Power Consolidation in Groups*

Freddie Papazyan†

December 20, 2024

Abstract

I develop an economic model of how a society's distribution of power and resources evolves over time. Multiple lineages of players compete by accumulating power, which is modeled as an asset that increases one's probability of winning conflicts over resources. This model provides sharp equilibrium predictions for how a society's distribution of power evolves and whether it approaches inclusivity, oligarchy, or dictatorship in the long run. My main result shows that power and resources *inevitably* fall into the hands of a few when political competition is left unchecked in large societies, which addresses a longstanding empirical puzzle.

JEL Codes: C73, D02, D63, D72, P00

Keywords: Inequality, power, dynamic games

*I am deeply grateful to Renee Bowen for her continued support, invaluable advice, and guidance throughout this project and as my advisor at UC San Diego. I also would like to thank Daron Acemoglu for his support, for taking the time to discuss this project with me throughout my Ph.D., and for providing the inspiration for this work. I am thankful for the many helpful conversations I have had with Joel Watson, Alexis Toda, Johannes Weiland, Lawrence Broz, Joel Sobel, Simone Galperti, Federica Izzo, Denis Shishkin, Aram Grigoryan, Prashant Bharadwaj, Karl Schlag, Debraj Ray, Garance Genicot, Germán Gieczewski, Pietro Spini, Danil Dmitriev, Giampaolo Bonomi, Carlos Góes, Alice Hallman, Kyung Hwan Baik, Remy Levin, Daniela Vidart, and David Lagakos. Finally, I would like to thank the participants of the ThReD 2024 Conference at MIT Sloan, Yale ISPS's Political Power in Democracies Conference, the WiET and MWIED conferences at the University of Chicago, the Midwest Economic Theory and and International Trade Conference at IU Indianapolis, the Stony Brook International Conference on Game Theory, the WEAI Conference, the Econometric Society's 2024 NASM (Vanderbilt) and AM (IIT Delhi) conferences, UConn's Microeconomics seminar, and UC San Diego's Theory, TBE, and GSRS seminars for their invaluable feedback.

†Texas Tech University, Department of Economics, PO Box 41014, Lubbock, Texas 79409. Email: freddie.papazyan@ttu.edu. Website: www.freddiepapazyan.com

1 Introduction

As inequality continues to rise in the United States, so have economists' concerns that it may be drifting towards *oligarchy*.[1] This trend is not exceptionally American: persistently rising political and economic inequality has been observed in several other nations in the OECD (2008, 2011, 2012, 2015, 2021). These developments coincide with worldwide trends of spreading authoritarianism (Freedom House, 2022; United Nations, 2023; *Reuters*, 2023) and mounting anxieties about the health of democracies (Hyde, 2020; Freedom House 2020, 2024; DW, 2024; Riedl et al., 2024). This has made the following age-old question as relevant as ever: what allows – or indeed prevents – power and resources from falling into the hands of a few in a society?

The dynamics of inequality are not fully understood. A rising number of leading economists[2] have called for more attention to a crucial piece of this puzzle: *power*. The reasoning is that since power and wealth are intimately linked (Stiglitz, 2011), fully understanding how wealth inequality evolves requires further development of our systematic understanding of how the distribution of power evolves in societies. Moreover, a complete understanding of most questions in economics may be difficult without understanding the underlying power dynamics (Acemoglu, 2024c; Deaton, 2024). Given all of this, developing a systematic approach to how a society's distribution of power evolves over time would be of general interest to economists as well as other social scientists.

To address these pressing issues, this paper constructs an economic framework of how a society's distribution of power evolves due to intergenerational competition over resources. I model a society that is populated by (non-overlapping generations of) players from multiple lineages. Each lineage is initially endowed with a stock of power, which is modeled as an asset that increases one's probability of winning conflicts over consumable resources.[3] Every period, players inherit and accumulate power and then engage in conflicts over resources. These individual strategic power accumulation decisions jointly determine how a society's distribution of power *endogenously* evolves over time.

Contributions This framework generates sharp, rich equilibrium predictions for how a society's distribution of power evolves (Propositions 2 and 3). Given any initial distribution of power, my model provides a unique prediction of the trajectory it will follow and the stable distribution it approaches in the long run. Three types of stable distributions can emerge in the long run: *inclusive* (power shared uniformly), *dictatorial* (only one player holds power), and *oligarchic* (power shared uniformly among only a subset of players). This is the first unified framework that can precisely predict the emergence of dictatorial, inclusive, and oligarchic regimes. Moreover, I also provide a novel characterization of the necessary and sufficient conditions for the stability of each type of regime.

My main results (Propositions 4 and 5) show that power and resources must fall into the hands of a few in sufficiently large societies. This addresses a long-standing open

[1]Krugman (2014), Piketty (2014, p. 514), Solow (2014a,b), Saez and Zucman (2019), and Acemoglu (2024b).

[2]Piketty (2014, 2015), Stiglitz (2011, 2016), Rausser et al. (2011), Krugman (2020), World Bank (2006, 2017), United Nations (2020), Callander et al. (2022), Deaton (2024), and Acemoglu (2012, 2024a).

[3]This is a standard contest-theoretic approach to modeling power; a detailed discussion is provided in the Related Literature section, below.

question: the tendency for power consolidation to take place in large groups of people was notably observed by Michels (1915), stating it as the *Iron Law of Oligarchy*, but over a century later there appears to be little agreement on why or when this Iron Law holds true (Leach (2005, 2015); Diefenbach (2019)). This paper provides a novel theoretical foundation for the Iron Law of Oligarchy and insights into the competitive forces that underlie it. Strikingly, my results show that this Iron Law not only holds under standard economic assumptions, but is fundamentally driven by two standard assumptions in particular: convex investment/adjustment costs and difference/logistic form contest success functions.[4]

Additional contributions include an analysis of institutional features that make societies more or less robust to power consolidation (in section 4). Finally, Proposition 6 analytically characterizes how larger populations induce stronger dictatorships.

Related Literature The framework I construct in this paper builds on Acemoglu and Robinson's (2022) model of state-society power dynamics.[5] I generalize and re-frame the analysis to societies made up of any finite number of players, who may be viewed as *individual* agents or as *representative* agents of a socioeconomic sub-group like in Becker (1983, p. 372). This allows the present paper to make several novel contributions (listed above) that are of general interest to economists and other social scientists.

In addition to these broader contributions, this paper provides several new insights specifically relating to Acemoglu and Robinson (2022), which will often be abbreviated as AR. One of AR's main conclusions was that competitive pressure is what allows inclusive societies to emerge. My paper shows that this effect of competitive pressure – while true in small societies – inverts in sufficiently large societies. That is: competitive pressure drives small societies towards inclusive regimes, but as societies become large, competitive pressure is also precisely what causes such regimes to destabilize while dictatorships and concentrated oligarchies remain stable. Second, my paper also sheds light on the key drivers societal power dynamics. AR assume a specific functional form for power accumulation costs, which are convex in the investment rate and have sufficiently weak players face an additional linear cost. The authors note that the latter property is a key driver of their results. In contrast, my paper considers a general cost function that is convex in the investment rate and *weakly* diminishes in inherited power (i.e. it is a generalization – not a restriction – of the standard case where costs do not depend on inherited power). My results demonstrate that for societies with more than two players, only the standard assumption of convex investment costs are required. In addition to the aforementioned generalizations, this paper also dispenses with a number of assumptions made in AR (all four parts of their Assumption 2 and parts 2 and 3 of Assumption 3). First, this sheds light on which assumptions are key. Moreover, this allowed for new types of regimes to be stable, even in the two player case. Another key finding in AR is that strong dictatorships are not possible due to lack of competitive pressure. My paper allows for both strong

[4]The former assumption is standard in asset accumulation models (Dixit and Pindyck, 1994) and contest theory (Fang, Noe, and Strack, 2020); the latter is standard in contest theory (Hirshleifer, 1989; Skaperdas, 1996; Ewerhart, 2021; Ewerhart and Sun, 2024).

[5]This in turn is related to Acemoglu (2005) – which provides its microfoundation – as well as Acemoglu and Robinson (2019) wherein the authors provide an extensive view of history through the lens their model.

and weak dictatorships, and that dictatorial strength grows without bound as societies become large.

Prevailing explanations for how dictatorial, oligarchic, or inclusive regimes emerge overwhelmingly hinge on the role played by *structural factors*, specific features of societies such as culture, geography, economic conditions, exposure to external threats, etc. As Acemoglu and Robinson (2022) point out, such explanations necessarily cannot account for how otherwise similar societies can arrive at vastly different power structures, which is a widespread occurrence (Acemoglu and Robinson, 2019). Their model addressed this gap in the emergence of dictatorial and inclusive regimes, but not of oligarchy, which has been the subject of academic debate for over a hundred years. Michels viewed inequality as an inevitable eventuality in large societies, positing *scalar stress* and bureaucratization as the key drivers thereof (Johnson, 1982; Perret et al., 2020). Leach (2005) notes that "[d]espite almost a century of scholarly debate ... there is still no consensus about whether and under what conditions Michels's claim holds true." This assessment appears to be supported by the thorough surveys of the modern literature provided by Leach (2015) and Diefenbach (2019).

The framework constructed in this paper draws on the long-standing *contest theory* approach to modeling power and conflict. It formalizes Weber's (1925) widely-adopted definition of power as "the probability that one actor within a social relationship will be in a position to carry out his own will despite resistance, regardless of the basis on which this probability rests," and the fact that it must be *accumulated* by, for instance, "expenditures of time and money on campaign contributions, political advertising, and other ways that exert political pressure" (Becker, 1983). As Hirshleifer (1989, 1991a, 1991b) discusses, both *military* conflicts (e.g. Lanchester (1916) and Ewerhart (2021)) and *political* conflicts (e.g. Becker (1983, 1985) and Tullock (1980)) can be modeled as a contest, which "is a game in which players compete for a prize by exerting effort so as to increase their probability of winning" (Skaperdas, 1996). The "prize" in this model is control over resources (a consumable good such as natural resources, public funds, etc.), which is seized by the victor of conflicts. Players accumulate power ("effort") to increase their probability of winning conflicts.

A Contest Success Function (CSF) defines the conditional probability that a player wins a conflict given the amount of power they hold and the amounts of power held by each of their opponents. This paper focuses on the commonly used difference-form/Hirshleifer CSF which, as its name suggests, only depends on power differences. The properties of this form were notably studied by Hirshleifer (1989), who discusses how it relaxes certain overly-idealized aspects of its counterpart, the ratio-form/Tullock CSF (Tullock, 1980). Both forms of CSF are axiomized in Skaperdas (1996). In the context of this framework, the most important property of the difference-form CSF is that each player's marginal benefit of accumulating power is increasing in how closely matched they are with their opponents.

As mentioned above, power is modeled as an *asset* that is accumulated at a *cost*. I assume that investment costs are increasing and convex, and that the marginal cost of accumulating power (at any fixed rate) is *weakly* diminishing in the amount of power one currently holds. Intuitively, this captures the notion that the more powerful one already is, the less costly it is to obtain more of it, at least weakly. However, it is important to emphasize the *weak* nature of this property, as it nests the standard case where investment costs do not depend at all on one's current stock of power. This also sheds more light on the theoretical insights provided by Acemoglu and Robinson (2022), who emphasized the importance of both (i) the aforementioned property of difference-form CSF's and (ii) the "increasing returns" property of power accumulation costs; this paper shows that the former plays the key role.

As will be seen below, the equilibrium dynamics of this model feature what is known as the *discouragement effect*. This phenomenon was notably observed in the Harris and Vickers (1987) model of patent races and in a variety of other dynamic contests (Konrad, 2012).[6] In the present paper, the discouragement effect emerges on the paths towards dictatorship and oligarchy, where relatively weak players' power accumulation incentives erode as they are increasingly outpaced by stronger players. Experimental evidence of the discouragement effect is reviewed in Dechenaux et al. (2015). Indirectly related to this work are Berry (1993), Clark and Riis (1996), and Chowdhury and Kim (2014) who study multi-winner contests. While the model herein explicitly uses a *single*-winner contest mechanism, one may interpret there being multiple "effective" winners at the oligarchic and inclusive power structures in this model.

The form of power considered here has parallels with the notion of *personal power*, which was first systematically studied by Bowen et al. (2022). Their notion is similar in that it increases the probability of actualizing one's ideal outcome by asserting one's will. However, it is qualitatively different in that personal power derives from one's personal characteristics, and its effectiveness must be learned by others. This paper instead focuses on the forms of power that are accumulated and inherited and whose effectiveness is common knowledge.

Jeon and Hwang (2020) resembles the present paper in motivation but not in approach. In contrast to the dynamic *contest* setup considered here, they work in a dynamic *bargaining* framework. Their model admits two classes of stable power structures that resemble the dictatorial and oligarchic power structures seen here. Another key difference is that Jeon and Hwang (2020) consider infinitely forward-looking agents while the agents considered here are short-lived, being replaced each period. In their model, dictatorial power structures are unstable given that agents are sufficiently forward-looking. Interestingly, while my model does admit a dictatorial class of stable power structures, it also admits inclusive and oligarchic classes despite agents' short lifespan; in fact, assuming longer-lived agents does not provide substantive additional insight in my model.

Acemoglu and Robinson (2006) also theoretically investigate the Iron Law of Oligarchy (Michels, 1915), but through a different mechanism than the one seen here. Namely, they work in a model with both *de jure* and *de facto* power. Motivated by their observation on p. 327, the present framework focuses on *de facto* power. A more recent related work by Prato and Invernizzi (2023) analyzes the distribution of power within political parties in a moral hazard framework and also uncovers a novel explanation for why (and when) the Iron Law of Oligarchy holds. They find that the degree of power consolidation among parties and within parties tend to be inversely related. Perret et al. (2020) analyze the role

[6]The discouragement effect has also been observed in other settings such as the innovation investment models of Aghion et al. (2005) and Aghion (2005).

played by scalar stress in the Iron Law of Oligarchy.

This paper is also related to Bowen and Zahran (2012). In the present model, dictatorial and oligarchic power structures are reached by trajectories that originate sufficiently nearby. These classes respectively have qualitative similarities to the compromise and no-compromise classes in Bowen and Zahran (2012), which are reached in an analogous fashion. The dynamics in this paper also have parallels to those seen in Genicot and Ray (2017), who analyze how inequality and aspirations evolve and interact over time. Namely, the magnitude of agents' aspirations in their model has a similar encouraging/discouraging effect as power differences do in my model. Another related paper is Li et al. (2017), who study power dynamics in organizations using a principal-agent framework.

In a relatively recent seminal work on oligarchy, Winters (2011) provides an extensive study of how oligarchies emerge and persist in a variety of societies around the world, where he also notes that a "consistent pattern in human history is for very small minorities to amass great wealth and power." Rather than focusing on how particular institutional structures allow or preclude the formation of oligarchies, he instead argues that wealth defence and the accumulation of material power are far more important factors in the formation of oligarchies. The notion of power modeled in my paper is similar to the notion of material power in the sense that power in my paper can be accumulated. Note however that my notion of power is not limited to material power. Moreover, he emphasizes how oligarchies can emerge even in the presence of democratic norms and institutions, and the possibility of having democracies only in name. This is also stressed in Winters and Page (2009), an empirical study on the distribution of material power in the United States.

The rest of this paper is organized as follows. Section 2 constructs the model. Section 3 provides equilibrium predictions for how a society's distribution of power and resources evolves, and how dictatorial, oligarchic, and inclusive regimes each emerge in the long run. Section 4 characterizes the properties of stable power structures in large societies and contains the aforementioned main result of this paper. Section 5 concludes. The Online Appendix contains the proofs of the results in the main text of this paper (section A) and auxiliary technical results (section B). Section C shows how the main results extend to the case where societies' resource endowments scale with population size.

2 Model

Time $t \in \{0, \Delta, 2\Delta, ...\}$ has an infinite horizon and is initially[7] taken to be discrete with period length, $\Delta > 0$. There are $N \geq 2$ lineages of risk neutral, short-lived[8] players that are replaced each period. Lineage $i \in \{1, ..., N\}$ is formally defined as $\mathbb{i} \equiv \{i_0, i_\Delta, i_{2\Delta}, ...\}$, where i_t denotes the generation-t player from lineage i.

[7]Period length is later made arbitrarily small when attention is brought to model dynamics, which are more tractably characterized in continuous time.

[8] The main insights of this paper remain intact when players are forward-looking and not exceedingly patient. The results do qualitatively change when players are very patient, although not for an insightful reason; this is discussed in Remark 3, as soon as the necessary context is established.

Players compete by accumulating and passing along stocks of *power*, an asset that increases one's probability of winning conflicts over resources (in a way made precise below). The amount of power held by the lineage-i player at time t is given by $x_{it} \in [0, \chi]$, where $\chi > 0$ is arbitrarily fixed.[9]

Lineage i is initially endowed with x_{i0} units of power; this is held by player i_0, who is assumed to remain inactive for the entirety of period 0 and simply serves to initialize the game. Play then proceeds as follows: at the beginning of period $t \in \{\Delta, 2\Delta, \ldots\}$, player i_t inherits their predecessor's power $x_{i,t-\Delta}$ which linearly[10] depreciates at *rate*, $\delta > 0$ (hence by *amount* $\delta\Delta$ each period). Players then simultaneously choose how much to invest in their own power. Formally, player i_t commits to accumulating power at *rate*, $I_{it} \geq 0$ throughout the period, which adds $I_{it}\Delta$ newly-created units of power to i_t's stock by the end of the period. This yield's the following law of motion for player i_t's power:

$$\underset{\text{period-}t\text{ power}}{x_{it}} = \max\{\overset{\text{inherited power}}{x_{i,t-\Delta}} - \overset{\text{period length}}{\Delta} \cdot \underset{\text{depreciation } rate}{\delta}, 0\} + \overset{\text{period length}}{\Delta} \cdot \underset{\text{accumulation } rate}{I_{it}}$$

Note that the presence of the "$\max\{\cdot, 0\}$" term in the law of motion above reflects the fact that only inherited power depreciates. The instantaneous flow cost of investing at rate I_{it} given inherited power $x_{i,t-\Delta}$ is given by $C(I_{it}, x_{i,t-\Delta})$.[11] The *marginal* cost of investment is denoted by $C_I(I_{it}, x_{i,t-\Delta}) \equiv \frac{\partial}{\partial I_{it}} C(I_{it}, x_{i,t-\Delta})$.

After players accumulate power, society endows a lump-sum unit of resources, a consumable good.[12] Players compete over these resources through a winner-takes-all[13] conflict whose victor is randomly chosen according to the conditional probability distribution,

$$H(x_{it}, \boldsymbol{x}_{-i,t}; N) \equiv \mathbb{P}\{\text{Player } i_t \text{ wins the conflict} \,|\, \text{Power structure is } \boldsymbol{x}_t\}. \tag{1}$$

That is, each players' probability of victory depends not only on how much power they hold (x_{it}), but also that held by others ($\boldsymbol{x}_{-i,t} \equiv (x_{jt})_{j \neq i}$). At the end of period t, player i_t earns

[9]Assuming that power takes values in $[0, \chi]$ simplifies exposition without qualitatively altering the main results of this model. This is explored in Section 4.1.

[10]This paper assumes linear depreciation to facilitate comparison with Acemoglu and Robinson (2022); assuming geometric depreciation yields similar (albeit less transparent) results.

[11]A model with heterogeneous costs is possible, but this generalization does not provide substantive insight and complicates the statement of the results of this paper.

[12] Section C in the Online Appendix relaxes the normalization to unity. To provide a benchmark model for how societies' distributions of power evolve over time, it is appropriate to concentrate on an endowment economy to focus attention on players' power accumulation choices. I consider the case where players can also accumulate productive capital in the follow-up work to this paper.

[13]The results of this paper do not depend on the assumptions that players *necessarily* engage in conflict and that said conflict is *winner-takes-all*. Another version of this model can allow players to peacefully split resources through bargaining: first, players simultaneously make public announcements, first of their desired share of resources, then of their acceptance or rejection the proposed split. If a feasible split is unanimously accepted, players receive their desired shares; otherwise conflict ensues as before. Since this bargaining stage does not affect results in any way, it is omitted for parsimony.

an expected payoff of

$$\pi_i\left(x_{it}, I_{it}, \boldsymbol{x}_{-i,t}, x_{i,t-\Delta}\right) \equiv H(x_{it}, \boldsymbol{x}_{-i,t}; N) - \Delta \cdot C(I_{it}, x_{i,t-\Delta}), \tag{2}$$

where $C(\cdot,\cdot)$ is weighted by period length – while $H(\cdot,\cdot;\cdot)$ is not – since the former is a *flow* cost while the latter is a *lump sum* benefit.[14]

Two assumptions are made in this paper: one on the cost C and marginal cost C_I of power accumulation (Assumption 1) and one on the benefit H of power accumulation (Assumption 2).

Assumption 1. $C : [0,\infty)^2 \to [0,\infty)$ satisfies the following:

1. $C(\cdot, x_{i,t-\Delta})$ is strictly increasing and strictly convex for any fixed $x_{i,t-\Delta}$.
2. $C(I_{it},\cdot)$ and $C_I(I_{it},\cdot)$ are weakly decreasing for any fixed I_{it}.
3. C_I is continuously differentiable in its first argument and continuous in its second argument.

Assumption 1.1 (increasing, convex[15] power accumulation costs) is standard. Assumption 1.2 states that the cost – and marginal cost – of power accumulation weakly[16] diminish with how much power one currently holds. This captures the idea that the more powerful one already is, the less costly it is to further accumulate power, both in absolute terms and on the margin. Finally, Assumption 1.3 is a mild smoothness assumption which is essentially a relaxation of the assumption that C is twice continuously differentiable that permits $C_I(I_{\cdot},\cdot)$ to have "kinks," given any fixed $I_{\cdot} \geq 0$.

The second assumption made in this paper concerns the *benefit H* of holding power. I assume it takes a standard form that is commonly referred to as the Difference-Form Contest Success Function (CSF), whose properties were notably studied by Hirshleifer (1989) and Skaperdas (1996).

Assumption 2. Given $\boldsymbol{x}_{\cdot}$, the lineage-i player wins the conflict with probability

$$H(x_{i\cdot}, \boldsymbol{x}_{-i\cdot}; N) \equiv \frac{e^{\lambda x_{i\cdot}}}{\sum_{j=1}^{N} e^{\lambda x_{j\cdot}}} = \frac{1}{1+\sum_{j\neq i} e^{-\lambda(x_{i\cdot}-x_{j\cdot})}}, \quad (\lambda > 0). \tag{3}$$

Beyond assuming that H is continuous and only directly depends on power *differences*, $(x_i - x_j)_{j\neq i}$, assuming the above functional form is equivalent to assuming that it satisfies a collection of natural axioms (Skaperdas, 1996, Theorem 3).[17] As Corchón and Dahm (2010)

[14]See Dixit and Pindyck (1994) and Acemoglu and Robinson (2022, p. 412) for more details about this this conventional modeling approach in asset accumulation models.

[15]Intuitively, this standard assumption says that larger investments in power are increasingly expensive within a given period. Loosely put, this implies that smooth, gradual power accumulation is less costly than rapid, large adjustments.

[16]Notice that the standard assumption that costs only depend on accumulation rate I_{it} (and not on inherited power $x_{i,t-\Delta}$) is a special case of the cost function considered herein.

[17]The relevant axioms are summarized as follows: they require that victory probabilities are given by a valid conditional probability distribution that only directly depends on the power held by each participating player and that each participating player's probability of victory is increasing (resp. decreasing) in the amount of power they (resp. each opposing player) holds. The one axiom not covered in this summary (Axiom 4) is similarly natural but concerns "breakaway" conflicts, which are not featured in my paper.

note, it is standard to interpret $e^{\lambda x_{i\cdot}}$ as the *effectivity* of player $i\cdot$'s power, which corresponds to how effectively player $i\cdot$'s power influences their victory probability.

The main implication of this assumption is that the marginal benefit of power accumulation ("contest incentives")

$$h(x_{i\cdot}, \boldsymbol{x}_{-i\cdot}; N) \equiv \frac{\partial}{\partial x_{i\cdot}} H(x_{i\cdot}, \boldsymbol{x}_{-i\cdot}; N) = \frac{\lambda \sum_{j\neq i} e^{-\lambda(x_{i\cdot}-x_{j\cdot})}}{\left[1+\sum_{j\neq i} e^{-\lambda(x_{i\cdot}-x_{j\cdot})}\right]^2} = \frac{\lambda \frac{e^{\lambda x_{i\cdot}}}{\sum_{j\neq i} e^{\lambda x_{j\cdot}}}}{\left(1+\frac{e^{\lambda x_{i\cdot}}}{\sum_{j\neq i} e^{\lambda x_{j\cdot}}}\right)^2} \tag{4}$$

is increasing in how closely-matched one is with other players in terms of power. This captures the idea that gains over a closely-matched opponent are more valuable than those made against a much weaker (or much stronger) one. This property is formalized in the final equality of (4): the closer the *relative effectivity* $e^{\lambda x_{i\cdot}} / \sum_{j\neq i} e^{\lambda x_{j\cdot}}$ of player $i\cdot$'s power is to 1, the larger their marginal benefit of power accumulation, as shown in Figure 1. Note that the dependence of H and h on N will henceforth be suppressed when there is little risk of confusion.

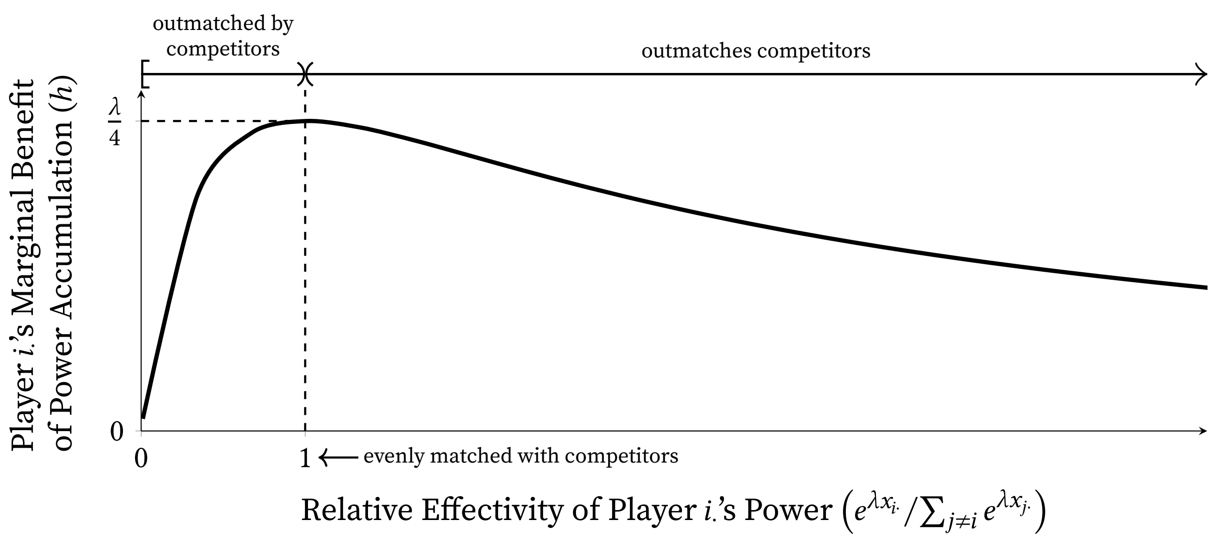


Figure 1: Player $i\cdot$'s marginal benefit of power accumulation h as a function the relative effectivity $e^{\lambda x_{i\cdot}} / \sum_{j\neq i} e^{\lambda x_{j\cdot}}$ of their power.

Remark 1. Parameter λ provides a tractable, systematic way to analyze the role played by the institutional constraints on the effectivity of power in reduced form. Larger λ increase the effectivity $e^{\lambda x_{\cdot}}$ of any given level of power $x_{\cdot}$. This is because larger λ correspond to conflicts that are less noisy in that their outcome depends more heavily on players' relative powers (Hirshleifer, 1989). To illustrate, note that as $\lambda \to 0$, the victor is essentially decided by a fair N-sided dice roll. As $\lambda \to \infty$, (one of) the strongest player(s) win with probability 1, like in an all-pay auction.[18]

[18]Lanchester (1916) and Hillman and Riley (1989) consider the latter limiting case.

I focus on Markov perfect equilibrium (Maskin and Tirole, 2001). The state variable in period t is $\boldsymbol{x}_{t-\Delta} \in [0,\chi]^N$, the previous period's power structure; the initial power structure $\boldsymbol{x}_0 \in [0,\chi]^N$ is exogenously fixed. Given $\boldsymbol{x}_{t-\Delta}$, player i_t's action set is $X_{it}(\boldsymbol{x}_{t-\Delta}) \equiv [\max\{x_{i,t-\Delta} - \delta\Delta, 0\}, \chi]$.[19] A *strategy* $x_{it} : [0,\chi]^N \to X_{it}$ for player i_t maps each state $\boldsymbol{x}_{t-\Delta}$ to an action x_{it} in $X_{it}(\boldsymbol{x}_{t-\Delta})$. The sequence $\{(x^*_{1t}, ..., x^*_{Nt})\}_{t\in\{\Delta,2\Delta,...\}}$ is a (*Markov perfect*) *equilibrium* – henceforth simply referred to as "equilibrium" – if at each t, $x^*_{it}(\boldsymbol{x}^*_{t-\Delta})$ is a best response to $\boldsymbol{x}^*_{-i,t}(\boldsymbol{x}^*_{t-\Delta})$ $\forall i \in \{1, ..., N\}$.

3 Equilibrium Power Structures

The problem faced by the lineage-i player in period $t \in \{\Delta, 2\Delta, ...\}$ is given by

$$\begin{cases} \max\limits_{x_{it}, I_{it}} & H(x_{it}, \boldsymbol{x}_{-i,t}) - \Delta \cdot C(I_{it}, x_{i,t-\Delta}) \\ \text{s.t.} & x_{it} = I_{it}\Delta + \max\{x_{i,t-\Delta} - \delta\Delta, 0\} \\ & 0 \leq x_{it} \leq \chi \\ & I_{it} \geq 0 \end{cases} \tag{5}$$

The equilibrium of the game described above can be characterized using (5). Before doing so, it is important to note the following:

Proposition 1. *Given any initial power structure $\boldsymbol{x}_0$, the equilibrium of this game is unique for all sufficiently small period length Δ.*

Proof. Found in Online Appendix subsection A.1. ■

Proposition 1 guarantees that each initial power structure $\boldsymbol{x}_0$ yields a unique equilibrium path $\{\boldsymbol{x}^*_0, \boldsymbol{x}^*_\Delta, \boldsymbol{x}^*_{2\Delta}, ...\}$ when Δ becomes small. Intuitively, this uniqueness stems from the fact that increasing power by any fixed amount $I_\cdot > 0$ becomes prohibitively expensive as Δ becomes small. As Δ approaches zero, players will differentially adjust their stocks of power in equilibrium (as opposed to making infrequent "large" adjustments). This will be seen explicitly in Proposition 2, where I solve the above game and make period length Δ arbitrarily small so that that the equilibrium dynamics of $\boldsymbol{x}^*_t$ can be studied in continuous time, which is far more tractable for analysis.

3.1 Equilibrium Power Dynamics

Let $\dot{x}^*_{it} \equiv \lim\limits_{\Delta\to 0} \frac{x^*_{it} - x^*_{i,t-\Delta}}{\Delta}$ denote the equilibrium power accumulation rate and let $C_I^{-1}(\cdot,\cdot)$ denote the inverse function of $C_I(\cdot,\cdot)$ with respect to its first argument, keeping its second argument fixed. With this notation in hand, the equilibrium dynamics of $\boldsymbol{x}^*_t$ are characterized in continuous time as follows.

[19]Notice that since x_{it} and I_{it} "pin down" one another, the number of choice variables can be reduced to one. For the purposes of defining strategies and equilibria, x_{it} is considered the only choice variable of player i_t.

Proposition 2. *As* $\Delta \to 0$*,* $\dot{x}^*_{it}$ *obeys the following law of motion for each i:*

$$\dot{x}^*_{it} = \begin{cases} -\delta \mathbb{1}_{(0,\chi]}(x^*_{it}), & \text{if } h(x^*_{it}, \boldsymbol{x}^*_{-i,t}) < C_I(0, x^*_{it}) \\ 0, & \text{if } h(x^*_{it}, \boldsymbol{x}^*_{-i,t}) > C_I(\delta, x^*_{it}) \text{ and } x^*_{it} = \chi \\ C_I^{-1}\big(h(x^*_{it}, \boldsymbol{x}^*_{-i,t}), x^*_{it}\big) - \delta \mathbb{1}_{(0,\chi]}(x^*_{it}), & \text{otherwise,} \end{cases} \tag{6}$$

Proof. Found in appendix subsection A.1. ■

The first two parts of (6) correspond to the corner solutions of (5) while the third corresponds to the interior solution. The first part states that when the marginal benefit $h(x^*_{it}, x^*_{-i,t})$ of power accumulation is less than the marginal cost $C_I(I_{it}, x^*_{it})$ of accumulating power at any $I_{it} \geq 0$, then it is optimal for player i_t to not add any power to their stock so that it depreciates unabated ($\dot{x}^*_{it} = -\delta$) or remains at zero. The second equation implies that player i_t optimally maintains the maximum level of power ($x^*_{it} = \chi$) when the net marginal gain ($h(\chi, \boldsymbol{x}^*_{-i,t} - C_I(\delta, \chi))$ of doing so is positive. Otherwise the third equation applies, and the optimal $\dot{x}^*_{it}$ equalizes the marginal benefit and marginal cost of power accumulation:

$$C_I^{-1}\big(h(x^*_{it}, \boldsymbol{x}^*_{-i,t}), x^*_{it}\big) - \delta \mathbb{1}_{(0,\chi]}(x^*_{it}) \Leftrightarrow \underbrace{h(x^*_{it}, \boldsymbol{x}^*_{-i,t})}_{\text{marginal benefit}} = \underbrace{C_I(\dot{x}^*_{it} + \delta \mathbb{1}_{(0,\chi]}(x^*_{it}), x^*_{it})}_{\text{marginal cost}}. \tag{7}$$

This characterization of players' equilibrium power accumulation behavior provides a natural, intuitive foundation for the main results of this paper, rooted solely in standard assumptions on the cost (C) and benefit (H) of power accumulation.[20] These two economic forces underlie players' equilibrium power accumulation decisions, and thus how the society's distribution of power $\boldsymbol{x}^*_t$ evolves over time. Intriguingly, depending on how unequal the society currently is, each force can either reinforce inequality or push toward greater balance. This will be illustrated explicitly in Section 3.3, below.

Notice that the law of motion in equation (6) is *time-invariant*; how a group's power structure evolves in equilibrium depends only on its current power structure ($\boldsymbol{x}^*_t = \boldsymbol{x}^*_{t'} \Leftrightarrow \dot{x}^*_{it} = \dot{x}^*_{it'}\ \forall i, t, t'$). This is indicative of the results of the next section, which characterizes the asymptotic behavior of $\boldsymbol{x}^*_t$: as will soon be seen, the group's initial power structure $\boldsymbol{x}_0$ will solely determine the trajectory of x^*_t and the power structure it approaches in the limit, in the absence of shocks. Moving forward, notation will often be simplified by suppressing time subscripts and asterisks: "$\dot{x}_i$" and "x_i" should henceforth be taken to mean "$\dot{x}^*_{it}$" and "x^*_{it}," respectively. Moreover, I will often refer to "player i_t" simply as "player i."

[20] Recall that these familiar assumptions were, respectively, the convex accumulation costs (Assumption 1a) and difference-form CSF (Assumption 2). The weak – possibly constant – dependence of C on inherited power (Assumption 1b) and the mild smoothness assumption (Assumption 1c) are not key drivers of the results in this paper; the former serves as a generalization of the standard case and the latter provides analytical tractability.

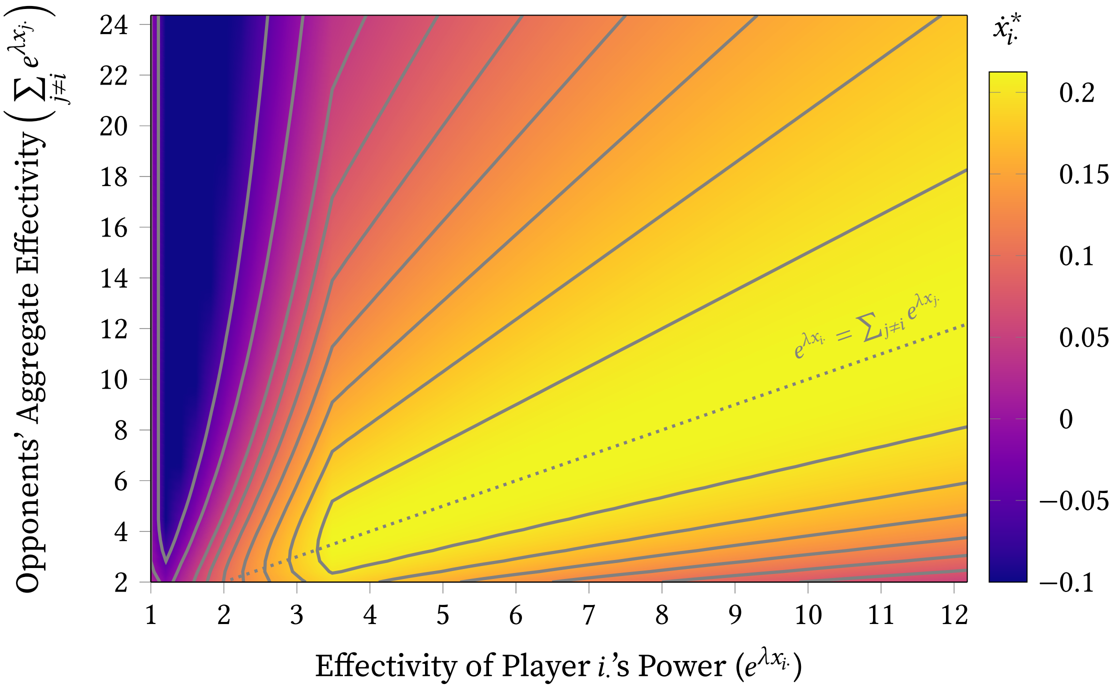


Figure 2: Player $i_\cdot$'s equilibrium power accumulation rate ($\dot{x}^*_{i\cdot}$) as a function of their effectivity ($e^{\lambda x_{i\cdot}}$) and their opponents' aggregate effectivity ($\sum_{j\neq i} e^{\lambda x_{j\cdot}}$). This figure is produced assuming $N = 3$ players, depreciation rate $\delta = 0.1$, institutional constraint parameter $\lambda = 2.5$, power cap $\chi = 1$, and cost function $C(I_\cdot, x_\cdot) = I_\cdot^2 + \max\{0.5 - x_\cdot, 0\}I_\cdot$.

3.2 Stable Power Structures

Now that the equilibrium dynamics of $\boldsymbol{x}_t$ have been fully characterized, attention is turned to the the *stable* power structures that can arise in the long run.

Definition 1. A power structure $\bar{\boldsymbol{x}} \in [0, \chi]^N$ is (asymptotically) *stable* if

a. $\dot{x}_i = 0 \; \forall \, i$ at $\bar{\boldsymbol{x}}$, and

b. $\forall \varepsilon > 0$, $\exists \rho > 0$ such that if $\|\boldsymbol{x}_0 - \bar{\boldsymbol{x}}\| < \rho$, then $\|\boldsymbol{x}_t - \bar{\boldsymbol{x}}\| < \varepsilon \; \forall \, t \geq 0$ and $\lim_{t\to\infty} \|\boldsymbol{x}_t - \bar{\boldsymbol{x}}\| = 0$, where $\|\cdot\|$ denotes the Euclidean norm.

Part a of this definition requires the system to be at rest at $\bar{\boldsymbol{x}}$; when this is satisfied $\bar{\boldsymbol{x}}$ is often referred to as a *steady state*. Part b requires that all trajectories that start near $\bar{\boldsymbol{x}}$ not only remain near $\bar{\boldsymbol{x}}$, but also converge to $\bar{\boldsymbol{x}}$. Proposition 3 fully characterizes the stable power structures that can arise under Assumptions 1 and 2, which will turn out to always take one of the following forms:

1. *Inclusive*, where all players hold zero power or all hold χ units power. That is,

$$\bar{\boldsymbol{x}} \in \big\{(0, \dots, 0), (\chi, \dots, \chi)\big\} =: \mathscr{I}. \tag{8}$$

I refer to $(0, \dots, 0)$ as *de-escalated inclusive*, and to $(\chi, \dots, \chi)$ as *escalated inclusive*.

2. *Oligarchic*, where $k \in \{2, ..., N-1\}$ players ("the oligarchs") hold χ units of power, and the remaining $N-k$ players are powerless. That is

$$\bar{\boldsymbol{x}} \in \left\{\boldsymbol{x} \in \{0, \chi\}^N \;:\; \textstyle\sum_{i=1}^{N} \mathbb{1}_{\{\chi\}}(x_i) = k\right\} =: \mathcal{O}_k. \tag{9}$$

Given $k \in \{2, \ldots, N-1\}$, I refer to the elements of $\mathcal{O}_k$ as *k-archic* power structures. I define the set of oligarchic power structures as $\bigcup_{k=2}^{N-1} \mathcal{O}_k$.

3. *Dictatorial*, where only one player ("the dictator") holds a strictly positive amount $d \in (0, \chi]$ of power. That is,

$$\bar{\boldsymbol{x}} \in \left\{(d, 0, \ldots, 0), \ldots, (0, \ldots, 0, d)\right\} =: \mathcal{D}_d. \tag{10}$$

I refer to $\bar{\boldsymbol{x}}$ as *strong dictatorial* if $d = \chi$ and *weak dictatorial* otherwise.

With this terminology in hand, the following proposition characterizes the necessary and sufficient conditions (labeled with roman numerals) under which inclusive, oligarchic, and dictatorial power structures are stable (parts 1-5) and establishes that power structures outside these classes are never stable (part 6).

Proposition 3. *$\bar{\boldsymbol{x}}$ is stable if and only if it is inclusive, oligarchic, or dictatorial. More specifically:*

1. *The escalated inclusive power structure $(\chi, ..., \chi)$ is stable if and only if*

$$h\big(\chi, (\chi, ..., \chi)\big) > C_I(\delta, \chi) \tag{I}$$

2. *The de-escalated inclusive power structure $(0, ..., 0)$ is stable if and only if*

$$h\big(0, (0, \ldots, 0)\big) \le C_I(0, 0) \tag{II}$$

3. *Let $k \in \{2, ..., N-1\}$. Each k-archic power structure $\bar{\boldsymbol{x}} \in \mathcal{O}_k$ is stable if and only if*

$$h\big(\chi, (\overbrace{\chi, ..., \chi}^{k-1}, \overbrace{0, ..., 0}^{N-k})\big) > C_I(\delta, \chi) \text{ and } h\big(0, (\overbrace{\chi, ..., \chi}^{k}, \overbrace{0, ..., 0}^{N-k-1})\big) \le C_I(0, 0). \tag{III}$$

4. *Let $d \in (0, \chi)$. Each weak dictatorial power structure $\bar{\boldsymbol{x}} \in \mathcal{D}_d$ is stable if and only if*

$$\begin{aligned} &h\big(\cdot, (0, ..., 0)\big) \text{ intersects } C_I(\delta, \cdot) \text{ from above at } d,^{21} \text{ and} \\ &h\big(0, (d, 0, ..., 0)\big) \le C_I(0, 0) \end{aligned} \tag{IV}$$

5. *Each strong dictatorial power structure $\bar{\boldsymbol{x}} \in \mathcal{D}_\chi$ is stable if and only if*

$$\begin{aligned} &h\big(x., (0, ..., 0)\big) > C_I(\delta, x.) \; \forall x. \in (\chi - \varepsilon, \chi) \text{ for some } \varepsilon > 0, \text{ and} \\ &h\big(0, (\chi, 0, ..., 0)\big) \le C_I(0, 0). \end{aligned} \tag{V}$$

6. *No other stable power structures are possible.*

Proof. Found in Online Appendix subsection A.2. ■

The intuition behind each of the above conditions is quite natural. Condition I says that each player has a strictly positive net marginal gain of maintaining χ units of power. This makes it optimal for each player to maintain χ units of power when $\boldsymbol{x} = (\chi, \ldots, \chi)$ and – by the continuity of h and C_I – guarantees that players accumulate power ($\dot{x}_i > 0\ \forall i$) when $\boldsymbol{x}$ is sufficiently near to $(\chi, \ldots, \chi)$. The necessity of Condition I for the stability of $(\chi, \ldots, \chi)$ is most easily seen in the case where $h(\chi, (\chi, \ldots, \chi)) < C_I(\delta, \chi)$: convex investment costs make it optimal for each player to allow their power to depreciate when $\boldsymbol{x} = (\chi, \ldots, \chi)$, so that it fails to be a steady state.[22]

Condition II implies that when all players are powerless, it is not optimal for any player to accumulate power (so that $\dot{x}_i = 0\ \forall i$ at $\boldsymbol{x} = 0$). When the inequality in Condition II is strict, then $h(x_i, \boldsymbol{x}_{-i}) < C_I(0, \boldsymbol{x}_i)\ \forall i$ when $\boldsymbol{x}$ is sufficiently close to $(0, \ldots, 0)$, since h and C_I are continuous. At all such $\boldsymbol{x}$, every player optimally allows their power to depreciate at rate δ, eventually causing each to hold no power.[23] When Condition II fails, it follows from the convexity of investment costs that each player begins to accumulate power (so that $\dot{x}_i > 0\ \forall i$ at $\boldsymbol{x} = (0, \ldots, 0)$).

Remark 2. The de-escalated inclusive power structure $(0, \ldots, 0)$ is approached by $\boldsymbol{x}$ in equilibrium only when each player lets their power fully depreciate; this represents in a certain sense the trivial case of the model (which is not ruled out by Assumptions 1 and 2).

The first (resp. second) part of Condition III plays a similar role as Condition I (resp. II). The first part says that the net marginal gain of maintaining χ units of power is positive when faced with $k-1$ other players who also hold χ units of power, and $N-k$ players who hold no power. The second part implies that it is not optimal for powerless players to accumulate power at k-archic power structures. Notice that Condition V is essentially the $k = 1$ analogue of Condition III.

Condition IV is necessary and sufficient for the stability of weak dictatorships, where a single player i ("the dictator") holds $x_i = d \in (0, \chi)$ units of power, and all other players hold no power. The second part of this condition makes power accumulation sub-optimal for powerless players when the dictator holds d units of power. The first part of Condition IV says that the dictator player i's marginal cost $C_I(\delta, d)$ of maintaining d units of power is equal to its marginal benefit $h(d, (0, \ldots, 0))$, so that maintaining d units of power is optimal for the dictator. Furthermore, it says that $h(\cdot, (0, \ldots, 0))$ intersects $C_I(\delta, \cdot)$ from *above*, which is crucial for the second part of Definition 1 to be satisfied. When this holds, then

[21] I.e. $h(d, (0, \ldots, 0)) - C_I(\delta, d) = 0$ and $\exists \varepsilon > 0$ s.t. $h(x_\cdot, (0, \ldots, 0)) - C_I(\delta, x_\cdot) > 0\ \forall x_\cdot \in (d - \varepsilon, d)$ and $h(x_\cdot, (0, \ldots, 0)) - C_I(\delta, x_\cdot) < 0\ \forall x_\cdot \in (d, d + \varepsilon)$.

[22] When $h(\chi, (\chi, \ldots, \chi)) = C_I(\delta, \chi)$, each player maintains their power level when $\boldsymbol{x} = (\chi, \ldots, \chi)$, so that it is a steady state. However, if one takes $\varepsilon > 0$ units of power from each player, the power structure will not return to $\boldsymbol{x}$ in equilibrium, so that the second part of the definition of stability is violated.

[23] The explanation in the case where Condition II holds with equality is more involved; for more details, please see the proof of Proposition 3.

decreasing (resp. increasing) the dictator's power by any "small" amount causes them to optimally accumulate power (resp. let their power depreciate) until they return to holding d units of power. However, when $h(\cdot, (0,\ldots,0))$ intersects $C_I(\delta,\cdot)$ from *below*, any such perturbation will cause x_i to drift away from d in equilibrium. The intuition behind the first parts of Conditions IV and V is illustrated in Figure 3. Note that unlike the two player case, dictatorship can be stable in the standard case where $C_I(I_\cdot, x_\cdot)$ is constant $x_\cdot$, as is permitted by the assumptions in this paper; the stability of inclusive and oligarchic regimes can similarly remain stable in this standard case.

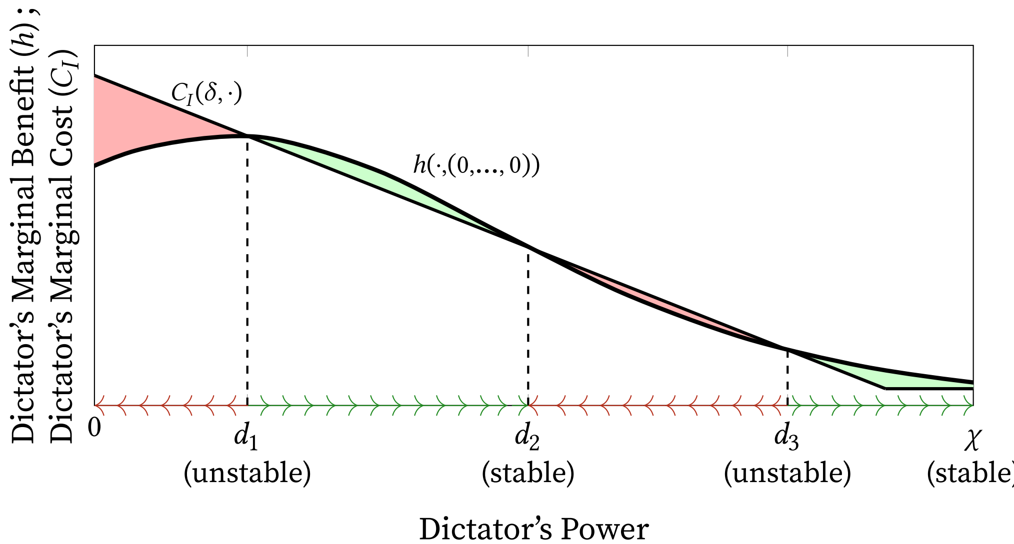


Figure 3: Power level $d \in (0,\chi)$ is sustained in a stable *weak* dictatorship only if the dictator's marginal benefit $h\big(\cdot,(0,\ldots,0)\big)$ of power intersects marginal cost $C_I(\delta,\cdot)$ at d from above. Power level $d = \chi$ is sustained in a stable *strong* dictatorship only if a dictator's marginal benefit $h(\chi,(0,\ldots,0)$ of maintaining χ units of power strictly outweighs the marginal cost $C_I(\delta,\chi)$ of maintaining that power level. Notice that this intuition still holds if costs do not depend on inherited power (e.g. when $C_I(\delta, x_\cdot) = \gamma\ \forall \boldsymbol{x}_\cdot$ for some $\gamma \in (h(d_1,(0,\ldots,0)), h(\chi,(0,\ldots,0)))$.

Finally, the last part of Proposition 3 shows that if $\bar{\boldsymbol{x}}$ is not inclusive, oligarchic, or dictatorial, it cannot be stable. First consider the case where $0 < \bar{x}_i < \bar{x}_j < \chi$ for some $i \neq j$ (i.e. two players hold different interior power levels at $\bar{\boldsymbol{x}}$). If such an $\bar{\boldsymbol{x}}$ is a steady state ($\dot{x}_\cdot = 0$ for all players), then (6) implies that for each player $\ell \in \{i,j\}$, the marginal cost $C_I(\delta,\bar{x}_\ell)$ of maintaining $\bar{x}_\ell$ units of power is equal to its marginal benefit $h(\bar{x}_\ell, \bar{\boldsymbol{x}}_{-\ell}) = C_I(\delta,\bar{x}_\ell)\ \forall \ell \in \{i,j\}$. This leads to a contradiction since $C_I(\delta,x_i) \geq C_I(\delta,x_j)$ (Assumption 1.2) and $h(\bar{x}_i,\bar{\boldsymbol{x}}_{-i}) < h(\bar{x}_j,\bar{\boldsymbol{x}}_{-j})$ (shown in the proof). In the remaining cases, $\bar{\boldsymbol{x}}$ may be a steady state, but not a stable one. Intuitively, this is because arbitrarily small perturbations can cause $\boldsymbol{x}_t$ to not return to a neighborhood of $\bar{\boldsymbol{x}}$. Specifically, this is achieved by providing each interior player an equal, sufficiently small, windfall of power. This has an identical, non-negative effect on each such player's net marginal gain of power accumulation (and thus their respective accumulation rates) without affecting the behavior of other players (since this windfall was chosen to be sufficiently small, and since h and C_I are continuous).

Remark 3. The result in Proposition 3 remains qualitatively similar when players are forward-looking but sufficiently impatient. If players are exceedingly patient, the escalated inclusive power structure is the only stable power structure, since in this case the cost of accumulating the maximum level of power χ (which is endured over a finite number of periods) is outweighed by the high benefit of holding χ units of power (which is enjoyed over an infinite time horizon). This is proven in virtually the same way as Proposition 3 in Acemoglu and Robinson (2022), with only trivial differences. This forward-looking extension is not among the several significant differences between this paper and Acemoglu and Robinson (2022), which were discussed in the Related Literature.

3.3 Three Player Illustration

This section illustrates the global equilibrium dynamics in the case with $N = 3$ players. In the interest of clearly visualizing these results, I focus on the case where the set of stable power structures is

$$\Big\{ \underbrace{(\chi,\chi,\chi)}_{\substack{\text{escalated}\\ \text{inclusive}}}, \underbrace{(d,0,0),(0,d,0),(0,0,d)}_{\text{dictatorial}}, \underbrace{(\chi,\chi,0),(\chi,0,\chi),(0,\chi,\chi)}_{\substack{\text{oligarchic}\\ \text{(2-archic)}}} \Big\},$$

where $d \in (0,\chi]$. Figure 4 visualizes the equilibrium dynamics in this case. The intuition behind how these novel dynamics arise from standard economic assumptions are discussed below. Interestingly, these very same incentives and costs act as equalizing forces when players are sufficiently equal, but otherwise act as forces towards inequality.

When players' powers are initially close to one another, the *(escalated) inclusive* power structure is reached through competition. As a result of the familiar difference-form CSF, players' power accumulation incentives are strongest when they are evenly matched with one another (in the sense discussed below equation (4)). Thus, when players begin sufficiently closely matched, they each begin with similarly strong power accumulation incentives. Moreover, convex accumulation costs prevent each player from outrunning the rest, since rapid power accumulation is increasingly expensive. Consequently, players begin accumulating power at similar rates in equilibrium, resulting in each becoming more powerful in *absolute* terms while their *relative* powers remain similar. This cycle repeats until each player reaches χ units of power.

Dictatorial power structures are reached through a qualitatively different process. When one player – say, player 1 – begins significantly more powerful than the rest, that player will eventually be the only one holding power (the "dictator"). Players' power accumulation incentives are weaker than before due to this power disparity, but by construction, player 1's incentives are strong enough for power accumulation to be optimal. Other players may also find it initially optimal to accumulate power, but not so fast as to close the gap with player 1, since convex investment costs make rapid accumulation relatively expensive. Consequently, this power disparity widens, further discouraging players 2 and 3 from catching up with player 1. This process continues until weaker players 2 and 3 find power accumulation sub-optimal, allowing their respective powers to depreciate to zero,

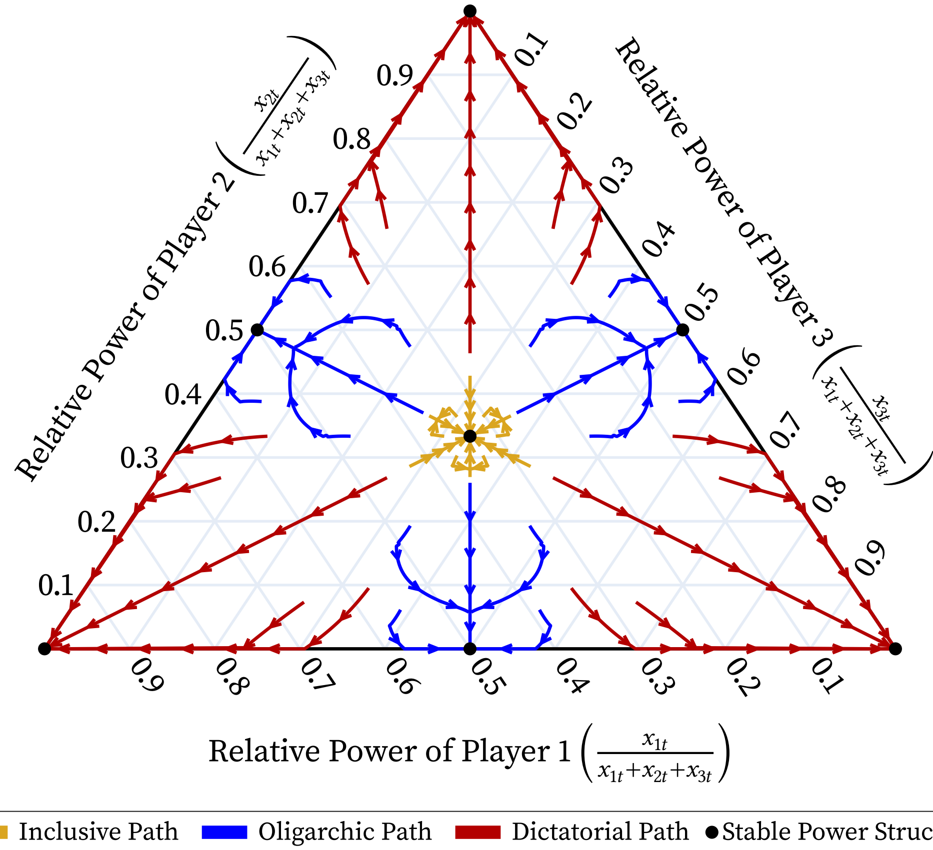


Figure 4: Simplex plot representation of simulated equilibrium paths produced using cost function $C(I_\cdot, x_\cdot) = I_\cdot^2 + (1 - x_\cdot)I_\cdot$, institutional constraint parameter $\lambda = 5.5$, depreciation parameter $\delta = 0.1$, and power cap $\chi = 1$.

at which point player 1 has fully consolidated power. Notice that this process did not rely on player 1 having a cost advantage (i.e., this intuition holds even when $C(I_\cdot, x_\cdot)$ is constant in $x_\cdot$).

Finally, *oligarchic* (here, 2-archic) power structures are reached through a combination of the above two processes. These are reached when two players (say, players 1 and 2) begin closely matched to each other but outmatch the rest (here, player 3). Players 1 and 2 compete with one another, each driving the other player's power up in the same way that the escalated inclusive power structure is reached. This causes these players to "outrun" (in terms of power accumulation) player 3, who eventually allows their power to fully depreciate, like in the dictatorial case.

4 Properties of Stable Power Structures

4.1 Stable Power Structures in Large Societies

I now turn to the main result of this paper: in the following proposition, I show that there always exists a finite group size past which the *escalated inclusive* power structure ceases to be stable.

Proposition 4. *The escalated inclusive power structure* $(\chi,\ldots,\chi)$ *is not stable in groups larger than*

$$\bar{N}_{\chi}^{I} \equiv \left\lceil \frac{\lambda + \sqrt{(\lambda - 4C_I(\delta,\chi))\lambda}}{2C_I(\delta,\chi)} \right\rceil \cdot \mathbb{1}_{\left[4C_I(\delta,\chi),\infty\right)}(\lambda) \tag{11}$$

Proof. Found in Online Appendix subsection A.3. ■

This result implies that in a sufficiently large society, unchecked political competition[24] will *inevitably* leave a subset of its population marginalized. The takeaway of this result should not be the sense of resignation evoked by Michels's (1915) *Iron Law of Oligarchy*, but instead a sense of urgency: political competition must be regulated to make inclusivity possible to achieve in large societies; failing to do so guarantees its impossibility.

What kinds of interventions can keep the escalated inclusive power structure stable in large societies? Before turning to this matter, it is first important to understand why the escalated inclusive power structure destabilizes in sufficiently large groups. Recall that by Proposition 3.1, the escalated inclusive power structure is stable if and only if Condition I holds:

$$(N-1)\lambda/N^2 = h(\chi,(\chi,\ldots,\chi);N) > C_I(\delta,\chi).$$

The formula for N_{χ}^{I} in (11) is derived using the above condition.[25]

Notice that each player's marginal benefit $h\big(\chi,(\chi,\ldots,\chi);N\big)$ at $\boldsymbol{x} = (\chi,\ldots,\chi)$ is decreasing in N and decays to zero as N grows large. Eventually – after the group grows larger than N_{χ}^{I} – this marginal benefit becomes outweighed by the marginal cost $C_I(\delta,\chi) > 0$ of maintaining χ units of power. This makes it suboptimal for all players to maintain χ units of power, and hence the escalated inclusive power structure is not stable.

To see the intuition behind this, recall that players' power accumulation incentives h are increasing in how closely matched they are with their competitors (in the sense discussed after equation (4)). At the escalated inclusive power structure, every player faces $N-1$ opponents that are each as strong as they are. This is why each player has strong power accumulation incentives when N is small; in fact they are as strong as possible when $N = 2$. However, as N grows large, players become overwhelmed by their aggregate competition at $\boldsymbol{x} = (\chi,\ldots,\chi)$. This is somewhat ironic, since – as was discussed in

[24]By "unchecked political competition," I precisely mean that the game modeled in this paper proceeds without intervention originating externally to the game.

[25]The presence of the indicator function in equation (11) reflects the fact that Condition I fails to hold at any $N \geq 2$ when $\lambda/4 < C_I(\delta,\chi)$.

subsection 3.3 – competitive pressure is what drives $\boldsymbol{x}$ towards this power structure; when N grows large, competitive pressure is also what snuffs it out.

I now analyze the institutional features that make the escalated inclusive power structure more (or less) robust to population size. The following Corollary characterizes how this population bound $\bar{N}^I_\chi$ varies with λ and $C_I(\delta,\chi)$; namely it is increasing in the former and decreasing in the latter.

Corollary 1. *When $\lambda/4 \geq C_I(\delta,\chi)$, $\bar{N}^I_\chi$ is strictly increasing in λ and strictly decreasing in $C_I(\delta,\chi)$; otherwise it is increasing in λ and decreasing in $C_I(\delta,\chi)$.*

Proof. Found in Online Appendix section A.4. ■

Lowering $C_I(\delta,\chi)$ corresponds to decreasing the marginal cost of maintaining χ units of power; increasing λ corresponds to loosening institutional constraints on the effectivity $e^{\lambda x.}$ of power $x.$ (Remark 1). The above corollary suggests that either change increases $\bar{N}^I_\chi$, so that the escalated inclusive power structure is more robust to group size. This may seem counter-intuitive at first, since these policies seem to favor those who are already powerful.

To see the intuition behind this result, observe that interventions that increase λ (resp. decrease $C_I(\delta)$) simply serve to increase (resp. decrease) the left-hand (resp. right-hand) side of Condition I, mentioned just above. Recalling the discussion of Proposition 3, Condition I is necessary and sufficient for the stability of $(\chi,\ldots,\chi)$ because it ensures that players' net marginal gain of power accumulation is positive when they are all sufficiently powerful. Moreover, any finite increase (resp. decrease) in λ (resp. $C_I(\delta)$) is only a temporary solution in growing societies, since they only increase $\bar{N}^I_\chi$ by a finite amount.

The hard upper bound χ on power is also a policy lever. Recalling Assumption 1, notice that increasing χ raises $\bar{N}^I_\chi$ only by diminishing $C_I(\delta,\chi)$. However, this policy lever will only get you so far: as I now show in Proposition 5.1, if one makes χ arbitrarily large, $\bar{N}^I_\chi$ remains finite in all but a knife-edge case where the marginal cost of maintaining an arbitrarily large amount of power becomes arbitrarily close to zero. The remainder of this proposition shows that dictatorships and oligarchies are far more robust to population size.

Proposition 5. *Suppose that $\lim\limits_{\chi\to\infty} C_I(\delta,\chi) = \gamma > 0$. If χ is made arbitrarily large, then*

1. *The group size past which the escalated inclusive power structure is unstable remains finite.*

2. *Apart from the trivial case when $\lambda/4 \leq \gamma$,[26] dictatorial power structures remain stable at arbitrarily large group sizes.*

3. *k-archies remain stable at arbitrarily large group sizes if k is no greater than*

$$\bar{k} \equiv \left\lfloor \frac{\lambda + \sqrt{(\lambda-\gamma)\lambda}}{\gamma} \right\rfloor \cdot \mathbb{1}_{(\gamma,\infty)}(\lambda), \tag{12}$$

[26]As seen in Figure 1, when $\lambda/4 \leq \lim_{\chi\to\infty} C_I(\delta,\chi)$, the maximum attainable marginal benefit $\lambda/4$ is less than the marginal cost of maintaining any positive level of power (because of Assumption 1.2) $\forall N \geq 2$.

and k-archies with $k \in \{\bar{k}+1, \bar{k}+2, \ldots\}$ are not stable at any group size.

Proof. Found in Online Appendix section A.5. ■

When $\chi \to \infty$, the restriction of power to $[0, \chi]$ becomes relaxed by an arbitrarily large amount. The first part of this proposition shows that in all but a knife-edge case,[27] the escalated inclusive power structure still becomes unstable past a finite group size, and for the same reason as before: players become overwhelmed by their aggregate competition. However, notice in equation (11) that relaxing institutional constraints on the *effectivity* of power (i.e. making λ arbitrarily large) allows the escalated inclusive power structure to remain stable in arbitrarily large group sizes. Recalling Remark 1, this is effectively amounts to turning conflict into an all-pay auction. On the other hand, if $\lambda = 0$, so that there is no benefit to holding power, only the de-escalated inclusive power structure $\boldsymbol{x} = (0, \ldots, 0)$ is stable. Thus, extreme values of λ appear to make inclusivity robust to group size.

Proposition 5 provides another interesting implication in its latter two parts: dictatorships and oligarchies with sufficiently few oligarchs are robust to group size.[28] As I discussed in the Related Literature section of the Introduction, this provides a theoretical foundation for a stylized fact that is far from fully understood: power tends to fall into the hands of a few in large groups of people. In contrast to prevailing explanations, the one provided here does not rely on the particular details of political institutions; it simply stems from the nature of incentives in power accumulation competitions.

To see the intuition for why k must be sufficiently small for a k-archy to be robust to group size, recall that in k-archic power structures, each oligarch[29] is individually equally matched with $k-1$ other players. When $k \leq \bar{k}$, oligarchs face more than one – but not too many – closely matched opponents, which ensures strong competition incentives. Otherwise, the oligarchs become overwhelmed like the players in the escalated inclusive power structure.

This leaves one final mystery: why do dictatorships remain stable at arbitrarily large group sizes? Since dictators have no closely matched competitors, shouldn't their contest incentives be weak? This is indeed the case when N is small, but as N becomes large this story qualitatively shifts. While the technical details for why this is true are found in the proof of Proposition 5, the intuition behind this result is best understood after first considering the effect of group size on the amount of power held in stable dictatorships.

[27] The supposition in Proposition 5 says that $C_I(\delta, \cdot)$ is bounded below by some γ>0, which may be arbitrarily small. This just leaves the knife-edge case where $C_I(\delta, \chi)$ decays to *exactly* zero as $\chi \to \infty$.

[28] Note that keeping χ fixed will artificially cause dictatorships and oligarchies to become unstable past a finite group size (details are provided in Propositions 8 and 9 in Online Appendix B). Moreover, note that since the trivial case of this model is not ruled out by its assumptions, the de-escalated inclusive power structure may remain stable at arbitrarily large group sizes (Remark 2).

[29] Recall that the strongest players in k-archic power structures are termed "oligarchs."

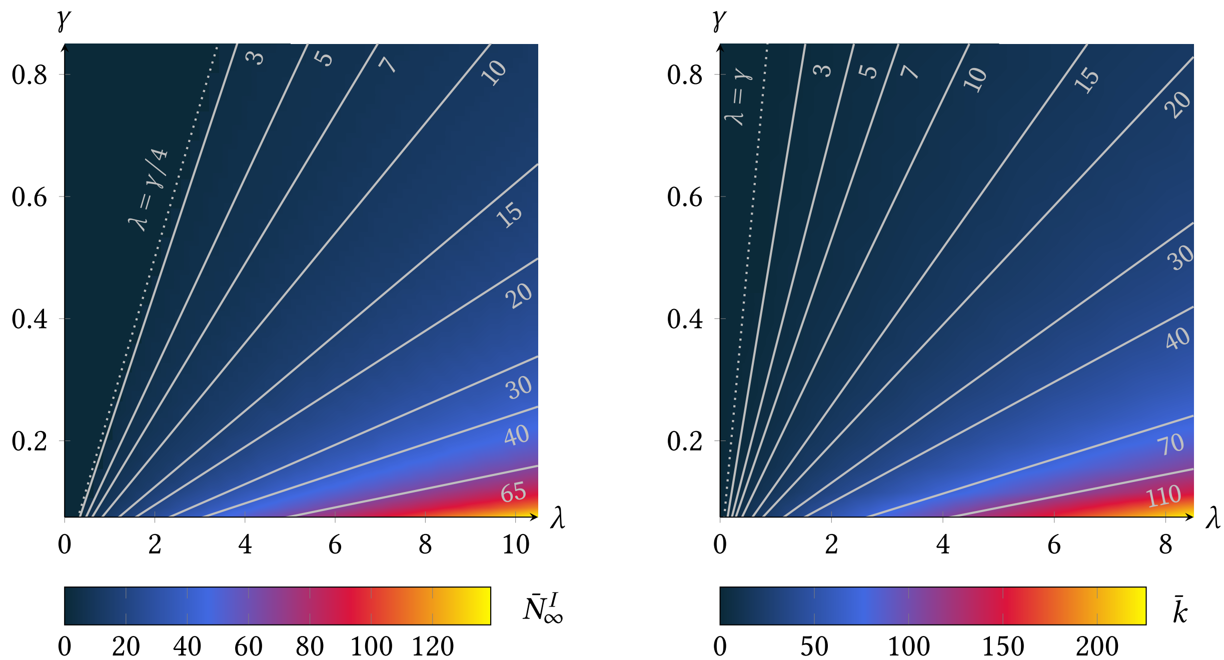


Figure 5: How $\bar{N}^I_\infty \equiv \lim_{\chi\to\infty} \bar{N}^I_\chi$ and $\bar{k}$ vary with λ and $\gamma \equiv \lim_{\chi\to\infty} C_I(\delta, \chi)$

4.2 Comparative Statics of Stable Dictatorial Power

This section characterizes the comparative statics of stable dictatorial power, the amount of power held by the strongest player ("the dictator") in a stable dictatorship. Recall that Proposition 3 established that weak dictatorships (with dictatorial power $d \in (0, \chi)$) are stable if and only if Condition IV holds and strong dictatorships (with dictatorial power $d = \chi$) are stable if and only if Condition V holds. Depending on model primitives, it is possible for no dictatorships to be stable or for *multiple* levels of power to be sustained in stable dictatorships as in Figure 3 in subsection 3.2.[30] For ease of exposition, assume throughout this subsection that *exactly one* level of power d is sustained in a stable dictatorship.[31] Analogous results hold when multiple types of dictatorship are stable, but are substantially more cumbersome to state, and offer insubstantial additional insight.

Proposition 6. *The amount of power held by dictators in stable dictatorships increases in group size N.*

Proof. Found in Online Appendix subsection A.6. ■

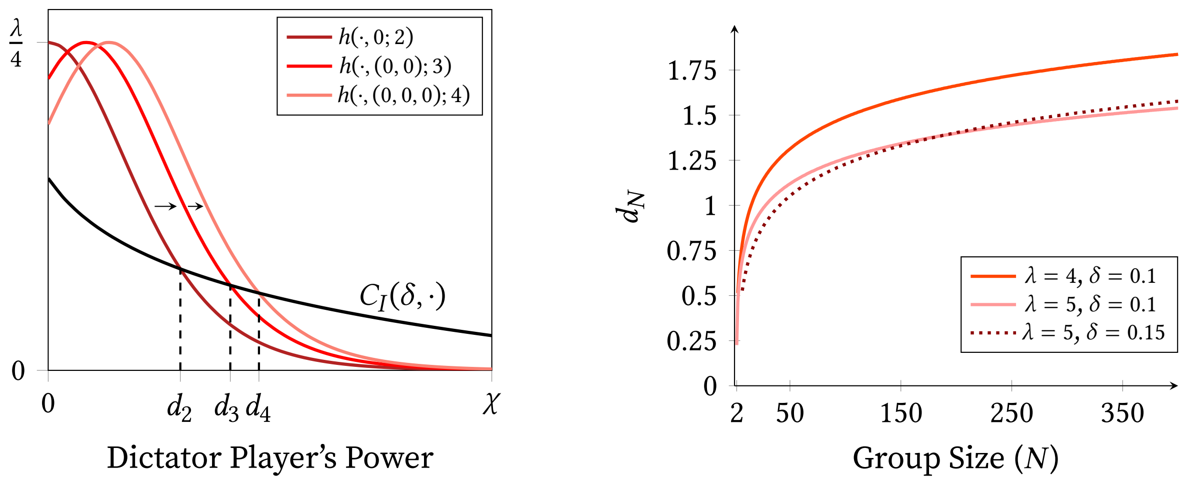
It is natural to expect that larger group sizes lead to stronger dictators. Mechanically, this is because increasing group size N translates the dictator's marginal benefit $h(\cdot, (0, \dots, 0); N)$ rightward (by equation (32) in the proof of Proposition 6). This is illustrated in Figure 6A, below. Intuitively, this is because under Assumption 2, powerless

[30]Recall that the equilibrium dynamics in (6) are always unique, hence even when multiple "kinds" of dictatorships are stable.

[31]Formally put: assume that either (1) Condition IV holds for exactly one $d \in (0, \chi)$ and Condition V fails or (2) IV fails at all $d \in (0, \chi)$ and Condition V holds.

players have a small but non-zero probability of victory.[32] As a result, powerless players collectively exert competitive pressure on the dictator player. This pressure grows with the number of powerless players, thereby inducing the dictator to hold an increasingly high level of power in stable dictatorships. Note that when this level of power is in the interior of $(0, \chi)$, it is strictly increasing in group size N. Hence, if χ is made arbitrarily large, the amount of power held in stable dictatorships grows without bound with group size.

As N becomes large, dictators' contest incentives – and hence optimal behavior – start to resemble those of oligarchs. The way in which dictators optimally respond to increases in group size is what ultimately causes their contest incentives to grow with N and hence allows dictatorships to be robust to group size in Proposition 5. Strong dictators emerge when (the rest of) society is *collectively* strong. While this resembles Acemoglu and Robinson's (2022) main result, there is an added twist: non-dictator players are individually powerless, having only collective strength in numbers.



(A) Larger N induce higher d_N because it shifts the dictator's marginal benefit of power accumulation $h(\cdot, (0,\ldots,0); N)$ rightward.

(B) Simulated relationship between N and d_N, assuming cost function $C(I_\cdot, x_\cdot) = 3.25 I_\cdot^2 + \max\{0.5 - x_\cdot, 0\} I_\cdot$.

Figure 6: How larger group size N induces higher levels of power d_N held by the strongest player ("dictator") in stable dictatorships.

Other comparative statics properties of the amount of power d held in stable dictatorships are given in the result below.

Proposition 7.

1. *Uniformly increasing the marginal cost of investment $C_I(\cdot, \cdot)$ decreases d.*
2. *d is decreasing in δ.*
3. *d increasing in λ if and only if $\frac{\lambda d - 1}{\lambda d + 1} e^{\lambda d} < N - 1$.*

[32] As discussed in Hirshleifer (1989), this reflects the inherent noisiness of conflicts.

Proof. Found in Online Appendix subsection A.7. ■

The first two parts of this proposition consider the negative relationship between the amount of power d held by dictators in stable dictatorships and the marginal cost $C_I(\delta, d)$ of maintaining d units of power. If the latter value were to increase – say, due to an unexpected shock or a policy intervention that makes accumulating power more costly – the dictator's marginal cost of maintaining d units of power would outweigh its marginal benefit $h(d, (0, \dots, 0))$. The dictator consequently lets their power depreciate until stabilizing at a new, lower level of power.

Conflict noise parameter λ has a less straightforward relationship with stable dictatorial power d. Increasing λ induces an increase in d if and only if they are both sufficiently small, a requirement that becomes less stringent as N increases. As λ becomes large, simply surpassing the other players – rather than the amount by which one surpasses – becomes the dominant influencing factor in winning conflicts. The role played by group size is also natural: larger N correspond to more powerless players, who always have a strictly positive probability of winning conflicts when $\lambda < \infty$. Thus, dictators in larger groups face more pressure to maintain higher levels of power in parallel fashion to Proposition 6.

5 Conclusion

This paper developed an economic framework of how a society's distribution of power and resources evolves over time. Constructed using conventional tools from contest theory and asset accumulation models, this framework provided several novel insights into the emergence of inclusive, oligarchic, and dictatorial regimes, and the competitive forces that underlie them. This was studied using an intergenerational power accumulation contest among multiple lineages of players, where power was modeled as an asset that increases one's chances of winning conflicts over resources. Given any initial distribution of power, this model makes a *unique* equilibrium prediction[33] of how it will evolve over time and whether it will tend toward inclusivity, dictatorship, or oligarchy in the long run.

The main result of this model makes a far more concerning prediction, showing how power and resources generically fall into the hands of a few in large societies, in the absence of external intervention. This not only addresses a century-long open question by providing a robust theoretical mechanism for Michels's (1915) Iron Law of Oligarchy, it also shows that this Iron Law is driven by standard economic assumptions.

This paper generates new insights not only on the nature of political inequality in large societies, but also on the main conclusion of its foundation, Acemoglu and Robinson (2019, 2022). While I confirm that competitive pressure is indeed what allows strong, inclusive regimes to emerge, it is also precisely what causes it to destabilize in sufficiently large societies.

[33] In the absence of *shocks* to the group's power structure, the number of players, or any other model primitives that affect the costs or benefits of accumulating power. The takeaway from this prediction should not be fatalistic, as indeed these aforementioned shocks can an do occur. Rather, the implication of this prediction is that the deeper a society is in the basin of attraction of one type of regime (e.g. dictatorship), larger shocks are required to divert it towards another type of regime (e.g. inclusivity).

Michels (1915) asserted that "[h]istorical evolution mocks all the prophylactic measures that have been adopted for the prevention of oligarchy" (p. 406). In order to escape Michels's grim portent, we must systematically understand how societies' distributions of power evolve over time. This paper provides a benchmark model towards that end.

References

Acemoglu, Daron, "Politics and economics in weak and strong states," *Journal of monetary Economics*, 2005, *52* (7), 1199–1226.

—, "The Rise and Decline of Oligarchic Regimes," 5 2012. Zeuthen Lectures. https://economics.mit.edu/sites/default/files/inline-files/The%20Rise%20and%20Decline%20of%20Oligarchic%20Regimes.pdf.

—, "Escaping the New Gilded Age," *Project Syndicate*, 9 2024.

—, "The Fall and Rise of American Democracy," *Project Syndicate*, 12 2024.

—, "Reinvigorating Antitrust: Citizens, not just Consumers," 3 2024. Center for Economic Policy Research (CEPR). Interview by Lina Khan.

— **and James A Robinson**, "De facto political power and institutional persistence," *American economic review*, 2006, *96* (2), 325–330.

— **and** —, *The narrow corridor: States, societies, and the fate of liberty*, Penguin Press, 2019.

— **and James A. Robinson**, "Weak, Despotic, or Inclusive? How State Type Emerges from State versus Civil Society Competition," *American Political Science Review*, 2022, p. 1–14.

Aghion, Philippe., *Competition and growth : reconciling theory and evidence* Zeuthen lecture book series, Cambridge, Mass: MIT Press, 2005.

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt, "Competition and Innovation: An Inverted-U Relationship," *The Quarterly Journal of Economics*, 2005, *120* (2), 701–728.

Becker, Gary S., "A Theory of Competition Among Pressure Groups for Political Influence," *The Quarterly Journal of Economics*, 1983, *98* (3), 371–400.

—, "Public policies, pressure groups, and dead weight costs," *Journal of Public Economics*, 1985, *28* (3), 329–347.

Berry, S. Keith, "Rent-Seeking with Multiple Winners," *Public Choice*, 1993, *77* (2), 437–443.

Bowen, Renee and Zaki Zahran, "On dynamic compromise," *Games and Economic Behavior*, 2012, *76* (2), 391–419.

—**, Ilwoo Hwang, and Stefan Krasa**, "Personal power dynamics in bargaining," *Journal of Economic Theory*, 2022, *205*, 105530.

Callander, Steven, Dana Foarta, and Takuo Sugaya, "Market Competition and Political Influence: An Integrated Approach," *Econometrica*, 2022, *90* (6), 2723–2753.

Chowdhury, Subhasish M. and Sang-Hyun Kim, "A note on multi-winner contest mechanisms," *Economics Letters*, 2014, *125* (3), 357 – 359.

Clark, Derek J. and Christian Riis, "A Multi-Winner Nested Rent-Seeking Contest," *Public Choice*, 1996, *87* (1/2), 177–184.

Corchón, Luis and Matthias Dahm, "Foundations for contest success functions," *Economic Theory*, 2010, *43* (1), 81–98.

Deaton, Angus, "Rethinking my economics," 3 2024. International Monetary Fund, Finance & Development.

Dechenaux, Emmanuel, Dan Kovenock, and Roman M Sheremeta, "A survey of experimental research on contests, all-pay auctions and tournaments," *Experimental Economics*, 2015, *18* (4), 609–669.

Diefenbach, Thomas, "Why Michels' 'iron law of oligarchy' is not an iron law – and how democratic organisations can stay 'oligarchy-free'," *Organization Studies*, 2019, *40* (4), 545–562.

Dixit, Avinash K. and Robert S. Pindyck, *Incremental Investment and Capacity Choice*, Princeton University Press, Chapter 11 of Investment under Uncertainty.

Ewerhart, Christian, "A typology of military conflict based on the Hirshleifer contest," *University of Zurich, Department of Economics, Working Paper*, 2021, (400).

— **and Guang-Zhen Sun**, "The n-player Hirshleifer contest," *Games and Economic Behavior*, 2024, *143*, 300–320.

Fang, Dawei, Thomas Noe, and Philipp Strack, "Turning Up the Heat: The Discouraging Effect of Competition in Contests," *Journal of Political Economy*, 2020, *128* (5), 1940–1975.

Freedom House, "A leaderless struggle for democracy," Technical Report, Freedom House 2020. by Sarah Repucci.

— , "Freedom in the World 2022: The Global Expansion of Authoritarian Rule," Technical Report, Freedom House February 2022. By Sarah Repucci and Amy Slipowitz.

— , "Freedom in the World 2024: The Mounting Damage of Flawed Elections and Armed Conflict," Technical Report, Freedom House 2024.

Genicot, Garance and Debraj Ray, "Aspirations and Inequality," *Econometrica*, 2017, *85* (2), 489–519.

Harris, Christopher and John Vickers, "Racing with uncertainty," *The Review of Economic Studies*, 1987, *54* (1), 1–21.

Hillman, Arye L. and John G. Riley, "Politically Contestable Rents and Transfers," *Economics & Politics*, 1989, *1* (1), 17–39.

Hirshleifer, Jack, "Conflict and Rent-Seeking Success Functions: Ratio vs. Difference Models of Relative Success," *Public Choice*, 1989, *63* (2), 101–112.

— , "The paradox of power," *Economics & Politics*, 1991, *3* (3), 177–200.

— , "The Technology of Conflict as an Economic Activity," *The American Economic Review*, 1991, *81* (2), 130–134.

Hyde, Susan D., "Democracy's backsliding in the international environment," *Science*, 2020, *369* (6508), 1192–1196.

Jeon, Jee Seon and Ilwoo Hwang, "The emergence and persistence of oligarchy: A dynamic model of endogenous political power," *Available at SSRN 3177771*, 2020.

Johnson, Gregory A, "Organizational structure and scalar stress," *Theory and explanation in archaeology*, 1982, pp. 389–421.

Konrad, Kai A, "Dynamic contests and the discouragement effect," *Revue d'économie politique*, 2012, *122* (2), 233–256.

Krugman, Paul, *What the 1% Don't Want Us to Know*, Public Broadcasting Service, Apr 2014. Interview by Bill Moyers.

— , "Why Do the Rich Have So Much Power?," *The New York Times*, July 2020.

Lanchester, Frederick William, *Aircraft in warfare: The dawn of the fourth arm*, Constable limited, 1916.

Leach, Darcy K., "The Iron Law of What Again? Conceptualizing Oligarchy across Organizational Forms," *Sociological Theory*, 2005, *23* (3), 312–337.

— , "Oligarchy, Iron Law of," in James D. Wright, ed., *International Encyclopedia of the Social & Behavioral Sciences*, 2 ed., Oxford: Elsevier, 2015, pp. 201–206.

Li, Jin, Niko Matouschek, and Michael Powell, "Power Dynamics in Organizations," *American Economic Journal: Microeconomics*, 2017, *9* (1), 217–241.

Lyapunov, Aleksandr Mikhailovich, "The general problem of the stability of motion (English translation of original 1892 manuscript)," *International journal of control*, 1992, *55* (3), 531–534.

Maskin, Eric and Jean Tirole, "Markov perfect equilibrium: I. Observable actions," *Journal of Economic Theory*, 2001, *100* (2), 191–219.

Michels, Robert, *Political parties: A sociological study of the oligarchical tendencies of modern democracy*, Hearst's International Library Company, 1915.

OECD, *Growing unequal?: Income distribution and poverty in OECD countries* 2008.

— , *Divided We Stand: Why Inequality Keeps Rising* 2011.

— , "Income inequality and growth: The role of taxes and transfers," *OECD Economics Department Policy Notes*, January 2012, *No. 9*.

— , *In It Together: Why Less Inequality Benefits All* 2015.

— , *Does Inequality Matter?* 2021.

Papazyan, Freddie, "The Intertwined Dynamics of Power and Wealth," Research in progress, 2024.

Perret, Cedric, Emma Hart, and Simon T Powers, "From disorganized equality to efficient hierarchy: how group size drives the evolution of hierarchy in human societies," *Proceedings of the Royal Society B*, 2020, *287* (1928), 20200693.

Piketty, Thomas, *Capital in the Twenty-First Century*, Harvard University Press, 2014.

— , "Putting distribution back at the center of economics: Reflections on capital in the twenty-first century," *Journal of Economic Perspectives*, 2015, *29* (1), 67–88.

Prato, Carlo and Giovanna Maria Invernizzi, "Bending the Iron Law," *Working Paper*, 2023. Available at SSRN: https://ssrn.com/abstract=3462071.

Rausser, Gordon C., Johan Swinnen, and Pinhas Zusman, *Political power and economic policy : theory, analysis, and empirical applications*, Cambridge: Cambridge University Press, 2011. *Reuters*

Reuters, "Democracy under threat around the world -intergovernmental watchdog," 11 2023.

Riedl, Rachel Beatty, Paul Friesen, Jennifer McCoy, and Kenneth Roberts, "Democratic backsliding, resilience, and resistance," *World Politics*, 2024.

Saez, Emmanuel and Gabriel Zucman, "Alexandria Ocasio-Cortez's Tax Hike Idea Is Not About Soaking the Rich," *The New York Times*, January 2019.

Skaperdas, Stergios, "Contest success functions," *Economic Theory*, Jun 1996, *7* (2), 283–290.

Solow, Robert, "Robert Solow: Are we becoming an oligarchy?," 2014. Economic Policy Institute. Interview.

__ , "Thomas Piketty Is Right," *The New Republic*, 2014.

Stiglitz, Joseph E., "Of the 1%, by the 1%, for the 1%," *Vanity Fair*, May 2011.

__ , "8. Inequality and Economic Growth," *The Political Quarterly*, 2016, *86* (S1), 134–155.

Thurau, Jens, "Democracies under threat around the globe," *DW*, 3 2024.

Tullock, Gordon, "Efficient rent-seeking," in James M. Buchanan, Robert D. Tollison, and Gordon Tullock, eds., *Toward a Theory of the Rent-Seeking Society*, Texas A&M University Press, 1980, chapter 5, pp. 97–112.

United Nations, "With Rising Authoritarianism, Promoting, Respecting Human Rights More Important Than Ever, Secretary-General Stresses in Message on International Day," 11 2023. SG/SM/22057.

United Nations Department of Economic and Social Affairs, "World Social Report 2020: Inequality in a Rapidly Changing World," Technical Report, United Nations 2020.

Weber, Max, *Wirtschaft und gesellschaft*, JCB Mohr (P. Siebeck), 1925.

Winters, Jeffrey A., *Oligarchy*, Cambridge University Press, 2011.

Winters, Jeffrey A and Benjamin I Page, "Oligarchy in the United States?," *Perspectives on politics*, 2009, *7* (4), 731–751.

World Bank, *World Development Report 2006: Equity and Development* 2006. : World Bank; New York: Oxford University Press.

__ , *World Development Report 2017: Governance and the Law* 2017. : World Bank. doi:10.1596/978-1-4648-0950-7. License: Creative Commons Attribution CC BY 3.0 IGO.

Appendix

“Power Consolidation in Groups”

by Freddie Papazyan

(For Online Publication)

A Proofs

This section contains the proofs to the results in the main text (Propositions 1-7). Note that in the main text, resources are normalized to unity, for simplicity. Section C of this Online Appendix considers an extension where the size of resources is $y_N \geq 0$, which can vary with population size $N \in \{2, 3, \ldots\}$. In preparation for this section, Propositions 1-3 are proven assuming the aforementioned extended model. This proves the results in the main text (since it is the special case where $y_N = 1\ \forall N$) and avoids redundancy in Section C.

A.1 Proof of Propositions 1 and 2

Arbitrarily fix $\boldsymbol{x}_{t-\Delta} \in [0, \chi]^N$. Notice that the optimization problem in (5) is equivalent to the optimization problem

$$\begin{cases} \max\limits_{x_{it}} & H(x_{it}, \boldsymbol{x}_{-i,t}) - \Delta \cdot C\left(\frac{x_{it}-x_{i,t-\Delta}}{\Delta} + \min\left\{\frac{x_{i,t-\Delta}}{\Delta}, \delta\right\}, x_{i,t-\Delta}\right) \\ \text{s.t.} & 0 \leq x_{it} \\ & x_{it} \leq \chi \\ & x_{i,t-\Delta} - \min\{x_{i,t-\Delta}, \delta\Delta\} \leq x_{it} \end{cases} \tag{5'}$$

wherein x_{it} is the only choice variable. This is achieved by rearranging the equality constraint of (5)

$$x_{it} = \Delta \cdot I_{it} + \max\{x_{i,t-\Delta} - \delta, 0\} \;\Leftrightarrow\; I_{it} = \frac{x_{it} - x_{i,t-\Delta}}{\Delta} + \min\left\{\frac{x_{i,t-\Delta}}{\Delta}, \delta\right\}$$

and substituting out I_{it}. Optimization problem (5′) is now solved. I form the Lagrangian

$$\begin{aligned} \mathcal{L} = \;& H(x_i, \boldsymbol{x}_{-i}) - \Delta \cdot C\left(\frac{x_{it}-x_{i,t-\Delta}}{\Delta} + \min\left\{\frac{x_{i,t-\Delta}}{\Delta}, \delta\right\}, x_{i,t-\Delta}\right) \\ & + x_{it}\mu_1 + (\chi - x_{it})\mu_2 + (x_{it} - x_{i,t-\Delta} + \min\{x_{i,t-\Delta}, \delta\Delta\})\mu_3 \end{aligned} \tag{13}$$

where μ_1, μ_2, and μ_3 respectively denote the Lagrange multiplier of the first, second, and third inequality constraints of (5′). The first order condition is given by

$$h(x_{it}, \boldsymbol{x}_{-i,t}) = C_I\left(\frac{x_{it} - x_{i,t-\Delta}}{\Delta} + \min\left\{\frac{x_{i,t-\Delta}}{\Delta}, \delta\right\}, x_{i,t-\Delta}\right) - \mu_1 + \mu_2 - \mu_3. \tag{FOC}$$

Recall that $h(x_{it}, \boldsymbol{x}_{-i,t}) \equiv \frac{\partial}{\partial x_{it}} H(x_{it}, \boldsymbol{x}_{-i,t})$ (equation (4)) and C_I denotes the partial derivative of C with respect to its first argument. The Karush Kuhn-Tucker (KKT) conditions are given by (CS_1), (CS_2), and (CS_3), which respectively correspond to the first, second, and third inequality constraints of (5′).

$$x_{it} \geq 0,\ \mu_1 \geq 0,\ x_{it}\mu_1 = 0 \tag{CS$_1$}$$

$$x_{it} \leq \chi,\ \mu_2 \geq 0,\ (\chi - x_{it})\mu_2 = 0 \tag{CS$_2$}$$

$$x_{i,t-\Delta} - \min\{x_{i,t-\Delta}, \delta\Delta\} \leq x_{it},\ \mu_3 \geq 0,\ \left[\frac{x_{it} - x_{i,t-\Delta} + \min\{x_{i,t-\Delta}, \delta\Delta\}}{\Delta}\right]\mu_3 = 0 \tag{CS$_3$}$$

Case 1: Suppose the first two constraints of (5′) are both binding. This implies that $x_{it} = 0$ and $x_{it} = \chi$, which is a contradiction since $\chi > 0$. Therefore it is never possible for the first two constraints of (5′) to both be binding.

Case 2: Now suppose that only the first constraint of (5′) is slack. This implies that $x_{it} = \chi$ and $x_{it} = x_{i,t-\Delta} - \min\{x_{i,t-\Delta}, \delta\Delta\}$, which in turn jointly imply that $x_{i,t-\Delta} = \chi + \min\{x_{i,t-\Delta}, \delta\Delta\}$. This equation yields a contradiction when $x_{i,t-\Delta} > 0$ (since it implies that $x_{i,t-\Delta} = \chi + \delta\Delta > \chi$, which is impossible) and when $x_{i,t-\Delta} = 0$ (since it implies that $x_{i,t-\Delta} = \chi \neq 0$). Thus it is never possible for only the first constraint of (5′) to be slack.

Case 3: I now turn to the case where only the second constraint of (5′) is binding. In this case, $\mu_1 = \mu_3 = 0$, $\mu_2 \geq 0$, $x_{it} = \chi$, and $x_{it} > x_{i,t-\Delta} - \min\{x_{i,t-\Delta}, \delta\Delta\}$. In this case, it follows from (FOC) that

$$h(\chi, \boldsymbol{x}_{-i}) \geq C_I\left(\frac{\chi - x_{i,t-\Delta}}{\Delta} + \min\left\{\frac{x_{i,t-\Delta}}{\Delta}, \delta\right\}, x_{i,t-\Delta}\right) \tag{14}$$

Since h is a bounded function and C_I is strictly increasing in its first argument, it follows that there exists a sufficiently small $\tilde{\Delta}_1 > 0$ such that $\forall \Delta < \tilde{\Delta}_1$ the above inequality holds only at $x_{i,t-\Delta} = \chi$ and $\chi/\Delta > \delta$.

Case 4: Now consider the case where only the third constraint of (5′) is binding. In this case $\mu_1 = \mu_2 = 0$, $\mu_3 \geq 0$, $x_{it} \in (0, \chi)$, and $x_{it} = x_{i,t-\Delta} - \min\{x_{i,t-\Delta}, \delta\Delta\}$. If $x_{i,t-\Delta} \in [0, \delta\Delta]$, then we must have $x_{it} = x_{i,t-\Delta} - x_{i,t-\Delta} = 0$, which contradicts the fact that the first constraint of (5′) is slack in the present case. Therefore we must have $x_{i,t-\Delta} \in (\delta\Delta, \chi]$. In this case, it follows from (FOC) that

$$h(x_{it}, \boldsymbol{x}_{-i,t}) \leq C_I\left(\frac{x_{it} - x_{i,t-\Delta}}{\Delta} + \delta, x_{i,t-\Delta}\right) = C_I(0, x_{i,t-\Delta}) \tag{15}$$

where the equality follows from the fact that the third constraint of (5′) is binding. Therefore it is established that $(x_{it} - x_{i,t-\Delta})/\Delta = -\delta$ if both $h(x_{it}, \boldsymbol{x}_{-i,t}) \leq C_I(0, x_{i,t-\Delta})$ and $x_{i,t-\Delta} \in (\delta\Delta, \chi]$ are true.

Case 5: Now suppose that only the second constraint of (5′) is slack. In this case, we have $\mu_1 \geq 0$, $\mu_2 = 0$, $\mu_3 \geq 0$, $x_{it} = 0$, and $x_{it} = x_{i,t-\Delta} - \min\{x_{i,t-\Delta}, \delta\Delta\}$. In this case, we must have $x_{i,t-\Delta} \in [0, \delta\Delta]$. This is because when $x_{i,t-\Delta} \in (\delta\Delta, \chi]$, the binding first and third constraints of (5′) imply that $x_{i,t-\Delta} = \delta\Delta$, which is a contradiction. When $x_{i,t-\Delta} \in [0, \delta\Delta]$, these binding constraints imply that $x_{it} = x_{i,t-\Delta} - x_{i,t-\Delta} = 0$, which is true. It follows from this and from (FOC) that $x_{it} = 0$ if $h(0, \boldsymbol{x}_{-i,t}) \leq C_I(0, \delta\Delta)$ and $x_{i,t-\Delta} \in [0, \delta\Delta]$.

Case 6: Next, consider the case where only the first constraint of (5′) is binding. Here, we have $\mu_1 \geq 0 = \mu_2 = \mu_3$, $x_{it} = 0$, and $x_{it} > x_{i,t-\Delta} - \min\{x_{i,t-\Delta}, \delta\Delta\}$. The binding first constraint and slack third constraint of (5′) imply that $x_{i,t-\Delta} < \min\{x_{i,t-\Delta}, \delta\Delta\}$. This inequality yields a contradiction when $x_{i,t-\Delta} \leq \delta\Delta$ (since it implies that $x_{i,t-\Delta} < x_{i,t-\Delta}$) and when $x_{i,t-\Delta} > \delta\Delta$ (since it implies that $x_{i,t-\Delta} < \delta\Delta$). Thus it is impossible for only the first constraint of (5′) to bind.

Case 7 (Interior): Finally, attention is turned to the *interior* case where all constraints of (5′) are slack. Here, $\mu_i = 0\ \forall i \in \{1, 2, 3\}$, $x_{it} \in (0, \chi)$, and $x_{it} > x_{i,t-\Delta} - \min\{x_{i,t-\Delta}, \delta\Delta\}$. Equation (FOC) then implies that x_{it} satisfies the following equality:

$$h(x_{it}, \boldsymbol{x}_{-i,t}) = C_I\left(\frac{x_{it} - x_{i,t-\Delta}}{\Delta} + \min\left\{\frac{x_{i,t-\Delta}}{\Delta}, \delta\right\}, x_{i,t-\Delta}\right). \tag{16}$$

Note that $h(\cdot, \boldsymbol{x}_{-i,t})$ and its partial derivative with respect to its first argument are both bounded given any $\boldsymbol{x}_{-i,t} \in [0, \chi]^{N-1}$ and recall that $C_I(\cdot, x_{i,t-\Delta})$ is strictly increasing given any $x_{i,t-\Delta} \in [0, \chi]$. It then follows that either the solution is never interior at any period length or there exists a sufficiently small $\tilde{\Delta}_2 > 0$ such that given any fixed $\Delta < \tilde{\Delta}_2$, a unique interior value of x_{it} satisfies (16) and that at said x_{it}

$$\frac{\partial}{\partial x_{it}}\left[h(x_{it}, \boldsymbol{x}_{-i,t}) - C_I\left(\frac{x_{it} - x_{i,t-\Delta}}{\Delta} + \min\left\{\frac{x_{i,t-\Delta}}{\Delta}, \delta\right\}, x_{i,t-\Delta}\right)\right] < 0. \tag{17}$$

Note that the "$\min\left\{\frac{x_{i,t-\Delta}}{\Delta}, \delta\right\}$" term in (16) is equal to δ when $x_{i,t-\Delta} \in [\delta\Delta, \chi]$ and equal to zero when $x_{i,t-\Delta} = 0$. When $x_{i,t-\Delta} \in (0, \delta\Delta)$, the right-hand side of (16) is $C_I(x_{it}/\Delta, x_{i,t-\Delta})$. Note that $(0, \delta\Delta)$ converges to the empty set as Δ is made arbitrarily small in the sense that for any sequence $\{\hat{\Delta}_k\}_{k\in\mathbb{N}}$ such that $\hat{\Delta}_k \to 0$ as $k \to \infty$, we have $\cap_{k\in\mathbb{N}}(0, \delta\hat{\Delta}_k) = \emptyset$.

Given the above, for all sufficiently small Δ the maximizer x_{it}^* of equation (5′) is unique and characterized as follows: $x_{it}^* = (x_{i,t-\Delta} - \delta\Delta)\mathbb{1}_{(\delta\Delta,\chi]}(x_{i,t-\Delta})$ if $h(x_{it}^*, \boldsymbol{x}_{-i,t}) \leq C_I(\delta, x_{i,t-\Delta})$, $x_{it}^* = \chi$ if $x_{i,t-\Delta} = \chi$ and $h(\boldsymbol{x}_{it}*, \boldsymbol{x}_{-i,t}) \geq C_I(\delta, x_{i,t-\Delta})$, and otherwise x_{it}^* satisfies (16). Making Δ arbitrarily small yields the autonomous system in (6). Note that under Assumption 1, $(C_I)^{-1}$ is guaranteed to be a well-defined function (Kumagai, 1980). ■

A.2 Proof of Proposition 3

Parts 1, 2, 3, 4, and 5, respectively establish the necessary and sufficient conditions under which escalated inclusive, de-escalated inclusive, oligarchic, weak dictatorial, and strong dictatorial power structures are stable. Part 6 then shows that all other power structures will fail to be stable. In what follows, let N be arbitrarily fixed. Let $\mathbf{1}_k$ and $\mathbf{0}_k$ respectively denote $(1, \ldots, 1) \in \mathbb{R}^k$ and $(0, \ldots, 0) \in \mathbb{R}^k$ ($k = 1, \ldots, N$). Let $\boldsymbol{e}_i \in \mathbb{R}^N$ denote the i^{th} standard basis vector ($i = 1, \ldots, N$). Finally, let $B_\varepsilon(\boldsymbol{x})$ denote the ε-ball centered at $\boldsymbol{x}$, where $\varepsilon > 0$ and $\boldsymbol{x} \in \mathbb{R}^N$.

To establish the stability of inclusive and oligarchic power structures, I use Lyapunov's Direct/Second Method (Lyapunov, 1992). According to this method, $\bar{\boldsymbol{x}} \in [0, \chi]^N$ is stable if there exists a continuous, differentiable function $\Lambda : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ and an ε-ball of $\bar{\boldsymbol{x}}$, $B_\varepsilon(\bar{\boldsymbol{x}})$, such that the following hold:

1. $\Lambda(\boldsymbol{x}_t = \bar{\boldsymbol{x}}; \bar{\boldsymbol{x}}) = 0$ and $\Lambda(\boldsymbol{x}_t; \bar{\boldsymbol{x}}) > 0\ \forall \boldsymbol{x}_t \in B_\varepsilon(\bar{x}) \setminus \{\bar{\boldsymbol{x}}\}$.

2. $\frac{d}{dt}\Lambda(\boldsymbol{x}_t; \bar{\boldsymbol{x}}) < 0\ \forall \boldsymbol{x} \in B_\varepsilon(\bar{x}) \setminus \{\bar{\boldsymbol{x}}\}$

Λ is often referred to as a *Lyapunov function* and thought of as an "energy function." Intuitively, Lyapunov's Direct Method amounts to showing that the energy of the system *strictly* decreases to zero along all trajectories starting sufficiently close to a steady state $\bar{\boldsymbol{x}}$. In this proof, I use the Lyapunov function,

$$\Lambda(\boldsymbol{x}_t; \bar{\boldsymbol{x}}) \equiv \frac{1}{2} \sum_{i=1}^{N} (\bar{x}_i - x_{it})^2 \tag{18}$$

whose time-derivative is

$$\dot{\Lambda}(\boldsymbol{x}_t; \bar{\boldsymbol{x}}) \equiv \frac{d}{dt} \Lambda(\boldsymbol{x}_t) = -\sum_{i=1}^{N} (\bar{x}_i - x_{it}) \dot{x}_{it}. \tag{19}$$

Notice that by construction, given any $\bar{\boldsymbol{x}} \in [0, \chi]^N$, $\Lambda(\bar{\boldsymbol{x}}; \bar{\boldsymbol{x}}) = 0$ and $\Lambda(\boldsymbol{x}_t; \bar{\boldsymbol{x}}) > 0$ when $\boldsymbol{x}_t \neq \bar{\boldsymbol{x}}$. I now proceed with the proof of Proposition 3, proving each part in turn.

Proof of Part 1 Here I show that $\bar{\boldsymbol{x}} = \chi \mathbf{1}_N$ is stable if and only if Condition I

$$h(\chi, \chi \mathbf{1}_{N-1}) > C_I(\delta, \chi)$$

holds. First suppose that Condition I holds. Then $\dot{x}_i = 0 \forall i$ at $\boldsymbol{x} = \bar{\boldsymbol{x}}$ by (6). Continuity ensures that $\exists \varepsilon > 0$ s.t. $h(x_i, \boldsymbol{x}_{-i}) > C_I(x_i, \delta)$ holds $\forall i$ at every $\boldsymbol{x}$ in $B_\varepsilon(\bar{\boldsymbol{x}}) \cap [0, \chi]^N =: \mathbb{B}_1$. It then follows from Lemma 1 that $\dot{x}_i > 0$ $\forall i$ at every $\boldsymbol{x} \in \mathbb{B}_1$. It then follows that

$$\dot{\Lambda}(\boldsymbol{x}; \bar{\boldsymbol{x}}) = -\sum_{i=1}^{N} (\chi - x_i) \dot{x}_i < 0 \quad \forall \boldsymbol{x} \in \mathbb{B}_1 \setminus \{\bar{\boldsymbol{x}}\}.$$

Thus $\bar{\boldsymbol{x}}$ is stable.

Now suppose that $h(\chi, \chi \mathbf{1}_{N-1}) \leq C_I(\delta, \chi)$. It then follows that $h(\alpha, \alpha \mathbf{1}_{N-1}) \leq C_I(\delta, \alpha)$ for all $\alpha \in [0, \chi]$ since $h(\alpha, \alpha \mathbf{1}_{N-1})$ is constant in α and $C_I(\delta, \cdot)$ is weakly decreasing. Thus at every $\boldsymbol{x} \in \cup_{\alpha \in [0,\chi]} (\alpha \mathbf{1}_N)$ $\dot{x}_i = \dot{x}_j = (C_I)^{-1}(h(\alpha, \alpha \mathbf{1}_{N-1}), \alpha) - \delta \mathbb{1}_{(0,\chi)}(\alpha)$ for all i and j (by (6)) and $\dot{x}_i \leq 0$ $\forall i$ (by Lemma 1). This implies that $\bar{\boldsymbol{x}}$ is not stable because for any $\alpha \in [0, \chi)$ trajectories starting at $\boldsymbol{x} = \alpha \mathbf{1}_N$ will remain bounded away from $\bar{\boldsymbol{x}}$ for all $t \geq 0$.

Proof of Part 2 Here I show that $\bar{\boldsymbol{x}} = \mathbf{0}_N$ is stable if and only if Condition II, $h(0, \mathbf{0}_{N-1}) \leq C_I(0, 0)$, holds. First suppose that Condition II holds. Then $\dot{x}_i = 0$ $\forall i$ at $\boldsymbol{x} = \bar{\boldsymbol{x}}$ by the first piece of (6). Condition II implies that $h(0, \mathbf{0}_{N-1}) < C_I(\delta, 0)$ since $C_I(\cdot, 0)$ is strictly increasing. Continuity ensures that $\exists \varepsilon > 0$ such that $h(x_i, \boldsymbol{x}_{-i}) < C_I(\delta, x_i)$ holds $\forall i$ at every $\boldsymbol{x} \in B_\varepsilon(\bar{\boldsymbol{x}}) \cap [0, \chi]^N =: \mathbb{B}_2$. The first piece of equation (6) then implies that $\dot{x}_i = -\delta \mathbb{1}_{(0,\chi]}(x_i)$ $\forall i$ at every $\boldsymbol{x} \in \mathbb{B}_2$. Therefore $\dot{\Lambda}(\boldsymbol{x}; \bar{\boldsymbol{x}}) = \sum_1^N x_i \dot{x}_i < 0$ $\forall \boldsymbol{x} \in \mathbb{B}_2$. Thus $\bar{\boldsymbol{x}}$ is stable.

If Condition II fails then $h(\chi, \chi \mathbf{1}_{N-1}) > C_I(\delta, \chi)$. It then follows from 2 that $\dot{x}_i > 0$ $\forall i$ at $\boldsymbol{x} = \bar{\boldsymbol{x}}$, so $\bar{\boldsymbol{x}}$ cannot be stable.

Proof of Part 3 Arbitrarily fix $k \in \{2, \ldots, N-1\}$. Here I show that $\bar{\boldsymbol{x}} = (\chi \mathbf{1}_k, \mathbf{0}_{N-k})$ is stable if and only if Condition III holds. Focusing on this case is without loss of generality. Recall that Condition III requires that the following two equations hold:

$$h(\chi, (\chi \mathbf{1}_{k-1}, \mathbf{0}_{N-k})) > C_I(\delta, \chi), \tag{III.1}$$

$$h(0,(\chi\mathbf{1}_k,\mathbf{0}_{N-k-1})\leq C_I(0,0). \tag{III.2}$$

First assume that Condition III holds. Then at $\boldsymbol{x}=\bar{\boldsymbol{x}}$, $\dot{x}_i=0\ \forall\leq k$ by the second piece of (6) and $\dot{x}_i=0\ \forall i\geq k+1$ by Lemma 2. Notice that (III.2) implies that $h(0,(\chi\mathbf{1}_k,\mathbf{0}_{N-k-1})<C_I(\delta,0)$ since $C_I(\cdot,0)$ is strictly increasing. Continuity then ensures that $\exists\varepsilon>0$ such that $h(x_i,\boldsymbol{x}_{-i})>C_I(\delta,x_i)\ \forall i\leq k$ and $h(x_i,\boldsymbol{x}_{-i})<C_I(\delta,x_i)\ \forall i\geq k+1$ hold at all $\boldsymbol{x}$ in $B_\varepsilon(\bar{\boldsymbol{x}})\cap[0,\chi]^N=:\mathbb{B}_3$. It then follows from Lemma 1 that for each fixed $i\leq k$, $\dot{x}_i>0$ at every $\boldsymbol{x}\in\mathbb{B}_3\cap\{\boldsymbol{x}:x_i<\chi\}$; it follows from the first piece of (6) that for each $i\geq k+1$, $\dot{x}_i=-\delta$ at every $\boldsymbol{x}\in\mathbb{B}_3\cap\{\boldsymbol{x}:x_i>0\}$. Therefore at every $\boldsymbol{x}\in\mathbb{B}_3\setminus\{\bar{\boldsymbol{x}}\}$,

$$\dot{\Lambda}(\boldsymbol{x};\bar{\boldsymbol{x}})=-\textstyle\sum_{i=1}^N(\bar{\boldsymbol{x}}_i-x_i)\dot{x}_i=-\left[\sum_{i=1}^k(\chi-x_i)\dot{x}_i+\sum_{i=k}^N(-x_i)\dot{x}_i\right]<0.$$

Thus $\bar{\boldsymbol{x}}$ is stable.

Now assume that $h(0,(\chi\mathbf{1}_k,\mathbf{0}_{N-k-1})>C_I(0,0)$. It then follows from Lemma 2 that $\dot{x}_i>0\ \forall i\geq k+1$ at $\boldsymbol{x}=\bar{\boldsymbol{x}}$, so $\bar{\boldsymbol{x}}$ cannot be stable. Now suppose that $h(\chi,(\chi\mathbf{1}_{k-1},\mathbf{0}_{N-k}))<C_I(\delta,\chi)$. Then by (6), $\dot{x}_i$ satisfies $h(x_i,\boldsymbol{x}_{-i})=C_I(\dot{x}_i+\delta,x_i)\ \forall i\leq k$ at $\boldsymbol{x}=\bar{\boldsymbol{x}}$. In this case, $\dot{x}_i<0\ \forall i\leq k\ \boldsymbol{x}=\bar{\boldsymbol{x}}$ because $C_I(\cdot,x_i)$ is strictly increasing for any fixed x_i. It then follows that $\bar{\boldsymbol{x}}$ in this case. Finally, I consider the case where $h(\chi,(\chi\mathbf{1}_{k-1},\mathbf{0}_{N-k}))=C_I(\delta,\chi)$ and (III.2) holds. Consider the following subset $\mathbb{L}$ of $\mathbb{B}_3$:

$$\mathbb{L}:=\left\{\boldsymbol{x}\in\mathbb{B}_3:\boldsymbol{x}=\alpha\textstyle\sum_{i=1}^k\boldsymbol{e}_i\text{ for some }\alpha\in(\chi-\varepsilon,\chi)\right\}$$

Notice that by construction $\dot{x}_i=0\ \forall i\geq k+1$ at every $\boldsymbol{x}\in\mathbb{L}$. Furthermore, at any $\boldsymbol{x}\in\mathbb{L}$ we also have $\dot{x}_i=\dot{x}_j<0\ \forall i,j\in\{1,\dots,k\}$. This is because

$$\frac{\partial}{\partial\alpha}h(\alpha,(\alpha\mathbf{1}_{k-1},\mathbf{0}_{N-k}))=\frac{\lambda^2e^{\alpha\lambda}(N-k)[(k-2)e^{\alpha\lambda}+(N-k)]}{[N+(e^{\alpha\lambda}-1)k]^3}>0 \tag{20}$$

with the inequality following from the facts that $N>k\geq 2$, $\lambda>0$, and the fact that $e^{\alpha\lambda}>1\ \forall\alpha,\lambda$. In words: at all states in $\mathbb{L}$, powerless lineages ($k+1,\dots,N$) maintain zero power, and all other lineages let their power depreciate at the same rate. It then follows that if $\boldsymbol{x}_0\in\mathbb{L}$, then there exists a $\tau>0$ such that $\boldsymbol{x}_\tau\notin B_\varepsilon(\bar{\boldsymbol{x}})$. At every $t\in[0,\tau)$ $\dot{x}_{it}=\dot{x}_{jt}<0$ $\forall i,j\in\{1,\dots,k\}$ and $\dot{x}_{it}=0\ \forall i\in\{k+1,\dots,N\}$, so that at future time $t'\in(t,\tau)$, $\boldsymbol{x}_{t'}\in\mathbb{L}$. $\boldsymbol{x}_t$ moves along $\mathbb{L}$ in this fashion at all $t\in[0,\tau)$ until it leaves the ε-ball of $\bar{\boldsymbol{x}}$ at time τ. Therefore $\bar{\boldsymbol{x}}$ cannot be stable.

Proof of Part 4 Arbitrarily fix $d\in(0,\chi)$. Here I show that $\bar{\boldsymbol{x}}=d\boldsymbol{e}_i$ is stable $\forall i\in\{1,\dots,N\}$ if and only if Condition IV holds given d. I focus on the case where $i=1$ (i.e. $\bar{\boldsymbol{x}}=(d,\mathbf{0}_{N-1})$) as this is without loss of generality and it simplifies notation in this proof. Recall that the first part of Condition IV says that $h(d,\mathbf{0}_{N-1})=C_I(\delta,d)$ and that $\exists\alpha_1>0$ such that both of the following inequalities hold:

$$h(x_i,\mathbf{0}_{N-1})>C_I(\delta,x_i)\forall x_i\in(d-\alpha_1,d),$$

$$h(x_i,\mathbf{0}_{N-1})<C_I(\delta,x_i)\forall x_i\in(d,d+\alpha_1).$$

The second part of Condition IV says that $h(0,(d,\mathbf{0}_{N-2})<C_I(0,0)$.

First assume that Condition IV holds. It then follows from Lemma 1 (resp. Lemma 2) that $x_1 = 0$ (resp. $x_j = 0\ \forall j \geq 1$) at $\boldsymbol{x} = \bar{\boldsymbol{x}}$. The first part of Condition IV and Lemma 1 jointly imply that $\dot{x}_1 > 0\ \forall \boldsymbol{x} \in (d - \alpha_1, d) \times \{0\}^{N-1}$ and $\dot{x}_1 < 0\ \forall \boldsymbol{x} \in (d, d + \alpha_1) \times \{0\}^{N-1}$. Continuity guarantees that $\exists \alpha_2 > 0$ such that $\forall i \geq 2\ h(x_i, x_{-i}) < C_I(0, x_i)$ holds at all $\boldsymbol{x} \in B_{\alpha_2}(\bar{\boldsymbol{x}}) \cap [0,\chi]^N$. It then follows from the third piece of (6) that for each fixed $i \geq 2$, $\dot{x}_i = -\delta \mathbb{1}_{(0,\chi]}(x_i)$ at every $\boldsymbol{x} \in B_{\alpha_2}(\bar{\boldsymbol{x}}) \cap [0,\chi]^N$. Letting $\alpha := \min\{\alpha_1, \alpha_2\}$, we know that in $B_\alpha(\bar{\boldsymbol{x}}) \cap [0,\chi]^N$, trajectories away from the x_1 axis approach the x_1 axis; those on the x_1 axis do not depart from there and approach $\bar{\boldsymbol{x}}$ as $t \to \infty$.

To complete the proof that $\bar{\boldsymbol{x}}$ is stable, it just needs to be shown that trajectories that start in some neighborhood of $\bar{\boldsymbol{x}}$ always remain therein. For some $\gamma > 0$, the locus of $\{\boldsymbol{x} \in B_\gamma(\bar{\boldsymbol{x}}) \cap [0,\chi]^N : h(x_1, \boldsymbol{x}_{-1}) = C_I(\delta, x_1)\}$ is given by a hypersurface $x_1 = \mu(\boldsymbol{x}_{-1})$, where $\mu : \{\boldsymbol{x}_{-1} : x \in B_\gamma(\bar{\boldsymbol{x}}) \cap [0,\chi]^N\} \to [0,\chi]$ is continuous, strictly increasing in $x_j\ \forall j \geq 2$, and $\mu(\mathbf{0}_{N-1})$. This is because $h(d, \mathbf{0}_{N-1}) = C_I(\delta, d)$ (by the first part of Condition *IV*) and because increasing x_j translates $h(\cdot, \boldsymbol{x}_{-1})$ rightward $\forall j \geq 2$, given any $\boldsymbol{x}$. Since $x_1 = \mu(\boldsymbol{x}_{-1})$ bisects $B_\gamma(\bar{\boldsymbol{x}}) \cap [0,\chi]^N$, it follows that all trajectories originating in $B_{\min\{\alpha,\gamma\}}(\bar{\boldsymbol{x}}) \cap [0,\chi]^N$ always remain therein and approach $\bar{\boldsymbol{x}}$ as $t \to \infty$. Therefore $\bar{\boldsymbol{x}}$ is stable.

When $h(0, (d, \mathbf{0}_{N-2})) > C_I(0,0)$ then $\dot{x}_i > 0\ \forall i \geq 2$ at $\bar{\boldsymbol{x}}$ by Lemma 2, so $\bar{\boldsymbol{x}}$ cannot be stable. For the remainder of this proof, assume that $h(0, (d, \mathbf{0}_{N-2})) \leq C_I(0,0)$. If $h(d, \mathbf{0}_{N-1}) \neq C_I(\delta, d)$ then $\dot{x}_1 \neq 0$ by Lemma 1, so $\bar{\boldsymbol{x}}$ cannot be stable. Suppose that $h(\cdot, \mathbf{0}_{N-1}) - C_I(\delta, \cdot)$ equals zero at $x_1 = d$ and is constant in some ε_1-ball of d. It then follows that at each $\boldsymbol{x} \in (d, d + \varepsilon_1) \times \{0\}^{N-1}$, $\dot{x}_1 = 0$ because of Lemma 1. Moreover, at each $\boldsymbol{x} \in (d, d + \varepsilon_1) \times \{0\}^{N-1}$, $\dot{x}_i = 0\ \forall i \geq 2$ because $h(0, (x_1, \mathbf{0}_{N-2})) < h(0, (d, \mathbf{0}_{N-2})) \leq C_I(0,0)$; the first inequality follows from the fact that $\frac{\partial}{\partial d} h(0, (d, \mathbf{0}_{N-2})) < 0$, and the second inequality follows from the assumption mentioned in the second sentence of this paragraph. It follows that $\bar{\boldsymbol{x}}$ cannot be stable under the present supposition. Finally suppose that $h(\cdot, \mathbf{0}_{N-1}) - C_I(\delta, \cdot)$ equals zero at $x_1 = d$ and is increasing in some ε_2-ball of d. In this case, $\dot{x}_1 > 0$ at every $\boldsymbol{x} \in (d, d + \varepsilon_2) \times \{0\}^{N-1}$ because of Lemma 1. Moreover, for each $i \geq 2$, $\dot{x}_i = 0\ \forall \boldsymbol{x} \in (d, d + \varepsilon_2) \times \{0\}^{N-1}$ for the same reason as in the previous case. Thus $\bar{\boldsymbol{x}}$ cannot be stable.

Proof of Part 5 Here I show that $\bar{\boldsymbol{x}} = \chi \boldsymbol{e}_i$ is stable $\forall i$ if and only if Condition V holds. Recall that the first part of this condition requires that $\exists \varepsilon > 0$ such that $h(x_., \mathbf{0}_{N-1}) > C_I(\delta, x_.)$ $\forall x_. \in (\chi - \varepsilon, \chi)$ and the second part requires $h(0, (\chi, \mathbf{0}_{N-2}) \leq C_I(0,0)$. Here I focus on the case where $i = 1$ (i.e. $\bar{\boldsymbol{x}} = (\chi, \mathbf{0}_{N-1})$) without loss of generality.

First suppose that Condition V holds. Notice that continuity and the first part of Condition V jointly guarantee that $h(\chi, \mathbf{0}_{N-1}) \geq C_I(\chi, \mathbf{0}_{N-1})$. Thus $\dot{x}_1 = 0$ at $\bar{\boldsymbol{x}}$ because of the second piece of (6). The second part of Condition V implies that $\dot{x}_j = 0\ \forall j \geq 2$ at $\bar{\boldsymbol{x}}$ because of the first piece of (6). Notice that the second part of Condition V implies that $h(0, (\chi, \mathbf{0}_{N-2}) < C_I(\delta, 0)$ since $C_I(\cdot, 0)$ is strictly increasing. Continuity ensures that $\exists \varepsilon > 0$ such that $h(x_1, \boldsymbol{x}_{-1}) > C_I(\delta, x_1)$ and $h(x_j, \boldsymbol{x}_{-j}) < C_I(\delta, x_j)\ \forall j \geq 2$ hold at every state $\boldsymbol{x} \in B_\varepsilon(\bar{\boldsymbol{x}}) \cap [0,\chi]^N =: \mathbb{B}_5$. It then follows from Lemma 1 that $\dot{x}_1 > 0$ at every state in $\{\boldsymbol{x} \in \mathbb{B}_5 : x_1 < \chi\}$ and that for every $j \geq 2$ $\dot{x}_j < 0$ at every state in $\{\boldsymbol{x} \in \mathbb{B}_5 : x_j > 0\}$. Thus,

$$\dot{\Lambda}(\boldsymbol{x}; \bar{\boldsymbol{x}}) = -\textstyle\sum_{j=1}^N (\bar{x}_j - x_j)\dot{x}_j = -\left[(\chi - x_1)\dot{x}_1 + \textstyle\sum_{j=2}^N (-x_j)\dot{x}_j\right] < 0$$

at every $\boldsymbol{x} \in \mathbb{B}_5 \setminus \{\bar{\boldsymbol{x}}\}$. Therefore $\bar{\boldsymbol{x}}$ is stable.

If $h(0,(\chi,\mathbf{0}_{N-2}) > C_I(0,0)$, then it follows from Lemma 2 that $\dot{x}_j > 0\ \forall j \geq 2$ at $\bar{\boldsymbol{x}}$. Thus $\bar{\boldsymbol{x}}$ cannot be stable. Assume that the second part of Condition V holds for the remainder of this part. If $h(\chi,\mathbf{0}_{N-1}) < C_I(\delta,\chi)$, then $\dot{x}_1 < 0$ at $\bar{\boldsymbol{x}}$, as was shown above in Part 3 of this proof. It then follows that $\bar{\boldsymbol{x}}$ cannot be stable in this case. Finally, suppose that $\forall\varepsilon > 0$, $\exists x_\cdot \in (\chi-\varepsilon,\chi)$ such that $h(x_\cdot,\mathbf{0}_{N-1}) \leq C_I(\delta,x_\cdot)$. Continuity and Assumption 1.1 guarantee that $\exists\zeta > 0$ such that $h(0,(\chi,\mathbf{0}_{N-2}) < C_I(\delta,0)\ \forall \boldsymbol{x} \in B_\zeta(\bar{\boldsymbol{x}}) \cap [0,\chi]^N$. This implies that $\dot{x}_j = 0\ \forall j \geq 2$ at every $\boldsymbol{x} \in (\chi-\zeta,\chi)\times\{0\}^{N-1}$ because of the first piece of (6). Moreover, in this case $\dot{x}_1 \leq 0$ at every $\boldsymbol{x} \in (\chi-\zeta,\chi)\times\{0\}^{N-1}$ because of Lemma 1. Thus, trajectories that start in $(\chi-\zeta,\chi)\times\{0\}^{N-1}$ will always remain bounded away from $\bar{\boldsymbol{x}}$; thus $\bar{\boldsymbol{x}}$ is not stable in this case.

Proof of Part 6 The previous parts of this proposition established conditions on model primitives under which stable power structures in

$$\{0,\chi\}^N \sqcup \{d\boldsymbol{e}_i \,:\, d \in (0,\chi], i \in \{1,\dots,N\}\}$$

exist. I now show that no $\bar{\boldsymbol{x}}$ outside this set is ever stable under Assumptions 1 and 2. To prove this, it is useful to first define

$$\mathcal{K}_z(\bar{\boldsymbol{x}}) \equiv \{i \in \{1,\dots,N\} \,:\, \bar{\boldsymbol{x}}_i = z\}\ (z = 0,\chi); \quad \mathcal{K}_{int}(\bar{\boldsymbol{x}}) \equiv \{i \in \{1,\dots,N\} \,:\, \bar{\boldsymbol{x}}_i \in (0,\chi)\}.$$

Moreover, let $k_z(\bar{\boldsymbol{x}}) \equiv \#(\mathcal{K}_z(\bar{\boldsymbol{x}}))$ for each $z \in \{0,\chi\}$ and $k_{int}(\bar{\boldsymbol{x}}) \equiv \#(\mathcal{K}_{int}(\bar{\boldsymbol{x}}))$, where $\#(\cdot)$ outputs the cardinality of its input. Note that when there is little risk of confusion, $\mathcal{K}_z$, k_z, $\mathcal{K}_{int}$, and k_{int} may have their inputs suppressed.

First, let us consider an arbitrarily fixed $\bar{\boldsymbol{x}}$ where at least two players have interior levels of power (i.e. $k_{int}(\bar{\boldsymbol{x}}) \geq 2$). I now suppose that $\bar{\boldsymbol{x}}$ is stable and proceed to demonstrate that this yields a contradiction. If this supposition is true, then at $\boldsymbol{x} = \bar{\boldsymbol{x}}$ we have $\dot{x}_i = 0\ \forall\, i$ (by the first part of Definition 1). The following are then immediately implied by (6):

$$h(\bar{x}_i,\bar{\boldsymbol{x}}_{-i}) = C_I(\dot{x}_i+\delta,\bar{x}_i)\big|_{\dot{x}_i=0}\ \forall i \in \mathcal{K}_{int}(\bar{\boldsymbol{x}}), \tag{21}$$

$$h(\bar{x}_i,\bar{\boldsymbol{x}}_{-i}) > C_I(\delta,\bar{x}_i)\ \forall i \in \mathcal{K}_\chi(\bar{\boldsymbol{x}}), \tag{22}$$

$$h(\bar{x}_i,\bar{\boldsymbol{x}}_{-i}) < C_I(0,\bar{x}_i)\ \forall i \in \mathcal{K}_0(\bar{\boldsymbol{x}}). \tag{23}$$

First consider the case where for two players $j,j' \in \mathcal{K}_{int}(\bar{\boldsymbol{x}})$ we have $0 < \bar{x}_j < \bar{x}_{j'} < \chi$. It then follows from Assumption 1 that $C_I(\delta,\bar{x}_{j'}) \leq C_I(\delta,\bar{x}_j)$. It is also straightforward to verify that $h(\bar{x}_{j'},\bar{\boldsymbol{x}}_{-j'}) > h(\bar{x}_j,\bar{\boldsymbol{x}}_{-j})$: let $a = e^{\lambda\bar{x}_j}$, $b = e^{\lambda\bar{x}_{j'}}$, and $y = \sum_{\ell\in\{i\}_1^N\setminus\{j,j'\}} e^{\lambda\bar{x}_\ell}$. Then,

$$h(\bar{x}_j,\bar{\boldsymbol{x}}_{-j}) = \frac{\lambda a(y+b)}{(a+b+y)^2}, \tag{24a}$$

$$h(\bar{x}_{j'},\bar{\boldsymbol{x}}_{-j'}) = \frac{\lambda b(y+a)}{(a+b+y)^2}. \tag{24b}$$

Elementary algebra verifies that $h(\bar{x}_{j'},\bar{\boldsymbol{x}}_{-j'}) > h(\bar{x}_j,\bar{\boldsymbol{x}}_{-j})$.

The inequalities $C_I(\delta,\bar{x}_{j'}) \leq C_I(\delta,\bar{x}_j)$ and $h(\bar{x}_{j'},\bar{\boldsymbol{x}}_{-j'}) > h(\bar{x}_j,\bar{\boldsymbol{x}}_{-j})$, yield a contradiction in light of (21):

$$h(\bar{x}_{j'},\bar{\boldsymbol{x}}_{-j'}) = C_I(\delta,\bar{x}_{j'}) \leq C_I(\delta,\bar{x}_j) = h(\bar{x}_j,\bar{\boldsymbol{x}}_{-j}) < h(\bar{x}_{j'},\bar{\boldsymbol{x}}_{-j'}).\ \lightning$$

Therefore, there exist no stable $\bar{\boldsymbol{x}}$ in the set

$$\left\{\boldsymbol{x} \in [0,\chi]^N \;:\; \exists j \in \{1,...,N\},\; j' \in \{1,...,N\} \setminus \{j\} \text{ s.t. } 0 < x_j < x_{j'} < \chi\right\},$$

under Assumptions 1 and 2.

Now suppose that $\bar{x}_j = \alpha \;\forall j, j' \in \mathcal{K}_{int}(\bar{\boldsymbol{x}})$ for some $\alpha \in (0,\chi)$. It is possible to choose a sufficiently small ε so that the inequalities in (22) and (23) hold in the following subset of an ε ball of $\bar{\boldsymbol{x}}$:

$$\mathscr{A} \equiv \{\boldsymbol{x} \in [0,\chi]^N \;:\; x_i = \bar{x}_i \;\forall i \in \mathcal{K}_0(\bar{\boldsymbol{x}}) \cup \mathcal{K}_\chi(\bar{\boldsymbol{x}}), |\boldsymbol{x} - \bar{\boldsymbol{x}}| < \varepsilon\}$$

This follows from the continuity of h and the fact that within this neighborhood only interior components of $\bar{\boldsymbol{x}}$ vary in $\mathscr{A}$. Now consider $\hat{\boldsymbol{x}} = \bar{\boldsymbol{x}} + \rho \sum_{i \in \mathcal{K}_{int}(\bar{\boldsymbol{x}})} \boldsymbol{e}_i$, where $\rho > 0$ is chosen so that $\hat{\boldsymbol{x}} \in \mathscr{A}$. Then, we have for each $i \in \mathcal{K}_{int}(\bar{\boldsymbol{x}})$

$$h(\hat{x}_i, \hat{\boldsymbol{x}}_{-i}) > h(\bar{x}_i, \bar{\boldsymbol{x}}_{-i}) = C_I(\delta, \bar{x}_i) \geq C_I(\delta, \hat{x}_i),$$

where the first inequality follows from the fact that

$$\frac{\partial}{\partial \alpha} h(\alpha, (\alpha \mathbf{1}_{k_{int}(\bar{\boldsymbol{x}})-1}, \chi \mathbf{1}_{k_\chi(\bar{\boldsymbol{x}})}, \mathbf{0}_{k_0}(\bar{\boldsymbol{x}}))) > 0,$$

while the last inequality follows from Assumption 1 that $C_I(I,\cdot)$ is weakly decreasing for every fixed $I \geq 0$. Therefore if one perturbs $\boldsymbol{x}_t$ from $\bar{\boldsymbol{x}}$ to $\hat{\boldsymbol{x}}$, it follows from (6) that $\dot{x}_j = \dot{x}_{j'} \geq 0 \;\forall j, j' \in \mathcal{K}_{int}(\bar{\boldsymbol{x}})$ at the perturbed point. Note that by construction, $\dot{x}_{i\tau} = 0 \;\forall i \in \mathcal{K}_0(\bar{\boldsymbol{x}}) \sqcup \mathcal{K}_\chi(\bar{\boldsymbol{x}})$ at every $\tau \in [t, t']$. This implies that $\boldsymbol{x}_{t'}$ will leave the ε-ball of $\bar{\boldsymbol{x}}$ at some time $t > t'$, thereby ruling out stability.

The case that remains to be considered are the $\bar{\boldsymbol{x}} \in [0,\chi]^N$ such that $k_{int}(\bar{\boldsymbol{x}}) = 1$ and $k_0(\bar{\boldsymbol{x}}) < N - 1$.[34] By symmetry, I can focus on the case where $\bar{\boldsymbol{x}} = (\alpha, \chi \mathbf{1}_{k_\chi(\bar{\boldsymbol{x}})}, \mathbf{0}_{k_0(\bar{\boldsymbol{x}})})$ (for an arbitrarily fixed $\alpha \in (0,\chi)$) without loss; as before, rearranging the components of $\bar{\boldsymbol{x}}$ does not affect the proof (besides notation).

Suppose that $\dot{x}_i = 0 \;\forall i \in \{1,\ldots,N\}$ at $\bar{\boldsymbol{x}}$ (if this were not true, then the first part of Definition 1 has already been violated). I will now show that the second part of Definition 1 is violated. By continuity, (22) and (23) respectively hold $\forall i \in \mathcal{K}_\chi(\bar{\boldsymbol{x}})$ and $\forall i \in \mathcal{K}_0(\bar{\boldsymbol{x}})$ when $\boldsymbol{x}$ is inside some sufficiently small ε-ball of $\bar{\boldsymbol{x}}$. Since $\frac{\partial}{\partial\alpha} h(\alpha, (\chi \mathbf{1}_{k_\chi(\bar{\boldsymbol{x}})}, \mathbf{0}_{k_0(\bar{\boldsymbol{x}})})) > 0$ and $C_I(\delta,\cdot)$ is weakly decreasing, it follows that at $\boldsymbol{x} = (\alpha + \rho, \chi \mathbf{1}_{k_\chi(\bar{\boldsymbol{x}})}, \mathbf{0}_{k_0(\bar{\boldsymbol{x}})})$, we have $\dot{x}_1 > 0 \;\forall \rho \in (0,\varepsilon)$. Define the following subset of the ε-ball of $\bar{\boldsymbol{x}}$:

$$\mathbb{L} = \{\boldsymbol{x} \in B_\varepsilon(\bar{\boldsymbol{x}}) \;:\; x_1 \in (\alpha, \alpha + \varepsilon),\; x_i = 0 \;\forall i \in \mathcal{K}_0(\bar{\boldsymbol{x}}),\; x_j = \chi \;\forall i \in \mathcal{K}_\chi(\bar{\boldsymbol{x}})\}$$

Given the above, it has been established that at every $\boldsymbol{x} \in \mathbb{L}$ $\dot{x}_1 > 0$ and $\dot{x}_j = 0 \;\forall j \in \{2,\ldots,N\}$. It then follows that for any $\rho \in (0,\varepsilon)$, if the initial power structure is $\boldsymbol{x}_0 = (\alpha + \rho, \chi \mathbf{1}_{k_\chi(\bar{\boldsymbol{x}})}, \mathbf{0}_{k_0(\bar{\boldsymbol{x}})}) \in \mathbb{L}$, then $\boldsymbol{x}_t \in B_\varepsilon(\bar{\boldsymbol{x}})$ at some $t > 0$, thus violating the second part of Definition 1. ■

[34] Hence $k_\chi(\bar{\boldsymbol{x}}) \geq 1$; note that if $k_\chi(\bar{\boldsymbol{x}}) = 0$, this would correspond to the weak dictatorial case, which was covered in Part 4 of this proof.

The remaining proofs in this section directly assume that $y_N = 1\ \forall N \in \{2, 3, \ldots\}$, as in the main text. The analysis of the general case where $\{y_N\}_{N=2}^{\infty} \subseteq \mathbb{R}_+$ is continued in section C. The following remark discusses how a qualitatively very similar (but technically much more complicated) version of Proposition 3 can be achieved with more relaxed versions of Conditions II, III, and IV.

Remark 4. Oligarchic and dictatorial power structures remain stable in a technically weaker – but visually indistinguishable – Filippov (2013) sense if one relaxes the second parts of Conditions III, IV, and V by replacing "$\leq C_I(0,0)$" with "$< C_I(\delta, 0)$." The same is true for the de-escalated inclusive power structure if II is relaxed in the same way. If the left-hand side of each condition is strictly between $C_I(0,0)$ and $C_I(\delta,0)$, the corresponding power structure $\bar{x}$ is technically no longer a steady state (since part (a) of Definition 1 is violated) but still behaves very similarly to before (since part (b) of Definition 1 is still satisfied). In this case, x_t still approaches $\bar{\boldsymbol{x}}$ within finite time (and remains arbitrarily close forever afterwards). The only difference from before is that once $\boldsymbol{x}_t$ reaches $\bar{\boldsymbol{x}}$, it does not remain *exactly* there; it proceeds to *infinitesimally* oscillate around $\bar{\boldsymbol{x}}$. Formally, this is known as a type of *Zeno behavior*. In physical terms, the phenomenon just described is similar to a rubber ball that is arbitrarily close to being at rest (i.e. it is vibrating by an imperceptibly small amount).

The economic intuition is quite natural. This is illustrated in the case of a strong dictatorship $\bar{\boldsymbol{x}} = \chi \boldsymbol{e}_i$, where $i \in \{1, \ldots, N\}$ is the dictator player; the intuition is the same in the cases of weak dictatorship, oligarchy, and the de-escalated inclusive steady state. Suppose that $h(\chi, \mathbf{0}_{N-1}) > C_I(\delta, \chi)$, $h(0, (\chi, \mathbf{0}_{N-1})) \in (C_I(0,0), C_I(\delta, 0))$, and that $\boldsymbol{x}_t = \chi \boldsymbol{e}_i$ at some fixed time $t \in \mathbb{R}_+$. The former inequality guarantees that it is optimal for the dictator player i to maintain χ units of power at time t . The latter inequality implies that powerless players $j \neq i$ will accumulate power at some positive level, since their marginal benefit of power accumulation $h(0, (\chi, \mathbf{0}_{N-1}))$ exceeds the marginal cost of investment $C_I(I, 0)$ when $I = 0$ and because C is convex in I. However, these players will then proceed to let their powers depreciate, because the net marginal gain of maintaining any positive level of power $(h(x_j, x_{-j}) - C_I(\delta, x_j))$ is locally negative. Formally, this is because – under the aforementioned supposition – $h(0, (\chi, \mathbf{0}_{N-1})) < C_I(\delta, 0)$, and because $h(\cdot,\cdot)$ and $C_I(\cdot,\cdot)$ are assumed continuous.[35] The behavior of these weak players has a natural, concrete interpretation: they repeatedly try to "rise up" against the dictator by accumulating power before quickly acquiescing.

A.3 Proof of Proposition 4

Recall that by the proof of part 1 of Proposition 3, the $(\chi, \ldots, \chi)$ is stable if and only if

$$h(\boldsymbol{x}, (\chi, \ldots, \chi); N) > C_I(\delta, \chi) \tag{25}$$

[35]To make this intuition fully complete, recall that due to convex adjustment costs, it is optimal for players to differentially accumulate power, not in a discontinuous ("lumpy") fashion. Furthermore, note that everything stated in "The later inequality implies ... are assumed continuous" is formally a direct result of (6).

which is equivalent to

$$\frac{(N-1)\lambda}{N^2} > C_I(\delta,\chi). \tag{26}$$

Rearranging the above yields the quadratic inequality

$$0 > C_I(\delta,\chi)N^2 - \lambda N + \lambda \tag{27}$$

When $\frac{\lambda}{4} \leq C_I(\delta,\chi)$, the escalated inclusive power structure is not stable for any N; this quickly follows from Condition I.[36] Otherwise, solving the above quadratic inequality for N yields

$$\frac{\lambda - \sqrt{\big(\lambda - 4C_I(\delta,\chi)\big)\lambda}}{2C_I(\delta,\chi)} < N < \frac{\lambda + \sqrt{\big(\lambda - 4C_I(\delta,\chi)\big)\lambda}}{2C_I(\delta,\chi)}. \tag{28}$$

It is easily verified that the left-most term is always less than two:

$$\frac{\lambda - \sqrt{\big(\lambda - 4C_I(\delta,\chi)\big)\lambda}}{2C_I(\delta,\chi)} - 2 = \frac{\left[\sqrt{\lambda - 4C_I(\delta,\chi)} - \sqrt{\lambda}\right]\sqrt{\lambda - 4C_I(\delta,\chi)}}{2C_I(\delta,\chi)} \leq 0.$$

Noting that N is a natural number implies that no escalated inclusive steady state exist for group sizes larger than

$$\bar{N}^I_\chi \equiv \left\lfloor \frac{\lambda + \sqrt{\big(\lambda - 4C_I(\delta,\chi)\big)\lambda}}{2C_I(\delta,\chi)} \right\rfloor. \tag{29}$$

■

A.4 Proof of Corollary 1

Suppose that $C_I(\delta,\chi) = q$ for some $q > 0$ and that $\lambda \geq 4q$. Note that

$$\frac{\partial}{\partial\lambda}\left[\frac{\lambda + \sqrt{\lambda\cdot(\lambda-4q)}}{2q}\right] = \frac{1}{2q}\left[\frac{(\lambda-2q) + \sqrt{\lambda\cdot(\lambda-4q)}}{\sqrt{\lambda\cdot(\lambda-4q)}}\right], \tag{30}$$

and that

$$\frac{\partial}{\partial q}\left[\frac{\lambda + \sqrt{\lambda\cdot(\lambda-4q)}}{2q}\right] = \frac{-\sqrt{\lambda}}{2q^2\sqrt{\lambda-4q}}\left[(\lambda-2q) + \sqrt{\lambda\cdot(\lambda-4q)}\right]. \tag{31}$$

Observe that equations (30) and (31) are respectively positive and negative if $(\lambda - 2q) + \sqrt{\lambda\cdot(\lambda-4q)}$ is positive, which is always the case under the aforementioned supposition:

$$(\lambda-2q) + \sqrt{\lambda\cdot(\lambda-4q)} > \underbrace{(\lambda-4q)}_{\geq 0} + \underbrace{\sqrt{\lambda}}_{>0}\,\underbrace{\sqrt{\lambda-4q}}_{\geq 0} \geq 0$$

■

[36] Notice that if $h(\chi,\chi;2) = \lambda/4 < C_I(\delta,\chi)$, then Condition I ($h(\boldsymbol{x},(\chi,\ldots,\chi);N) \geq C_I(\delta,\chi)$) fails at all $N \geq 2$ since $h(\boldsymbol{x},(\chi,\ldots,\chi);N) = \frac{(N-1)\lambda}{N^2}$ is decreasing in N on $\{2,3,\ldots\}$.

A.5 Proof of Proposition 5

Arbitrarily fix $\varepsilon > 0$; suppose that $C_I(\delta, \cdot)$ is bounded by ε from below. By equation (28) in Proposition 4, the escalated inclusive power structure is stable only in group sizes smaller than

$$\bar{N}_\chi^I \equiv \left\lfloor \frac{\lambda + \sqrt{(\lambda - 4C_I(\delta, \chi))\lambda}}{2C_I(\delta, \chi)} \right\rfloor$$

when $\lambda > 4C_I(\delta, \chi)$, and is otherwise not stable in any group size permitted in this model ($N \in \{2, 3, \ldots\}$). If $\lambda \leq 4\varepsilon$, Proposition 5.1 trivially follows because of Assumption 1.2. Otherwise, $\bar{N}_\chi^I$ clearly remains finite as $\chi \to \infty$ given the aforementioned supposition.

Proposition 8 in section B of this Online Appendix showed in equation (39) that for each $k \in \{2, 3, \ldots\}$, k-archic power structures are stable only in groups smaller than

$$\bar{N}_\chi^{O_k} \equiv \left\lfloor k + e^{\lambda\chi} \left(\frac{\lambda + \sqrt{(\lambda - C_I(\delta, \chi))\lambda}}{C_I(\delta, \chi)} - k \right) \right\rfloor$$

when $\lambda > C_I(\delta, \chi)$ and not stable at any N otherwise. Similarly to before, if $\lambda < \varepsilon$ then Proposition 5.2 trivially follows because of Assumption 1.2. In the remaining case where $\exists z > 0$ s.t. $\lambda > C_I(\delta, \chi)\ \forall \chi > z$, notice that the limit of $\bar{N}_\chi^{O_k}$ as $\chi \to \infty$ only depends on the sign of

$$\lim_{\chi \to \infty} \frac{\lambda + \sqrt{(\lambda - C_I(\delta, \chi))\lambda}}{C_I(\delta, \chi)} - k.$$

$\bar{N}_\chi^{O_k} \to \infty$ as $\chi \to \infty$ if the above is strictly negative and $\lim_{\chi \to \infty} \bar{N}_\chi^{O_k} \leq 0$ otherwise.

Proposition 9 in section B of this Online Appendix showed in equation (42) that dictatorships are stable only in groups smaller than

$$\bar{N}_\chi^D \equiv \left\lfloor e^{\lambda\chi} \cdot \left\lfloor \frac{\lambda - 2C_I(\delta, \chi) + \sqrt{\lambda}\sqrt{\lambda - 4C_I(\delta, \chi)}}{2C_I(\delta, \chi)} \right\rfloor \right\rfloor$$

which becomes arbitrarily large as $\chi \to \infty$ given the aforementioned supposition.

■

A.6 Proof of Proposition 6

As in the above proofs, let $\boldsymbol{e}_i$ denote the i^{th} standard basis vector, and for each $n \in \mathbb{N}$ let $\mathbf{0}_n \in \mathbb{R}^n$ denote the vector of zeros. Recall that by Proposition 3, a dictatorial power structure where the strongest player has $d \in (0, \chi)$ units of power is stable if and only if

$$\begin{aligned} &h\big(\cdot, (0, \ldots, 0); N\big) \text{ intersects } C_I(\delta, \cdot) \text{ from above at } d, \text{ and} \\ &h\big(0, (d, 0, \ldots, 0); N\big) < C_I(0, 0), \end{aligned} \tag{Condition IV}$$

holds, and that strong dictatorial power structures are stable if and only if

$$h\big(\chi,(0,\ldots,0);N\big) > C_I(\delta,\chi) \text{ and } h\big(0,(\chi,0,\ldots,0;N)\big) < C_I(0,0). \qquad \text{(Condition V)}$$

holds. Arbitrarily fix $N \in \{2,3,\ldots\}$ and all other model primitives (χ, δ, λ, and C) such that for each $M\{N,N+1\}$, either (1) Condition IV holds at exactly one $d_M \in (0,\chi)$ and condition V fails or (2) Condition IV fails at all $d \in (0,\chi)$ and Condition V holds. The result of the proof is immediate in case where Condition V holds when the group size is $N+1$.

Now suppose that for each $M \in \{N,N+1\}$, Condition IV holds at exactly one $d_M \in (0,\chi)$ and condition V fails. Note that for all $n \in \{2,3,\ldots\}$ and $\ell \in \{0,1,2,\ldots\}$,

$$h\left(x_i - \frac{1}{\lambda}\ln\left(\frac{n+\ell-1}{n-1}\right), \mathbf{0}_{n-1}; n\right) = h\left(x_i, \mathbf{0}_{n+\ell-1}; n+\ell\right). \tag{32}$$

That is, given an initial group size of n, adding ℓ more players is equivalent to translating $h(\cdot,\mathbf{0}_{n-1};n)$ rightward by $\frac{1}{\lambda}\ln\left(\frac{n+\ell-1}{n-1}\right)$. Moreover, observe that $h(\cdot,\mathbf{0}_{N-1};N)$ can only intersect $C_I(\delta,\cdot)$ from above after the former attains its global maximum at $x_i = \frac{\ln(N-1)}{\lambda}$ as it is assumed that both functions are continuous and $C_I(I,x)$ is weakly decreasing in its second argument. It then follows that $d_M \in \left(\frac{\ln(M-1)}{\lambda},\chi\right)$ $\forall M \in \{N,N+1\}$. If $d_N \in \left(\frac{\ln(N-1)}{\lambda},\frac{\ln(N)}{\lambda}\right)$, then $d_N < d_{N+1}$ follows from the fact that $d_{N+1} > \frac{ln(N)}{\lambda}$. If instead $d_N \in \left(\frac{\ln(N)}{\lambda},\chi\right)$, note that $h(\cdot,\mathbf{0}_{N-1};N)$ and $h(\cdot,\mathbf{0}_N;N+1)$ are strictly decreasing on $\left(\frac{\ln(N)}{\lambda},\chi\right)$. Since the latter is a rightward translation of the former, and since $C_I(\delta,\cdot)$ is decreasing, it follows that $d_N < d_{N+1}$. Note that when restricting attention to the interval $\left(\frac{\ln(N)}{\lambda},\chi\right]$, translating $h(\cdot,\mathbf{0}_{N-1};N)$ rightward is equivalent to translating it upward; this immediately yields a contradiction upon supposing that Condition V holds for N and Condition IV holds for $N+1$ at exactly one $d_{N+1} \in (0,\chi)$. Noting that $h(0,(d,\mathbf{0}_{N-2});N)$ is decreasing in N and d $\forall (N,d,\lambda) \in \{2,3,\ldots\}\times(0,\infty)^2$, the proof is complete. ■

A.7 Proof of Proposition 7

I consider without loss of generality case where player 1 is a (weak) dictator: $\hat{\boldsymbol{x}} = (d,\mathbf{0}_N)$, where $d \in (0,\chi)$ is as in the first part of Condition IV, which is reproduced and discussed in the proof of Proposition 6, found immediately above. Assume that $C(\cdot,\cdot)$ complies with Assumption 1. Fix some small $\varepsilon > 0$ and choose some $\tilde{C}$ such that $\tilde{C}_I(\cdot,\cdot) = C_I(\cdot,\cdot)+\varepsilon$. Since $h(\cdot,\mathbf{0}_N)$ intersects $C_I(\delta,\cdot)$ from above at $d < \chi$, it follows from Assumptions 1 and 2 that $h(\cdot,\mathbf{0}_N) - C_I(\delta,\cdot)$ is locally decreasing around d. Hence, for sufficiently small but positive ε, $h(\cdot,\mathbf{0}_N)$ intersects $\tilde{C}_I(\delta,\cdot)$ at $\tilde{d} < d$. This completes the proof for part (i) of this proposition. Note that since C is assumed convex in its first argument, part (ii) immediately follows.

Turning to part (iii), we consider the effect of an increase in λ on d. Note that

$$\frac{\partial}{\partial\lambda}h(x_i,\mathbf{0}_{N-1}) = \underbrace{-\frac{(N-1)e^{\lambda x_i}}{\left[(N-1)+e^{\lambda x_i}\right]^3}}_{<0\ \because N\geq 2}(1 - e^{\lambda x_i} - N + \lambda x + \lambda x e^{\lambda x_i} - \lambda N x) \tag{33}$$

Setting the second term to less than zero and rearranging yields the inequality in part (iii) of this proposition, thereby completing its proof. ■

B Auxiliary Results

Lemma 1. *For each $i \in \{1, \dots, N\}$, sign$(\dot{x}_i)$ = sign$(h(x_i, \boldsymbol{x}_{-i}) - C_I(\delta, x_i))$ at every $\boldsymbol{x}$ in $\{\boldsymbol{x} \in [0,\chi]^N : x_i \in (0,\chi)\}$.*

Proof. Assume throughout that $x_i \in (0, \chi)$. First suppose that $h(x_i, \boldsymbol{x}_{-i}) < C_I(0, x_i)$. This implies that $h(x_i, \boldsymbol{x}_{-i}) < C_I(\delta, x_i)$ since $C_I(\cdot, x_i)$ is assumed to be strictly increasing. Then $\dot{x}_i = -\delta$ by the first piece of (6). Assume henceforth that $h(x_i, \boldsymbol{x}_{-i}) \geq C_I(0, x_i)$. It then follows from (6) that $\dot{x}_i$ satisfies

$$h(x_i, \boldsymbol{x}_{-i}) = C_I(\dot{x}_i + \delta, x_i) \tag{34}$$

since $0 < x_i < \chi$. If $h(x_i, \boldsymbol{x}_{-i}) = C_I(\delta, x_i)$, then we must have $\dot{x}_i = 0$ in order for (34) to hold. If $h(x_i, \boldsymbol{x}_{-i})$ is strictly greater (resp. less) than $C_I(\delta, x_i)$, $\dot{x}_i$ must be strictly positive (resp. negative) in order for (34) to hold because $C_I(\cdot, x_i)$ is assumed to be strictly convex. ■

Lemma 2. *Let $i \in \{1, \dots, N\}$ and arbitrarily fix $\boldsymbol{x} \in \{\boldsymbol{x} \in [0,\chi]^N : x_i = 0\}$. Then $\dot{x}_i > 0$ if $h(x_i, \boldsymbol{x}_{-i}) > C_I(0, x_i)$ and $\dot{x}_i = 0$ if $h(x_i, \boldsymbol{x}_{-i}) = C_I(0, x_i)$.*

Proof. Assume that $x_i = 0$ throughout. If $h(x_i, \boldsymbol{x}_{-i}) < C_I(0, x_i)$ then $\dot{x}_i = 0$ by the first piece of (6). Suppose that $h(x_i, \boldsymbol{x}_{-i}) = C_I(0, x_i)$. Since $x_i = 0$, (6) implies that $\dot{x}_i$ must satisfy

$$h(x_i, \boldsymbol{x}_{-i}) = C_I(\dot{x}_i, x_i). \tag{35}$$

In order for (35) to hold in the present case we must have $\dot{x}_i = 0$, since $C_I(\cdot, x_i)$ is strictly increasing. Finally, suppose that $h(x_i, \boldsymbol{x}_{-i}) > C_I(0, x_i)$. Then $\dot{x}_i$ must satisfy (35), which implies that $\dot{x}_i > 0$ since $C_I(\cdot, x_i)$ is strictly increasing. ■

Proposition 8. *Let $k \in \{2, \dots, N-1\}$. k-archies are never stable past group size*

$$\bar{N}_\chi^{O_k} = \left\lceil k + e^{\lambda\chi}\left(\frac{\lambda + \sqrt{(\lambda - C_I(\delta,\chi))\lambda}}{C_I(\delta,\chi)} - k\right)\right\rceil \tag{36}$$

Proof. To simplify notation, let q_χ denote $C_I(\delta, \chi)$ Recall that in the proof of Part 3 of Proposition 3, it was established that

$$h(\chi, (\chi \mathbf{1}_{k-1}, \mathbf{0}_{N-k}); N) > q_\chi \text{ and } h(0, (\chi \mathbf{1}_k, \mathbf{0}_{N-k-1}; N) \leq C_I(0,0) \tag{37}$$

are the necessary and sufficient conditions for the stability of each element of

$$\left\{\boldsymbol{x} \in \{0,\chi\}^N : \sum_{i=1}^N x_i = k\chi\right\}.$$

Note that the first inequality in (37) is equivalent to

$$\frac{\lambda\left[k - 1 + (N-k)e^{-\lambda\chi}\right]}{\left[k + (N-k)e^{-\lambda\chi}\right]^2} > q_\chi, \tag{38}$$

which yields the following quadratic inequality in N. Letting $\alpha = e^{-\lambda\chi}$ and $\beta = \frac{\lambda}{q_\chi}$, this is as follows:

$$\alpha^2 N^2 + \alpha[2(1-\alpha)k - \beta]N + \left\{\left[(1-\alpha)k - \frac{\beta}{2}\right]^2 + \left(1 - \frac{\beta}{4}\right)\beta\right\} < 0$$

Note that the coefficient of N^2 is positive; by the formula for the vertex of a parabola, it follows that no N satisfies this inequality if $\beta \cdot \left(1 - \frac{\beta}{4}\right) > 0$ ($\Leftrightarrow \lambda < 4q_\chi$). Otherwise, solving the above quadratic inequality yields the following:

$$k + e^{\lambda\chi}\left[\frac{\lambda - \sqrt{(\lambda - q_\chi)\lambda}}{q_\chi} - k\right] < N < k + e^{\lambda\chi}\left[\frac{\lambda + \sqrt{(\lambda - q_\chi)\lambda}}{q_\chi} - k\right] \tag{39}$$

Therefore, k-archies are never stable if N is greater than or equal to

$$\bar{N}_\chi^{O_k} = \left\lceil k + e^{\lambda\chi}\left(\frac{\lambda + \sqrt{(\lambda - q_\chi)\lambda}}{q_\chi} - k\right)\right\rceil$$

Finally note that if the second inequality of (37) holds for some $N = \tilde{N} \in \{2, 3, ...\}$ then it holds for all $N \geq \tilde{N}$ since $\frac{\partial}{\partial N} h(0, (\chi \mathbf{1}_k, \mathbf{0}_{N-k-1}; N) < 0$.

■

Proposition 9. *Suppose Condition IV holds for some N and χ, then there exist finite $\underline{N}_\chi^{D_W}, \bar{N}_\chi^{D_W}, \bar{N}_\chi^{D_S}$ such that*

1. *Weak dictatorships are stable if $\underline{N}_\chi^{D_W} \leq N < \bar{N}_\chi^{D_W}$.*
2. *Only strong dictatorships are stable if $\bar{N}_\chi^{D_W} \leq N \leq \bar{N}_\chi^{D_S}$*
3. *Weak and strong dictatorships are unstable if $N > \bar{N}_\chi^{D_S}$.*

Proof. Recall that the marginal benefit of investment for player i when her power is $x_i \in [0, \chi]$ and all other players have zero power is given by

$$h(x_i, \mathbf{0}_{N-1}; N) \equiv \frac{\lambda(N-1)e^{-\lambda x_i}}{(1 + (N-1)e^{-\lambda x_i})^2} \quad (\lambda > 0). \tag{40}$$

Recall that by (32) given an initial group size of N, adding K more players shifts marginal benefit $h(\cdot, \mathbf{0}_{N-1}; N)$ rightward.

Suppose that $d_N \boldsymbol{e}_i$ ($d_N \in (0, 1)$) is stable when group size is $N \in \{2, 3, ...\}$. This is only possible if $h(\cdot, \mathbf{0}_{N-1}; N)$ intersects $C_I(\delta, \cdot)$ from above at d_N. I demonstrate this via proof by contrapositive. If $h(\cdot, \mathbf{0}_{N-1}; N)$ is strictly greater than (strictly less than) $C_I(\delta, \cdot)$ at d_N, then the player i's marginal benefit of investment is strictly greater than (strictly less than) her marginal cost when $\boldsymbol{x} = d_N \boldsymbol{e}_i$, hence $\dot{x}_i > 0$ ($\dot{x}_i < 0$) at this point. Therefore if $h(d_N, \mathbf{0}_{N-1}; N) \neq C_I(\delta, d_N)$ then $d_N \boldsymbol{e}_i$ is not a steady state. If the intersection is from below, then $d_N \boldsymbol{e}_i$ is not stable. Let $\varepsilon > 0$. Perturbing x_i to $d_N + \varepsilon$ ($d_N - \varepsilon$) causes the marginal benefit of investment to become strictly greater than (strictly less than) the marginal cost for player i, thereby inducing $\dot{x}_i > 0$ ($\dot{x}_i < 0$) at this perturbed point. $d_N \boldsymbol{e}_i$ is not stable if $h(\cdot, \mathbf{0}_{N-1}; N)$ is tangent to $C_I(\delta, \cdot)$ at d_N. This is shown through similar reasoning. Finally, we consider the case where $h(x_i, \mathbf{0}_{N-1}; N) = C_I(\delta, x_i) \ \forall \ x_i \in (d_N - \varepsilon, d_N + \varepsilon)$ for some $\varepsilon > 0$. (That is, $h(\cdot, \mathbf{0}_{N-1}; N)$ and $C_I(\delta, \cdot)$ overlap in some ε-neighborhood of $x_i = d_N$.) Note that if $d_N \boldsymbol{e}_i$ is a steady state, we must have that $h(0, (d_N, \mathbf{0}_{N-2}); N) < C_I(\delta, 0)$ Otherwise $\dot{x}_j > 0 \, \forall j \neq i$ at this point. By the continuity of h and C_I, there must be some η-neighborhood of $d_N \boldsymbol{e}_i$ throughout which this strict inequality holds. Consider the perturbation to $\boldsymbol{x}' = (d_N + \nu) \boldsymbol{e}_i$, where $0 < \nu < \min\{\varepsilon, \eta\}$. By construction $h(d_N + \nu, \mathbf{0}_{N-1}; N) = C_I(\delta, d_N + \nu)$, so $\dot{x}_i = 0$ at this point. Similarly, $h(0, (d_N + \nu, \mathbf{0}_N); N) < C_I(\delta, 0)$, so $\dot{x}_j = 0 \ \forall j \neq i$. Therefore a trajectory that begins at $\boldsymbol{x}_0 = \boldsymbol{x}'$ does *not* approach $d_N \boldsymbol{e}_i$ in the limit, thereby ruling out its stability. Note that $\max\limits_{x_i \in \mathbb{R}} h(x_i, 0; 2) = \frac{\lambda}{4}$; this global maximum is attained at $x_i = 0$. Since (32) implies that $h(\cdot, \mathbf{0}_{N-1}; N)$ is a rightward horizontal translation of $h(\cdot, 0; 2)$ ($N = 2, 3, \ldots$), it follows that $\max\limits_{x_i \in \mathbb{R}} h(x_i, \mathbf{0}_{N-1}; N) = \frac{\lambda}{4}$ for every such N. Recall that $h(x_i, \mathbf{0}_{N-1}; N)$ attains its global maximum (about which it is unimodal) at $x_i = \frac{\ln(N-1)}{\lambda}$. Notice that this is monotonically increasing in N when $N \geq 2$. Since $C_I(I, x)$ is weakly decreasing in its second argument, it follows that $\min\limits_{x_i \in [0, \chi]} C_I(\delta, x_i) = C_I(\delta, \chi)$. Finally, notice that $\lim\limits_{x_i \to \infty} h(x_i, \mathbf{0}_N; N) = 0$. The desired result is immediate.

If $\lambda < 4C_I(\delta, \chi)$ then $\bar{N} = 2$. Now assume $\lambda > 4C_I(\delta, \chi)$ throughout the remaining duration of this proof. Since $h(x, \mathbf{0}_{N-1}; N)$ is unimodal about $\frac{\ln(N-1)}{\lambda}$ it follows that there exist exactly two values of N that solve $h(\chi, \mathbf{0}_{N-1}; N) = C_I(\delta, \chi)$. These are

$$\widehat{N}_1 = 1 + \left(\frac{e^{\lambda \chi}}{2C_I(\delta, \chi)} \right) \left(\lambda - 2C_I(\delta, \chi) - \sqrt{\lambda} \sqrt{\lambda - 4C_I(\delta, \chi)} \right) \tag{41}$$

and

$$\widehat{N}_2 = 1 + \left(\frac{e^{\lambda \chi}}{2C_I(\delta, \chi)} \right) \left(\lambda - 2C_I(\delta, \chi) + \sqrt{\lambda} \sqrt{\lambda - 4C_I(\delta, \chi)} \right). \tag{42}$$

Let $\bar{N}_\chi^{D_W} = \lceil \widehat{N}_1 \rceil$. If $\widehat{N}_2 \in \mathbb{N}$, then let $\bar{N}_\chi^{D_S} = \widehat{N}_2 - 1$; otherwise let $\bar{N}_\chi^{D_S} = \lfloor \widehat{N}_2 \rfloor$. When $\mathrm{N} = \bar{N}_\chi^{D_S}$, we know that $h(x_i, \mathbf{0}_{\bar{N}_\chi^{D_S} - 1}; \bar{N}_\chi^{D_S}) > C_I(\delta, x_i) \ \forall \ x_i \in (\chi - \varepsilon, \chi]$ for some $\varepsilon > 0$ and the reverse inequality holds in $[0, \chi - \varepsilon]$. Therefore the only stable dictatorships that exist are $\{\chi \boldsymbol{e}_i\}_1^{\bar{N}_\chi^{D_S}}$. It follows from (32) that for all $N > \bar{N}_\chi^{D_S}$, $h(x_i, \mathbf{0}_{N-1}; N) < C_I(\delta, x_i) \ \forall \ x_i \in [0, \chi]$. Therefore no dictatorial steady state can exist at any such N. By construction $h(x_i, \mathbf{0}_{\bar{N}_\chi^{D_W} - 1}; \bar{N}_\chi^{D_W}) > C_I(\delta, x_i) \ \forall \ x_i \in \left(\frac{\ln\left(\bar{N}_\chi^{D_W} - 1\right)}{\lambda}, \chi \right]$. It then follows that the only dictatorial steady states that exist are $\{\chi \boldsymbol{e}_i\}_1^M$. It follows from (32) that the same is true for all $N \in \{\bar{N}_\chi^{D_W}, \ldots, \bar{N}_\chi^{D_S}\}$. ∎

Remark 5. It is natural to expect that – given a *fixed* χ – dictatorships also become unfeasible once the group surpasses a certain size: powerless players, in sufficiently large numbers, overwhelm all dictators. This subject was considered in Proposition 5, which explores what happens when χ is made arbitrarily large. Note that a non-trivial *lower* bound $\underline{N}^W_\chi$ is possible; this follows from the fact mentioned earlier: powerless players always have a chance of winning conflicts.

C Extension to Population-Varying Resource Endowments

For simplicity, the baseline model in the main text assumes a fixed resource endowment size normalized to unity. This section explores the extension where the size of a society's resource endowment $y_N \geq 0$ can vary with its population size $N \in \{2, 3, \ldots\}$. The only assumption made on $\{y_N\}_{N=2}^{\infty} \subseteq \mathbb{R}_+$ is that it is a non-negative sequence. As discussed in Footnote 12 in the main text, focusing on an endowment economy is appropriate given the purpose of this paper; the case with a production economy will be considered in the follow-up work to this paper (Papazyan, 2024).

Mechanically, this extension generalizes one specific part of the setup: player $i_{\cdot}$'s benefit and marginal benefit from power (given power structure $\boldsymbol{x}$) are now respectively given by $y_N \cdot H(x_{i\cdot}, \boldsymbol{x}_{-i,\cdot})$ and $y_N \cdot h(x_{i\cdot}, \boldsymbol{x}_{-i,\cdot})$; the main text focused on the special case where $y_N = 1 \ \forall N$. The effect of this extension is trivial before Proposition 4: everything[37] in the main text remains identical apart from replacing each instance of "$H(\cdot)$" (resp. "$h(\cdot)$") with "$y_N H(\cdot)$" (resp. "$y_N h(\cdot)$").

The more intriguing finding from this extension regards the main results in the main text (Propositions 4 and 5), which showed that in sufficiently large societies, escalated inclusive regimes become unstable while dictatorships and concentrated oligarchies remain stable in arbitrarily large societies. It is reasonable to suspect that this result may be sensitive to this extension, since an increasingly large number of players compete over a fixed economic pie.[38] Interestingly, this turns out not to be the case. This is demonstrated by Proposition 10, below. This result also provides a tight analytical relationship between the stability of inclusive, oligarchic, and dictatorial regimes and the growth rate of y_N.

Proposition 10. *As population size N grows arbitrarily large, the inclusive power structure remains stable if and only if the size of the society's economy y_N grows sufficiently quickly in N; dictatorial and oligarchic regimes remain stable if and only if y_N grows sufficiently quickly, but not too quickly. More precisely:*

a. The inclusive power structure $\boldsymbol{x} = \chi \mathbf{1}_N$ is stable at arbitrarily large N if and only if the

[37] I.e. the model, exposition, results (i.e. Propositions 4 and 5), corresponding proofs, etc.

[38] To see why, recall why regimes with sufficiently many powerful players – that is, the escalated inclusive regime and oligarchies with more than $\bar{k}$ oligarchs (characterized in (12)) – eventually became unstable past some finite population size N in the main text: as N grows large, the marginal benefit of maintaining high levels of power in such regimes shrinks, eventually falling below its corresponding marginal cost, rendering the regime unstable thereafter.

following inequality holds at arbitrarily large N:

$$y_N \geq \frac{C_I(\delta,\chi)}{\lambda}\frac{N^2}{N-1} =: \underline{y}_N^I. \tag{43}$$

b. *For any fixed $k \in \{2,3,...\}$, all k-archies $\boldsymbol{x} \in \mathcal{O}_k$ are stable at arbitrarily large N if and only if the following holds at arbitrarily large N:*

$$\underline{y}_N^{O_k} := \frac{C_I(\delta,\chi)}{\lambda}\frac{[(1-e^{-\lambda\chi})k+e^{-\lambda\chi}N]^2}{(1-e^{-\lambda\chi})k-1+e^{-\lambda\chi}N} < y_N \leq \frac{C_I(0,0)}{\lambda}\frac{[(e^{\lambda\chi}-1)k+N]^2}{(e^{\lambda\chi}-1)k-1+N} =: \bar{y}_N^{O_k}. \tag{44}$$

c. *All strong dictatorships $\boldsymbol{x} \in \mathcal{D}_\chi$ are stable at arbitrarily large N if and only if*

$$\underline{y}_N^D := \frac{C_I(\delta,\chi)}{\lambda}\frac{[1+(N-1)e^{-\lambda\chi}]^2}{(N-1)e^{-\lambda\chi}} < y_N \leq \frac{C_I(0,0)}{\lambda}\frac{[e^{\lambda\chi}+N-1]^2}{e^{\lambda\chi}+N-2} =: \bar{y}_N^D \tag{45}$$

holds for arbitrarily large N.

Proof. Power structure $\boldsymbol{x} = \chi\mathbf{1}_N$ is stable if and only if $y_N h(\chi,\mathbf{1}_N;N) > C_I(\delta,\chi)$, which is equivlent to $y_N > [h(\chi,\mathbf{1}_N;N)]^{-1}C_I(\delta,\chi)$; part a of the result then immediately follows from equation (4).

For any fixed $k \in \{2,3,...\}$ and $N \in \{k+1,k+2,...\}$, all k-archies $\boldsymbol{x} \in \mathcal{O}_k$ are stable if and only if inequalities $y_N h(\chi,(\chi\mathbf{1}_{k-1},\mathbf{0}_{N-k});N) > C_I(\delta,\chi)$ and $y_N h(0,(\chi\mathbf{1}_k,\mathbf{0}_{N-k-1});N) \leq C_I(0,0)$ both hold; these inequalities are easily rearranged as $C_I(\delta,\chi)[h(\chi,(\chi\mathbf{1}_{k-1},\mathbf{0}_{N-k});N)]^{-1} < y_N \leq C_I(0,0)[h(0,(\chi\mathbf{1}_k,\mathbf{0}_{N-k-1});N)]^{-1}$. Part b then immediately follows from equation (4).

All strong dictatorships $\boldsymbol{x} \in \mathcal{D}_\chi$ are stable if and only if $y_N h(\chi,\mathbf{0}_{N-1};N) > C_I(\delta,\chi)$ and $y_N h(0,(\chi,\mathbf{0}_{N-2});N) \leq C_I(0,0)$ both hold. These inequalities are simply rearranged as $C_I(\delta,\chi)[h(\chi,\mathbf{0}_{N-1};N)]^{-1} < y_N \leq C_I(0,0)[h(0,(\chi,\mathbf{0}_{N-2});N)]^{-1}$. Part c immediately follows from equation (4). ■

All parts of the above characterization are non-trivial, as there is no universal observed relationship between population size and the size of a society's economy in terms of monotonicity, curvature, etc.[39] Note that all upper and lower bounds ($\underline{y}_N^{\cdot}$ and $\bar{y}_N^{\cdot}$, respectively) in Proposition 10 are monotonically increasing and convex, and are asymptotically affine.

Part a of Proposition 10 shows that the escalated inclusive power structure can only remain stable in large societies if and only if y_N persistently grows sufficiently rapidly in N; specifically, it must persistently outpace $\underline{y}_N^I$. Parts b and c show that the stability of oligarchic and dictatorial regimes rely on y_N growing sufficiently fast, but not too fast. Specifically, in order for k-archies (resp. strong dictatorships) to remain stable, y_N must persistently remain above $\underline{y}_N^{O_k}$ and below $\bar{y}_N^{O_k}$ (resp. above $\underline{y}_N^D$ and below $\bar{y}_N^D$). This is illustrated in Figure 7.

[39] Becker et al. (1999); Alesina et al. (2005); Acemoglu (2009); Peterson (2017); Bucci and Prettner (2020); Jones (2022).

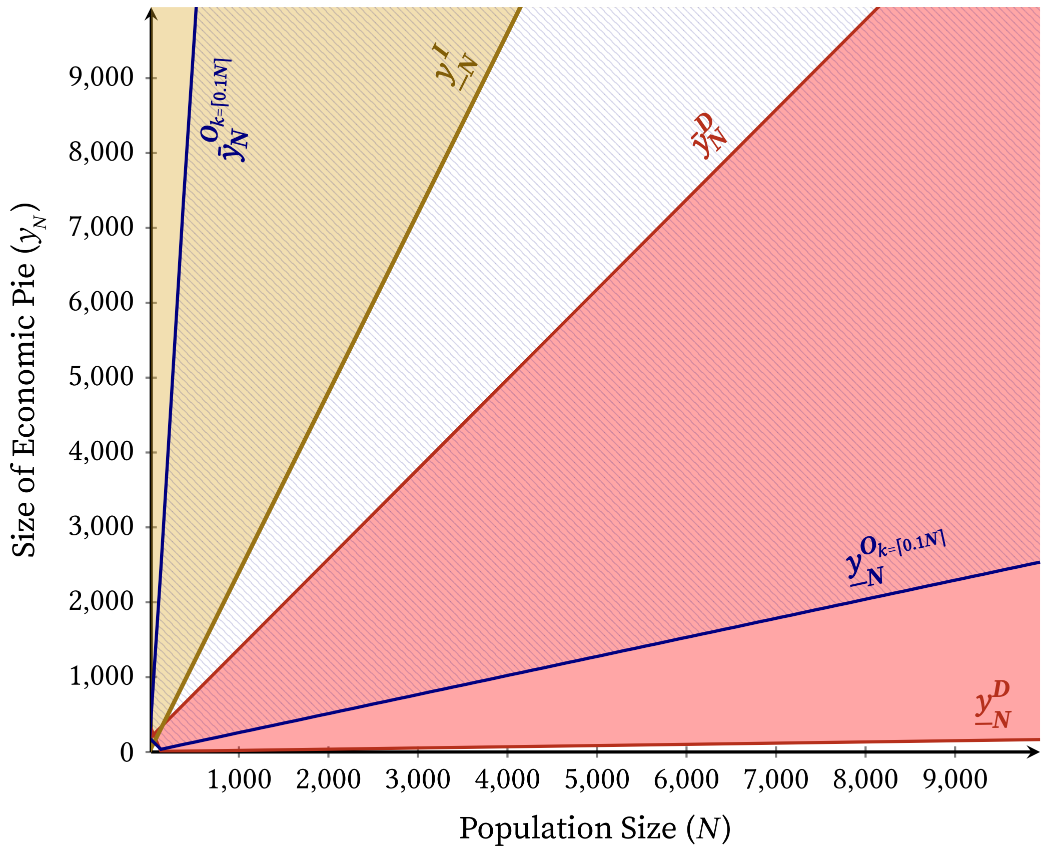


Figure 7: This figure assumes $\lambda = 5$, $C(I, z) = 6(1 + I)I$, $\delta = 0.5$, and $\chi = 1$.

Appendix References

Acemoglu, Daron, *Introduction to Modern Economic Growth*, Princeton University Press,

Alesina, Alberto, Enrico Spolaore, and Romain Wacziarg, "Chapter 23 - Trade, Growth and the Size of Countries," in Philippe Aghion and Steven N. Durlauf, eds., *Philippe Aghion and Steven N. Durlauf, eds.*, Vol. 1 of *Handbook of Economic Growth*, Elsevier, 2005, pp. 1499–1542.

Becker, Gary S., Edward L. Glaeser, and Kevin M. Murphy, "Population and Economic Growth," *The American Economic Review*, 1999, *89* (2), 145–149.

Bucci, Alberto and Klaus Prettner, "Endogenous education and the reversal in the relationship between fertility and economic growth," *Journal of Population Economics*, 2020, *33* (3), 1025–1068.

Filippov, Aleksei Fedorovich, *Differential equations with discontinuous righthand sides: control systems*, Vol. 18, Springer Science & Business Media, 2013.

Jones, Charles I, "The end of economic growth? Unintended consequences of a declining population," *American Economic Review*, 2022, *112* (11), 3489–3527.

Kumagai, Sadatoshi, "An implicit function theorem: Comment," *Journal of Optimization Theory and Applications*, 1980, *31* (2), 285–288.

Peterson, E. Wesley F., "The Role of Population in Economic Growth," *SAGE Open*, 2017, *7.*