external resilience. The latter suggests that a ruling coalition that plunders society intensively benefits more from facing a more powerful and wealthier opposition than a weaker and poorer one.

Lastly, although the direction of change in external resilience—driven by change in plundering technology—generally depends on the realization of power and resources inside the ruling coalition, we identify a wide range of political environments where this is not the case. That is, external resilience is robust with respect to changes in the internal configuration of power and resources. In these political environments, the plundering process is "relatively weak"; for instance, it is endowed with better protections of prop-

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

erty rights. [3] 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 external resilience of the ruling coalition, since it renders the inclusive alternatives less beneficial for the players. Thus, in relatively weak plundering environments, there exists a trade-off between the external and internal resilience of the ruling coalition with respect to 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—potentially hinder the ruling coalition from achieving both internal and external stability when plundering technology changes. This contrasts with "relatively intensive plundering" environments, wherein a change in plundering technology could alleviate both internal and external threats to a ruling coalition.

1.1 Relevant Literature

Our paper is relevant to a few strands of literature. The literature on coalition formation largely focuses on characterizing the equilibrium coalition (Acemoglu et al. (2008); Ray and Vohra (2015b); Battaglini (2021)) or defines stability mainly by incorporating the notion of âfarsightednessâ (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 numerous substantive insights.

We bring together the two strands of literature on coalition formation and organizational economics of hierarchy. Existing literature usually focuses on how a hierarchy may improve incentive-alignment or the division of labor (Qian (1994); Qian et al. (2006); Mookherjee (2013); Garicano (2000); Garicano and Rossi-Hansberg (2015); Halac et al. (2021); Halac et al. (2024)). 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 it. Our novel justification is relevant to many hierarchies where the characteristics of outsiders are a first order concern, such as armies and fiscal bureaucra-

[3] As we will discuss, this would imply that hierarchical coalitions are less justified in political environments where property rights protections are relatively better.

cies (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, 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 the 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); Jia et al. (2015); Zakharov (2016); Bai and Zhou (2019); Li (2023); 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 that dictators resolve while tackling external threats, for instance, the one between "informational openness" and "security" (Lorentzen et al. (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 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 the characteristics of political environments where this trade-off does not hold and into the characteristics of oppositions (i.e., outsiders) that enable autocratic states to achieve greater stability.

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 resilience in Section 4. Section 5 concludes.

2 Environment

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

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

We refer to $p_i := p(i)$, and $x_i := x(i)$, as, respectively, the political power and economic resources of individual $i \in N$. Our aim is to highlight how differences in the power and resources of individuals map into political decisions. 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 respectively denoted by

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

Coalition I is called a winning coalition if $P_I \geq \beta P_N$, where $\beta \in [1/2, 1]$ is a fixed degree of super-majority. Denote the set of all winning coalitions as $\mathcal{W}$. There is a baseline pay-off function $U : N \times \mathcal{W} \to \mathbb{R}$ that assigns to any player $i \in N$ the pay-off $U_i(I)$ when the winning coalition $I \in \mathcal{W}$ becomes the ruling coalition. By abusing the notation, we denote $U(i, I) := U_i(I)$.

A ruling coalition of our model is necessarily a winning coalition. As a key new feature 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 anymore. This key trade-off is formally captured by Assumption 1(1). The rest of the assumption is more straightforward, enforcing that inclusion in the ruling coalition is strictly better for the individuals.

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

1. *(Trade-off) If $I \in \mathcal{W} \backslash \{N\}$ and $i \in I$, we have $w_i(I) = G_i(P_I, X_I) > 0$, where*

$G_i(.,.) : [\beta P_N, P_N) \times [0, X_N) \to \mathbb{R}_{++}$ *is a continuous function that satisfies:*[4]

(a) *(more resources, less plundering) For all* $I, I' \in \mathcal{W} \backslash \{N\}$*, where* $P_I = P_{I'}$*, if* $i \in I, I'$*, then* $G_i(P_I, X_I) > G_i(P_{I'}, X_{I'})$ *if and only if* $X_I < X_{I'}$*.*

(b) *(more power, more plundering) For all* $I, I' \in \mathcal{W} \backslash \{N\}$*, where* $X_I = X_{I'}$*, if* $i \in I, I'$*, then* $G_i(P_I, X_I) > G_i(P_{I'}, X_{I'})$ *if and only if* $P_I > P_{I'}$*.*

2. *If* $i \notin I$*, then* $w_i(I) < 0$*.*

3. $\forall i \in N$*,* $w_i(N) = 0$*.*

Assumption 1 establishes the key aspects of the model. The function $G_i(\cdot)$ in part 1 ranks the share of any individual across different non-trivial ruling coalitions of which she is a member.[5] Part 1(a) says that between ruling coalitions with equal powers, players prefer the one with fewer resources, allowing access to more external resources for plundering. In addition, between ruling coalitions with equal resources, players prefer the one with larger power (part 1(b)), as it enhances resource extraction. Both parts 1(a) and part 1(b) say that when the ruling coalition is not the grand coalition (i.e., there exist resources outside it to be plundered), the payoffs of the insiders of the ruling coalition from the plundered resources are strictly positive. This, along with part 2, ensures that the game is zero-sum with respect to inclusion in the ruling coalition. This means that any player not included in the ruling coalition will also receive a payoff that is strictly lower than their initial resources. Part (c) states that when the ruling coalition is the grand coalition, the players' payoff from the plundered resources is zero, since there are no outsiders to plunder.

Furthermore, Assumption 1(1) imposes that the only important property of a ruling coalition is its pair of aggregate power and resources. This simplifies the model by ignoring the complexities that arise when the combination of players inside the ruling coalition, achieving a fixed aggregate power and resources, is also important. Accordingly, we can denote $G_i(I) := G_i(P, X)$ in the rest of the paper. Assumption 1(1) immediately results in the following Lemma.

[4] The continuity here is with respect to the aggregate power and resources of the coalition.

[5] For example, a function $G_i(.,.)$ can be viewed 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., $G_i(I) := \Pi(i, I)F(I)$ is the share allocated to individual i within the coalition I from plundered resources $F(I)$.

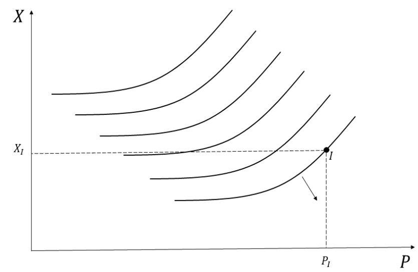


Figure 1: Identical indifference curves under Assumption 2.

Suppose Assumption 1 holds. Then, any player $i \in N$ has strictly increasing and continuous indifference curves in the subspace $[\beta P_N, P_N) \times [0, X_N) \subset \mathbb{R}^2_{++}$ across the set $\mathcal{W}_i := \{I \in \mathcal{W} | i \in I\}$.

Lemma 1 states any player has strictly increasing and continuous indifference curves across the set of non-trivial winning coalitions containing them. Importantly, this Lemma does not imply that the indifference curves are identical for different individuals. The following assumption introduces consistency in the players' preferences, which is crucial for establishing the main results of the paper.

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

Assumption 2 states that the players' preferences regarding ruling coalitions containing them are specified by two components: an idiosyncratic component $g(i)$, which relies on their personal endowments (p_i, x_i), and a common component $G(I)$, which depends on the aggregate power and resources of the coalition, (P_I, X_I). Consequently, since the power and resources of individuals are fixed throughout the game, this assumption implies that for all players, the indifference curves over the coalitions containing them are the same and determined by the function $G(.)$ (Figure 1). Specifically, under Assumption 2, if $i, j \in I, I'$, then $U_i(I) \geq U_i(I')$ if and only if $U_j(I) \geq U_j(I')$, i.e., the preferences of players regarding any pair of ruling coalitions containing them are identical.

Definition 1. *Fix a function $G(.)$ that satisfies Assumption 1. For any ruling coalition I, and for any fixed $\overline{G(I)}$, let denote the indifference curves corresponding to ruling coalition I as $\overline{G(I)}$ as $X = G^{-1}(\overline{G(I)}, P) := H_I(P)$.*

Throughout the paper, we assume that the joint power and resources mapping is generic in the sense that $\forall I, I' \in \mathcal{W}$, we have $P_I \neq P_{I'}$, or $X_I \neq X_{I'}$.[6] The following assumption is essential for establishing the uniqueness results in the subsequent section.

Assumption 3. *Fix the power and resource mappings. Then, $\forall I, I' \in \mathcal{W}$, we have: $G(I) \neq G(I')$.*[7]

This Assumption imposes that players receive strictly different payoffs from different ruling coalitions involving them.

3 Preliminary Analysis of Coalition Formation Game

3.1 Basic game form

We next define the extensive-form complete information game $\Gamma = (N, I_0, p(.), x(.), \{U_i(.)\}_{i\in N}, \beta)$, where N is the set of players, I_0 is the initial winning coalition, $p(.)$ and $x(.)$ are the power and resource mappings, $\{U_i(.)\}_{i\in N}$ are the payoff functions satisfying Assumption 1 and Assumption 2, and $\beta \in [1/2, 1]$ is the degree of the super-majority. The game starts with an initial winning coalition $I_0 \in \mathcal{W}$, and the steps are as follows:

1. Nature randomly picks an agenda-setter from the initial winning coalition, $a_q \in I_0$ for $q = 1$.

2. The agenda setter a_q makes a proposal to a subset of players $I_q \subseteq N$ to form a ruling coalition (possibly empty). If $I_q \notin \mathcal{W}$ and $q < |I_0|$, step 1 begins again with the new agenda setter a_{q+1} from $I_0 \backslash \{a_1, a_2, \ldots, a_q\}$. If $I_q \notin \mathcal{W}$ and $q = |I_0|$, I_0 becomes the ruling coalition, and the payoff of any player $i \in N$ is given by $U_i(I_0) = x_i + w_i(I_0)$. Otherwise, if $I_q \in \mathcal{W}$, Nature chooses the order of votes for the proposal I_q and the game proceeds to step 3.

3. The voting process begins, and the coalition I_q forms if and only if it is accepted by *all* the players in I_q. In such a case, I_q becomes the ruling coalition, and the payoff

[6] 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 finite hyper-planes so that it has Lebesgue measure 0.

[7] This Assumption is also made without much loss of generality, as the set of functions from $\mathbb{R}^2$ to $\mathbb{R}$, where for a finite number of distinct pairs of inputs, the outputs are not necessarily distinct, forms a measure-zero set in the space of all functions from $\mathbb{R}^2$ to $\mathbb{R}$.

of any player $i \in N$ is given by $U_i(I_q) = x_i + w_i(I_q)$. Otherwise, after the rejection of I_q by the first voter, the game proceeds to step 4.

4. If $q < |I_0|$, the game proceeds to step 1 where Nature randomly picks a new agenda setter from the initial winning coalition, $a_{q+1} \in I_0 \backslash \{a_1, a_2, \ldots, a_q\}$. Otherwise, if $q = |I_0|$, I_0 becomes the ruling coalition, and the payoff of any player $i \in N$ is given by $U_i(I_0) = x_i + w_i(I_0)$.

The equilibrium concept we use is subgame perfect equilibrium (SPE). The extensive-form game delineates the strategies of players in any such equilibrium. The 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 their behavior at each decision node h; where function $v_i(h, I)$ specifies player i's vote (either 'Yes' or 'No') in any history h where Nature selects her to vote on a proposal $\mathcal{P}$ involving her made before; and $\mathcal{P}_i(h)$ determines 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 $i \in N \backslash I_0$, player i might only be a voter throughout the game.[8] Thus, the strategy of $i \in N \backslash I_0$ is a function $v_i(h, \mathcal{P})$ assigning either 'Yes' or 'No' to any proposed ruling coalition $\mathcal{P}$ containing i in history h that she is chosen by the Nature to vote to $\mathcal{P}$.

3.2 Axiomatic analysis

In this section, three axioms are introduced, motivated by the game's structure. Just as in Acemoglu et al. (2012) and Acemoglu et al. (2008), the axiomatic analysis will show that our results are independent of the details in the agenda setting and voting protocols. The axiomatic analysis also will help us characterize the equilibrium of the non-cooperative game in Section 3.1. In theorem 1, a mapping $\phi : \mathcal{W} \longrightarrow C$ is characterized, satisfying these axioms and identifying the set of ruling coalitions corresponding to each initial winning coalition. Formally, with fixed power and resource mappings, pay-off functions, and degree of super-majority, the following axioms are adopted on ϕ:

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

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

[8] Notice that all the results of the paper hold when we change the structure of the game such that all players can both be voters and proposers.

Axiom 3 (Rationality). *For any* $I \in \mathcal{W}$, $I' \in \phi(I)$, *and* $I'' \in \mathcal{W}$, *we have:* $I'' \notin \phi(I) \iff G(I'') < G(I')$.

These axioms are natural and capture the salient economic forces that give rise to the pure strategy SPE of the game. Axiom 1 implies that ϕ maps any initial ruling coalition to a non-trivial subset of players. Axiom 2 requires the ruling coalition to be a winning coalition. Axiom 3 implies that if I' is a ruling coalition for an initial winning coalition I, and it brings a strictly higher payoff than another winning coalition I'', then I'' cannot be the ruling coalition for the initial winning coalition I, and vice versa. Theorem 1 establishes that these natural axioms pin down a unique mapping under Assumptions 1-2 that is single-valued under Assumptions 1-3.

Theorem 1. *Fix a set of players* N, *a power and resource mapping,* $p(\cdot)$, *and* $x(\cdot)$, *and a degree of super-majority* $\beta \in \left[\frac{1}{2}, 1\right]$. *Then:*

1. *(Existence) Under Assumptions 1-2, the mapping* $\phi : I \longrightarrow \arg\max_{W \in \mathcal{W}} G(W)$ *is the unique mapping that satisfies axioms 1-3.*
2. *(Uniqueness) Under Assumptions 1-3, the mapping* ϕ *is single-valued.*

Theorem 1 characterizes the ruling coalition. The ruling coalition is a winning coalition that maximizes the plundering of outsiders' resources (i.e., the ruling coalition must maximizes $G(W)$ among all possible winning coalitions $W \in \mathcal{W}$). The ruling coalition is then unique under Assumption 3.

3.3 Non-cooperative characterization

The main results of this section characterize the conditions that guarantee the existence, and uniqueness of the ruling coalition for the non-cooperative coalition formation game. Furthermore, we establish that the subgame perfect equilibrium of the coalition formation game coincides with the ruling coalition derived from the axiomatic approach in Section 3.2. To achieve this, we first need to define a set of "potential" ruling coalitions.

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

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

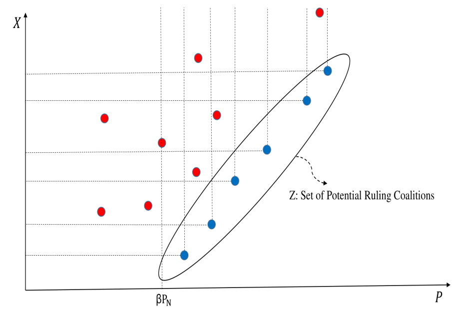


Figure 2: Set of potential ruling coalitions.

Fixing the power and resource mappings, the set of potential ruling coalitions is the set of all winning coalitions for which there is no winning coalition with *both* higher powers and fewer resources. For instance, any red point in Figure 2 is a coalition that cannot be in the set of potential ruling coalitions because a blue point (another coalition) exists with higher power and fewer resources.[9]

We next proceed with Theorem 2. Part 1 of this theorem establishes that, under Assumptions 1-2, the game described in Section 3.1 will essentially have an equilibrium in which the ruling coalition is within the set of potential ruling coalitions. Part 2 states that under Assumptions 1-3, any equilibrium would result in a unique ruling coalition.

Theorem 2. *Suppose:*

1. (Existence) Assumptions 1-2 hold and $\phi(I_0)$ satisfies Axioms 1-3. Then for any $I \in \phi(I_0)$, we must have $I \in Z$, and there exists a pure strategy profile σ_I that is SPE and leads to ruling coalition I. In this equilibrium, any player $i \in N$ receives payoff $U_i(I) = x_i + w_i(I)$.

2. (Uniqueness) Assumptions 1-3 hold, $\phi(I_0)$ satisfies Axioms 1-3, and $\phi(I_0) = \{I\}$. Then any SPE results in I as the ruling coalition. In equilibrium, any player $i \in N$ receives payoff $U_i(I) = x_i + w_i(I)$.

The intuitions behind both parts are straight-forward given Assumptions 1-3 and the specification of the mapping $\phi(.)$ in Theorem 1. For part (1), suppose $I \in \phi(I_0)$. According to Theorem 1(1), we have $I \in \arg\max_{W\in\mathcal{W}} G(W)$. Based on Assumptions 1-2, if $i \in I$, then $I \in \arg\max_{W\in\mathcal{W}} G(W)$ implies $I \in \arg\max_{W\in\mathcal{W}} w_i(W)$. This means that for any player

[9] Although a comprehensive characterization of the set of potential ruling coalitions is not feasible in general, Example 5 in the online Appendix illustrates a specific case where any potential ruling coalition corresponds to a power threshold such that all players whose power exceeds that threshold are included in the potential ruling coalition.

i in I, coalition I brings the highest payoff for i among the winning coalitions involving i (though not necessarily uniquely). Furthermore, since we have $I \in \arg\max_{W \in \mathcal{W}} w_i(W)$, Assumption 1(1) implies there does not exist a winning coalition I' such that $P_{I'} > P_I$ and $X_{I'} < X_I$. Therefore, by definition, we have $I \in Z$. Following this, constructing σ_I is straight-forward.

First, note that as $\beta \in [\frac{1}{2}, 1]$, we have $I_0 \cap I \neq \emptyset$. This ensures that during the course of the game, a player from I will inevitably be chosen by Nature to propose a ruling coalition. We then show that there would be no profitable deviation for players within I from the following strategy: always propose I and vote 'YES' for I, and vote 'NO' for any other proposal as long as someone from $I \cap I_0$ remains to propose. Given that $\beta \in [\frac{1}{2}, 1]$, any coalition I' proposed before I would include at least one player from I. The mentioned strategy then prevents any such I' from becoming the ruling coalition, as any ruling coalition forms according to the unanimity rule. This ensures that I would be the ruling coalition.[10]

For part (2), similarly, Assumptions 1-2 imply that if $i \in I$, then $I \in \arg\max_{W \in \mathcal{W}} G(W)$ is equivalent to $I \in \arg\max_{W \in \mathcal{W}} w_i(W)$. Furthermore, under Assumptions 1-3, the mapping $\phi(\cdot)$ is single-valued (Theorem 1(2)). Therefore, if $\phi(I_0) = I$, under Assumptions 1-3, coalition I is the *unique* ruling coalition that brings the highest payoff for all its members (Figure 3). Consider any SPE of the game. Given that $\beta \in \left[\frac{1}{2}, 1\right]$, a history h_a would necessarily reach in which the first proposer from the set $I \cap I_0$, namely a, is chosen by Nature to propose. Let us denote the subgame that starts from the history immediately after player a proposes I as Γ_a^I. Consider a pivotal voter on the proposal I made by a, denoted as v. If the ruling coalition is not I in the continuation of the game after v's vote, then the only optimal vote for player v would be to vote 'YES' to I.

Conversely, if the ruling coalition is I in the continuation of the game after v's vote, v would be indifferent between voting 'NO' and 'YES'. In both cases, I would become the ruling coalition in the subgame Γ_a^I. Thus, if player a proposes I, it would necessarily become the ruling coalition in the continuation of the game after her proposal. Finally, consider the subgames that start from the decision node after a proposes another coalition $I' \neq I$, denoted by $\Gamma_a^{\neg I}$. If I is not the ruling coalition in the subgame $\Gamma_a^{\neg I}$, the only optimal

[10] The off-path equilibrium actions are given by Equation 3 in the proof of Theorem 2(1) in the appendix. The idea is to construct the strategy profile σ_I such that, for any given set of remaining agenda setters at any stage of the game, any player would only accept and propose one of their favorite coalitions that is feasible under that set of remaining agenda setters.

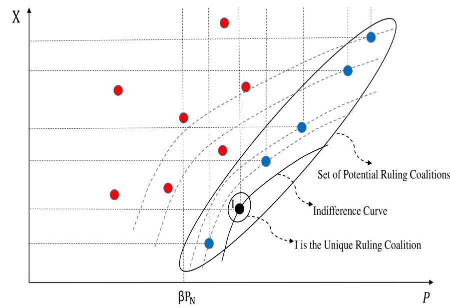


Figure 3: Unique ruling coalition of the game.

action for player a would be to propose I in the history h_a. Otherwise, player a would be indifferent between proposing I or not. In both cases, the ruling coalition would essentially be I in the subgame starting from h_a.

Moreover, given that $\beta \in \left[\frac{1}{2}, 1\right]$, for any proposal made before h_a, there always exists a voter from I within that proposal. As the ruling coalition forms according to the unanimity rule, the only optimal action for such a voter is to prevent any other coalition from becoming the ruling coalition before h_a. This implies that I would necessarily be the ruling coalition in *any* subgame perfect equilibrium of the game.

4 Resilience

This section studies the resilience of the ruling coalition.[11] Our strategy is as follows: We begin with Proposition 1, which is equivalent to the uniqueness results delineated in Theorem 1(2) and Theorem 2(2). Next, building on Proposition 1, we define internal and external resilience. Our ultimate objective is to determine how resilience is affected by shocks to the joint distribution of power and resources, as well as plundering intensity.

4.1 Internal and external resilience

There are two motivations behind Proposition 1. First, it establishes that a ruling coalition must have a relatively higher power-to-resource ratio compared to the alternatives. Furthermore, while it's clear that any ruling coalition must overcome all alternatives, Proposition 1 distinguishes between two types of alternatives confronting any ruling coali-

[11]The standard robustness of the ruling coalition to small perturbations in power and resources is shown in Proposition 4 in the online Appendix.

tion: sub-coalitions and other alternatives encompassing segments of the outsider society. These distinctions reflect the dual challenges faced by authoritarian regimes—authoritarian power-sharing and authoritarian control (Svolik (2012); Paine (2021); Meng (2020))—and enable us to distinguish between two notions of resilience. These challenges, the likelihood of revolutions from outsiders and coup threats from insiders, are major forces explaining the dynamic of coalition formation (Francois et al. (2015)). To achieve this, we introduce a new concept called the set of best sub-coalitions.

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

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

For any subset of players $I \subseteq N$, the set A_I comprises 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 of Definition 3 for potential ruling coalitions, only restricting our attention to the sub-coalitions of the coalition in question.

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

(i) $\forall A^{ins} \in (\mathcal{A}_I \backslash I) \cap \mathcal{W}$, we have $G(I) > G(A^{ins})$ [i.e., there is no profitable internal secession].

(ii) $\forall A^{ext} \in A_{N\backslash I}$ and $\forall A^{ins} \in \mathcal{A}_I$ where $A^{ins} \cup A^{ext} \in \mathcal{W}$, we have $G(I) > G(A^{ins} \cup A^{ext})$ [i.e., there is no profitable external secession].

Conditions of Proposition 1 distinguish two types of alternatives that the ruling coalition must defeat: its own subsets (condition (i)) and other coalitions comprising outsiders (condition (ii)). The first argument of Proposition 1 is that a necessary and sufficient condition for a ruling coalition to win against any of these types of alternatives is to be able to respectively overcome: (i) all its nontrivial best sub-coalitions, and (ii) the combination of its best sub-coalitions and the best sub-coalitions of outsiders. For instance, suppose $A_I = \{A_1^{ins}, A_2^{ins}, I\}$ where $A_1^{ins}, A_2^{ins} \in \mathcal{W}$, and $A_{N\backslash I} = \{A_1^{ext}, A_2^{ext}\}$, where $A_2^{ext} = N\backslash I$. Condition (i) states that I must be able to win against both A_1^{ins} and A_2^{ins}. Condition (ii) requires I to overcome any coalition of the form $A_j^{ins} \cup A_k^{ext}$, where $j, k \in \{1, 2\}$.

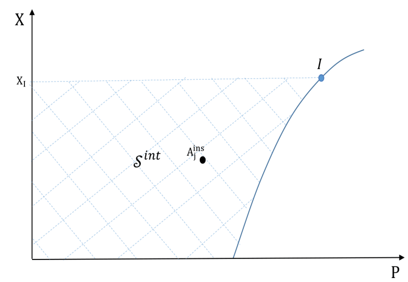


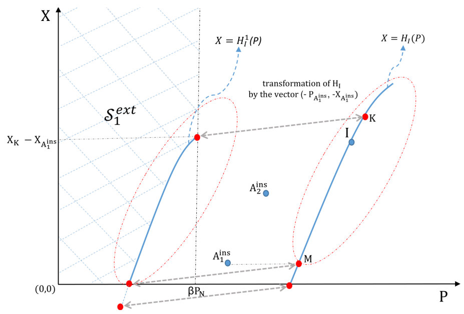
Figure 4: Condition (i) of proposition 1— $\mathcal{S}^{int}$ (the internal safe area): the blue area.

Specifically, condition (i) says that the ruling coalition must satisfy: $X_{A_j^{ins}} > H_I(P_{A_j^{ins}})$, for $j \in \{1, 2\}$. Fixing the power and resources of the ruling coalition and the function $G(.)$, this would specify an area in the (P, X) space (i.e., the area $\mathcal{S}^{int}$ in Figure 4) in which if A_1^{ins} and A_2^{ins} lie, condition (i) would hold. Formally, we have $\mathcal{S}^{int} = \{(P, X) \in \mathbb{R}^2_{++} | X > H_I(P)\}$. We refer to this region as the "internal safe area."

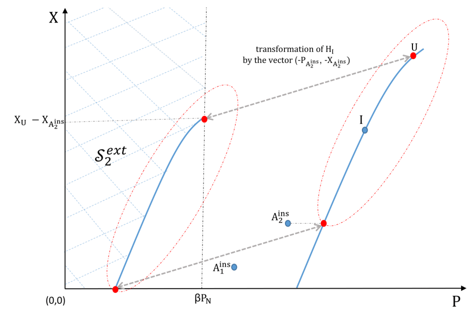
Condition (ii) requires $G(I) > G(A^{ins} \cup A^{ext})$, which is equivalent to $X_{A^{ins} \cup A^{ext}} > H_I(P_{A^{ins} \cup A^{ext}})$. As the total power and resources are the summation of power and resources, this condition implies $X_{A^{ins}} + X_{A^{ext}} > H_I(P_{A^{ins}} + P_{A^{ext}})$. To elucidate the geometrical interpretation of this condition, let first fix A_1^{ins}. We aim to determine the pairs of power and resources that the outsiders' subsets, such as A^{ext}, can have while ensuring that $A_1^{ins} \cup A^{ext}$ is *not* preferred over I. To do this, we transform the curve $X = H_I(P)$ by vector $(-P_{A_1^{ins}}, -X_{A_1^{ins}})$ and denote the transformed curve as $X = H_I^1(P)$. Then, let define $\mathcal{S}_1^{ext} = \{(P, X) \in \mathbb{R}^2_{++} | X > H_I^1(P)\}$ (Figure 5(a)). By definition, this is a subspace in which, if all the subsets of outsiders lie, there would be no alternative better than I that is constructed by A_1^{ins} and a non-empty subset of outsiders.[12] We refer to $\mathcal{S}_1^{ext}$ as the "external safe area corresponding to A_1^{ins}." Consider the external safe areas corresponding to different best sub-coalitions of the ruling coalition (for instance the area $\mathcal{S}_2^{ext}$ corresponding to A_2^{ins} in Figure 5(b) and the area $\mathcal{S}_I^{ext}$ corresponding to I in Figure 5(c)).[13] Condition (ii) would then hold if and only if all the best sub-coalitions of outsiders lie at the intersection of these regions, which we denote as $\mathcal{S}^{ext}$, i.e., the "external safe area" of ruling coalition I (i.e., $\mathcal{S}^{ext} = \mathcal{S}_1^{ext} \cap \mathcal{S}_2^{ext} \cap \mathcal{S}_I^{ext}$ in Figure 5(d)). This would ensure that there is no coalition containing a non-empty subset of outsiders that is more preferable for its members than I

[12] More rigorously, we will show that it is both necessary and sufficient that all of the best sub-coalitions of outsiders lie in this subspace.

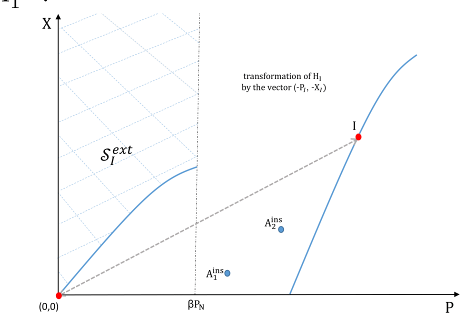
[13] Any coalition is always a best sub-coalition of itself by definition, i.e., $I \in A_I$.



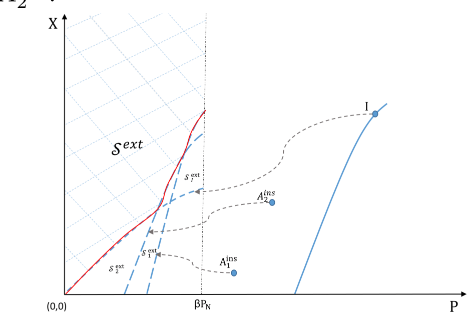
(a) Condition (ii) of Proposition 1 with respect to A_1^{ins}.



(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.

once become the ruling coalition.

Proposition 1 provides an important insight: it underscores the necessity for a ruling coalition to maintain a relatively higher power-to-resource ratio compared to the alternatives (as outlined in both internal and external safe areas).[14] For example, it offers a rationale for the voluntary destruction of wealth by a ruling coalition confronting a powerful alternative.

Another advantage of Proposition 1 is that the internal and external safe areas present natural candidates for defining the internal and external resilience of a ruling coalition to shocks impacting power and resources, respectively inside and outside it. The resilience notions we define are proportional to the size of these areas. This is because our interest is in studying the *direction* of change in the external and internal resilience of the ruling coalition. We thus focus on whether the magnitudes of the external and internal areas increase or decrease due to changes in individuals' characteristics and the underlying environment (i.e., the plundering technology). Accordingly, the resilience concepts we define

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

do not incorporate the specifications of shocks, without loss of generality.

Equally important, throughout the resilience analysis, we assume that the unique ruling coalition—characterized by Theorem 1(2), Theorem 2(2), and Proposition 1—remains unchanged. Specifically, when we address external resilience (i.e., resilience to shocks affecting the outsiders), we focus on cases where there are no internally profitable secessions for members of the ruling coalition, i.e., condition (i) of proposition 1 remains satisfied. Furthermore, for internal resilience, we restrict our attention to the internal shocks that do not change the aggregate properties of the ruling coalition, i.e., its power and resources. The unique ruling coalition in question is denoted by I in the rest of this section.

4.1.1 Internal resilience

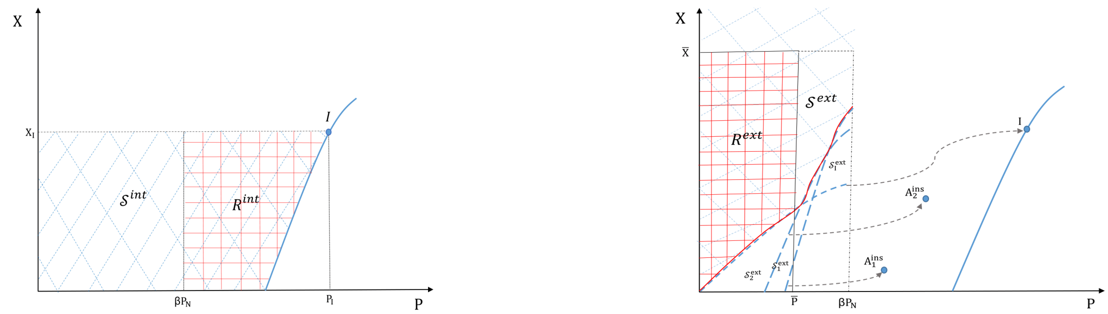
Suppose the power and resources of the players within I are due to exogenous perturbations that do not change the total power and resources of the ruling coalition. Let us assume that any best sub-coalition of I is realized within the area $[\beta P_N, P_I] \times [0, X_I] \subset \mathbb{R}^2_{++}$, by *any* posterior distribution of internal shocks. Let define $R^{int} := \mathcal{S}^{int} \cap ([\beta P_N, P_I] \times [0, X_I])$, where $\mathcal{S}^{int}$ is the internal safe area of ruling coalition I (Figure 6(a)). The internal resilience of I is then defined as $r^{int}(I) := \frac{||R^{int}||}{(P_I - \beta P_N).X_I}$. When the ruling coalition has one best sub-coalition which is realized by uniform posterior distribution of internal shock within the area $[\beta P_N, P_I] \times [0, X_I] \subset \mathbb{R}^2_{++}$, internal resilience would be the probability that the members of that sub-coalition would not prefer it over I.

4.1.2 External resilience

Suppose the power and resources of players outside I are due to exogenous shocks. Let assume that all best sub-coalitions of outsiders are realized within the region $[0, \overline{P}] \times [0, \overline{X}] \subset \mathbb{R}^2_{++}$ where $0 \leq \overline{P} \leq \beta P_N$,[15] and $\overline{X} \geq 0$, by any posterior distribution of external shocks.[16] Denote $R^{ext} = \mathcal{S}^{ext} \cap ([0, \overline{P}] \times [0, \overline{X}])$ where $\mathcal{S}^{ext}$ is the external safe area of ruling coalition I. The external resilience of I is then defined as $r^{ext}(I) := \frac{||R^{ext}||}{\overline{P}.\overline{X}}$ (Figure 6(b)). Once there is an external best sub-coalition realized by uniform distribution in

[15] This assumption is without much loss of generality because if the outside coalitions can win the majority of power, then I will not, which implies that regardless of the joint distribution inside I, I will be fragile with probability 1. But, as mentioned, we are interested in studying the cases where this does not occur and I is maintained unchanged.

[16] All of the results in this section are established regardless of the specification of this region.



(a) Internal resilience of the ruling coalition.

(b) External resilience of the ruling coalition.

Figure 6: Internal and external resilience of ruling coalition I.

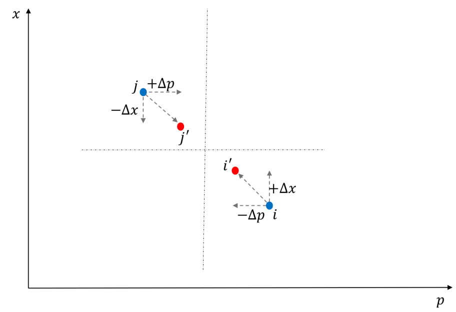


Figure 7: Exchange of power and resources between player i and player j—from blue to red.

the region $[0, \overline{P}] \times [0, \overline{X}]$, external resilience is the likelihood that no external secession—those containing some insiders and some outsiders—would be more profitable than I for its members.

4.2 Analysis: conditions of external resilience

Consider the ruling coalition I and suppose there are two players, $i, j \in I$, where $p_i > p_j$ and $x_i < x_j$. Holding the power and resources of other players fixed, let us take a portion of the power of player i, such that $0 < \Delta p \leq \frac{p_i - p_j}{2}$, or a portion of the resources of player j, such that $0 < \Delta x \leq \frac{x_j - x_i}{2}$, and transfer this power to player j, or this resource to player i (Figure 7). The following proposition establishes that performing this exchange within the ruling coalition (weakly) reduces the risk posed by stronger or poorer members to the ruling coalition. This (weakly) increases external resilience of the ruling coalition. This result is quite general and is obtained regardless of the precise specification of the plundering function $G(.)$, i.e., irrespective of the form of the indifference curves.

Proposition 2. *Suppose I is the unique ruling coalition of the game Γ, and $\exists i, j \in I$*

where $p_i > p_j$ *and* $x_i < x_j$*. Holding fixed the power and resources of players within* $I \setminus \{i, j\}$*, external resilience of* $(I \setminus \{i, j\}) \cup \{i', j'\}$ *is weakly higher than external resilience of* I*, where* $p_{i'} = p_i - \Delta p$*,* $x_{i'} = x_i + \Delta x$*,* $p_{j'} = p_j + \Delta p$*,* $x_{j'} = x_j - \Delta x$*,* $\forall 0 < \Delta p \leq \frac{p_i - p_j}{2}$*, and* $0 < \Delta x \leq \frac{x_j - x_i}{2}$.

The proof consists of three steps. In the first step, we demonstrate that two changes occur due to performing this exchange: (i) some best sub-coalitions of I will have lower power and higher resources (i.e, they become "less threatening" according to Assumption 1(1)); and (ii) some new best sub-coalitions might emerge.[17] In step 2, we argue that, first, the cases (i) would not lower external resilience of the ruling coalition according to the trade-off Assumption. Furthermore, if case (ii) occurs, there must have existed a sub-coalition before the exchange that was more threatening than the emerging best sub-coalitions (i.e., it had higher power and lower resources). This is shown by contradiction, as otherwise, the emerging best sub-coalition would have had to be a best sub-coalition before the exchange. As a result, the emerging best sub-coalitions are not more threatening than the best sub-coalitions before the exchange. This means that the emerging sub-coalitions could not decrease external resilience. Building upon steps 1-2, the final step argues that these changes in the set of best sub-coalitions would not result in a profitable internal secession. Intuitively, this is because once the ruling coalition has been able to withstand the threat of internal secession (condition (i) of Proposition 1) before the exchange, with a less threatening set of best sub-coalitions after the exchange, an internal secession would not occur as well.

A key implication of Proposition 2 is that if we repeat the exchanges of power and resources depicted in Figure 7, it results in a hierarchical structure within the ruling coalition. Under this arrangement, coalition members are categorized into distinct "ranks." Within each rank, members possess identical levels of power and resources, while across ranks, the distribution is proportional: the highest rank holds the most power and resources, followed by the second rank, and so forth. At each step of the exchange, external resilience weakly increases. Consequently, this conditionally proportional distribution demonstrates a weakly higher external resilience compared to the initial distribution before the exchanges. What

[17]It might also be the case that some best sub-coalitions do not remain the best sub-coalitions anymore. However, as we show in the formal proof in the appendix, it is evident from the definition of the external safe area in the proof of Proposition 1 that if a best sub-coalition does not remain a best sub-coalition after the exchange, it cannot decrease external resilience.

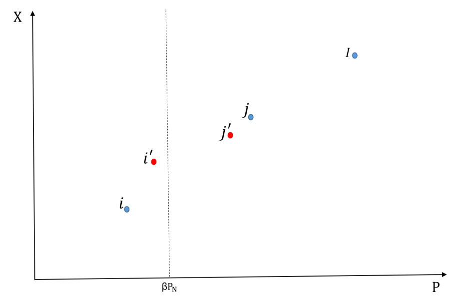


Figure 8: Ambiguity in comparing external resilience of different hierarchies.

emerges at the end of the process is a strict hierarchy, consisting of well-defined ranks. We will discuss in detail the significance of our model in thinking about hierarchies.[18]

It's important to note that further specification of the plundering function $G(.)$ is necessary to compare the external resilience of different hierarchical coalitions (i.e., various proportional distributions of power and resources within the ruling coalition). For instance, suppose we fix the total power and resources of a ruling coalition I consisting of two ranks of players, and consider two configurations of power and resources inside the ruling coalition: $\{i, j\}$ and $\{i', j'\}$, where $P_I = p_i + p_j = p_{i'} + p_{j'}$, and $X_I = x_i + x_j = x_{i'} + x_{j'}$ (Figure 8). Then, it can be simply shown that there is no *general* argument determining whether the joint distributions of power and resources $\{i, j\}$ or $\{i', j'\}$ result in greater external resilience under *any* indifference curves.

Remark 1. (A novel justification of hierarchy) As mentioned, our model offers a novel justification for the important organizational structure of hierarchies. Existing literature usually justifies hierarchy for its efficiency in incentive-alignment (Qian (1994); Mookherjee (2013); Halac et al. (2021); Halac et al. (2024)) or the division of labor (Garicano et al. (2013); Garicano and Rossi-Hansberg (2015)). As Garicano et al. (2013) highlight, the division of labor approach further includes knowledge hierarchies, team theory, and information processing. To our knowledge, our paper constructs the first model that justifies hierarchies through coalition games, bridging the two strands of largely separate literature in coalition formation and organizational economics.

Specifically, we uncover a special advantage of a strict hierarchy. Under such a structure, coalition members have minimal incentives to break away through an alliance with

[18] Notice that this analysis does not imply that, in general, a ruling coalition is most externally resilient when its members are absolutely equal. In the online appendix, we will discuss the conditions under which a single-class coalition is the most externally resilient.

outsiders. Our model formalizes this intuition by showing that a strict hierarchy is uniquely resilient against exogenous changes to outsiders. Our analysis therefore embodies another novelty in comparison with previous theories of hierarchies that predominantly focus on individuals inside the hierarchies, largely ignoring the significance of outsiders. By contrast, our justification of hierarchy focuses on how coalitions resist breakaway pressures that are generated by outsiders. This novel focus on outsiders is substantively important in many applications. For example, our analysis may justify the typically hierarchical structure of armies because these armies are uniquely powerful in prevailing against their opponents. Our analysis may also justify the historical emergence of bureaucracies, which helped rulers extract resources from subjects who were outside the bureaucracy (Weber (1978); Stasavage (2020)).

We will identify more precisely the condition for the emergence of a strict hierarchy in the next two sections. Specifically, we will characterize conditions where an exchange of power and resources in Figure 7 does not increase external resilience at all, and conditions where such an exchange strictly increases external resilience.

4.2.1 Relatively weak plundering

While performing the aforementioned exchange weakly increases external resilience, we demonstrate that when the indifference curves are convex, any exchange of power and resources within the ruling coalition does *not* increase its external resilience.[19] To understand the economic intuition behind this, we first need to understand the economic intuition of a convex indifference curve.

There is an inherent difference between concave and convex indifference curves. In environments with concave indifference curves, as the power of the ruling coalition increases, the marginal value of additional power added to the coalition decreases. Consequently, there is a relatively greater preference for a ruling coalition with lower aggregate power, or what we term "exclusive ruling coalitions." For instance, consider an environment where there is a lack of rule of law protecting outsiders from plundering by the state's insiders. In such a context, the players would be more inclined to form exclusive coalitions, aiming to keep more resources outside and secure them for extraction.

In contrast, in environments with convex indifference curves, as the power of the ruling

[19] Remember that we are restricting our attention to the exchanges that does not result in a violation of condition (i) of proposition 1.

coalition increases, the marginal value of additional power added to the coalition also increases. This leads to a tendency for the formation of an "inclusive ruling coalition" with high power. For instance, when institutions are in place that limit the extent of plundering, inclusive ruling coalitions have an advantage due to their more effective capability of overcoming these constraints.[20] Accordingly, throughout the rest of the paper, we refer to the concave and convex indifference curves as, respectively, environments with "relatively intensive plundering technology" and "relatively weak plundering technology."

The underlying economic contexts that facilitate weak and intensive plundering may be even more fundamental in nature. For example, consider a king who plunders a community of farmers, his principal source of wealth. Intensive plundering might lead to the destruction of farming infrastructures, which would be undesirable for a king needing a continuous supply of resources (Olson (1993); McGuire and Olson (1996)). Conversely, a king with access to abundant natural resources may intensively exploit society (Ross (2001)).

The above discussions point out why the internal configuration of power and resources does not impact external resilience under convex indifference curves (i.e., relatively weak plundering environments). To clarify, let's consider a ruling coalition I and any of its non-trivial best sub-coalitions $A_1^{ins} \in A_I \backslash I$. The inclination toward forming inclusive ruling coalitions in weak plundering environments means that for any best sub-coalition of society, such as $B \in A_{N \backslash I}$, the insiders of A_1^{ins} would prefer $B \cup I$ to $B \cup A_1^{ins}$. More intuitively, a higher tendency to form inclusive ruling coalitions implies that if a successful external secession occurs, it would necessarily encompass the entire ruling coalition, i.e., the ruling coalition would be more fragile with respect to itself, as a whole, compared to its non-trivial best sub-coalitions. As a result, since an exchange of powers or resources within the ruling coalition does not change the aggregate properties of the ruling coalition (i.e., its power and resources), it will have no effect on external resilience.

More formally, external resilience defined in section 4.1.1 is increasing with respect to the size of the external safe area of the ruling coalition. This area is the intersection of all internal safe areas corresponding to different best sub-coalitions of the ruling coalition. It is then straight-forward to establish that under convex indifference curves, the external safe area corresponding to the ruling coalition is a subset of those corresponding to non-trivial best sub-coalitions of the ruling coalition. For instance, in Figure 9, the external

[20] Example 3 in the online Appendix provides a more detailed explanation of environments with intensive and weak plundering.

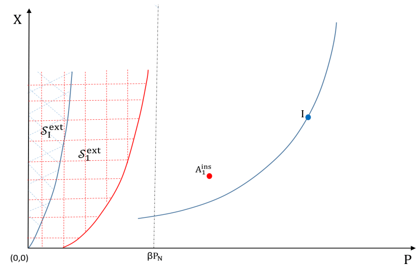


Figure 9: Comparing the external safe area for I and $A_1^{ins} \in A_I \backslash I$ under convex indifference curves.

safe area $\mathcal{S}^{ext}_{A_1^{ins}}$ corresponding to the best sub-coalition $A_1^{ins} \in A_I \backslash I$ encompasses $\mathcal{S}_I^{ext}$, i.e., the external safe area corresponding to I. By definition, I is always a best sub-coalition of itself. As a result, regardless of the power and resources of non-trivial best sub-coalitions of I, we have $\mathcal{S}^{ext} := \bigcap_{A_i^{ins} \in A_I} \mathcal{S}^{ext}_{A_i^{ins}} = \mathcal{S}_I^{ext}$ (Figure 9). The external resilience would then remain unchanged with respect to the exchange of power and resources within the ruling coalition.

Remark 2. (Fewer hierarchies with better protection of property rights) This suggests that when there are better protections of property rights in place, which may not even fully prevent plundering, a "specialized" coalition—wherein there is a separation of economic power and political power—could be highly stable against external shocks. Meanwhile, the analysis in this section shows that our previous justification for strict hierarchy does not apply when property rights are relatively well protected, even though properties are still subject to some risk of expropriation. Therefore, the model predicts that regimes with relatively good property rights are less likely to support a hierarchical bureaucracy, which is consistent with historical data showing that historical bureaucracies were more likely to emerge under autocracies (Ahmed and Stasavage (2020); Stasavage (2020)).

4.2.2 Relatively intensive plundering

We next show that when the plundering technology is relatively intensive (i.e., the indifference curves are concave), external resilience *strictly* increases once we transition to a hierarchy of well-defined ranks by repeating the aforementioned exchanges within the ruling coalition. To do this, suppose the exchange of power and resources has been repeated until only two players remain in the ruling coalition among whom this exchange can be

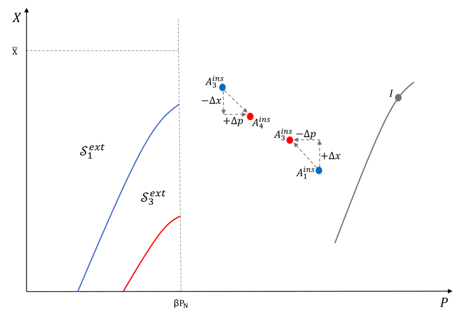


Figure 10: The increase in external resilience after repeating the exchange under a concave indifference curve.

operated. For example, there are players $i, j \in I$ where $p_i > p_j$ and $x_i < x_j$. If the exchanges thus far have already strictly increased external resilience, no further proof is needed, as the exchange between i and j will not decrease external resilience according to Proposition 2. If, however, external resilience has remained constant up to this point, we will show that the aforementioned exchange of power and resources between i and j would strictly increase external resilience of the ruling coalition.

To prove this, consider the sub-coalition $A_1^{ins} := I \setminus \{j\}$. It is straight-forward to establish that this is a best sub-coalition of I.[21] Next, as shown in Figure 10, following the exchange, A_1^{ins} would move to A_3^{ins}, which corresponds to a larger external safe area ($\mathcal{S}_3^{ext}$ compared to $\mathcal{S}_1^{ext}$). Moreover, no sub-coalition that includes j but not i—such as A_2^{ins} in Figure 10—could become a best sub-coalition after the exchange. We can simply prove this by contradiction. Suppose such a best sub-coalition, A_2^{ins}, exists. Since after the exchange we still have $p_{i'} > p_{j'}$ and $x_{i'} < x_{j'}$, replacing j' with i' in A_2^{ins} would result in a coalition with higher power and fewer resources than A_2^{ins}, which is a contradiction. Thus, although j' becomes more threatening than j, the sub-coalitions containing j' would not pose a greater threat to the ruling coalition. We have also shown in the proof of Proposition 2 that neither the coalitions that become a best sub-coalition nor the ones that are no longer a best sub-coalition after the exchange would decrease external resilience. Therefore, these imply that external resilience strictly increases as a result of this exchange under concave indifference curves.

[21] This is because, otherwise, there are two cases. First, there exists another player $l \in A_1^{ins}$ who can be replaced by j (either in $I \setminus \{j\}$ or any of its subsets, to reach a best sub-coalition with higher power and lower resources than A_1^{ins}. However, this requires $p_j > p_l$, and $x_j < x_l$. This is not possible as we initially assumed that, at this point, there are just i and j among whom we can perform this exchange. Second, a subset of $I \setminus \{j\}$ must have higher power and lower resources than it, which is not possible.

Remark 3. (More hierarchies with weak protection of property rights) This section suggests that hierarchical ruling coalitions are more likely to emerge in political environments lacking proper protections of property rights. As mentioned, this prediction is consistent with systematic historical data in Ahmed and Stasavage (2020) and Stasavage (2020).

4.3 Analysis: trade-off between internal and external resilience

Our final results investigate a prevailing argument in the literature on the political economy of non-democracies. Numerous studies have formalized the various trade-offs that authoritarian states face while addressing challenges to their rule. Some of these trade-offs, for instance the loyalty- competence trade-off, are particularly salient when confronting internal threats, such as palace coups or the collapse of their support coalition. Informational trade-offs, primarily between "informational openness" and security, deal with external threats. Our final goal is to show how the internal and external threats are related. To our knowledge, this is a novel attempt in the literature on the political economy of non-democracies, made possible through our coalition formation game and the resilience concepts we introduced. [22]

4.3.1 Relatively weak plundering

We begin by demonstrating that when plundering is relatively weak, a trade-off exists between the external and internal resilience of the ruling coalition with respect to the "intensity of plundering." Specifically, weaker plundering results in lower external resilience but higher internal resilience. This generally holds, regardless of the precise specification of shocks. To do this, we first need to formalize the concept of "intensity of plundering." As we discussed, if an environment exhibits a strictly higher marginal rate of substitution between power and resources—a higher value for power added to the ruling coalition relative to resources lost to plundering—we will call it a "lower intensity of plundering."

[22] The approach we employ in this section to tackle coalition resilience is similar to that of Pycia (2012). That is, we study resilience in a context where the preferences of players vary with an underlying and commonly known, state of nature, which is the plundering technology in this section. This is different from both our previous approach to defining the stability of coalitions which deals with changes in power and resource distributions, and the most common definition of stability in the literature on coalition formation, which is built upon the notion of "farsightedness" (Harsanyi (1974); Ray and Vohra (2015b)).

Definition 4. *The indifference curve* $H_I(.)$ *represents more intensive plundering than* $H'_I(.)$ *if and only if* $\forall P \in [\beta P_N, P_N]$, $\frac{d}{dP}[H_I(P)] < \frac{d}{dP}[H'_I(P)]$. *We also denote* $H \succ H'$.

For any fixed ruling coalition I, by abuse of notation, let us denote r_H^{ext}, r_H^{int} as, respectively, the external and internal resilience of I when the indifference curve passing I is $H_I(.)$.

Proposition 3. *Fix ruling coalition* I. *Suppose the indifference curves* $H_I(.)$ *and* $H'_I(.)$ *are convex and the distribution of internal and external shocks are independent. Then,* $H \succ H'$ *if and only if* $r_H^{int} < r_{H'}^{int}$ *and* $r_H^{ext} > r_{H'}^{ext}$.

As mentioned, in environments with relatively weak plundering, the external alternatives containing the ruling coalition are strictly more threatening than those involving non-trivial subsets of it. This result is then straightforward, given that when plundering weakens, there is a higher tendency for more inclusive coalitions. Thus, as the highly threatening alternatives necessarily involve the ruling coalition in such environments, the higher tendency for inclusive coalitions decreases external resilience. Conversely, the lower tendency for exclusive ruling coalitions increases internal resilience (Figure 11).[23]

Remark 4. (**Trade-off between resiliences with better property rights protection**) This implies that when property rights are relatively better protected, lower-intensity plundering reduces the likelihood of a coup d'etat but increases the probability of a successful popular uprising, and vice versa.[24] Thus, even imperfect property rights make the autocratic state overall more fragile; that is, the autocratic state faces a trade-off between resolving the problem of power-sharing (internal problem) and maintaining authoritarian control (external problem). This further underscores the difference between weak and intensive plundering we highlighted in the last section, suggesting that even limited differ-

[23]The formal proof of this proposition is straight-forward and is therefore omitted. As we showed, in environments with convex indifference curves, the external safe area corresponding to the ruling coalition as a whole is a subset of those corresponding to non-trivial best sub-coalitions of the ruling coalition. Thus, by definition, the external resilience only depends on the external safe area corresponding to the ruling coalition. One can then easily see that when the intensity of plundering decreases (as shown by the shift from the red to the blue indifference curve in Figure 11), the external safe area corresponding to the ruling coalition would shrink (area R_2^{ext} (blue) comparing to area R_1^{ext} (red) in Figure 11). It is also easy to see that the internal safe area would always expand under such a shift (area R_2^{int} (blue) comparing to area R_1^{int} (red) in Figure 11).

[24]Note that by "popular uprising," we are specifically referring to external secession—revolutions in which a segment of society and some insiders of the state are involved.

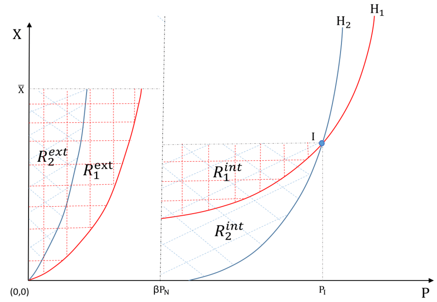


Figure 11: Trade-off between internal and external resilience under convex indifference curves.

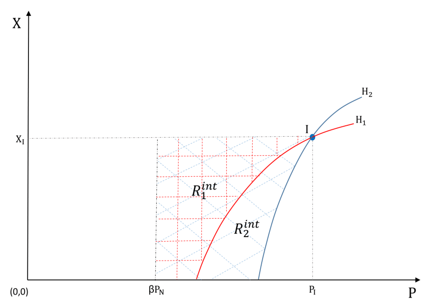


Figure 12: The increasing effect of weaker plundering on internal resilience under concave indifference curves.

ences in institutions could lead to notable differences in the state's resilience and political dynamics.

4.3.2 Relatively intensive plundering

Similar to the convex case, it is straightforward to establish that less intensive plundering increases internal resilience when indifference curves are concave (Figure 12). However, we demonstrate that the impact of changes in plundering intensity on external resilience is generally ambiguous. Example 1 illustrates how weaker plundering may result in either higher or lower external resilience, depending on the specifics of external perturbations.

Example 1. Consider a ruling coalition I and a plundering technology given by the indifference curve H_1 (that passes through I). Suppose plundering becomes weaker in the environment and shifts from H_1 to H_2. Furthermore, there is a best sub-coalition A_1^{ins} inside the ruling coalition that is realized with probability 1. Then, the effect of a move from H_1 to H_2 on external resilience is ambiguous (Figure 13), i.e., it is not clear whether the external safe area corresponding to A_1^{ins} expands or shrinks; that is, the specification of

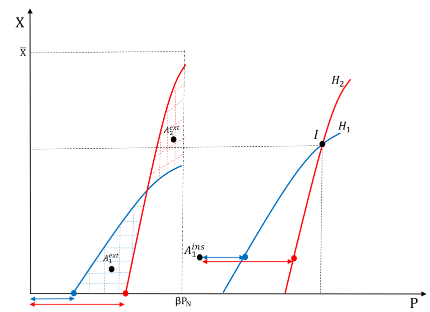


Figure 13: The ambiguous effect of a weaker plundering on external resilience.

shocks outside the ruling coalition could impact external resilience. For instance, consider a distribution of external shock $Pr_1^{ext}(.,.)$ that assigns a higher probability to the best sub-coalition A_2^{ext} than to A_1^{ext} compared to another distribution of external shock $Pr_2^{ext}(.,.)$ (Figure 13). Then, as the internal and external shocks are assumed to be independent, the probability that $A_2^{ext} \cup A_1^{ins}$ emerges is higher than $A_1^{ext} \cup A_1^{ins}$. When the plundering weakens as we shift from H_1 to H_2, the coalition $A_1^{ext} \cup A_1^{ins}$ is preferred over I, while this is not the case before the shift. This implies that external resilience is lower under the external distribution $Pr_1^{ext}(.,.)$ compared to $Pr_2^{ext}(.,.)$.

Alternatively, suppose the external shock $Pr_2^{ext}(.,.)$ gives a higher probability to A_1^{ext} than A_2^{ext} compared to Pr_1^{ext}, or equivalently, $A_1^{ext} \cup A_1^{ins}$ has a higher probability to appear than $A_2^{ext} \cup A_1^{ins}$ under $Pr_2^{ext}(.,.)$. The coalition $A_2^{ext} \cup A_1^{ins}$ is not preferred over I after the shift of the indifference curve from H_1 to H_2, while it is preferred over I before that shift. As a result, under external shock $Pr_2^{ext}(.,.)$, external resilience would be lower compared to the external shock $Pr_1^{ext}(.,.)$. Therefore, in this case, there would no longer be a trade-off between internal and external resilience. Instead, weakening plundering would increase both the external and internal resilience of the ruling coalition.

Remark 5. (**More overall resilience with weaker protection of property rights**) The above discussion suggests that in contexts with weak protections of property rights, a decrease in plundering intensity could alleviate the threats of both a coup and a revolution, particularly when economic and political power are "separated" within society. This means that an autocratic leadership seeking greater overall stability will be motivated to prevent economic elites in society from gaining political power and to hinder opposition figures from accumulating wealth. This enables the leadership to enhance both internal and external stability by reducing the intensity of plundering during periods of crisis, for instance, through "economic appeasement."

There is substantial historical evidence that dictators often attempt to mitigate dissent and consolidate mass loyalty by reducing expropriation and distributing tangible economic benefits among the general population during periods of crisis. Our theory predicts that such a policy is effective in promoting the overall stability of the state—both internal and external—only under specific conditions: first, when the protection of property rights is relatively weak within the political environment, and second, when economic elites in society are not politically powerful, and political oppositions are not wealthy. This provides a new rationale for the policy of imposing economic punishment on political opposition or activists, often adopted by autocratic regimes, from the perspective of state stability.

5 Concluding remarks

In this paper, we develop a theory of plundering coalition. We now comment 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 the nomadic tribes had a much higher level of "social cohesion" than the urban civilizations, and it is this strong social cohesion that facilitated the nomadic tribes to successfully conquer urban civilizations. Our model micro-founds the higher social cohesion of nomadic tribes by their very poverty in contrast to the urban civilizations, which produced a high power-to-resources ratio. It is therefore much 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 provides a clue to understanding the oligarchical tendencies of these plundering coalitions. Our model may explain why successful nomadic conquerors and communist parties, even when started as movements of radical equality, eventually evolved into strictly hierarchical structures.

In future research, our framework may be helpful as a methodological approach for studying the resilience of coalitions in response to exogenous changes in players' characteristics and the underlying environment of the coalition formation process. Furthermore, although we discussed the robustness of coalitions to dynamic exogenous changes in powers, resources, and plundering technology, these elements are exogenously given in our model. Endogenizing these variables could be insightful and lead to a few potential extensions. A primary extension could involve addressing the case where players endogenously determine their investment in power prior to the coalition formation process, shedding light on how the initial distribution of resources affects power investment and the ultimate coalition. It may also be intuitive to investigate the case where the ruling coalition receives exogenous fixed resources while also exploiting the outsider society. Another insightful extension of our model is to endogenize plundering technology, i.e., property rights protections, in a dynamic version of our framework where ruling coalitions at each stage could invest in changing the institutions. This contributes to the literature that formalizes how the protection of property rights emerges and evolves in various economics and political contexts (Andolfatto (2002); Hafer (2006); Guriev and Sonin (2009); Diermeier et al. (2017)), particularly from the new perspective of resilience.

6 Appendix

6.1 Proofs

Proof of Lemma 1. An indifference curve for a given utility level u is the set of combinations $I(u) = \{(P, X) \in \mathbb{R}^2 : G_i(P, X) = u\}$. To prove that the indifference curves are continuous, consider a sequence of points $\{(P_k, X_k)\}$ on an indifference curve corresponding to the utility level u. This means that for all k, $G_i(P_k, X_k) = u$. Now, suppose the sequence $\{(P_k, X_k)\}$ converges to some point (P, X). Due to the continuity of the utility function G_i, the limit of the utility levels $G_i(P_k, X_k)$ as $k \to \infty$ must be $G_i(P, X)$. Since $G_i(P_k, X_k) = u$ for all k, the limit of this constant sequence is u. Therefore, $G_i(P, X) = \lim_{k\to\infty} G_i(P_k, X_k) = u$, i.e., the point (P, X) also lies on the indifference curve corresponding to the utility level u. Thus, the indifference curves are continuous. Given Assumption 1(1), the indifference curves are also increasing for any player; as otherwise, by contradiction, either part (i) or (ii) of Assumption 1(1) must necessarily be violated. □

Proof of Theorem 1. Consider $\phi(I) := \arg\max_{W\in\mathcal{W}} G(W)$. For any power mapping $p(.)$,

$\mathcal{W}$ is not empty by definition. Moreover, we have finite players, thus $\mathcal{W}$ contains finite subsets. This implies that mapping $\arg\max_{W\in\mathcal{W}} G(W)$ is well-defined because it selects a maximum over a finite set of elements (i.e. finite winning coalitions). This implies that, first, $\arg\max_{W\in\mathcal{W}} G(W)$ is not empty so that the first part of Axiom 1 is satisfied ($\phi(I) \neq \emptyset$); Second, Axiom 2 holds since the maximization is taken over the set of winning coalitions $\mathcal{W}$. Moreover, if $I' \in \phi(I)$, then $I' \in \arg\max_{W\in\mathcal{W}} G(W)$. Therefore, if $I'' \notin \phi(I)$, we must have $G(I'') < G(I')$. Conversely, if $G(I'') < G(I')$ then $I'' \notin \arg\max_{W\in\mathcal{W}} G(W)$ which means $I'' \notin \phi(I)$. This implies that Axiom 3 is also satisfied. According to Assumptions 1-2, we have $G(N) = 0$, and $\forall W \in \mathcal{W}\backslash N, G(W) > 0$. Together, these imply that the second part of Axiom 1 is satisfied ($\phi(I) \neq N$).

For uniqueness of the mapping, by contradiction, suppose there exists another mapping ϕ' that satisfies Axiom 1-3. Consider $I \in \mathcal{W}$, and $I'' \in \phi'(I)$ such that $I'' \in \mathcal{W}$, and $I'' \notin \arg\max_{W\in\mathcal{W}} G(W)$. Now, take any $I' \in \arg\max_{W\in\mathcal{W}} G(W)$, which exists according to proof of part 1. Then, $I'' \notin \arg\max_{W\in\mathcal{W}} G(W)$ implies $G(I'') < G(I')$ which according to Axiom 3, implies that $I'' \notin \phi'(I)$. This is a contradiction as we assumed $I'' \in \phi'(I)$. Thus, we must have $I'' \in \phi(I)$. Conversely, suppose $I' \in \phi(I)$ but $I' \notin \phi'(I)$. Since $\phi'(.)$ satisfies Axiom 1, there exists a winning coalition $I'' \in \phi'(I)$. Then, Axiom 3 implies $G(I'') > G(I')$. This, however, implies $I' \notin \phi(I)$ according to definition of $\phi(.)$, which is a contradiction. Therefore, if $I' \in \phi(I)$, we must have $I' \in \phi'(I)$.

The proof of part 2 of theorem 1 is straightforward. According to Assumption 3, for all $I, I' \in \mathcal{W}$, $G(I) \neq G(I')$. Thus, the set $\arg\max_{W\in\mathcal{W}} G(W)$ is singleton, which implies the mapping ϕ is single-valued. This completes the proof of Theorem 1. □

Proof of Theorem 2(1). Under Assumptions 1-2, if $i, j \in I, I'$, then $U_i(I) \geq U_i(I')$ if and only if $U_j(I) \geq U_j(I')$. This implies that $\exists C \subseteq \mathcal{W}$ (possibly non-singleton) such that $\forall I \in C$, and $\forall i \in I$, we have $I \in \arg\max_{I\subseteq N, i\in I} U_i(I)$; and vice versa, $I \in \arg\max_{I\subseteq N, i\in I} U_i(I)$ implies $I \in C$. Moreover, under Assumption 2, $\forall i \in I$, $I \in \arg\max_{I\subseteq N} U_i(I)$ is equivalent to $I \in \arg\max_{I\subseteq N} G(I)$. This, along with the definition of ϕ in theorem 1, implies $\phi(I_0) = C$. Intuitively, C is a set of coalitions that brings the highest pay-off for its members if becomes the ruling coalition. Furthermore, according to Assumption 1(1), if $I \in C$, then $I \in Z$. This can be shown by contradiction as if $I \notin Z$, we would have $I \notin \arg\max_{I\subseteq N} U_i(I)$ according to Assumption 1(1). Thus C is a subset of Z, the set of potential ruling coalitions. By definition, Z cannot be empty. Moreover, the indifference curves are continuous according to Lemma 1. These, along with the the homogeneity of preferences implied by Assumption 2, imply that C is non-empty. We can then proceed with defining the profile σ_I that leads to the ruling coalition $I \in C$. For any $W \in \mathcal{W}$, let define $\mathcal{G}_i(W) = G(W)$ if $i \in W$, and $\mathcal{G}_i(W) = 0$ if $i \notin W$. Let consider any given history h and identify the player who is supposed to act in this history. If we are at an agenda-setting step in history h, we denote this player as $a = a(h)$. On the other hand, if we are at a voting step, where we decide on a proposal $\mathcal{P}^{a'}$ made by the agenda-setter a', we denote the voter in history h as $v = v(h)$. Furthermore, at any voting history h, let denote the set of agenda setters whose vote is already rejected, or their proposal is currently ongoing, as $A^-(h)$. Thus, the set of remaining agenda setters is denoted as $R(h) := I_0 \backslash A^-(h)$. By abuse of notation, we denote $A^-(h)$ and $R(h)$ respectively as A^- and R. In addition, for any player $i \in N$, let define $C_i = \arg\max_{I \in \mathcal{W}} \mathcal{G}_i(I)$ (i.e., $C_i = \{I \in C | i \in I\}$ is the set of ruling coalitions that brings the highest pay-off among all the ruling coalitions containing i). The strategy profile σ_I is then straightforward for players within i: to always propose I and vote 'YES' to I; and to vote 'No' to any other proposal as long as someone from $I \cap I_0$ has remained to propose. As the crucial aspect of constructing a SPE, we also need to take into account the behavior of the players off-equilibrium path. To achieve this, let define the mapping $\psi : N \backslash I \to C$, where $\psi(i) \in C_i$, i.e., $\psi(i)$ maps any player outside I to one of her favourite ruling coalitions. For any set of remaining agenda setters R, let define: $\mathcal{W}_R := \{W | \exists i \in R, W = \psi(i)\}$. Given any set of remaining agenda setters R and a player $i \in N$, let denote $m_i(R) := \arg\max_{R \cap I \neq \emptyset, \mathcal{P} \in \mathcal{W}_R} \mathcal{G}_i(\mathcal{P})$. The strategy profile σ_I that leads to the ruling coalition $I \in C$ is then defined as:

$$
\sigma_i(h) = \begin{cases} a \text{ proposes } \mathcal{P}^a(h) & \begin{cases} \mathcal{P}^a(h) = I & \text{if } a \in I \\ \mathcal{P}^a(h) = \psi(a) & \text{if } a \notin I \end{cases} \\ v \text{ votes on } \mathcal{P}^a & \begin{cases} \text{if } v \in I & \begin{cases} \text{YES} & \text{if } \mathcal{P}^{a'} = I \\ & \text{or } R \cap I = \emptyset, R \neq \emptyset \text{ and } \mathcal{G}_v(\mathcal{P}^{a'}) \geq m_v(R) \\ & \text{or } R = \emptyset \text{ and } \mathcal{G}_v(\mathcal{P}^{a'}) \geq \mathcal{G}_v(I_0) \end{cases} \\ \text{if } v \notin I & \begin{cases} \text{YES} & \text{if } \mathcal{P}^{a'} = \psi(i) \\ & \text{if } R \cap \psi(i) = \emptyset, R \neq \emptyset \text{ and } \mathcal{G}_v(\mathcal{P}^{a'}) \geq m_v(R) \\ & \text{or } R = \emptyset \text{ and } \mathcal{G}_v(\mathcal{P}^{a'}) \geq \mathcal{G}_v(I_0) \end{cases} \\ \text{NO,} & \text{Otherwise} \end{cases} \end{cases} \tag{6.1}
$$

We prove the strategy profile (3) is SPE in two steps. The only cases that a profitable deviation might exists among voting histories is when the votes are pivotal. Thus, throughout the proof, we only consider the histories where an agenda setter is picked to propose or there is a voter whose vote is pivotal. Moreover, as the stages are finite, it suffices to show that there is no one-shot profitable deviation for any player at any history of the game. By induction, we first show that $\{\sigma_i\}_{i\in N}$ is an SPE for the (potentially off-path equilibrium) subgames where no agenda setter in I remains to propose (step (i)). We then proceed by demonstrating that this holds true for all other subgames (step (ii)).

Step (i). First, let consider the subgame when the last voter v_0 is voting to the proposal $\mathcal{P}^{a_0}$ made by the last agenda-setter a_0. The action induces by the strategy $\sigma_{v_0}(h)$ for player v_0 is to accept this proposal if $\mathcal{G}_{v_0}(\mathcal{P}^{a_0}) \geq \mathcal{G}_{v_0}(I_0)$. Clearly, there is no profitable deviation from this action for player v_0 since reaching to this history means ruling coalition will be either I_0 or $\mathcal{P}^{a_0}$. By induction, the same argument is true for all the voters to the last proposal. Furthermore, given the actions played by the voters to the last proposal, there is no profitable deviation from proposing any $\mathcal{P}^{a_0} = \psi(a_0)$ for the last agenda setter a_0, which is the action induced by the strategy σ_{a_0}. The definition of $\psi(.)$ and the similarity of preferences then imply that $\psi(a_0)$ will be the ruling coalition in the subgame Γ_{a_0}, i.e., the subgame starting from the last agenda setter a_0.

Let us first consider the voting history h_{k+1} when there is a voting over the proposal of the player a_{k+1}, while the set of remaining agenda setters is $I_0 \backslash A^- := \{a_0, a_1, ..., a_k\}$

and $\forall 1 \leq i \leq k, a_i \notin I$, i.e., all the remaining agenda setter are outside I. Also suppose $a_{k+1} \notin I$. Let assume the strategy profiles $\{\sigma_i\}_{i\in N}$ is a SPE on all the subgames of Γ_{a_k} (induction's assumption). Let v_{k+1} be the last voter to the proposal $\mathcal{P}_{a_{k+1}}$ made by a_{k+1}. According to the induction's assumption, in the continuation of the game, strategies σ_i induces a SPE and a coalition $\mathcal{P}^{a_i} \in \mathcal{W}_{\{a_i|1\leq i\leq k\}}$ is the ruling coalition for some $1 \leq i \leq k$. Let denote this ruling coalition as $W := \mathcal{P}^{a_i}$. Notice that the strategy $\sigma_{v_{k+1}}$ induces voting YES to $\mathcal{P}^{a_{k+1}}$ when $\mathcal{G}_{v_{k+1}}(\mathcal{P}^{a_{k+1}}) \geq m_{v_{k+1}}(\{a_0, a_1, ..., a_k\})$. By definition, $m_{v_{k+1}}(\{a_0, a_1, ..., a_k\}) = \arg\max_{\{a_0,a_1,...,a_k\}\cap I\neq\emptyset, \mathcal{P}\in\mathcal{W}_R} \mathcal{G}_{v_{k+1}}(\mathcal{P})$, i.e., given that the set of remaining agenda setter is $\{a_0, a_1, ..., a_k\}$, $m_{v_{k+1}}(\{a_0, a_1, ..., a_k\}$ is the highest pay-off that is feasible for v_{k+1} from the coalitions containing v_{k+1} that are put forward shortly by the remaining agenda setters $\{a_0, a_1, ..., a_k\}$.

Now, there are two cases, if $v_{k+1} \in W$, due to the consistency of preferences imposed by Assumption 2, we have $\mathcal{G}_{v_{k+1}}(W) = m_{v_{k+1}}(\{a_0, a_1, ..., a_k\})$, i.e., the coalition W brings the highest payoff for v_{k+1} as well. Thus, in this case, there will be no profitable deviation from strategy $\sigma_{v_{k+1}}$ for v_{k+1}, in that the only optimal action for player v_{k+1} is to vote 'NO' to $\mathcal{P}^{a_{k+1}}$ if it does not bring the highest pay-off for v_{k+1}; and otherwise, both voting 'YES' and 'No' will be optimal for player v_{k+1}; which is the action induced by the strategy $\sigma_{v_{k+1}}$. On the other hand, when $v_{k+1} \notin W$, W brings the lowest pay-off for v_{k+1} by definition of $\mathcal{G}_{v_{k+1}}$. As a result, there will also be no profitable deviation from strategy $\sigma_{v_{k+1}}$ in this case for the player v_{k+1}. This follows because if $v_{k+1} \notin W$, either $\mathcal{G}_{v_{k+1}}(\mathcal{P}^{a_{k+1}}) > \mathcal{G}_{v_{k+1}}(W)$ or $\mathcal{G}_{v_{k+1}}(\mathcal{P}^{a_{k+1}}) = \mathcal{G}_{v_{k+1}}(W)$, there is no profitable deviation from voting 'YES' to $\mathcal{P}^{a_{k+1}}$ which is the action induced by the strategy $\sigma_{v_{k+1}}$ in history h_{k+1}. (Notice that it could not be $\mathcal{G}_{v_{k+1}}(\mathcal{P}^{a_{k+1}}) < \mathcal{G}_{v_{k+1}}(W)$ when $v_{k+1} \notin W$, according to definition of $\mathcal{G}_{v_{k+1}}(.)$.)

Finally, let's consider the agenda-setting step where the player a_{k+1} is chosen by Nature to propose. The discussion above showed that there is no profitable deviation from strategies (3) in subgames of Γ_{a_k}, and those starting from the voters on the proposal of a_{k+1}. The consistency of preferences then implies that if $\mathcal{P}^{a_{k+1}} = \psi(a_{k+1})$, the voters will vote YES to this proposal according to strategies (3). Given this, as the proposal $\mathcal{P}^{a_{k+1}} = \psi(a_{k+1})$ brings the highest payoff for the player a_{k+1} by the definition, there will be no profitable deviation from the strategies (3) for the agenda setter a_{k+1}. This completes the proof of step (i).

Step (ii). Now, suppose $\exists a \in I, a \in I_0\backslash A^-$, i.e., the set of remaining agenda setters has at least one player from I. Then we show that there is no profitable deviation from strategies (3) for the players in any such history of the game. First, note that for any player in I, the strategy profile (3) induces proposing I and only voting YES to I in all the subgames that the proposal of at least one agenda setter in I has not been rejected yet or is ongoing. Consider a history when a player $i \in I$ is picked by the Nature to act. Whether this history is voting or agenda setting, the strategies (3) make I the ruling coalition in the continuation of the game. As a result, any players in I does not have a profitable deviation from strategies (3) (whether by voting NO or proposing some other coalitions). Moreover, as $\beta \in [\frac{1}{2}, 1]$, any other proposal made by other agenda setters $\mathcal{P} \neq I$ will be rejected by strategies $\{\sigma_i\}_{i \in I}$ in any such stages, i.e., where there is an ongoing voting over $\mathcal{P} \neq I$, and $(I_0 \backslash A^-) \cap I \neq \emptyset$. This follows because $\beta \in [\frac{1}{2}, 1]$ implies $\mathcal{P} \cap I \neq \emptyset$. Let $j \in \mathcal{P} \cap I$. The unanimity rule then implies that a voting stage necessarily reaches that the Nature chooses j to vote to $\mathcal{P}$, whose vote is 'NO' to the proposal $\mathcal{P} \neq I$ according to the strategies (3). Clearly, as I brings the highest pay-off for its players, they will have no strictly profitable deviation from (3) at such histories. Intuitively, the degree majority greater than a half and unanimity rule enable the players within I to stop any other coalition to form. This completes the proof of step (ii). As a result, step (i)-(ii) implies there is no profitable deviation for strategies (3) in *any* history of the game, which completes the proof of theorem 2.

$\square$

Proof of Theorem 2(2). We prove this part in two steps. In step (i), we show that in the subgame that starts from the first agenda setter in $I \cap I_0$ proposing a coalition (who exists as $\beta \in [\frac{1}{2}, 1]$), the ruling coalition will necessarily be $\phi(I_0) = I$ regardless of the moves by Nature. In step (ii), we show that in the subgames where no player from $I \cap I_0$ has proposed yet, no new coalition forms in any SPE of the game until the first agenda setter in I proposes. These imply that I must be the ruling coalition in any SPE of the game.

Step (i). First, note that according to Theorem 1, $\phi(I_0) = \arg\max_{A \in \mathcal{W}} G(A)$. Thus, $\phi(I_0) = \{I\}$, along with Assumptions 1-2, imply that I is the unique coalition that brings the highest payoff for all its members. Now, consider the history h_a when the first proposer from the set $\in I_0 \cap I$, namely a, is picked by Nature to propose a coalition $\mathcal{P}_a$. Let us denote the subgame starting from this history as Γ_a. Note that player a exists as $\beta \in [\frac{1}{2}, 1]$ (which leads to $I_0 \cap I \neq \emptyset$). Furthermore, we distinguish two types of subgames: first, the subgame that start from the history after the action $\mathcal{P}_a = I$, and the subgames starting from the history after the action $\mathcal{P}_a \neq I$, made by player a. Let denote these respectively as Γ_a^I and $\Gamma_a^{\neg I}$. Consider the subgame Γ_a^I and any SPE that involves playing $\mathcal{P}_a = I$. Denote the player v be as last player that votes to I in history h_v while all the last voters has voted YES to I. There are two cases if v votes NO to I: whether I is the ruling coalition in the continuation of the game—in the subgame that starts from the agenda setter after a denoted as Γ_{-v}— or not. In the first case, the coalition I will be the ruling coalition in the subgame Γ_v regardless of the vote of v (where Γ_v is the subgame starting form history v). In the latter case, the only optimal vote of v in h_v must be "YES". This also makes I the ruling coalition of Γ_v. By induction, for the other voters to I who has voted earlier, there would not be a profitable deviation from voting YES to I regardless of the order of votes. This implies that the coalition I certainly becomes the ruling coalition of the subgame Γ_a^I. Now, let's consider the subgame $\Gamma_a^{\neg I}$. If I is the ruling coalition of $\Gamma_a^{\neg I}$, then player a regardless of the proposal of a, the coalition I would be the ruling coalition of Γ_a; since we showed I is the ruling coalition of the subgame Γ_a^I. Otherwise, if another coalition is the ruling coalition in $\Gamma_a^{\neg I}$, the only optimal action of the player a in any SPE of the game would be to propose I because I uniquely brings the highest payoff for the members (and would necessarily become the ruling coalition of the subgame Γ_a^I). This implies that in any SPE of the game, I must be the ruling coalition of the subgame Γ_a

Step (ii). Now, lets consider any history before the history h_a. As the ruling coalition of subgame Γ_a is necessarily I, then for any other proposal $\mathcal{P}_{a'} \neq I$, where $\mathcal{P}_{a'} \in \mathcal{W}$, made by any player $a' \in I_0 \setminus \{a\}$ before the history h_a, there will exist a player $i \in I \cap \mathcal{P}_{a'}$ (as $\beta \in [\frac{1}{2}, 1]$), whose only optimal decision in any SPE of the game is to vote 'No' to $\mathcal{P}_{a'} \neq I$ whenever her vote is pivotal.[25] Since any coalition forms according to the unanimity rule, the vote of any such player i would stop any proposal $\mathcal{P} \neq I$ from becoming the ruling coalition before the history h_a. Furthermore, if $\mathcal{P}_{a'} \notin \mathcal{W}$, the proposal cannot win the majority of power and becomes the ruling coalition. As a result, in *any* SPE of the game, no coalition can form until the first agenda setter in I, the player a, proposes. This completes the proof of Theorem 2(2).

$\square$

Proof of proposition 1. According to theorem 1, if I is the ruling coalition of the game, we have $I \in \arg\max_{W \in \mathcal{W} G(W)}$. Without loss of generality, suppose $A_I = \{A_1^{ins}, A_2^{ins}, .., A_m^{ins}\}$

[25]Otherwise, if her vote is not pivotal, $\mathcal{P}$ has already been rejected due to the unanimity rule.

where $\forall 1 \leq i \leq m, A_i^{ins} \in \mathcal{W}$ and $A_m^{ins} = I$; and $A_{N\backslash I} = \{A_1^{ext}, A_2^{ext}, .., A_s^{ext}\}$ where $A_s^{ext} = N\backslash I$. Condition (i) says $G(I) > G(A)$ for all $A \subset I$ (there is no profitable internal secession); and condition (ii) states $G(I) > G(A)$ for all $A \subseteq N$ (there is no external profitable deviation).

Condition (i). Consider $A \subset I$. First, $G_{A_i^{ins}} < G_I, \forall 1 \leq i \leq m$, implies $G_A < G_I$. This is because according to the definition 3, if $A \subset I$ and $A \notin A_I \backslash I$, then there exists $A_i^{ins} \in A_I$ such that we have $P_{A_i^{ins}} > P_A$ and $X_{A_i^{ins}} < X_A$. The pay-off Assumption will then imply $G_{A_i^{ins}} > G_A$, which implies $G_A < G_I$. Additionally, by definition of indifference curves, $G_{A_i^{ins}} < G_I$ is equivalent to $H_I(P_{A_i^{ins}}) < X_{A_i^{ins}}$. This implies that condition (i) holds if and only if $A_i^{ins} \in \mathcal{S}^{int}$, where $\mathcal{S}^{int} = \big\{(P, X) \in \mathbb{R}_{++}^2 | X > H_I(P)\big\}$.

Condition (ii). First, $G_{A_i^{ins} \cup A_j^{ext}} < G_I$, $\forall 1 \leq i \leq m$, $1 \leq j \leq s$, then $G_A < G_I$, $\forall A \subseteq N$ where $A \nsubseteq I$. Since $A \nsubseteq I$, there exist non-empty coalitions $A^{ins} \subseteq I$ and $A^{ext} \subseteq N\backslash I$ such that $A = A^{ins} \cup A^{ext}$. According to Definition 3, if $A \subseteq N, A \nsubseteq I$ and either $A^{ins} \notin A_I \backslash I$ or $A^{ext} \notin A_{N\backslash I}$, then there exist $1 \leq i \leq m$, $1 \leq j \leq s$ such that we have $P_{A_i^{ins}} + P_{A_j^{ext}} > P_A$ and $X_{A_i^{ins}} + X_{A_j^{ext}} < X_A$. The pay-off Assumption then implies $G_{A_i^{ins} \cup A_j^{ext}} > G_A$. Thus, $A \notin \arg\max_{W \in \mathcal{W}} G(W)$, which implies $G_A < G_I$ since $I \notin \arg\max_{W \in \mathcal{W}} G(W)$. As a result, for $G_A < G_I$ ($\forall A \subseteq N$ where $A \nsubseteq I$) to hold, it is sufficient that $G_{A_i^{ins} \cup A_j^{ext}} < G_I$, $\forall 1 \leq i \leq m$, $1 \leq j \leq s$. Now, fix i. Then, by definition of indifference curves, $G_{A_i^{ins} \cup A^{ext}} < G_I$ implies $X_{A_i^{ins}} + X_{A^{ext}} > H_I(P_{A_i^{ins}} + P_{A^{ext}})$. Let transform the curve $X = H_I(P)$ by vector $(-P_{A_i^{ins}}, -X_{A_i^{ins}})$ and denote the transformed curve as $X = H_I^i(P)$. Consider the region $S_i^{ext} = \big\{(P, X) \in \mathbb{R}_{++}^2 | X > H_I^i(P)\big\}$. by definition, we have $X_{A_i^{ins}} + X_{A^{ext}} > H_I(P_{A_i^{ins}} + P_{A^{ext}})$ if and only if $A^{ext} \in S_i^{ext}$. Let denote $\mathcal{S}^{ext} = \bigcap_{A_i^{ins} \in A_I} \mathcal{S}_i^{ext}$. Therefore, if $\forall A^{ext} \subseteq N\backslash I$, $A^{ext} \in \mathcal{S}^{ext}$, we have $X_{A^{ins}} + X_{A^{ext}} > H_I(P_{A^{ins}} + P_{A^{ext}})$, which implies $G_A < G_I$, $\forall A \subseteq N$ where $A \nsubseteq I$. □

Proof of proposition 2. We prove the proposition 2 in three steps. The prove is obtained irrespective of the form of the indifference curves.

Step 1: First, once we transfer of power and resources as outlined in Figure 7 within the society of outsiders, the player i remains more threatening than player j, i.e., her power would remain higher and her resources would be lower both before and after the exchange. Then, according to the definition of the best sub-coalitions of A_I, there would be no best sub-coalition that only contains j but not i. This is because, by contradiction, if this is the case and there is a best sub-coalition A' that contains j but not i, we can substitute j with

i before the exchange, or j' with i' after the exchange, in A' to get a coalition with higher power and lower resources than A'. This is a contradiction according to the definition of the set of best sub-coalitions (given by definition 3). This means that all of the best sub-coalitions before (after) the exchange would contain either both players i, j (players i' and j'), only the player i (player i'), or neither of them. This implies that there could be three potential changes in the set of best sub-coalitions after the exchange in Figure 7: (i) the best sub-coalitions that contain i but not j would move to the left and up due to the exchange; (ii) the best sub-coalitions that contain both i, j or neither of them would not move due to the exchange; (iii) some best sub-coalitions might disappear, and some new best sub-coalitions might emerge due to the exchange. Regarding (i), we will show in step 2 that such a move would increase external resilience of the ruling coalition regardless of the specifications of the function $G(.)$ in Assumption 1. This is because, loosely speaking, the move to a best sub-coalition with higher resources and lower power would make it a less profitable coalition for the members according to Assumption 1(1) (i.e., "less threatening"). Thus, external resilience of the ruling coalition increases with respect to such a move.

Regarding (ii), consider a best sub-coalition B before the exchange that contains i, j or neither of them. The power and resources of B would not change due to the exchange. This means that the external safe area corresponding to B would not change, $\mathcal{S}_B^{ext} = \{(P, X) \in \mathbb{R}^2_{++} | H_I(P_B + P) - X_B < X\}$. External resilience is defined as the intersection of all external safe areas corresponding to different best sub-coalitions of the society. Therefore, external resilience would not be impacted by changes in B. For (iii), according to the definition of external resilience, if a best sub-coalition does not remain a best sub-coalition after the exchange, this would weakly increase external resilience. Furthermore, notice that as mentioned, if a best sub-coalition C emerge after the exchange, it cannot only contain j' but not i'. Also, by contradiction, it can be easily show that C cannot contain i' but not j'. This is because, otherwise, C must be a best sub-coalition before the exchange. As a result, C either contains both i' and j' or neither of them. This implies that, necessarily, C emerges due to the move of a best sub-coalition such as A, which contains i but not j, to left and up; where before the exchange, we have: $P_A > P_C$, and $X_C > X_A$. Since $P_A > P_C$, and $X_C > X_A$ before the exchange, and also C does not move due to the exchange, according to step 2, we would have: $S_A^{ext} \subset S_C^{ext}$, where S_A^{ext} is the external safe area corresponding to A before the exchange, and S_C^{ext} is the external safe area corresponding to C both before and after the exchange (which are equal since we established that C does not move due to the exchange). This, by definition, implies that external resilience would not decrease after the exchange due to the emergence of C.

Step 2: Now, we aim to prove the argument we used in step 1; that is, external resilience would be lower with respect to a best sub-coalition with higher power and lower resources, regardless of the specification of the function $G(.)$ satisfying Assumption 1. To prove, notice that according to the proof of proposition 1 and definition of external resilience, for any best best sub-coalition of insiders such as $A_1^{ins} \in A_I$, there is an external safe area $\mathcal{S}_1^{ext}$ such that if all the external best sub-coalitions lie on $\mathcal{S}_1^{ext}$, there would be no profitable deviation from I to a coalition that contains A_1^{ins} and a best sub-coalition outside the ruling coalition. More Specifically, we have $G(I) > G(A_1^{ins} \cup A_1^{ext}), \forall A_1^{ext} \in N \backslash I$. By definition of indifference curves, this implies $H_I(P_{A_1^{ext}} + P_{A_1^{ins}}) < X_{A_1^{ext}} + X_{A_1^{ins}}, \forall A_1^{ext} \in N \backslash I$. Thus, the external safe area corresponding to A_1^{ins} is given by $\mathcal{S}_1^{ext} = \Big\{(P, X) \in \mathbb{R}_{++}^2 | H_I(P_{A_1^{ins}} + P) - X_{A_1^{ins}} < X\Big\}$. Let us now transform $X = H_I(P)$ by the vector $(-P_{A^{ins}}, -X_{A^{ins}})$, and denote the curve after the transformation as $X = H_I^1(P)$. Then, we have $\mathcal{S}_1^{ext} = \big\{(P, X) \in \mathbb{R}_{++}^2 | H_I^1(P) < X\big\}$. Now suppose A_1^{ins} moves to left and up in the (P, X) space, i.e., its power becomes lower and its resources increases. Let denote the coalition after the move as A_2^{ins}, and the external safe area corresponding to it as $\mathcal{S}_1^{ext} = \big\{(P, X) \in \mathbb{R}_{++}^2 | H_I^2(P) < X\big\}$. We next show that $H_I^1(P) = X$, and $H_I^2(P) = X$ could not cross in the space $(P, X) \in \mathbb{R}_{++}^2$. More formally, there does not exist a $P > 0$ such that $H_I^1(P) = H_I^2(P) = X > 0$. By contradiction, suppose there exists $P^* > 0$ such that $H_I^1(P^*) = H_I^2(P^*) = X^* > 0$. Then, by definition, we have $H_I(P^* + P_{A_1^{ins}}) - X_{A_1^{ins}} = H_I(P^* + P_{A_2^{ins}}) - X_{A_2^{ins}}$. Since $X_{A_2^{ins}} > X_{A_1^{ins}}$, we have $H_I(P^* + P_{A_1^{ins}}) < H_I(P^* + P_{A_2^{ins}})$. Also, we have: $P_{A_1^{ins}} > P_{A_2^{ins}}$, which means $P_{A_1^{ins}} + P^* > P_{A_2^{ins}} + P^*$. But this is a contradiction since according to Assumption 1(1), we must have $H_I(P^* + P_{A_1^{ins}}) > H_I(P^* + P_{A_2^{ins}})$ because $P_{A_1^{ins}} + P^* > P_{A_2^{ins}} + P^*$. [26] Finally, for all $P > 0$ such that $H_I^1(P) > 0$, and $H_I^2(P) > 0$, we have $H_I^2(P) < H_I^1(P)$. This is straight-forward according to the definition of H_I^1 and $H_I^2(P)$; which imply $H_I^1(P^* + P_{A_1^{ins}}) = H_I(P^* + P_{A_1^{ins}}) - X_{A_1^{ins}} > H_I(P^* + P_{A_1^{ins}}) - X_{A_2^{ins}} = H_I^2(P^* + P_{A_1^{ins}})$. Based on the definition of external safe area, we then have $\mathcal{S}_1^{ext} \subseteq \mathcal{S}_2^{ext}$. This completes the proof of this step.

Step 3. So far, we established that the aforementioned exchange weakly increases

[26]Intuitively, if the boundary of the areas $\mathcal{S}_1^{ext}$ and $\mathcal{S}_2^{ext}$ cross, the indifferent curves must cross which violates Assumption 1.

external resilience. The final step of the proof is to show that after performing the exchange inside the ruling coalition, no best sub-coalitions emerge that can violate condition (i) of the Proposition 1, i.e., does not result in an internal profitable secession. In step 1, we discussed that there are three changes to the set of best sub-coalitions due to the exchange: (i) the best sub-coalitions that contain i but not j would move to the left and up due to the exchange; (ii) the best sub-coalitions that contain both i, j or neither of them would not move due to the exchange; (iii) some best sub-coalitions might disappear, and some new best sub-coalitions might emerge due to the exchange. By definition of internal safe area, it is clear that the move of best sub-coalitions of I, could not result in condition (i) of proposition 1. Furthermore, it is clear that the best sub-coalitions that disappear due to this exchange could not not result in violation of this condition. It thus only remains to prove that emergence of new best sub-coalitions does not violate condition (i) of proposition 1. To prove, notice that in step 2, we demonstrated that corresponding to any emerged best sub-coalition such as C (that must only contain both i, j or neither of them according to step 1), there must exists a best sub-coalition containing i but not j such as A that has had a higher power and lower resources than C before the exchange. If I is a ruling coalition before the exchange, we cannot have $A \notin \mathcal{S}^{int}$, where $\mathcal{S}^{int}$ is the internal safe area defined in the proof of proposition 1. As according to Assumption 1(1), the indifference curves are increasing, $A \notin \mathcal{S}^{int}$ implies $C \in \mathcal{S}^{int}$. $\square$

References

Acemoglu, D., Egorov, G., and Sonin, K. (2008). Coalition formation in non-democracies. *The Review of Economic Studies*, 75(4):987–1009.

Acemoglu, D., Egorov, G., and Sonin, K. (2012). Dynamics and stability of constitutions, coalitions, and clubs. *American Economic Review*, 102(4):1446–1476.

Acemoglu, D. and Robinson, J. A. (2008). Persistence of power, elites, and institutions.

Acemoglu, D. and Robinson, J. A. (2013). Economics versus politics: Pitfalls of policy advice. *Journal of Economic perspectives*, 27(2):173–192.

Ahmed, A. T. and Stasavage, D. (2020). Origins of early democracy. *American Political Science Review*, 114(2):502–518.

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

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

Bai, Y. and Zhou, T. (2019). âmaoâs last revolutionâ: a dictatorâs loyalty–competence tradeoff. *Public Choice*, 180(3):469–500.

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

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

Besley, T. J. and Kudamatsu, M. (2007). Making autocracy work. *LSE STICERD Research Paper No. DEDPS48*.

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

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

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

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

Diermeier, D., Egorov, G., and Sonin, K. (2017). Political economy of redistribution. *Econometrica*, 85(3):851–870.

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

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

Egorov, G. and Sonin, K. (2024). The political economics of non-democracy. *Journal of Economic Literature*, 62(2):594–636.

Enikolopov, R., Makarin, A., and Petrova, M. (2020). Social media and protest participation: Evidence from russia. *Econometrica*, 88(4):1479–1514.

Francois, P., Rainer, I., and Trebbi, F. (2015). How is power shared in africa? *Econometrica*, 83(2):465–503.

Garicano, L. (2000). Hierarchies and the organization of knowledge in production. *Journal of political economy*, 108(5):874–904.

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

Garicano, L., Van Zandt, T., et al. (2013). Hierarchies and the division of labor. *The Handbook of Organizational Economics*, pages 604–654. edited by Robert Gibbons and John Roberts.

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

Guriev, S. and Sonin, K. (2009). Dictators and oligarchs: A dynamic theory of contested property rights. *Journal of Public Economics*, 93(1-2):1–13.

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

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

Halac, M., Kremer, I., and Winter, E. (2024). Monitoring teams. *American Economic Journal: Microeconomics*, 16(3):134–161.

Halac, M., Lipnowski, E., and Rappoport, D. (2021). Rank uncertainty in organizations. *American Economic Review*, 111(3):757–786.

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

Henn, S. J., Mastaki Mugaruka, C., Ortiz, M., Sánchez de la Sierra, R., and Qihang Wu, D. (2024). Monopoly of taxation without a monopoly of violence: The weak state's trade-offs from taxation. *Review of Economic Studies*, page rdae036.

Jia, R., Kudamatsu, M., and Seim, D. (2015). Political selection in china: The complementary roles of connections and performance. *Journal of the European Economic Association*, 13(4):631–668.

Khaldun, I. (1967). The muqaddimah, an introduction to history, tr. by franz rosenthal, 3 vols.

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

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

Li, W. (2023). Dual leadership and bureaucratic capacity. *Monash University Working Pap.*

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

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

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

McGuire, M. C. and Olson, M. (1996). The economics of autocracy and majority rule: The invisible hand and the use of force. *Journal of economic literature*, 34(1):72–96.

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

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

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

Olson, M. (1993). Dictatorship, democracy, and development. *American political science review*, 87(3):567–576.

Paine, J. (2021). The dictator's power-sharing dilemma: Countering dual outsider threats. *American Journal of Political Science*, 65(2):510–527.

Puga, D. and Trefler, D. (2014). International trade and institutional change: Medieval veniceâs response to globalization. *The Quarterly Journal of Economics*, 129(2):753–821.

Pycia, M. (2012). Stability and preference alignment in matching and coalition formation. *Econometrica*, 80(1):323–362.

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

Qian, Y., Roland, G., and Xu, C. (2006). Coordination and experimentation in m-form and u-form organizations. *Journal of Political Economy*, 114(2):366–402.

Ray, D. and Vohra, R. (2015a). Coalition formation. In Young, H. P. and Zamir, S., editors, *Handbook of Game Theory with Economic Applications*, pages 240–322. North Holland.

Ray, D. and Vohra, R. (2015b). Coalition formation. *Handbook of game theory with economic applications*, 4:239–326.

Ray, D. and Vohra, R. (2015c). The farsighted stable set. *Econometrica*, 83(3):977–1011.

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

Roemer, J. E. (1981). *Analytical foundations of Marxian economic theory*. Cambridge University Press.

Ross, M. L. (2001). Does oil hinder democracy? *World politics*, 53(3):325–361.

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

Shadmehr, M. (2018). Multiplicity and uniqueness in generalized regime change games. *Available at SSRN 3121241*.

Stasavage, D. (2020). *The Decline and Rise of Democracy: A Global History from Antiquity to Today*. Princeton University Press.

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

Weber, M. (1978). *Economy and Society: An Outline of Interpretive Sociology*. University of California Press. edited by Guenther Roth and Claus Wittich.

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

Wintrobe, R. (2000). *The political economy of dictatorship*. Cambridge University Press.

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

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

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

Supplement to "Plundering Coalitions"

Alireza Mohammadi Sepahvand* Weijia Li†

November 22, 2024

Appendix A

A.1 Single-class coalitions

The analysis of Section 4 does not imply that, in general, a ruling coalition is the most externally resilient if its members are absolutely equal. An interesting phenomenon to investigate is when transitioning to a single-class coalition—where all players within the ruling coalition possess equal power and resources—would enhance external resilience. Although a general argument cannot be made for this case without further specifying the internal joint distribution and the plundering function, one instance where this occurs is when the most powerful member of the ruling coalition has the lowest resources, the second-most powerful players has the second-lowest resources, etc. It is then straightforward to establish that in such a case, we can perform the exchange depicted in Figure 7 among the players in a way that leads to a single-class coalition.[1] The following examples elucidates a more general case where the exchange of power and resources depicted in Figure 1 within the ruling coalition also increases external resilience. This, along with Proposition 2, would give rise to a broad range of single-class coalitions.

Example 1. Suppose there is a ruling coalition $I = \{i, j\}$ where $p_i < p_j$ and $x_i < x_j$. Suppose the indifference curve passing form I is concave and denote it as $H_I(.)$. Suppose player i is more threatening for the ruling coalition than j, i.e., we have $\mathcal{S}_i^{ext} \subset \mathcal{S}_j^{ext}$.

*Monash University, Australia. Email: alireza.mohammadisepahvand@monash.edu.

†Monash University, Australia. E-mail: weijia.li@monash.edu.

[1]Precisely, we can start by performing the exchange between the two players with the highest power until they become equal in both power and resources, then continue with the player with the third-highest power until the three players become equal in power and resources, and so forth.

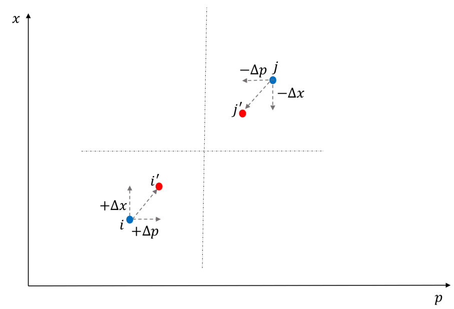


Figure 1: Exchange of power and resources between player i and player j—from blue to red.

This means that the external safe area of the ruling coalition is the external safe area of player i. Now, if we perform the exchange outlined in Figure 1 between i and j, and we have $\mathcal{S}_i^{ext} \subset \mathcal{S}_{i'}^{ext}$ (the red area covers the blue area in Figure 2), the external safe area of the ruling coalition expands due to this exchange. Loosely speaking, this occurs because player i is much more threatening than player j given the indifference curve H_I. Thus, after the exchange, although taking some resources from player j might potentially make her more threatening than before (moving from j to i'), this effect is not as strong as the decrease in the threat posed by player i due to having more resources after the exchange. This is reinforced when the plundering is highly intensive (i.e., the indifference curve is highly concave) and the players with low resources are highly threatening (i.e., they have relatively high power-to-resource ratios).

Remark 1. (**Single-class coalition with poor and powerful insiders**)

Although no general argument is made, Example 2 suggests an important intuition: when protections of property rights are severely lacking and there is a poor yet highly powerful group within the ruling coalition, a single-class coalition is the most externally resilient coalition. This provide new insights on why proletarian revolutions in various contexts have attempted to enforce a radical vision of equality.

A.2 Robustness

Next proposition states that when the power and resource distributions, and the plundering functions in two different games are sufficiently similar, these games will lead to the same ruling coalition. That is, the ruling coalition is not fragile with respect to small shocks in the underlying environment and distributions. To obtain this result, without much loss of generality, we restrict our attention to the case that $\beta \in (\frac{1}{2}, 1]$.

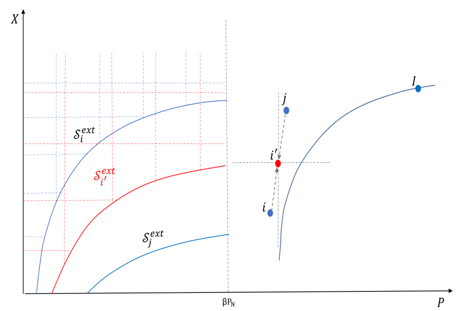


Figure 2: Increases in external resilience due to an exchange of power and resources outlined in Figure 11.

Proposition 1. *Consider* $\Gamma = (N, I_0, J, \{U_i(\cdot)\}_{i\in N}, \beta)$ *where* $\beta \in (1/2, 1]$. *Then, under Assumptions 1-3:*

(i) there exists $\delta > 0$ *such that if* $J' : N \to \mathbb{R}_{++}{}^2$ *lies within* δ*-neighbourhood of* J*, we have* $\phi(N, I_0, J, \{U_i(\cdot)\}_{i\in N}, \beta) = \phi(N, I_0, J', \{U_i(\cdot)\}_{i\in N}, \beta)$.

(ii) there exists $\delta > 0$ *such that if* $\forall (P, X)$ *with* $\beta P_N < P \leq P_N$, $X > 0$*, we have* $|G(P, X) - G'(P, X)| < \delta$*, then* $\phi(N, I_0, J, \{U_i(\cdot)\}_{i\in N}, \beta) = \phi(N, I_0, J, \{U_i'(\cdot)\}_{i\in N}, \beta)$*, where* $J := (p(.), x(.))$ *is the joint distribution of power and resources.*

Proof of Proposition 4. **Part (i):** We show that if condition (i), and (ii) of proposition 1 are satisfied under the joint distribution of power and resources J, these are satisfied under any joint distribution of power and resources J' that is in δ-neighbourhood of J. Throughout the proof, let $P_I(J)$, and $X_I(J)$ be the power and resources of the coalition I when the joint distribution of power and resources is $J = (P, X)$.

Let's start with the condition (i). Suppose for $A^{ins} \in \mathcal{A}_I \backslash I$ with $P_{A^{ins}} \geq \beta P_N$, we have $G(P_I(J), X_I(J)) > G(P_{A^{ins}}(J), X_{A^{ins}}(J))$. Equivalently, there exists $\epsilon > 0$ such that $G(P_I(J), X_I(J)) = \epsilon + G(P_{A^{ins}}(J), X_{A^{ins}}(J))$. In addition, G is continuous with respect to both power and resources, which means for any $0 < \epsilon_I < \frac{\epsilon}{2}$, there exists sufficiently small $\epsilon_I^1 > 0$ such that if $|P(J) - P_I(J)| < \epsilon_I^1$, and $|X(J) - X_I(J)| < \epsilon_I^1$, we have $|G(P_I(J), X_I(J)) - G(P(J), X(J))| < \epsilon_I$. Moreover, for any $0 < \epsilon_A < \frac{\epsilon}{2}$, there exists sufficiently small $\epsilon_A^1 > 0$ such that if $|P(J) - P_{A^{ins}}(J)| < \epsilon_A^1$, $|X(J) - X_{A^{ins}}(J)| < \epsilon_A^1$, we have $|G(P_{A^{ins}}(J), X_{A^{ins}}(J)) - G(P, X)| < \epsilon_A$. Now, by taking $0 < \delta < min\left\{\frac{\epsilon_A^1}{n}, \frac{\epsilon_I^1}{n}\right\}$, we will have $|G(P_{A^{ins}}(J), X_{A^{ins}}(J)) - G(P_{A^{ins}}(J'), X_{A^{ins}}(J'))| < \frac{\epsilon}{2}$, and $|G(P_I(J), X_I(J)) - G(P_I(J'), X_I(J'))| < \frac{\epsilon}{2}$. These imply $G(P_I(J'), X_I(J')) > G(P_{A^{ins}}(J'), X_{A^{ins}}(J'))$. Hence, if condition (i) is satisfied under the joint distribution of power and resources J, it would also be satisfied under any joint distribution of power and resources J' that is in δ-neighbourhood of J.

For condition (ii), first note that as mentioned in section 2, we restrict our attentions to joint distributions that for any $S, S' \subseteq N$ satisfy $P_S \neq P_{S'}$, or $X_S \neq X_{S'}$. Thus, $\forall S, S' \subseteq N$, $|P_S - P_{S'}| \neq 0$, or $|X_S - X_{S'}| \neq 0$. Let's denote $0 < \epsilon_p := \inf_{S,S' \subseteq N} |P_S(J) - P_{S'}(J)|$, and $0 < \epsilon_x := \inf_{S,S' \subseteq N} |X_S(J) - X_{S'}(J)|$. Take $0 < \delta < min\left\{\frac{\epsilon_p}{n}, \frac{\epsilon_x}{n}\right\}$. Now, suppose $J' : N \to {\mathbb{R}_{++}}^2$ lies within δ-neighbourhood of J. Then, as $\forall S, S' \subseteq N$, $|S|, |S'| \leq n$, J' would give the same order of power and resources to the subsets of N as J does. Precisely, if without loss of generality, for S, S', under the joint distribution of power and resources J, we have $P_S(J) > P_{S'}(J)$ (or $X_S(J) > X_{S'}(J)$), then we would have $P_S(J') > P_{S'}(J')$ (or $X_S(J') > X_{S'}(J')$). This means that, by definition, the set of best sub-coalitions of are "close‘' under J, and J', i.e., corresponding to any best sub-coalition of I under J, there exists a best sub-coalition under J' in close proximity to it, and vice versa. A same argument is true about the set of best sub-coalitions outside I, i.e., $A_{N\setminus I}$, under J and J'.

Now, suppose for $A^{ext} \in A_{N\setminus I}(J)$ and $A^{ins} \in \mathcal{A}_I(J)\setminus I$ where $P_{A^{ins}} + P_{A^{ext}} > \beta P_N$, we have $X_{A^{ins}}(J) + X_{A^{ext}}(J) > H_I(P_{A^{int}}(J) + P_{A^{ext}}(J))$. This means that there exists $\epsilon > 0$ such that $X_{A^{ins}}(J) + X_{A^{ext}}(J) = \epsilon + H_I(P_{A^{int}}(J) + P_{A^{ext}}(J))$. Then, continuity of indifference curves implies for any $0 < \epsilon_I < \frac{\epsilon}{2}$, there exists $\epsilon_g > 0$ such that if we have $|P_{A^{ins}}(J) + P_{A^{ext}}(J) - P_{A^{ext}}(J') - P_{A^{ins}}(J')| < \epsilon_g$, then $|H_I(P_{A^{ext}(J)} + P_{A^{ins}}(J)) - H_I(P_{A^{ext}}(J') + P_{A^{ins}}(J'))| < \frac{\epsilon}{2}$. Now, by taking $\delta < \frac{min\{\epsilon,\epsilon_g\}}{2n}$, we obtain both:

$$|X_{A^{ins}}(J) + X_{A^{ext}}(J) - X_{A^{ins}}(J') + X_{A^{ext}}(J')| < \epsilon_I < \frac{\epsilon}{2}$$

and:

$$|H_I(P_{A^{ext}(J)} + P_{A^{ins}}(J)) - H_I(P_{A^{ext}}(J') + P_{A^{ins}}(J'))| < \frac{\epsilon}{2}$$

Hence, $\delta < \frac{min\{\epsilon,\epsilon_g\}}{2N}$ implies $X_{A^{ext}}(J') + X_{A^{ins}}(J') > H_I(P_{A^{ext}}(J') + P_{A^{ins}}(J'))$. This means that if condition (ii) is satisfied under the joint distribution of power and resources J, it would also be satisfied under any joint distribution of power and resources J' that is in δ-neighbourhood of J. According to Proposition 1, these imply that if J' is in δ-neighbourhood of J, we have $\phi(N, I_0, J, \{U_i(.)\}_{i\in N}, \beta) = \phi(N, J', \{U_i(.)\}_{i\in N}, \beta)$.

Part (ii): The proof is straight-forward given the continuity of function $G(.)$. Condition (i) means that there exists a $\delta_1 > 0$ such that $G(P_I(J), X_I(J)) = \delta_1 + G(P_{A^{ins}}(J), X_{A^{ins}}(J))$. Since $G(.)$ is continuous, for any $0 < \epsilon < \frac{\delta_1}{2}$, there exists $\epsilon_p > 0$ such that if $|P_{A^{ins}}(J) - P_{A^{ins}}(J')| < \epsilon_p$, and $|X_{A^{ins}}(J) - X_{A^{ins}}(J')| < \epsilon_p$, then we have $|G(P_{A^{ins}}(J), X_{A^{ins}}(J)) -$

$G(P_{A^{ins}}(J'), X_{A^{ins}}(J'))| < \epsilon < \frac{\delta_1}{2}$. Moreover, for any $0 < \epsilon' < \frac{\delta_1}{2}$, there exists $\epsilon_g > 0$ such that $|P_I(J) - P_I(J')| < \epsilon_g$, and $|X_I(J) - X_I(J')| < \epsilon_g$ imply:

$$|G(P_I(J), X_I(J)) - G(P_I(J'), X_I(J'))| < \epsilon' < \frac{\delta_1}{2}$$

As a result, if we take $0 < \delta < \frac{\min\{\epsilon_g, \epsilon_p\}}{n}$, we will have $G(P_I(J'), X_I(J')) = \delta_1 + G(P_{A^{ins}}(J'), X_{A^{ins}}(J'))$, or equivalently $G(P_I(J'), X_I(J')) > G(P_{A^{ins}}(J'), X_{A^{ins}}(J'))$.

To prove part (ii), similar to the first part of the proof, note that condition (ii) means that there exists a $\delta_2 > 0$ such that $G(P_I(J), X_I(J)) = \delta_2 + G(P_{A^{ins}}(J) + P_{A^{ext}}(J), X_{A^{ins}}(J) + X_{A^{ext}}(J))$. Since $G(.)$ is continuous, for any $0 < \epsilon < \frac{\delta_2}{2}$, there exists $\epsilon_p > 0$ such that if $|P_{A^{ins}}(J) + P_{A^{ext}}(J) - P_{A^{ins}}(J') - P_{A^{ext}}(J')| < \epsilon_1$, and $|X_{A^{ins}}(J) + X_{A^{ext}}(J) - X_{A^{ins}}(J') - X_{A^{ext}}(J')| < \epsilon_1$, we would have $|G(P_{A^{ins}}(J) + P_{A^{ext}}(J), X_{A^{ins}}(J) + X_{A^{ext}}(J)) - G(P_{A^{ins}}(J') + P_{A^{ext}}(J'), X_{A^{ins}}(J') + X_{A^{ins}}(J'))| < \epsilon < \frac{\delta_2}{2}$. Moreover, for any $\epsilon' < \frac{\delta_2}{2}$, there exists $\epsilon_2 > 0$ such that $|P_I(J) - P_I(J')| < \epsilon_2$, and $|X_I(J) - X_I(J')| < \epsilon_2$ imply $|G(P_I(J), X_I(J)) - G(P_I(J'), X_I(J'))| < \epsilon' < \frac{\delta_2}{2}$. As a result, if $\delta < \frac{\min\{\epsilon_1, \epsilon_2\}}{n}$, we will have: $G(P_I(J'), X_I(J')) = \delta_2 + G(P_{A^{ins}}(J') + P_{A^{ext}}(J'), X_{A^{ins}}(J') + X_{A^{ins}}(J'))$ or equivalently $G(P_I(J'), X_I(J')) > G(P_{A^{ins}}(J') + P_{A^{ext}}(J'), X_{A^{ins}}(J') + X_{A^{ins}}(J'))$. This completes the proof.

□

A.3 Examples

The following example illustrate a typical function $G_i(.)$ that satisfies Assumption 1-3, further clarifying the distinction between environments that give rise to inclusive and exclusive ruling coalitions.

Example 2. This example illustrate a typical function $G_i(.)$ that satisfies Assumption 1-3, further clarifying the distinction between environments that give rise to inclusive and exclusive ruling coalitions. For any $i \in I \in \mathcal{W} \backslash N$, a typical $G(.)$ function satisfying Assumption 1 is:

$$w_i(I) = G_i(I) := \left(\frac{p_i}{P_I}\right)\left(\frac{P_I}{P_N}\right)^{\alpha+1}\left(\frac{X_N}{X_I}\right) \tag{0.1}$$

where $\alpha > 0$[2], $\frac{p_i}{P_I}$ is the share of player i from the plundered resources which is propor-

[2] Otherwise, if $\alpha < 0$, part (i)-(ii) of Assumption 1 are violated.

tional 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 different ruling coalitions in term of the resources they plunder. Without loss of generality, let's normalize the total power and resources of the players, i.e., $P_N = X_N = 1$. It then can be rewritten as: $G_i(I) := p_i P_I^\alpha \left(\frac{1}{X_I}\right)$. Now, let fix the level of plundered resources the coalition I gives to player $i \in I$, i.e., $G_i(I)$. The indifference curve of player i that passes from I would then be:

$$X = C_i(I)P^\alpha \tag{0.2}$$

where $C_i(I) := \frac{p_i}{G_i(I)}$. [3], and $P \in [\beta, 1]$. Here, for any coalition with power P and resources X, the marginal rate of substitution between power and resources of the ruling coalition is $MRS_{PX} = -\alpha(\frac{X}{P})$[4]. The parameter α then determines to what extent the power is valued within this environment relative to the resources [5]. A higher value of α indicates that these players value more the power of the ruling coalition, relative to the resources lost to plundering. This means that as α increases, players become more inclined to sacrifice resources to enhance the ruling coalition's power, which in turn give rise to a more inclusive ruling coalition. Moreover, given that $P \in [\beta, 1]$, an increase in α leads to a decrease in plundered resources. Thus, a higher value of α corresponds to an environment with both lower plundering intensity and favours a potentially more inclusive ruling coalition. Conversely, a lower value of α means that the the cost of losing resources is high for the ruling coalition, i.e., when α decreases, players within the ruling coalition become more inclined to sacrifice resources to enhance the ruling coalition's power. This means that the ruling coalition will generally be exclusive under such environments. Since $P \in [\beta, 1]$, a lower value of α means larger plundered resources, or more intensive plundering environment. Hence, a higher α results in a intensive plundering environment that give rise to exclusive ruling coalitions.

The following example shows that that there is no comprehensive way to characterize the composition of ruling coalition without restricting the distribution of power and re-

[3]Precisely, substituting $G_i(I)$ from (2.2) into the expression of $C_i(I)$, the set of pairs of power and resources of the winning coalitions (containing i) that give player $i \in I$ a share identical to the share that ruling coalition I gives to her, is given by $X = \left(\frac{X_I}{P_I^\alpha}\right) P^\alpha := CP^\alpha$.

[4]Since we normalized the aggregate power and resources of the players to 1, the power and resources of the ruling coalition here is the fraction of the power and resources of it.

[5]Remember the Assumption 1 in which we highlight that there is a trade-off between power and resources in our environment.

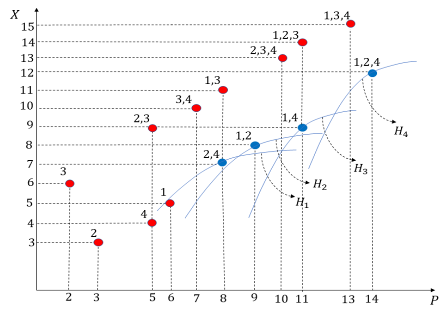


Figure 3: ruling coalition under different indifference curves.

sources and the plundering function. This arises first, from theorem 2 which establishes that any coalition in the set of potential ruling coalitions might be the ruling coalition for a specific range of indifference curves, second, there is no tight characterization of the set of potential ruling coalitions without specifying further the joint distribution of power and resources. As a result, there is no guarantee that the player with the highest power, the player with the fewest resources, or the player with the greatest power-to-resource ratio will be included in the ruling coalition (Figure 3)[6].

Example 3. Suppose $N = \{1, 2, 3, 4\}$, $p_1 = 6$, $x_1 = 5$, $p_2 = 3$, $x_2 = 3$, $p_3 = 2$, $x_3 = 6$, $p_4 = 5$, and $x_4 = 4$. Let $\beta = \frac{1}{2}$. Then the set of winning coalition is $\mathcal{W} = \{(1, 2), (1, 4), (2, 4), (1, 2, 4)\}$. As shown in Figure 3, when the indifference curve is H_1, the ruling coalition will be $(2, 4)$, which does not contain the player with the highest power, i.e., player #1. The indifference curve H_3 leads to ruling coalition $(1, 4)$, which excludes the player with the lowest resources. Finally, the ruling coalition does not contain the player with the highest power-to-resource ratio, i.e., player #4, if the indifference curve is H_2.

The following example demonstrates that when the resources are equal, the ruling coalition can be fully characterized by a power threshold in the sense that only players with a power above this threshold are within the ruling coalition. Hence, under such a joint distribution of power and resources, and regardless of the plundering function, any ruling coalition will encompass the player with the highest power (who also holds the highest power-to-resource ratio). A similar result is established when powers are equal, where the ruling coalition will solely be defined by a resource threshold, below which players fall

[6]Nonetheless, as we show in Example 4 in the appendix, a precise characterization of the ruling coalition's structure is feasible when the powers, resources, or both are distributed equally within society.

within the ruling coalition. Analogously, any such ruling coalition will inevitably include the player with the highest power-to-resource ratio, who is also the player with the lowest resources.

Example 4. Suppose the resources are equal. Hence, the power is increasing with respect to the power-to-resource ratio. Moreover, for simplicity, assume the maximum power of the players is bigger than the required power majority in that the player with the most power is a winning coalition. Then, the set of potential ruling coalitions Z can be proportional as the winning coalitions that are consecutively subset of each other, in that the first winning coalition I^1 is player with the highest power, the second winning coalition I^2 consists of the players with the highest power and the second highest power, etc. Precisely, if we denote $Z = \{1, 2, 3, .., |Z|\}$; where $k \in Z$ refers to I^k; then $k \leq |Z|$, $I^{k+1} - I^k = \left\{\arg\max_{i \in N \setminus I^k} p_i\right\}$.

Hence, any coalition in the set of potential ruling coalitions corresponds to a power threshold such that only players with powers higher than that threshold are within that coalition. Therefore, part 2 of Theorem 2 implies that any unique ruling coalition will be fully characterized by a power threshold. We prove this in two steps:

Claim 1: the set Z would be consisted of points that are consecutively subset of each other. Precisely, the point with the least power in A_I would be a subset of the point with second least power in Z, and the point with the second least power would be a subset of the point with third least power, and so on.

To prove, first, let denote a point in Z as l if it has the l_{th} least power among all other members (points) of Z and also define the set of players of point l as I_l. By contradiction, assume there is a $k \leq |Z|$ such that $I_k \not\subset I_{k+1}$, i.e., there exists a player $j \in I_k$ such that $j \notin I_{k+1}$. Now, if there exists a player $i \in I_{k+1}$ such that $p_i < p_j$, then as $p_i < p_j$ implies $x_i > x_j$, we can replace j with i in I_{k+1} to obtain a coalition with power higher than I_{k+1}, and resources lesser than I_{k+1}; which is a contradiction as we assumed I_{k+1} is on the frontier of potential ruling coalitions. On the other hand, if there is no player $i \in I_{k+1}$ with $p_i < p_j$, then we must have $I_{k+1} \subset I_k$ because otherwise, if there exists a player $h \in I_{k+1}$ such that $h \notin I_k$, we can replace h with j to obtain a coalition with power higher than the power of I_k and resources lesser than I^k. This is a contradiction as we assumed I^k is on the frontier of potential ruling coalitions. hence, we must have $I_{k+1} \subset I_k$. Nevertheless, this is also impossible as we assumed power of coalition I_k is lesser than power of the coalition I_{k+1}.

Claim 2: for any $k \leq |Z|$, $I_{k+1} - I_k = \left\{\text{argmax}_{i \in Z - I_k} p_i\right\}$.

To prove, as shown in the proof of claim 1, if $i \in I_{k+1} - I_k$, then for all $j \in I_k$, we have $p_i < p_j$. First, suppose $|I_{k+1}| - |I_k| > 1$ and $h \in I_{k+1} - I_k$. Then, $I_k \cup \{h\}$ would be on the frontier of potential ruling coalitions as for any $A \in Z - \{I_k, I_{k+1}\}$, either $P_A < P_{I_k \cup \{h\}}$ or $X_A > X_{I_k \cup \{h\}}$ as the power is increasing with respect to the power-to-resource ratio. Hence, we have $I_k \subset (I_k \cup \{h\}) \subset I_{k+1}$; which is a contradiction as we assumed there is no points on frontier of potential ruling coalitions with power between I_k and I_{k+1}. Therefore, $|I_{k+1}| - |I_k| = 1$. Now by contradiction, suppose $\text{argmax}_{i \in Z - I_k} p_i \notin I_{k+1}$. Then we can replace the player $\text{argmax}_{i \in Z - I_k} p_i$ with the player in $I_{k+1} - I_k$ to obtain a coalition with higher power and less resources which is a contradiction.

Claim 1, and claim 2 together imply that the members of Z are consecutively constructed by adding the unused player in Z with the highest power. Moreover, we assumed the maximum power is higher than the required majority, i.e., βP_N. These imply that $I^1 \subset I^2 \subset I^3 \subset ... \subset I^{|Z|}$; where I^1 is consisted of the player with the highest power, I^2 is consisted of the players with highest power and second highest power, and so on.