

# CAMBRIDGE WORKING PAPERS IN ECONOMICS

## JANEWAY INSTITUTE WORKING PAPERS

### The Emergence of Enforcement

Luca Anderlini Georgetown University	Leonardo Felli University of Cambridge	Michele Piccione London School of Economics
---	---	--

#### Abstract

We ask how enforcement can endogenously emerge in a landscape in which only raw power, iron fists, govern the interaction of agents. If two agents are ranked in terms of power, the more powerful one can expropriate, at a cost, the less powerful one. Alternatively, both agents can engage in surplus-augmenting cooperation (e.g. trade).

If expropriation is not too costly and cooperation is not overwhelmingly productive, for any pair of ranked agents the possibility of expropriation prevents cooperation. The more powerful agent finds it profitable to expropriate the less powerful one. However, if expropriating agents who are net expropriators of others is cheaper, then a more powerful agent may endogenously become an “enforcer” for lower ranked agents. In equilibrium, the more powerful agent expropriates the less powerful ones by smaller amounts, and the less powerful ones cooperate and refrain from expropriating agents below them. This is because if they do not the more powerful agent will find it cheaper to expropriate only them by a larger amount.

Surprisingly, the details of the power structure are irrelevant for enforcement to emerge as an equilibrium phenomenon provided that the original jungle is inhabited by a sufficiently large number of agents and by one that dominates all others.

#### Reference Details

2250	Cambridge Working Papers in Economics
2222	Janeway Institute Working Paper Series
Published	30 <sup>th</sup> August 2022
Key Words	Jungle, Power Structures, Enforcement, Rule of Law
JEL Codes	C79, D00, D01, D31, K19, K40, K49
Websites	<a href="http://www.econ.cam.ac.uk/cwpe">www.econ.cam.ac.uk/cwpe</a> <a href="http://www.janeway.econ.cam.ac.uk/working-papers">www.janeway.econ.cam.ac.uk/working-papers</a>

# The Emergence of Enforcement\*

LUCA ANDERLINI  
(Georgetown University)

LEONARDO FELLI  
(University of Cambridge)

MICHELE PICCIONE  
(London School of Economics)

August 2022



**Abstract.** We ask how enforcement can endogenously emerge in a landscape in which only raw power iron fists, govern the interaction of agents. If two agents are ranked in terms of power, the more powerful one can expropriate, at a cost, the less powerful one. Alternatively, both agents can engage in surplus-augmenting cooperation (e.g. trade).

If expropriation is not too costly and cooperation is not overwhelmingly productive, for any pair of ranked agents the possibility of expropriation prevents cooperation. The more powerful agent finds it profitable to expropriate the less powerful one. However, if expropriating agents who are net expropriators of others is cheaper, then a more powerful agent may endogenously become an “enforcer” for lower ranked agents. In equilibrium, the more powerful agent expropriates the less powerful ones by smaller amounts, and the less powerful ones cooperate and refrain from expropriating agents below them. This is because if they do not the more powerful agent will find it cheaper to expropriate only them by a larger amount.

Surprisingly, the details of the power structure are irrelevant for enforcement to emerge as an equilibrium phenomenon provided that the original jungle is inhabited by a sufficiently large number of agents and by one that dominates all others.

JEL CLASSIFICATION: C79, D00, D01, D31, K19, K40, K49

KEYWORDS: Jungle, Power Structures, Enforcement, Rule of Law.

ADDRESS FOR CORRESPONDENCE: Leonardo Felli, Faculty of Economics, University of Cambridge, Austin Robinson Building, Sidgwick Avenue, Cambridge CB3 9DD, United Kingdom. [lf454@cam.ac.uk](mailto:lf454@cam.ac.uk)

---

\*A previous version of this paper was entitled “The Emergence of the Rule of Law.” We are grateful to Roger Lagunoff and Josh Teitelbaum for helpful comments and to Helen Watson for bibliographic references. Leonardo Felli gratefully acknowledges support from the Keynes Fellowship at the University of Cambridge. The current version of this paper can be obtained by clicking on the date above, or by scanning the QR code next to it.

## 1. Introduction

In virtually all economic transactions agents can enter agreements that will then be “enforced” for all parties concerned. In any model in which relations are regulated by contracts, it is implicitly assumed that the contracts will somehow be enforced even when ex-post one or more agents have an obvious incentive to renege. Nevertheless, the issue of how enforcement mechanisms come about in a landscape in which they are not available to begin with has been largely neglected in the literature.<sup>1</sup> Enforcers, especially in one-off transactions, are implicitly assumed to exist and to operate largely in the background at least until one of the parties decides to bring the other party to Court. This is seldom (if ever) the case in standard models, at least “on the equilibrium path.”

We ask what seems *prima facie* a difficult yet basic question. How do enforcers of rules emerge when *none* are present in the first place? Once some agents become enforcers, one can envisage how enforcement mechanisms might evolve to be stronger and spread across societies given the obvious advantages for all involved — more and more gains from rule-based activities such as “gains from trade” are generated. But how does a society take the “first step” towards the enforcement of rule-based interactions? To our minds, the issue of language acquisition has, in a very loose sense, a similar flavor:<sup>2</sup> when there is *no* language to build on, how do the first steps towards meaningful linguistic communication take place? Or, in our context, how do institutions that allow the invisible hand to operate arise in a world in which only iron fists govern the interaction between agents?

A classic view posits that the members of a society agree to a “social contract” and submit to the authority of selected individuals in exchange for maintaining a rule-based order. While such a view is appealing and has been hugely influential as a moral foundation of a social order,<sup>3</sup> it does not provide an explanation of its actual origin, indeed it begs the question of how the original social contract is enforced. In this paper we offer an explanation: a sharp division between the rulers and the ruled can emerge even in a society in which power is the only primitive ingredient in all relations among agents, including among the ruled.

---

<sup>1</sup>Of course, this is far from universally true. For instance the literature on “relational contracts” — see Gibbons and Roberts (2013) and Malcomson (2013) for extensive surveys — addresses a similar question although in a radically different environment. We return to a discussion of this literature in Section 2 below.

<sup>2</sup>We are, perhaps improperly, scratching the surface of a vast body of research, that goes back centuries and in more recent times at least to the debate between Skinner (1957) and Chomsky (1959).

<sup>3</sup>Starting with the classical, the inevitable references are Plato (1943) Hobbes (1968), Locke (1952) and Rousseau (2004).

A preview of our main result is that, in a society in which all relations are characterised only by sheer power, under plausible but by no means general conditions, enforcement mechanisms can emerge “in equilibrium.” Even in such a non-cooperative environment some agents seize wealth and effectively act as enforcers by deterring other agents from using their own power to appropriate wealth.

The main feature of our primitives is a non-cooperative environment governed by force.<sup>4</sup> Piccione and Rubinstein (2007) term this environment the “jungle,” and we do the same here.<sup>5</sup> In the jungle, agents can use their power to seize the assets of other agents.<sup>6</sup> No other tools to regulate economic relationships are available. In particular, this precludes the possibility of any agreement that is not self-enforcing.

There are, of course, a variety of routes that could be followed when defining the “power” of agents over other agents. These could include, for instance, notions of “graduated power” where one agent can have degrees of power over other agents, possibly spanning a continuum of values. For simplicity, as in Piccione and Rubinstein (2007), we opt for a binary notion. When any two agents are compared either one is more powerful than the other, so that a strict ranking obtains, or the two are not comparable. An agent who has power over another agent will be able (at a cost) to seize the assets of the other.<sup>7</sup>

In many jungles, two individuals in isolation will not be able to trade or undertake other surplus-creating activities. The more powerful of the two individuals will be able to expropriate the less powerful one and appropriate the proceeds as well as the initial assets of the other agent. This destroys his incentives to engage in surplus-creating cooperation.<sup>8</sup>

In our jungle, expropriation is costly. Stealing more is, other things being equal, more costly at an increasing rate so that it is not always the case that an expropriating agent wants to appropriate *all* available assets from less powerful agents. In addition, stealing costs contain a *strategic component*: stealing from less powerful agents who have suffered from less

---

<sup>4</sup>We do not wish to pursue the analogy in any detailed sense, but this is akin to the state of “anarchy” described by Nozick (1974).

<sup>5</sup>We return to the key differences between their jungle of ours in Section 2 below.

<sup>6</sup>Throughout, we use the terms “seize,” “appropriate,” “expropriate,” and “steal” in a completely interchangeable way.

<sup>7</sup>Given that our main results involve agents *not* seizing the assets of less powerful agents in equilibrium, it is arguable that the binary (as opposed to gradual) notion of power we use here qualitatively strengthens our results.

<sup>8</sup>This, of course, is far from universally true. For instance, stealing could be a very costly activity and/or the surplus generated could be large enough that all involved would find it profitable to participate.

expropriation or have gained more from expropriating other agents is comparatively cheaper. Thus, stealing may entail a strategic disadvantage for an agent: accumulating stolen wealth makes him a more likely target for expropriation by more powerful agents.

Based on these two ingredients, the simplest version of our results runs as follows. Two individuals who fail to trade with each other for the reasons we have described above may in fact succeed to do so when a *third* agent more powerful than both of them is present. In particular, the most powerful agent expropriates *both agents but each by a smaller amount*, thus giving them “enough room” to generate the potential surplus that they can achieve. The intermediate agent (the more powerful of the original pair) does not find it profitable to deviate and expropriate the least powerful one because by doing so he would face more expropriation by the most powerful of the three agents. The top agent would find it less costly to extract resources from him rather than from both of the agents ranked below him.

We think it is reasonable to conceptualize this type of equilibrium as one in which the most powerful agent effectively acts as an “enforcer” of the trading, or other surplus-creating relationship, between the two original agents. In this sense, enforcement emerges endogenously from the jungle as an equilibrium phenomenon.

As we mentioned,<sup>9</sup> the parametric conditions (costs and surplus gains) under which the type of equilibria with endogenous enforcement emerge are far from universal. However, they are by no means knife-edge, as will become apparent below. Surprisingly, we find that the phenomenon of endogenous enforcement generalizes considerably with respect to the power structure that characterizes the jungle. We find that equilibria with endogenous enforcement are present not only in simple jungles with three or more players in a “linear order” of power but also in considerably more general power structures in which power only constitutes a partial order of the players. Somewhat counterintuitively, the details of the power structure are irrelevant for the emergence of enforcement provided that the original jungle is inhabited by a sufficiently large number of agents and by one more powerful than all the others.

## 2. Related Literature

We have mentioned already that we borrow both terminology and some of the basic ingredients from the jungle of Piccione and Rubinstein (2007).<sup>10</sup> It should be noted that while raw power being the only original ingredient of interaction between agents is the same, there are

---

<sup>9</sup>See footnote 8 above.

<sup>10</sup>See also Piccione and Rubinstein (2004) and the more recent Rubinstein and Yildiz (2022).

key differences between their jungle and ours. First of all, their environment is one with heterogeneous goods and their focus is on whether efficient allocations can emerge in the jungle. In our set up there is only one good (“wealth”), which makes the patterns of expropriation transparent. Second, in their jungle there are no expropriation costs. Our results instead revolve around the presence of expropriation costs and require that these have a strategic component. In view of these key differences, from now on we will just refer to a jungle environment without any further reference to the work of Piccione and Rubinstein (2007).

A point to clear up first, is that we do *not* pursue the self-enforcement avenue that opens up once one considers long-run relationships (repeated or open-ended). Even an attempt to provide an exhaustive list of references here would be out of place. For the more game-theoretic side we refer to Mailath and Samuelson (2006), while for the more contract-theory angle we rely on the surveys by Gibbons and Roberts (2013) and Malcomson (2013).<sup>11</sup> We clearly differ from this approach since our model is a stylized two-period game. We do not rely on repeated-game features.

The effects of the availability of enforcement and the type of legal system it is embedded in have also been the subject of extensive studies. In an influential paper Acemoglu and Johnson (2005) study the role of enforcement stemming from property rights and their historical role on a global basis. In another paper that has had a large impact on subsequent literature, Glaeser and Shleifer (2002) examine the role of legal origins — “how” enforcement is deployed, common versus civil law — on economic and social outcomes. Both these literature streams have concentrated on the role of enforcement mechanisms and their origins on subsequent outcomes.<sup>12</sup> In a sense, here we try and delve one step deeper and ask how enforcement mechanisms may have emerged in a landscape where none were available.

A series of studies have examined a key example of “privately provided” enforcement, namely the Sicilian Mafia. This literature notably takes off with a sociological study by Gambetta (1996) documenting how the private provision of protection and enforcement is the main business of the Mafia, with Bandiera (2003) following in a very similar vein. In a sequence of papers Dixit (2003a,b, 2009, 2011) investigates theoretical models of private provision of protection and enforcement and their effects on economic governance. Here, we also delve a step deeper asking where the actual enforcement mechanism may have come from in the first place.

---

<sup>11</sup>See also the recent Calzolari, Felli, Koenen, Spagnolo, and Stahl (2021) for a specific empirical case study.

<sup>12</sup>See also Anderlini, Felli, and Riboni (2020).

We think of our contribution as being related to some of the recent literature on the “neolithic revolution:” the transition from hunter-gatherers to agriculture that eventually led to complex hierarchical states, complete with enforcement mechanisms including the ones that enable tax-levying. This literature is vast and we do not even attempt to survey it here. The dominant view is that following the adoption of farming the newfound surplus of food somehow led to complex hierarchical societies.

One empirical study is closely related to our theoretical model: a recent paper by Mayshar, Moav, and Pascali (2022) to which we will return shortly. One of their main points of departure is in turn well expounded by Scott (2017). Their key observation is that *not all* types of agricultural crops led to the same outcomes. In particular Scott (2017) writes:

*It is surely striking that virtually all classical states were based on grain [...] History records no cassava states, no sago, yam, taro, plantain, breadfruit, or sweet potato states.*

The difference between grains and the rest being that they are *storable* and easier to transport because of their lower water content. As Mayshar, Moav, and Pascali (2022) point out

*Cereal grains can be stored and, because they are harvested seasonally, have to be stored so that they can be drawn on for year-round subsistence. The relative ease of confiscating stored cereals, their high energy density, and their durability enhance their appropriability, thereby facilitating the emergence of tax-levying elites. Roots and tubers, in contrast, are typically perennial and do not have to be reaped in a particular period, and once harvested they are rather perishable.*

The upshot of the empirical tour de force by Mayshar, Moav, and Pascali (2022) is that

*In summary, our empirical analysis provides repeated evidence that the cultivation of cereals had a significant causal effect on the development of complex hierarchies and states, consistent with the appropriability theory. It also shows that the correlation between land productivity and hierarchy disappears when the cultivation of cereals is controlled for, consistent with our critique of the conventional productivity theory. Moreover, the finding that it is unlikely that complex hierarchies would emerge when productive roots and tubers are available supports both the appropriability theory and the critique of the productivity theory.*

The distinction between appropriable, storable grains and other crops has a notable incarnation in the *strategic component* of the cost of stealing that we postulate in our model. That a higher net balance of stolen wealth should make it easier to be the subject of stealing makes no sense if the wealth is immediately perishable and/or does not need to be stored as is the case with perennial crops.

As we outlined already above, this feature of the cost of stealing is key to our results that display equilibria with centralized enforcement arising in a landscape where only brute force is available. It is noteworthy that in our independent theoretical quest for “the emergence of enforcement” we should home in on a feature of the cost of stealing function that tallies so naturally with the storability and appropriability attributes of grains versus tubers and other crops.

### 3. Simple Environments

We begin by illustrating with a simple numerical example how the emergence of enforcement that we investigate in this paper can take place.

Consider a jungle populated by two agents who find it impossible to cooperate and create surplus because one is more powerful than the other. Surplus creation does not take place because the more powerful agent will find it more profitable to expropriate the less powerful (in fact powerless in this case) agent than to engage in surplus-creating cooperation.

The relationship between these two agents changes dramatically when we add a third agent to the jungle more powerful than both the original ones. The more powerful agent in the two-agent set up now becomes the intermediate agent. In equilibrium the more powerful agent expropriates both the original agents, but by a smaller amount because of expropriation costs. The intermediate agent refrains from expropriating the powerless agent and hence the surplus-creating activity can be sustained. The intermediate agent refrains from expropriation because if he did not the more powerful agent would find it to his advantage to expropriate only the intermediate agent by a larger amount since this will save on his expropriation costs.

#### 3.1. A Jungle With Two Agents

To avoid having to rename them later, the two agents in the initial case are named  $i = 2, 3$ . Agent 2 is more powerful than agent 3, written as  $2 \succ 3$ . At a cost, 2 can expropriate 3.

In the absence of surplus-creating cooperation, the endowments of the two agents are  $\omega_2 = \omega_3 = 80$ . Cooperation enhances the endowments: they become  $\tilde{\omega}_2 = \tilde{\omega}_3 = 90$ .

Throughout the paper, we consider a framework with two time periods,  $t = 0$  and  $t = 1$ . At  $t = 0$  the agents, both 2 and 3, decide simultaneously and independently whether to cooperate or not.<sup>13</sup> Once taken, the decisions to cooperate or not become observable to all. The choice not to cooperate can be interpreted as undertaking an (observable) activity that will enable an agent to engage in costly expropriation (for instance investing in offensive hardware). An agent who chooses to cooperate cannot later on (at  $t = 1$ ) engage in costly expropriation — he lacks the “offensive means” to do so.

If both agents decide to cooperate then surplus grows and no expropriation takes place at  $t = 1$ . So, in this case the agents both receive payoff  $\tilde{\omega}_i = 90$  and nothing further takes place. If instead at  $t = 0$  at least one of the two agents decides not to cooperate then we enter  $t = 1$  with the agents having their original endowments  $\omega_i = 80$ . If 2 has decided not to cooperate he will then be able to expropriate at a cost (part or all of) the endowment of the weaker agent 3.

Anticipating the notation we will use below, we let the amount by which 2 expropriates 3 be denoted by  $s_{23}$ . To keep things as simple as possible in this example we assume that the cost of expropriation that 2 incurs as a function of  $s_{23}$  is such that the optimal stolen amount is 80,<sup>14</sup> so that the entire endowment of 3 is stolen by 2. The expropriation cost that 2 incurs when stealing the endowment of 3 is  $\mathcal{C}(s_{23}) = 60$ .

It now follows easily that 2 will decide not to cooperate at  $t = 0$  and then proceed to expropriate 3 at  $t = 1$ . Indeed, if he cooperates he will end up with a payoff of at most  $\tilde{\omega}_2 = 90$ .<sup>15</sup> If instead he decides not to cooperate he will end up with a payoff of  $\Pi_2 = \omega_2 + s_{23} - \mathcal{C}(s_{23}) = 80 + 80 - 60 = 100$ .<sup>16</sup>

In this two-agent jungle even though surplus-augmenting cooperation is possible in principle it will not take place because of the expropriation power that 2 holds over 3.

---

<sup>13</sup>As will be apparent, the fact that 3 has something to decide is completely immaterial here. In fact it only generates the possibility that cooperation may not arise because of an equilibrium “coordination failure” between 2 and 3. To avoid this “spurious” possibility, in the full-fledged model we describe in Section 4 (Subsection 4.2) below powerless players (like 3 in this example) will be assumed to take no decision at all.

<sup>14</sup>All the quantities we use here and below in Subsection 3.2 are consistent with the actual cost function we describe in Section 5. See footnote 27 below where this correspondence is made explicit.

<sup>15</sup>Of course the payoff might be even lower if 3 does not cooperate, in that case it will be 80.

<sup>16</sup>Notice that the fact that 2 decides to expropriate is not true for all numerical values. However here the *net* benefit to 2 from expropriation is equal to  $80 - 60 = 20$ , which is greater than the up-front benefit from cooperation, namely  $90 - 80 = 10$ . To make the problem interesting, we will always assume that this is the case. This is the purpose of Assumption 2 below.

### 3.2. *A Jungle With Three Agents — The Emergence of Enforcement*

We now add a third agent to the jungle — agent 1, who is more powerful than both 2 and 3. So, we now have a jungle populated by  $1 \succ 2 \succ 3$ .

The endowments of 2 and 3 in case of cooperation or non cooperation are the same as before (namely 90 and 80 respectively). The simplest way to proceed to make the basic point is to assume that 1's endowment is not affected by his decision to cooperate or not. Hence, we just set  $\omega_1 = \tilde{\omega}_1$ , the value of which is immaterial to the rest of our example.

Given that 1's decision whether to cooperate or not does not affect endowments, the effect of  $t = 0$  cooperation decisions on endowments is just as in Subsection 3.1 above. If both 2 and 3 decide to cooperate their endowments grow from 80 to 90. If either decides not to cooperate both of their endowments remain at 80.<sup>17</sup>

Expropriation costs are as in Subsection 3.1.<sup>18</sup> So, the optimal amount that a non-cooperating agent expropriates from less powerful agents remains at 80, and the expropriation cost also stays at 60.

There are three additional features of the expropriation costs that come into play in the three-agent jungle. The first is that if an agent is expropriating multiple agents below him who have accumulated the same “net expropriation” (the amount they have expropriated from others minus any amount that has been taken from them),<sup>19</sup> he is indifferent as to whom he in fact expropriates — the cost is the same. So, for instance 1 could (as will be the case in the equilibrium we construct) expropriate 40 each from 2 and 3, and incur a cost of 60. Second, agents who have accumulated more net expropriation are comparatively cheaper to expropriate. Third, the expropriation process cannot leave any agent with a negative amount of resources.

Our candidate equilibrium is that at  $t = 0$  agents 2 and 3 decide to cooperate while 1 decides not to cooperate. At  $t = 1$ , agent 1 then expropriates agents 2 and 3 in equal amounts adding up to 80 so that  $s_{12} = s_{13} = 40$  at a cost of 60. In this candidate equilibrium 1 ends up with a payoff of  $\omega_1 + 80 - 60$ . Agents 2 and 3 both end up with a payoff equal to their cooperating endowment minus what agent 1 expropriates from them, namely  $90 - 40 = 50$ .

Checking for deviations by 1 is trivial. The only one to consider is that he could decide

---

<sup>17</sup>See footnote 13.

<sup>18</sup>See footnote 14.

<sup>19</sup>Obviously, in the three-agent jungle this is only possible when 1 expropriates 2 and 3.

to cooperate at  $t = 0$  and forfeit his ability to expropriate at  $t = 1$ .<sup>20</sup> If 1 deviates in this way he ends up with a payoff simply equal to his endowment, namely  $\tilde{\omega}_1 = \omega_1$ . This is a net loss of 20 relative to the proposed equilibrium.

Checking for deviations by agent 3 is also simple enough. Whether he decides to cooperate or not, he is a powerless agent and cannot expropriate anyone in any case. So, if at  $t = 0$  he were to deviate and decide not to cooperate, his endowment would stay at 80 and 40 of this endowment would still be expropriated by 1. So, he would end up with a payoff of 40, with a net loss of 10 relative to the proposed equilibrium.

Next, consider agent 2. If at  $t = 0$  he decided not to cooperate, his optimal decision at  $t = 1$  would be to expropriate 80 from agent 3 at a cost of 60. However, following the decisions of both agents 1 and 2 not to cooperate the behavior of agent 1 would also change at  $t = 1$ .<sup>21</sup>

As we mentioned above, we assume that the cost of expropriation is such that it is preferable (less costly) to expropriate an agent who has accumulated more net expropriated wealth from others. Viewed from the position of agent 1, agent 2 now has a net expropriation balance of 80 while the powerless 3 of course has balance of  $-80$ . Hence, after the deviation by 2, agent 1 will find it optimal to expropriate the full amount of 80 from 2 alone. Indeed following 2's expropriation of 3 there is clearly "nothing left" to be expropriated from the endowment of agent 3.<sup>22</sup> Hence following a deviation to not cooperate at  $t = 0$ , agent 2 ends up with a payoff composed as follows. His endowment is  $\omega_2 = 80$ , he expropriates  $s_{23} = 80$  from agent 3 at a cost of  $\mathcal{C}(s_{23}) = 60$ , and he is expropriated  $s_{12} = 80$  by agent 1. Hence his overall payoff is  $80 + 80 - 80 - 60 = 20$ , a net loss of 30 relative to the proposed equilibrium.

Recall that in the jungle populated by only agents 2 and 3, no surplus-generating cooperation was possible in equilibrium. This is the reason why we interpret the equilibrium in which agents 2 and 3 engage in surplus-generating cooperation while getting expropriated by 1 as a situation in which agent 1 plays the role of *enforcer of the cooperation* between 2 and 3.

---

<sup>20</sup>If he decides not to cooperate we are already considering "optimal" choices at  $t = 1$ , namely to expropriate a total of 80 from 2 and 3 at a cost of 60.

<sup>21</sup>Play would be in a different  $t = 1$  subgame.

<sup>22</sup>This is clearly a particular feature of the numbers chosen in this example.

## 4. A General Environment

### 4.1. Agents and Power Structures

There is a set  $\mathcal{P} = \{1, \dots, N\}$  of agents, with  $N \geq 3$ . The power structure according to which these agents are arranged is one of the key primitives of the model. We opt for a binary notion of power for which when two agents are compared either one is more powerful than the other, so that a strict ranking obtains, or the two are not comparable. An agent who has power over another agent will be able (at a cost) to seize resources belonging to the other agent.

Formally, we have a binary relation  $\succ$  that is a *strict partial order* on  $\mathcal{P}$ , so that  $\succ$  is irreflexive and transitive. In addition  $\succ$  must satisfy two conditions. First,  $1 \succ i$  for every  $i \neq 1$ . Second, for any  $i, j, h \in \mathcal{P}$  such that  $j \succ i$  and  $h \succ i$ , either  $j \succ h$  or  $h \succ j$ , so that all players have exactly one immediate predecessor. A pair  $(\mathcal{P}, \succ)$  that satisfies the above conditions is known as an *arborescence with a root* (in our case player 1).

The set of players  $j$  such that  $i \succ j$  is denoted by  $\mathcal{B}_i$  and the set of players  $j$  such that  $j \succ i$  is denoted by  $\mathcal{A}_i$ . We will often refer to players in  $\mathcal{B}_i$  and  $\mathcal{A}_i$  as players “below” and “above”  $i$  respectively. If for some  $i$  we have that  $\mathcal{B}_i = \emptyset$  we say that  $i$  is a “powerless” player. We will rule out the case where all players but player 1 are powerless. The following assumption will hold throughout.

**Assumption 1.** *General Power Structures:* The players  $\mathcal{P}$  together with the binary relationship  $\succ$  form an arborescence with a root, namely player 1. Furthermore,  $\mathcal{B}_i \neq \emptyset$  for some  $i \neq 1$ .

We refer to a pair  $(\mathcal{P}, \succ)$  satisfying Assumption 1 as a *General Power Structure*. In Section A.1 in the Appendix we report a few properties of arborescences and their “subsets” known as sub-arborescences that prove to be useful in our analysis.

### 4.2. Endowments, Timing, Cooperation and Expropriation Costs

The players make decisions in two periods,  $t = 0, 1$ . In the first period, each player  $i$  is endowed with an amount of resources  $\omega_i$  and decides whether to “cooperate” or not cooperate and thus “activate” his power to expropriate players below him. The decision by player  $i$  to cooperate is interpreted as an irreversible commitment *not* to subsequently expropriate any of the players in  $\mathcal{B}_i$ . Obviously, if player  $i$  is powerless such commitment is inevitable and

thus any powerless player will always be assumed to cooperate. Cooperation decisions are taken simultaneously and independently at  $t = 0$  and once taken are observable by all players. Hence a set of  $t = 0$  decisions to cooperate or not by all players identifies a  $t = 1$  *subgame*.

The decisions to cooperate or activate one's power could be viewed as observable choices to engage in time consuming activities that preclude or enable the use of a player's power to expropriate other players. For instance, some players may plant wheat and other acquire offensive weapons that enable expropriation.

In the second period,  $t = 1$ , if a set  $\mathcal{D}$  of players has chosen to cooperate in the first period their available resources *grow* while the resources of the players who do not cooperate remain constant. The state of player  $i$ 's resources at the beginning of the second period is denoted by  $\omega_i^{\mathcal{D}}$  where  $\omega_i^{\mathcal{D}} = \tilde{\omega}_i > \omega_i$  if  $i \in \mathcal{D}$ , and  $\omega_i^{\mathcal{D}} = \omega_i$  if  $i \notin \mathcal{D}$ . Often, it will also be convenient to refer to the set of non-cooperating players  $\mathcal{Z} = \mathcal{P}/\mathcal{D}$ .<sup>23</sup>

In the second period, after the set  $\mathcal{D}$  and all the  $\omega_i^{\mathcal{D}}$  are realized, the expropriation decisions are taken. Suppose that  $i$  is in  $\mathcal{Z}$ . We denote by  $s_i$  the vector with elements  $s_{ij}$  each indicating the amount that player  $i$  chooses to expropriate from  $j \in \mathcal{B}_i$ .<sup>24</sup> If  $i$  instead is in  $\mathcal{D}$  then  $s_i$  is a vector of zeroes of length  $\|\mathcal{B}_i\|$ . Let  $\mathbf{s}$  be the array  $\{s_i\}_{i=1}^N$  that describes all expropriation decisions as constrained by the general power structure  $(\mathcal{P}, \succ)$  and by the cooperation decisions. As standard, we will denote by  $s_{-i}$  the array  $\mathbf{s}$  with the vector  $s_i$  removed. Any non-cooperating player  $i$  takes  $s_{-i}$  as given when making his own decision about  $s_i$ .

**Remark 1.** *Ranking of Cooperating Players:* Notice that we are positing that a power structure is unaltered by the decision to cooperate. Cooperating players cannot be expropriated by players that are ranked below them and who have activated their power. This modeling choice may be more or less appealing, depending on how "hard-wired" into the players one wishes to interpret to be the notion of power.

We proceed in this way since it simplifies the notation and the analysis. With the exception of Proposition 4, our results remain valid if all players who choose to cooperate join one of their powerless successors in the power ranking while the players who activate their power maintain their rank according to the original power structure.

Expropriation is assumed to be costly. When a non-cooperating player  $i$  selects a vector

---

<sup>23</sup>All powerless players are assumed to be automatically in  $\mathcal{D}$ . See footnote 13 above.

<sup>24</sup>The quantities  $s_{ij}$  need to satisfy certain inequality constraints that will be specified shortly.

of expropriations  $s_i$ , she will bear an expropriation cost denoted by  $\mathcal{C}_i(s_i, s_{-i})$  where the dependence on both  $s_i$  and all other players expropriation choices  $s_{-i}$  is made explicit.

Given the array  $\mathbf{s}$  of expropriation decisions, the payoff of a non-cooperating player  $i$  is equal to her net final resources obtained and surrendered via expropriation, minus the expropriation costs:

$$\pi_i = \omega_i + \sum_{j \in \mathcal{B}_i} s_{ij} - \sum_{q \in \mathcal{A}_i} s_{qi} - \mathcal{C}_i(s_i, s_{-i}) \quad (1)$$

If player  $i$  is cooperating, she will enjoy a payoff equal to her resources from cooperation minus the amount that is expropriated from her:<sup>25</sup>

$$\pi_i = \tilde{\omega}_i - \sum_{q \in \mathcal{A}_i} s_{qi} \quad (2)$$

## 5. Expropriation Costs

### 5.1. Preliminaries

Consider the second period  $t = 1$  after all cooperating decisions have been taken and observed by all players and a non-cooperating player  $i$ . Since he takes  $s_{-i}$  as given and his payoff is as in (1) above, the choice problem of  $i$  is to select a vector  $s_i$  so as to maximize

$$\sum_{j \in \mathcal{B}_i} s_{ij} - \mathcal{C}_i(s_i, s_{-i}) \quad (3)$$

We are now ready to define the class of cost functions that we will work with throughout the paper. Consider a pair  $(i, j)$  with  $j \in \mathcal{B}_i$  and let

$$h_{ij} = \sum_{\substack{q \in \mathcal{A}_j \\ q \neq i}} s_{qj} - \sum_{z \in \mathcal{B}_j} s_{jz} \quad (4)$$

so that  $h_{ij}$  are the net expropriation *losses* to player  $j \in \mathcal{B}_i$  that player  $i$  takes as given since they exclude the action of  $i$  himself. Indeed, these are the resources that have been expropriated from player  $j$  by players above him other than  $i$ , minus the resources that player

---

<sup>25</sup>The amounts expropriated need to be consistent since no player can be left with a negative amount of resources. We return to this issue below in Subsection 5.2.

$j$  has expropriated from players below him and represent the net balance of expropriated wealth of  $j$  from the point of view of player  $i$ .

Player  $i$ 's expropriation costs are of the form

$$\mathcal{C}(s_i, s_{-i}) = \mathcal{C}^1 \left[ \sum_{j \in \mathcal{B}_i} s_{ij} \right] + \sum_{j \in \mathcal{B}_i} s_{ij} \mathcal{C}^2(h_{ij}) \quad (5)$$

where  $\mathcal{C}^1(\cdot)$  and  $\mathcal{C}^2(\cdot)$  are strictly increasing,  $\mathcal{C}^1$  is strictly convex and satisfies standard Inada conditions,  $\mathcal{C}^1(0) = 0$ , and  $\mathcal{C}^2(\cdot) \geq 0$ .

The first term is a standard “effort” cost which depends only on the total amount that player  $i$  steals.<sup>26</sup> The second term in (5) is the strategic crowding out term that we have referred to multiple times above. The more resources other players expropriate from player  $j$  and the less resources player  $j$  expropriates from other players, the higher the *marginal* cost of expropriating player  $j$ .

In our view, to properly interpret the strategic crowding out term in our cost of stealing function requires that the resources that are the subject of expropriation are storable and hence somewhat durable as well. If it is easier to steal from a player who has a net expropriation surplus, presumably this is because the resources to be stolen are (at least in part) those that he has actually stolen from others. It is almost tautological that since the cost of stealing from a player depends on how much he has accumulated in (net) stealing from others, the resources must be of an “accumulable” nature. As we noted above, this is a point that figures prominently in the findings of Mayshar, Moav, and Pascali (2022) that we discussed at length in Section 2 above.

A natural observation at this point would be to say that a player that has accumulated more resources in net stealing could use them to increase his defense capabilities and hence make it more costly for other players to steal from him. We do not explicitly consider this effect since we view it as less “primitive” than the “liquidity” effect that we emphasize here that has the opposite sign. This is because our main purpose is precisely to explain the emergence of a social structure that may enable such complex and stratified defense strategies in the first place.

Before moving on, it is helpful to point out that a “canonical” class of cost functions that

---

<sup>26</sup>Or “attempts” to steal. See Remark 2 below, and Section A.2 in the Appendix.

fits (5) is

$$\mathcal{C}(s_i, s_{-i}) = A \left[ \sum_{j \in \mathcal{B}_i} s_{ij} \right]^\gamma + B \sum_{j \in \mathcal{B}_i} s_{ij} (\bar{h} + h_{ij}) \quad (6)$$

where  $\gamma$ ,  $\bar{h}$ ,  $A$  and  $B$  are constants that satisfy  $\gamma > 1$ ,  $A > 0$  and  $0 < B < 1/\bar{h}$ . The numerical examples that we discussed in Section 3 were constructed using this class.<sup>27</sup>

### 5.2. Rationing and Interiority

It is clear that the expropriation vectors in the subgames that follow the first period decisions could be inconsistent since the total amount expropriated from a player  $j$  could exceed the player's resources.

A possible way to resolve this inconsistency would be to include rationing rules that constrain the actual amount that is expropriated in such cases. Following this route, each  $s_{ij}$  can be viewed as the effort of player  $i$  intended for expropriating  $s_{ij}$  from player  $j$ , which, depending on the effort allocated by other players and the resources of player  $j$ , may or may not result in player  $i$  expropriating  $s_{ij}$  units of resources. The attempt of player  $i$  to expropriate  $j$  may succeed or only partially succeed.

**Remark 2.** *Rationing Excess Expropriation:* If we pursued the route of rationing rules to ensure consistency of the expropriation “requests” or “efforts,” the payoff of a non-cooperating and a cooperating player would be respectively defined as

$$\omega_i + \sum_{j \in \mathcal{B}_i} \alpha_{ij}(\mathbf{s}) - \sum_{q \in \mathcal{A}_i} \alpha_{qi}(\mathbf{s}) - \mathcal{C}_i(s_i, \alpha_{-i}(\mathbf{s})) \quad (7)$$

and

$$\tilde{\omega}_i - \sum_{q \in \mathcal{A}_i} \alpha_{qi}(\mathbf{s}) \quad (8)$$

where each  $\alpha_{ij}(\mathbf{s}) \leq s_{ij}$  satisfies suitable (but inevitably somewhat arbitrary) constraints to ensure aggregate consistency of the expropriation outcomes.

---

<sup>27</sup> In particular, the values we used in the numerical examples involving two and three agents in Sections 3.1 and 3.2 can be recovered using the cost function in (6) and setting  $A = B = 1/320$ ,  $\bar{h} = 160$  and  $\gamma = 2$ .

In what follows, we choose to pursue a straightforward alternative to arbitrary rationing rules which also simplifies the analysis and in our view makes the results considerably more transparent. We will assume that, given the cost of stealing, the players' endowments are sufficiently large so as to guarantee that the expropriation choices are mutually compatible in the sense that the constraint that no player is left with negative resources is simply not binding. The expropriation choices are "interior" since endowments are large enough. Introducing rationing rules that satisfy a mild set of "monotonicity" restrictions would leave the qualitative nature of our results unaffected.

We relegate the formalities of the interiority of solutions when endowments are large enough to Section A.2 in the Appendix. The formalities require a surprising amount of detail, and they would just constitute a distraction if presented here. The following remark suffices to clarify.

**Remark 3.** *Large Endowments and Interiority:* Fix two functions  $\mathcal{C}^1(\cdot)$  and  $\mathcal{C}^2(\cdot)$  as above. We can then find a lower bound  $\underline{\omega}$  on endowments such that if  $\omega_i \geq \underline{\omega}$  for every  $i$  the following is true.<sup>28</sup>

Consider any set of  $t = 0$  cooperating choices. Let  $\mathcal{D}$  be the set of cooperating players, and consider any  $i \notin \mathcal{D}$ . Consider any set of "undominated" expropriation choices  $s_{-i}$  for all players except for  $i$ .<sup>29</sup> Then the optimal expropriation choices  $s_i$  for  $i$  are consistent with the arbitrary choices  $s_{-i}$  in the sense that they do not leave any player with a negative amount of resources.<sup>30</sup>

For the remainder of the paper, we assume that endowments are sufficiently large in the sense we have just specified.

### 5.3. The Problem Of Player $i$

It is now convenient to sum up again the problem that any non-cooperating player  $i$  faces in the second stage of the game. Using (5), this is given by<sup>31</sup>

---

<sup>28</sup>What follows is an informal restatement of Proposition A.1.

<sup>29</sup>See Definition A.3.

<sup>30</sup>See Subsection 5.3 below.

<sup>31</sup>We are ignoring the constraints that stipulate that expropriation choices should not leave any player with negative resources. We do this on the basis of our assumption of sufficiently large endowments. See Remark 3 and the discussion in Subsection 5.2 as well as Section A.2 in the Appendix.

$$\max_{s_i = \{s_{ij}\}_{j \in \mathcal{B}_i}} \sum_{j \in \mathcal{B}_i} s_{ij} - \mathcal{C}^1 \left[ \sum_{j \in \mathcal{B}_i} s_{ij} \right] + \sum_{j \in \mathcal{B}_i} s_{ij} \mathcal{C}^2(h_{ij}) \quad (9)$$

where  $h_{ij}$  is as in (4) above.

In general Problem (9) admits multiple solutions. However, it is easy to see that the solution is unique if there are no “ties” among the values of the  $h_{ij}$ . Proposition 1 below fully characterizes the solutions. Define

$$\underline{h}_i = \min_{q \in \mathcal{B}_i} \{h_{iq}\} \quad \text{and} \quad \mathcal{M}_i = \{q \in \mathcal{B}_i \mid h_{iq} = \underline{h}_i\} \quad (10)$$

so that  $\underline{h}_i$  is the minimal value of  $h_{iq}$  among players  $q \in \mathcal{B}_i$ , and  $\mathcal{M}_i$  is the set of players in  $\mathcal{B}_i$  who have such minimal value. Hence  $\mathcal{M}_i$  is the set of players below  $i$  who have the highest balance of net stolen wealth, that  $i$  takes as given. Given any  $h$  define also

$$z(h) = \operatorname{argmax}_x \{x - \mathcal{C}^1(x) - x \mathcal{C}^2(h)\} \quad (11)$$

The behavior of a non-cooperating player  $i$  in period  $t = 1$  is then characterized by his best response to  $s_{-i}$  which is the object of Proposition 1 that follows.

**Proposition 1.** *Best Response:* Let  $s_i^* = \{s_{ij}^*\}_{j \in \mathcal{B}_i}$  be a solution to Problem (9). Then the quantities  $s_{ij}^*$  are fully characterized by the following conditions.

- (i) The total quantity stolen  $\sum_{j \in \mathcal{B}_i} s_{ij}^*$  by player  $i$  is equal to  $z(\underline{h}_i)$ .
- (ii) If  $j \notin \mathcal{M}_i$  then  $s_{ij}^* = 0$
- (iii)  $\sum_{j \in \mathcal{M}_i} s_{ij}^* = z(\underline{h}_i)$

Hence, whenever  $\|\mathcal{M}_i\| = 1$  we also know that the solution to Problem 9 is unique.

**Proof:** See Section A.3 in the Appendix.

The content of Proposition 1 is easy to summarize intuitively. A non cooperating player  $i$  will best respond to  $s_{-i}$  in period  $t = 1$  as follows. He will only steal from those players who have the maximum net stolen wealth balance among the players in  $\mathcal{B}_i$ . If there are two or more players that have the maximum balance  $i$  will be indifferent as to whom he steals from. Only in this case  $i$  will have multiple best responses. The total amount that  $i$  steals

from players in  $\mathcal{B}_i$  is determined once we know the actual maximum net stolen wealth balance among players in  $\mathcal{B}_i$  and is increasing in this amount.

#### 5.4. Gains From Stealing and Cooperation

Consider the extreme example of a player  $i$  who does not cooperate when all others do. By doing so, he will forgo the gain from cooperating, namely  $\tilde{\omega}_i - \omega_i$ , and he will gain the amount  $z(0)$  specified in (11). If the gain from expropriation does not exceed the direct loss from lack of cooperation, we are clearly in a case that is not particularly interesting. Even under the most conducive circumstances, player  $i$  will not choose to expropriate players below him. Our next assumption rules out that this is the case for any  $i$ .

**Assumption 2.** Gains from Expropriation and Cooperation Condition: For every player  $i$  we have that

$$z(0) - \mathcal{C}^1(z(0)) - z(0)\mathcal{C}^2(0) > \tilde{\omega}_i - \omega_i \quad (12)$$

We assume that this is the case for the remainder of the paper. Notice that if Assumption 2 were violated we would be in the case that Mayshar, Moav, and Pascali (2022) refer to as the “curse of plenty.” The low hanging fruit from cooperation is so plentiful that (costly) expropriation simply does not take place.

For completeness, we state the following without proof.

**Remark 4.** No Stealing Equilibrium: Suppose that (12) is violated for every player  $i$ . Then there is an equilibrium in which all players choose to cooperate at  $t = 0$  and hence no stealing takes place at  $t = 1$ .

Since (12) is violated for every  $i$ , no player will find it profitable to deviate at  $t = 0$  from a situation in which all players are cooperating. The amount gained from stealing, net of the cost of stealing, does not exceed the endowment loss following the decision to not cooperate.

## 6. General Power Structures

As we have mentioned before, our main result consists in identifying an equilibrium of the model that can be interpreted as one in which there is “Centrally Enforced Cooperation” among the players.

This result holds, in our view surprisingly, for any power structure that satisfies Assumption 1, provided that Assumption 2 also holds and that  $N$  is large enough.

A preliminary concern regards the possible use of mixed strategies. It would seem odd that to “model” the endogenous emergence of enforcement, randomized actions were essential. In society the actions of enforcers are subject to noise, that seems certain. The noise, however, naturally stems from informational imperfections and actual mistakes of various kinds. Our model is stylized in the extreme so that mistakes and informational imperfections are ruled out. Given these premises, to home in on an enforcement arrangement that is intrinsically random would seem counterintuitive.

Our first result asserts that pure strategy equilibria exist and hence recourse to mixed strategies is unnecessary.

**Proposition 2.** *Existence of Pure Strategy Equilibria: Consider a General Power Structure as in Assumption 1 and let Assumption 2 hold. Then the game has a subgame perfect equilibrium in pure strategies.*

**Proof:** See Sections A.4 and A.5 in the Appendix.

In view of our remarks above, and of Proposition 2, throughout the rest of the paper we focus exclusively on pure strategy equilibria.

We are now ready to state formally our main result. If  $N$  is large enough our model has an equilibrium with Centrally Enforced Cooperation.

**Proposition 3.** *Equilibrium with Centrally Enforced Cooperation: Consider a General Power Structure as in Assumption 1 and let Assumption 2 hold.*

*Then there exists an  $\underline{N}$  such that whenever  $N \geq \underline{N}$  the game has an equilibrium in which player 1 does not cooperate while all other players cooperate. The lower bound  $\underline{N}$  does not depend on the power structure of the players.*

**Proof:** See Section A.6 in the Appendix.

One of the appeals of Proposition 3 is that, remarkably in our view, the result does *not depend* on the power structure at hand. Arborescences come in many “varieties.” For instance  $\succ$  could be a complete (or linear) order, so that the arborescence does not “branch out” at all. On the other hand the arborescence could at various points “branch out” (or “fan out”) with one player having a relatively large number of immediate successors. This in turn could

happen at many points in the arborescence, or just a few. Branching could be symmetric or highly asymmetric.

Proposition 3 of course does not say much beyond the existence of the particular equilibrium it singles out. As it turns out the other equilibria that may be available and their features *do depend* on the features of the arborescence at hand, and on the class of equilibria that one chooses to focus on. This is what we investigate in Section 7 below.

## 7. The Structure of Power Structures

### 7.1. Cascading Equilibria

The set of equilibria for general power structures is rich and is not limited to the equilibrium with centrally enforced cooperation singled out in Proposition 3. Players are allowed to steal from players below them even if they are not their direct successors, but perhaps only distant ones via many intermediate agents. This generates the possibility of complex equilibria in which some layers in the arborescence are perhaps skipped in the expropriation process while others are not. While this is interesting, in our quest to further characterize the set of equilibria we choose to focus on a class in which this does *not* take place.

There is a natural class of equilibria that encompasses the one singled out on Proposition 3 and allows further characterization of the set of equilibria. Besides making our characterization task more tractable, we believe this to be a natural class to investigate. In essence we will restrict attention to equilibria in which those players who do steal in fact do so only from the players below “nearest” to them in the arborescence who also steal, if there are any.

We term these as *cascading equilibria*. Notice that besides “proximity” in and by itself, the stealing set on which these equilibria rely is a natural one. From the point of view of a given player  $i$ , stealing players who are at the top of a “chain of stealing” in  $\mathcal{B}_i$  will in general be the ones that have accumulated a larger net stolen wealth balance, and hence are more advantageous to steal from given the class of cost of stealing functions we have postulated. This is the reason cascading equilibria are not only analytically convenient but are in fact pervasive. Indeed the proofs of Proposition 2 and 3 are constructive and they both rely on cascading equilibria.

Cascading equilibria are best defined by looking at a specific  $t = 1$  subgame, after cooperation choices have been made at  $t = 0$  and then observed by all. As we noted above, a  $t = 1$  subgame is identified by a subset  $\mathcal{Z} \subset \mathcal{P}$  of non-cooperating players. Without loss

of generality  $\mathcal{Z}$  is a strict subset of  $\mathcal{P}$  since all powerless players are considered cooperating players.

Given a  $\mathcal{Z}$ , the original partial order  $\succ$  defines one or more *sub-arborescences* on  $\mathcal{Z}$ , which we denote  $(\mathcal{Z}, \succ)$ .<sup>32</sup> The sub-arborescences  $(\mathcal{Z}, \succ)$  each have a root,<sup>33</sup> moreover a player is a successor of another player in the relevant sub-arborescence if and only if it is a successor of the given player in the original partial order  $\succ$ . Armed with these observations we can proceed with a formal definition.

**Definition 1.** *Cascading Equilibria:* Fix a  $t = 1$  subgame, identified by the set  $\mathcal{Z}$  of non-cooperating players. An equilibrium in subgame  $\mathcal{Z}$  is a cascading equilibrium if and only if the following is true.

Any player in  $\mathcal{Z}$  who has one or more immediate successors in  $\mathcal{Z}$  only steals from his immediate successors in  $\mathcal{Z}$ .

An equilibrium is a cascading equilibrium if and only if it is a cascading equilibrium in every subgame, on or off the equilibrium path.

As mentioned above, cascading equilibria do not take us outside the scope of Propositions 2 and 3:

**Remark 5.** *Cascading Equilibria With Enforcement:* The equilibria used to prove Proposition 2 constructively are Cascading Equilibria.

The equilibrium with Centrally Enforced Cooperation of Proposition 3 is a Cascading Equilibrium.

Focusing on cascading equilibria allows a further characterization of the set of equilibria in our model via a key observation that is embodied in Proposition 4 below.

It is not hard to outline intuitively what the observation is. Suppose that in some cascading equilibrium we found three non cooperating players — say  $i_1$ ,  $i_2$  and  $i_3$  — “ranked in a row” in the arborescence, with no other non cooperating player in between. Since we are in a cascading equilibrium players  $i_1$  would be stealing only from  $i_2$  while  $i_2$  would be stealing only from  $i_3$ , and  $i_3$  would be stealing from some player in  $\mathcal{B}_{i_3}$  that could be cooperating if there are no non cooperating players in  $\mathcal{B}_{i_3}$ .

---

<sup>32</sup>See Definition A.1. Notice also that this is a slight abuse of notation since the partial order restricted to  $\mathcal{Z}$  is a different object than  $\succ$ . This will not cause any ambiguity here.

<sup>33</sup>If  $1 \in \mathcal{Z}$  the sub-arborescence is unique. If  $1 \notin \mathcal{Z}$  there are multiple *disjoint* sub-arborescences.

Using Proposition 1 and (11) it is then not hard to see that the amounts stolen by these three players would be as follows. Player  $i_3$  would steal some amount  $z_{i_3} \geq z(0)$  from one or more players in  $\mathcal{B}_{i_3}$ . Player  $i_2$  would steal some amount  $z_{i_2} > z_{i_3}$  from player  $i_3$ , while player  $i_1$  would steal some amount  $z_{i_1} > z_{i_2}$  from player  $i_2$ . However, player  $i_2$  would then have an incentive to deviate and cooperate at  $t = 0$ . Indeed if he did so, in the ensuing subgame the player in  $\mathcal{B}_{i_1}$  with the largest stolen wealth balance would no longer be  $i_2$ , but would instead be  $i_3$ . Hence after the deviation, player  $i_1$  would “skip” player  $i_2$  and only steal from player  $i_3$ .<sup>34</sup> Therefore after the deviation player  $i_2$  would enjoy a payoff of  $\tilde{\omega}_{i_2}$ . This is obviously greater than his payoff before the deviation.<sup>35</sup> Hence we conclude that the configuration we hypothesized is *not* possible in a cascading equilibrium.

To make this key observation formal, we first proceed with a definition.

**Definition 2.** *Attack Triple:* Consider a General Power Structure as in Assumption 1 and let Assumption 2 hold.

Consider a subset  $\mathcal{Z} \subset \mathcal{P}$  of players that contains no powerless players in  $(\mathcal{P}, \succ)$ .

We say that  $\mathcal{Z}$  contains an “attack triple” if we can find a set of three players  $i_1 \succ i_2 \succ i_3$  in  $\mathcal{Z}$  such that for any  $j \in \mathcal{Z}$ ,  $j \neq i_1, i_2, i_3$ , we have that

$$i_1 \succ j \quad \Rightarrow \quad i_3 \succ j \tag{13}$$

In other words, given an attack triple  $i_1, i_2, i_3$ , within  $\mathcal{Z}/\{i_2, i_3\}$ , all successors of  $i_1$  are also successors of  $i_3$  and all successors of  $i_2$  are also successors of  $i_3$ .

Once we have Definition 2, we can use it to formalize the observation we have described.

**Proposition 4.** *No Attack Triples in a Cascading Equilibrium:* Consider a General Power Structure as in Assumption 1 and let Assumption 2 hold. Consider any Cascading Equilibrium and let  $\mathcal{Z}$  be the set of players who choose not to cooperate. Then  $\mathcal{Z}$  does not contain an attack triple.

**Proof:** See Section A.7 in the Appendix.

Proposition 4 has more or less “bite” depending on the power structure at hand. This is what we investigate in the next two Subsections of the paper.

<sup>34</sup>Using Proposition 1 and (11) it is fairly easy to see that  $i_1$  would in fact steal  $z_{i_2}$  from player  $i_3$ .

<sup>35</sup>Even without subtracting his stealing costs, before the deviation the payoff to  $i_2$  is  $\omega_{i_2} + z_{i_2} - z_{i_1} < \omega_{i_2} < \tilde{\omega}_{i_2}$

### 7.2. Complete Orders

There is a case in which the bite of Proposition 4 is completely straightforward. When  $\succ$  is a strict *complete order* so that any two players *must* be ranked we are in the “linear order” case. Without loss of generality (up to a relabeling of players)  $i \succ j$  if and only if  $j > i$ . Player 1 is the root and the other players are ranked in the order of their “names.” For any  $i$ , the set  $\mathcal{B}_i$  is the set of players  $j$  such that  $j > i$  and the set  $\mathcal{A}_i$  is the set of players  $j$  such that  $j < i$ .

In a linear order it is completely obvious that *any* set of three players must have the property (13) invoked by Definition 2. Hence, the next Corollary is a simple consequence of Proposition 4.

**Corollary 1.** *Complete Orders and Cascading Equilibria:* Consider a General Power Structure as in Assumption 1 and let Assumption 2 hold. Assume that  $(\mathcal{P}, \succ)$  is such that  $\succ$  is a complete order.

Consider any Cascading Equilibrium and let  $\mathcal{Z}$  be the set of players who choose not to cooperate. Then  $\mathcal{Z}$  contains at most two players.

When the players are arranged in a single power “chain” the absence of attack triples has maximum impact on the set of cascading equilibria.

Our next task is to examine the “polar” case to that of a complete order.

### 7.3. Highly Branching Power Structures

A critical feature of arborescences with a root that defines how “far” they are from a complete order is how highly branching they are. The direct successors of a given player are not ranked among themselves. Loosely speaking, the more direct successors the more, in a sense, partial (incomplete) is the strict order. In addition to this observation, we recall that our model is interesting when a chain of three or more players is in place.<sup>36</sup> So we focus attention on structures that are highly branching, but in which the direct successors of the branching player have themselves successors of some type.

Bearing in mind these informal observations, we begin with a general power structure in which player 1 has numerous direct successors, which is arguably a “canonical” highly branching power structure to study.

---

<sup>36</sup>Any set consisting of one powerful and one powerless player is “uninteresting” because of Assumption 2.

**Proposition 5.** *Many Direct Successors And Non-Cooperative Equilibria:* Consider a general power structure  $(\mathcal{P}, \succ)$  in which player 1 has  $M$  direct non-powerless successors and let Assumptions 1 and 2 hold.<sup>37</sup>

Then there exists a  $\underline{M}$  such that whenever  $M \geq \underline{M}$  the game has a Cascading Equilibrium equilibrium in which

- (i) Player 1 does not cooperate and neither does each of his direct successors.
- (ii) All other players choose to cooperate.

**Proof:** See Section A.8 in the Appendix.

Before delving into the intuition behind Proposition 5, we pause to make two remarks.

First of all the equilibrium identified in Proposition 5 is a Cascading Equilibrium according to Definition 1. This is comforting in that it reassures us that the comparison between the degree of cooperation in the equilibria identified in Corollary 1 and Proposition 5 is not due to our restricting attention to Cascading Equilibria, but instead to the structure of the power structure itself, namely how highly branching it is.

Second, the power structure posited in Proposition 5 belongs to the class to which Proposition 1 applies. It follows that if  $M$  is sufficiently large then Proposition 1 will guarantee that *another* equilibrium with Centrally Enforced Cooperation will exist as well. Multiple equilibria are guaranteed to obtain in this case.

We now come to the intuition behind the non-cooperative equilibrium singled out in Proposition 5, which is not hard to outline. Each direct successor to 1, say  $i$ , expropriates an amount  $z(0)$  from a powerless player in  $\mathcal{B}_i$ . If there are any non powerless players in  $\mathcal{B}_i$  they do not deviate from cooperating at  $t = 0$  since otherwise they would be expropriated by a larger amount by player  $i$ .

Player 1 expropriates an amount  $z_1/M$  from each of his direct successors. The total amount he steals is equally distributed among  $M$  players, and using Proposition 1 we know that the total stolen (namely  $z_1$ ) does not change with  $M$ . Hence the total amount stolen from each direct successor becomes arbitrarily small as  $M$  grows large. Therefore in this equilibrium each of the direct successors  $i$  of 1 receives a payoff that is arbitrarily close to what he would get if he alone stole from one of his powerless successors.<sup>38</sup> By deviating and

---

<sup>37</sup>Notice that each of these  $M$  players must have at least one powerless successor. This implies that  $N \geq 2M + 1$ .

<sup>38</sup>Hence a payoff arbitrarily close to  $\omega_i$  plus the left-hand side of (12).

cooperating at  $t = 0$ , player  $i$  would instead escape expropriation by player 1 and gain  $\tilde{\omega}_i - \omega_i$ . However, by Assumption 2, as  $M$  becomes large, this gain will be less than the loss from not expropriating his powerless successors. Hence the deviation to cooperation at  $t = 0$  is not profitable for  $i$ .

#### 7.4. *A Hub and Spoke Structure*

A structure that naturally fits Proposition 5 has been so prominent in previous studies that we are prompted to specifically point it out.

If each of the many direct successors of player 1 has one powerless successor, the power structure can be interpreted as a “trading post” consisting of a “hub” and many “spokes.” A point where many trading pairs meet.

Putting together Propositions 3 and 5 we know that, provided the number of participants is large enough, the equilibrium outcome can be of (at least) one of two varieties. One in which the central player, in equilibrium, “disciplines” the trading activity of the numerous pairs by exacting a (small) tax on all participants. All participants engage in surplus-creating trade in this case. The second equilibrium outcome is in a strong sense the outcome of a significant coordination failure. All non-powerless participants who can are locked into expropriation activities that prevent them from engaging in surplus creation.

Trading posts have figured prominently in the Economics literature. A highly incomplete list begins at least with the classics Wicksell (1898) and Walras (1874), and then includes the more modern Starr (2008), Shapley and Shubik (1977) and Hahn (1973).<sup>39</sup>

The main purpose of trading posts in the Economics tradition is roughly speaking to overcome the lack of “double-coincidence of wants” in the sense of Jevons (1875). Of course, this cannot be explicitly modeled here since wealth in our model is in effect a single homogenous good. The insight that we gain from the model we study here is that in order to function with appropriately enforced transactions, trading posts must reach a critical size in terms of the number of participants.

To fit the trading post structure into the model we have put forward, we posit that the hub is player 1 and there are a total of  $2K + 1$  players (including 1). The power structure is

---

<sup>39</sup>Scholars across a variety of fields have had an interest in trading posts. We simply hope to scratch the surface in a meaningful way mentioning Anthropologists (Brøgger, 2009, Kelley, 1985, McC. Adams, 1974), Archaeologists (Malinowski, 1974), Ethnographers (Turck and Turck, 1992) and Historians (Hämäläinen, 1998).

such that  $1 \succ i$  for all  $i \geq 2$ , and for each  $k = 1, \dots, K$ ,  $2k \succ 2k + 1$  while each player  $2k + 1$  is powerless.

While player 1 is the “hub,” each of  $(2k, 2k + 1)$  with  $k = 1, \dots, K$  is a trading pair (with one dominant partner) that is connected to 1 and to no other players directly.

The equilibrium in which all players but 1 engage in surplus-producing cooperation is a simple corollary of Proposition 3.

**Corollary 2.** *Hub Player As Enforcer: Consider  $N = 2K + 1$  players arranged on a hub and spoke structure as above. Let Assumptions 1 and 2 hold.*

*Then there exists an  $\underline{K}$  such that whenever  $K \geq \underline{K}$  the game has an equilibrium in which player 1 does not cooperate while all other non-powerless players cooperate.*

The equilibrium in which all non-powerless players are trapped in a coordination failure that prevents all surplus-producing activities is a simple corollary of Proposition 5.

**Corollary 3.** *Hub and Spoke Non-Cooperative Equilibrium: Consider  $N = 2K + 1$  players arranged on a hub and spoke structure as above. Let Assumptions 1 and 2 hold.*

*Then there exists a  $\underline{K}$  such that whenever  $K \geq \underline{K}$  the game has an equilibrium in which Player 1 does not cooperate and each Player  $2k$  for every  $k$  also chooses not to cooperate.*

Whether a trading post should flourish into a hub of surplus-producing cooperation activities with a central enforcer, or flounder (and perhaps thus disappear) under the weight of a coordination failure is clearly beyond the scope of a theory that admits the possibility of multiple equilibria.

## 8. Summary of Results

We study a jungle in which players are arranged in a strict partial order. Players above can expropriate those below in the partial order at a cost.

At time  $t = 0$  players choose whether to cooperate and engage in surplus generating activities (e.g. trade) or acquire the means to expropriate players below, and the decision is irreversible. Those who have acquired the offensive means then take expropriation decisions at  $t = 1$ .

The expropriation cost should be neither too large nor too small. Not too small so that a player does not want to expropriate all players below from all their wealth. Not too large

relative to the gains from cooperation so that cooperation should not become the optimal choice even when all other players are cooperating — the “curse of plenty.”

We posit a cost of stealing that captures the fact that stealing from a player that has a larger net wealth balance from being expropriated and stealing from others has a lower marginal cost. Critically, this fits a set up in which the object of stealing is storable and divisible. It does not fit a world in which the object of stealing is perishable since in this case stealing from others what they have stolen from others makes little sense.

Under these restricted (but not knife edge) conditions, we find that any jungle, in which a player is more powerful than all others and which is populated by sufficiently many players, has an equilibrium that we interpret as one with Centrally Enforced Cooperation. This is so regardless of the details of the power structure in that the condition on the number of players does not depend on them. In the equilibrium with Centrally Enforced Cooperation the most powerful player expropriates a “small” amount from all other players who in turn all cooperate and engage in surplus-producing activities. The temptation to deviate and steal from players below for all players aside from the enforcer is offset by the fact that if they did so then the enforcer would steal a larger amount from them.

Within an appealing class of equilibria that we term Cascading Equilibria we are able to characterize further the set of possible equilibria. The key feature of the power structure that drives the possibility of cooperative and non-cooperative outcomes in equilibrium is how “highly branching” it is. In the case of a non-branching structure (a linear, complete order) the number of players who do not cooperate in equilibrium is at most two. In the polar case in which one player has many direct successors in the power structure the number of players who do not cooperate in equilibrium can grow without bound.

Thus, in some important cases, like that of a trading post populated by sufficiently many agents, the model allows multiple equilibria. Some with Centrally Enforced Cooperation by a hub player and some highly non-cooperative ones in which a coordination failure prevents surplus generating trade.

## Appendix

### A.1. Arborescences

**Definition A.1.** *Arborescences and Sub-Arborescences:* Consider an arborescence  $(\mathcal{P}, \succ)$  with root 1. Let  $i \neq 1$  be a player such that  $\mathcal{B}_i \neq \emptyset$ . Let  $\mathcal{P}_i = i \cup \mathcal{B}_i$ . Then  $(\mathcal{P}_i, \succ)$ , where  $\succ$  is restricted to  $\mathcal{P}_i$ , is a sub-arborescence with root  $i$ .

Notice that, given an arborescence  $(\mathcal{P}, \succ)$  and a proper subset  $\mathcal{Z}$  of  $\mathcal{P}$ , the set  $\mathcal{Z}$  together with the restriction of  $\succ$  to  $\mathcal{Z}$  is either a sub-arborescence with a root, or is a collection of disjoint sub-arborescences each with a root.

We use the term sub-arborescence to indicate those obtained from  $(\mathcal{P}, \succ)$  restricting attention to a  $\mathcal{Z} \subset \mathcal{P}$ . With a slight abuse of notation we indicate by  $(\mathcal{Z}, \succ)$  the collection of sub-arborescences obtained by restricting attention to  $\mathcal{Z}$ .

If  $i \succ j$ ,  $i, j \in \mathcal{P}$  and, for no  $h \in \mathcal{P}$ ,  $i \succ h \succ j$ , we say that  $i$  is an immediate predecessor of  $j$  and  $j$  is an immediate successor of  $i$ .

**Definition A.2.** *Paths and Lengths:* A path in  $(\mathcal{P}, \succ)$  is a sequence  $(i_1, \dots, i_L)$  where  $i_{l-1}$  is an immediate predecessor of  $i_l$ ,  $l = 2, \dots, L$ . The length of the path is  $L$ .

When a path is entirely contained in a sub-arborescence we say that it is a path in that sub-arborescence.

### A.2. Sufficiently Large Endowments and Interiority

**Definition A.3.** *Undominated Choices:* Fix a  $\mathcal{C}(\cdot)$  and hence a  $\mathcal{C}^1(\cdot)$ . We say that  $s$  is an undominated expropriation choice iff

$$s \leq \bar{z} = \operatorname{argmax}_x \{x - \mathcal{C}^1(x)\} \tag{A.1}$$

The reason to define  $\bar{z}$  as in (A.1) and to term any expropriation choice not exceeding  $\bar{z}$  an undominated choice is that from (11) and the fact that by assumption  $\mathcal{C}^2(\cdot)$  is non-negative we must have that

$$\bar{z} \geq z(h) \quad \forall h \tag{A.2}$$

Hence no player will optimally choose to expropriate an amount exceeding  $\bar{z}$ , regardless of the choice of all other players.

Next, we define what is meant by endowments that are “sufficiently large.” It should be noted that the lower bound we state is by no means tight for our purposes. We proceed in this way for the sake of transparency and simplicity.

**Definition A.4.** *Large Endowments:* We say that the players’s endowments are sufficiently large if

$$\omega_i \geq (N - 1)\bar{z} \quad \forall i \geq 2 \tag{A.3}$$

**Proposition A.1.** *Interiority:* Assume that endowments are sufficiently large according to Definition A.4. Consider a non powerless player  $i$  and any expropriation choices  $s_{-i}$  by all players other than  $i$  that are undominated according to Definition A.3. Next, use (4) to compute  $h_{ij}$  for every  $j \in \mathcal{B}_i$ .

Then the problem of player  $i$  as stated in (9) with the added constraints

$$\omega_j - h_{ij} - s_{ij} \geq 0 \quad \forall j \in \mathcal{B}_i \tag{A.4}$$

has the same solution as the original one since none of the constraints in (A.4) are in fact binding.

**Proof:** Since the choices in  $s_{-i}$  are undominated we know that  $h_{ij} \leq (N - 2)\bar{z}$ . Since endowments are sufficiently large we know that  $\omega_j \geq (N - 1)\bar{z}$ . Hence (A.4) must be satisfied for any  $s_{ij}$  that is itself undominated and hence does not exceed  $\bar{z}$ .  $\square$

### A.3. The Proof of Proposition 1

**Proof of Proposition 1:** Consider Problem (9). This clearly may have multiple solutions. Let a typical solution be denoted by  $s_i^* = \{s_{ij}^*\}_{j \in \mathcal{B}_i}$ . By inspection, it is clear that

$$j \notin \mathcal{M}_i \Rightarrow s_{ij}^* = 0 \tag{A.5}$$

By strict concavity  $\sum_{j \in \mathcal{B}_i} s_{ij}^*$  is uniquely determined. Using (11) we obviously have that

$$z(\underline{h}_i) = \sum_{j \in \mathcal{M}_i} s_{ij}^* = \sum_{j \in \mathcal{B}_i} s_{ij}^* \tag{A.6}$$

and this suffices to prove the claim.  $\square$

### A.4. Equilibria In Subgames

The following two definitions allow us to recursively define expropriations along the longest paths for any arbitrary subset of players  $\mathcal{Z}$ . We begin with an arbitrarily chosen sub-arborescence.

**Definition A.5.** *Longest Path Expropriations in  $\mathcal{Z}$ :* Let  $\mathcal{Z}$  be a subset of players not containing any powerless players in  $(\mathcal{P}, \succ)$  and consider any one of the disjoint sub-arborescences with a root in  $(\mathcal{Z}, \succ)$ . Let this arbitrarily chosen sub-arborescence be denoted by  $(\mathcal{K}, \succ)$  with  $\mathcal{K} \subseteq \mathcal{Z}$ , and let  $k$  be its root. Let  $L$  be the length of the longest path  $Y^k$  in  $(\mathcal{K}, \succ)$  (which obviously originates with  $k$ ), and choose one such longest path arbitrarily if necessary. Let  $y^k(n)$  be the  $n^{\text{th}}$  player on this path. Obviously,  $y^k(L)$  is the last player on such path and has no successors in  $(\mathcal{K}, \succ)$ , but since powerless players are excluded from  $\mathcal{Z}$ , it must have one or more successors in  $(\mathcal{P}, \succ)$ .

Given such path  $Y^k$  we define recursively expropriation decisions  $s_{ij}^*$  for players along on the path. Let player  $y^k(L + 1)$  be an arbitrary successor of player  $y^k(L)$ . Next, define

$$s_{y^k(L), y^k(L+1)}^* = z(0) \tag{A.7}$$

and then recursively backwards for  $n = 1, \dots, L - 1$

$$s_{y^k(n), y^k(n+1)}^* = z(s_{y^k(n+1), y^k(n+2)}^*) \tag{A.8}$$

The procedure we have just defined can be repeated recursively to cover the entire set  $\mathcal{Z}$ .

**Definition A.6.** Recursive Expropriation Attribution for  $\mathcal{Z}$ : The procedure explicated in Definition A.5 is recursively repeated as follows.

Begin by replacing  $\mathcal{Z}$  with  $\mathcal{Z}/Y^k$  and then selecting a new arbitrary sub-arborescence with a root in  $\mathcal{Z}/Y^k$  and a longest path within it. For this path we define recursively expropriation decisions  $s_{ij}^*$  for players on the path as above. Repeating this step, we will eventually exhaust all elements of  $\mathcal{Z}$ . We call the array so constructed, in which  $s_{ij}^* = 0$  for  $i \notin \mathcal{Z}$ , a recursive expropriation attribution for  $\mathcal{Z}$ . We denote this by  $s^*(\mathcal{Z})$  and the associated longest paths used in its construction by  $\{Y^k\}$ , where each player  $k$  denotes the root of each path  $Y^k$ .

In a recursive expropriation attribution a player expropriates resources from at most one of his successors.

**Remark A.1.** Increasing Recursive Stolen Amounts: Let  $\mathcal{Z}$  be a subset of players not containing any powerless players. Consider a recursive expropriation attribution for  $\mathcal{Z}$  as in Definition A.6. Then, using Assumption 2, simple algebra and (A.7) and (A.8) suffices to show that along any of the longest paths  $Y^k$  used in its construction we must have that for  $n = 1, \dots, L - 1$ .

$$s_{y^k(1), y^k(2)}^* > \dots > s_{y^k(n), y^k(n+1)}^* > \dots > s_{y^k(L), y^k(L+1)}^* > \tilde{\omega}_i - \omega_i \tag{A.9}$$

for every  $i$ . Furthermore when endowments are sufficiently large according to Definition A.4 (see Proposition A.1)

$$s_{y^k(n), y^k(n+1)}^* < s_{y^k(n+1), y^k(n+2)}^* + \omega_i$$

and

$$s_{y^k(L), y^k(L+1)}^* < \tilde{\omega}_i.$$

We now proceed to prove a preliminary result on the paths associated with recursive expropriation attributions.

**Lemma A.1:** Let  $\mathcal{Z}$  be a subset of players not containing any powerless players. Consider a recursive expropriation attribution  $s^*(\mathcal{Z})$  for  $\mathcal{Z}$  and the associated collection of longest paths used in its construction, with typical element  $Y^k$ . If  $i \in Y^k$  and  $j \in Y^{k'}$ , where  $k \neq k'$  and  $j$  is a successor of  $i$ , then every player in  $Y^{k'}$  is a successor of  $i$ . Furthermore, the length of  $Y^{k'}$  does not exceed the number of successors of player  $i$  in  $Y^k$ .

**Proof:** First notice that by the definition of sub-arborescence, either  $i \succ k'$  or  $k' \succ i$ . If the former, the claim is proved. If the latter, then either  $k \succ k'$  or  $k' \succ k$ . If  $k \succ k'$ ,  $Y^k$  was constructed before  $Y^{k'}$  and adding  $k'$  would have yielded a longer path. Conversely, if  $k' \succ k$  then  $Y^{k'}$  was constructed before  $Y^k$  and adding  $k$  would have yielded a longer path. Hence, we have shown that  $k'$  is a successor of  $i$ . Therefore,  $Y^k$  was constructed before  $Y^{k'}$ . The latter part of the claim follows noting that if the length of  $Y^{k'}$  exceeds the number of successors of player  $i$  in  $Y^k$ , then  $Y^k$  is not a longest path.  $\square$

We now state and prove a central lemma. Let  $\mathcal{Z}$  be a subset of players not containing any powerless players. The list of players in  $\mathcal{Z}$  identifies a subgame in which all players in  $\mathcal{Z}$  do not cooperate. Let this subgame be denoted by  $\tau^1(\mathcal{Z})$ .

**Lemma A.2:** *Let  $\mathcal{Z}$  be a subset of players not containing any powerless players. Any recursive expropriation attribution  $s^*(\mathcal{Z})$  for  $\mathcal{Z}$  is a Nash equilibrium of the subgame  $\tau^1(\mathcal{Z})$ .*

**Proof:** The proof of the claim consists in checking that the behavior in the statement of the Lemma constitutes a best response to the behavior of others for all players in  $\mathcal{Z}$ , with the best responses as characterized in Proposition 1.

Let an  $s^*(\mathcal{Z})$  as in the statement of the Lemma be given and consider the associated collection of longest paths used in its construction, with typical element  $Y^k$ . Select a player  $i \in \mathcal{Z}$  and  $Y^k = \{y^k(n)\}_{n=1}^L$  such that  $y^k(\hat{n}) = i$ . Then, by our construction and Lemma A.1 when  $\hat{n} = L$ ,  $j \notin \mathcal{Z}$  for any successor  $j$  of player  $i$ . Thus,  $s_{y^k(L), y^k(L+1)}^* = z(0)$  for one arbitrary successor  $y^k(L+1)$  is a best reply by Lemma 1 since  $\underline{h}_{y^k(L)} = 0$ . If  $\hat{n} < L$ , again by our construction and Lemma A.1 setting

$$s_{y^k(\hat{n}), y^k(\hat{n}+1)}^* = z(s_{y^k(\hat{n}+1), y^k(\hat{n}+2)}^*)$$

and  $s_{y^k(\hat{n}), j}^* = 0$  for any  $j \neq y^k(\hat{n}+1)$  is a best reply by Lemma 1 since  $\underline{h}_{y^k(\hat{n})} = -s_{y^k(\hat{n}+1), y^k(\hat{n}+2)}^*$ .  $\square$

**Remark A.2.** *Recursive Expropriation Attributions and Cascading Equilibria:* By inspection, it is immediate to see that the pure strategy equilibria identified in Lemma A.2 are in fact Cascading Equilibria as in Definition 1 in Subsection 7.1.

Finally, we notice that the equilibria identified in Lemma A.2 may not be the only equilibria in a given subgame.

**Remark A.3.** *Multiple Equilibria in Subgames:* In constructing the recursive expropriation arrays we arbitrarily assumed that a player expropriates only one of his successors even when multiple longest paths may be available. He may obtain the same payoff expropriating several of his successors. The total amount expropriated will not vary, but the amounts seized from each successor may. Proposition 1 suffices to pin down the aggregate amount, which can then be apportioned arbitrarily across the successors  $q$  of player  $i$  for whom  $h_{iq}$  is minimal as in (10).

## A.5. The Proof of Proposition 2

The proof of Proposition 2 is constructive. In particular, we will show that a subgame perfect equilibrium exists in which a Cascading Equilibrium, as in Definition 1 (Subsection 7.1) is played in every subgame.

**Proof of Proposition 2:** To begin with, let

$$\hat{C} = C^1[z(0)] + z(0) C^2(0) \tag{A.10}$$

Next, we distinguish two cases.

**Case 1:** Suppose that for some non-powerless  $i^* \neq 1$  we have that

$$\omega_{i^*} + z(0) - z[z(0)] - \hat{C} \geq \tilde{\omega}_{i^*} - z(0) \tag{A.11}$$

We then construct a SPE as follows. At time  $t = 1$ , players 1 and  $i^*$  choose not to cooperate while all other players choose to cooperate.

In the subgame following these  $t = 1$  choices, the following equilibrium is played. Player  $i^*$  expropriates the amount  $z(0)$  from a powerless successor, while player 1 expropriates the amount  $z[z(0)]$  from player  $i^*$ . Whether player  $i^*$  expropriates one or several powerless successors is inessential as long as the total amount that is seized is  $z(0)$ .

Let's now consider subgames that follow deviations. We will specify the corresponding expropriation arrays and verify that unilateral deviations do not increase payoffs.

(i) If player  $i^*$  deviates unilaterally and chooses to cooperate then player 1 seizes  $z(0)$  from player  $i^*$ . Therefore in this case player  $i^*$  obtains a payoff equal to the right-hand side of (A.11). Since the payoff if he does not deviate is equal to the left-hand of the same inequality, the deviation is not profitable.

(ii) Using Assumption 2, it is immediate to check that player 1 does not want to deviate unilaterally and cooperate.

(iii) If a successor  $i$  of player  $i^*$  who is not powerless chooses to not cooperate a cascading equilibrium, as in Definition 1, is played. The payoff of player  $i$  is equal to  $\omega_i + z(0) - z[z(0)] - \hat{C}$ . Since  $z(0) - z[z(0)] < 0$  and  $\omega_i < \tilde{\omega}_i$  such deviation is unprofitable since the payoff of player  $i$  when cooperating is  $\tilde{\omega}_i$ .

(iv) If a player  $i$  that is not a successor of player  $i^*$  chooses to activate his power an equilibrium is played in which player 1 expropriates  $z[z(0)]$  from player  $i$  and nothing from player  $i^*$ . Players  $i$  and  $i^*$  each expropriate  $z(0)$  from one of their successors, which are of course distinct. After the deviation, the payoff of player  $i$  is equal to  $\omega_i + z(0) - z[z(0)] - \hat{C}$ . As in (iii) this is smaller than his payoff before the deviation, namely  $\tilde{\omega}_i$ .

(v) In any other subgame a Cascading Equilibrium, as in Definition 1, of that subgame is played. Any such subgame entails deviations that are not unilateral and hence no further considerations about payoffs are needed.

Hence, a pure strategy subgame perfect equilibrium exists in this case. Notice that a Cascading Equilibrium is played in all subgames, on or off the equilibrium path.

**Case 2:** Suppose that (A.11) is false in the sense that for every  $i \neq 1$  we have that

$$\omega_i + z(0) - z[z(0)] - \hat{C} < \tilde{\omega}_i - z(0) \tag{A.12}$$

It is then obvious that in this case we have

$$\omega_i + z(0) - z[z(0)] - \hat{C} < \tilde{\omega}_i - \frac{z(0)}{N-1} \quad \forall i \neq 1 \quad (\text{A.13})$$

In this case, we construct an SPE as follows. At time  $t = 1$  player 1 decides not to cooperate and every other player chooses to cooperate. Player 1 steals an amount

$$\frac{z(0)}{N-1} \quad (\text{A.14})$$

from each of his  $N - 1$  successors. Notice that in the subgame that ensues, no player other than player 1 chooses a positive expropriation amount and that by Proposition 1, the choice of player 1 is optimal. Finally notice that player 1 does not wish to deviate unilaterally by Assumption 2.

In any other subgame following a unilateral deviation by a player  $i \neq 1$ , player 1 then expropriates  $z[z(0)]$  from player  $i$ , and player  $i$  expropriates  $z(0)$  from one of his successors. The payoff of player  $i$  is equal to the left-hand side of (A.13) and is smaller than the payoff from cooperating which is of course is equal to the right hand side of (A.13).

Since, Case 1 and Case 2 are exhaustive of all possibilities, the proof is now complete.  $\blacksquare$

### A.6. The Proof of Proposition 3

**Proof of Proposition 3:** The claim follows from case (2) in the proof of Proposition 2 noting that since  $\omega_i < \tilde{\omega}_i$  and  $z(0) - z[z(0)] < 0$  the inequality (A.13) holds when  $N$  is sufficiently large.  $\blacksquare$

### A.7. The Proof of Proposition 4

**Proof of Proposition 4:** The argument is by contradiction. Suppose that an equilibrium with an attack triple exists. Let  $\mathcal{Z} \subseteq \mathcal{P}$  be the set of players who choose not to cooperate on the equilibrium path. Let  $(i_1, i_2, i_3)$  be the attack triple within  $\mathcal{Z}$  that is assumed to exist. We will show that  $i_2$  has an incentive to deviate unilaterally and choose to cooperate.

We begin by noticing that (13) on any longest path within  $\mathcal{Z}$  starting with  $i_1$  has  $i_2$  next and then  $i_3$ , and the longest path within  $\mathcal{Z}$  starting with  $i_2$  has  $i_3$  next.

By Remark A.1 the equilibrium expropriation array  $s^*$  is such that

$$s_{i_1, i_2}^* > s_{i_2, i_3}^*$$

We can then conclude that the payoff to  $i_2$  in this equilibrium is smaller than  $\tilde{\omega}_{i_2}$ . If  $i_2$  were instead to unilaterally deviate and cooperate, in the subgame that follows this deviation any longest path within  $\mathcal{Z}/i_2$  starting with  $i_1$  has  $i_3$  next. Since after this deviation  $i_2$  earns  $\tilde{\omega}_{i_2}$ , the deviation is profitable and the proof is now complete.  $\blacksquare$

### A.8. The Proof of Proposition 5

**Proof of Proposition 5:** Without loss of generality suppose that the  $M$  direct successors of 1 are players  $2, \dots, M + 1$ .

We construct a SPE as follows. At time  $t = 1$ , players  $1, \dots, M + 1$  choose not to cooperate while all other players choose to cooperate.

In any subgame where  $\mathcal{Z} \neq \{1, \dots, M + 1\}$  a Cascading Equilibrium is played.

In the subgame where  $\mathcal{Z} = \{1, \dots, M + 1\}$ , each player in  $\mathcal{Z}$  different from player 1 expropriates  $z(0)$  from his powerless successors, and Player 1 steals an amount  $z[z(0)]/M$  from each of his immediate successors.

It is easy to check that by Assumption 2 player 1 does not want to deviate and receive  $\tilde{\omega}_1$  by cooperating. Also by Assumption 2, the same holds for sufficiently large  $M$  for players  $i = 2, \dots, M$  since their payoff is

$$\omega_i + z(0) - \frac{z[z(0)]}{M} - \mathcal{C}^1(z(0)) - z(0) \mathcal{C}^2(0)$$

and the amount  $z[z(0)]/M$  can be made arbitrarily small.

Consider now any non-powerless player  $i$  who is not an immediate successor to 1. Notice that there may or may not be any, and if there are none there is nothing more to prove. Let  $m_i$  be the player that is an immediate successor of 1 and also a predecessor of  $i$ . In the subgame with  $\mathcal{Z} = \{1, \dots, M + 1\}$  the payoff to  $i$  is simply  $\tilde{\omega}_i$  as  $m_i$  expropriates from powerless successors. If  $i$  deviates unilaterally, in the subgame with  $\mathcal{Z} = \{1, \dots, M + 1, i\}$ , in the cascading equilibrium that is played  $m_i$  will steal  $z[z(0)]$  from  $i$ , and  $i$  will steal  $z(0)$  from his successors. Hence, by deviating player  $i$  gets

$$\omega_i + z(0) - z[z(0)] - \hat{\mathcal{C}} \tag{A.15}$$

whereas if he does not deviate he gets  $\tilde{\omega}_i$ . Since  $\tilde{\omega}_i > \omega_i$  we conclude that this unilateral deviation cannot be profitable for  $i$ .  $\square$

### References

- ACEMOGLU, D., AND S. JOHNSON (2005): “Unbundling Institutions,” *Journal of Political Economy*, 113, 949–995.
- ANDERLINI, L., L. FELLI, AND A. RIBONI (2020): “Legal Efficiency and Consistency,” *European Economic Review*, 121, 103323.
- BANDIERA, O. (2003): “Land Reform, the Market for Protection, and the Origins of the Sicilian Mafia: Theory and Evidence,” *Journal of Law, Economics, & Organization*, 19, 218–244.
- BRØGGER, B. (2009): “Economic Anthropology, Trade and Innovation,” *Social Anthropology*, 17, 318 – 333.
- CALZOLARI, G., L. FELLI, J. KOENEN, G. SPAGNOLO, AND K. O. STAHL (2021): “Relational Contracts and Trust in a High-Tech Industry,” *Janeway Institute, Working Paper No. JIWP2101*.
- CHOMSKY, N. (1959): “A Review of Verbal Behavior,” *Language*, 35(1), 26–58.
- DIXIT, A. (2003a): “On Modes of Economic Governance,” *Econometrica*, 71, 449–481.
- (2003b): “Trade Expansion and Contract Enforcement,” *Journal of Political Economy*, 111(6), 1293–1317.
- (2009): “Governance Institutions and Economic Activity,” *American Economic Review*, 99(1), 5–24.
- (2011): “A Game-theoretic Perspective on Diego Gambetta’s Codes of the Underworld,” *Global Crime*, 12, 134–145.
- GAMBETTA, D. (1996): *The Sicilian Mafia: The Business of Private Protection*. Cambridge, MA: Harvard University Press.
- GIBBONS, R., AND J. ROBERTS (2013): “Economic Theories of Incentives in Organizations,” in *The Handbook of Organizational Economics*, ed. by R. Gibbons, and J. Roberts, pp. 56–99. Princeton, NJ: Princeton University Press.
- GLAESER, E. L., AND A. SHLEIFER (2002): “Legal Origins,” *Quarterly Journal of Economics*, 117, 1193–1229.

- HAHN, F. (1973): "On Transaction Costs, Inessential Sequence Economies and Money," *Review of Economic Studies*, 40(4), 449–461.
- HÄMÄLÄINEN, P. (1998): "The Western Comanche Trade Center: Rethinking the Plains Indian Trade System," *The Western Historical Quarterly*, 29(4), 485–513.
- HOBBS, T. (1968): *Leviathan, 1651*. Baltimore: Penguin Books.
- JEVONS, W. S. (1875): *Money and the Mechanisms of Exchange*. London: C. Kegan Paul.
- KELLEY, K. B. (1985): "Ethnoarchaeology of Navajo Trading Posts," *Kiva*, 51(1), 19–37.
- LOCKE, J. (1952): *The Second Treatise of Government, 1689*. Indianapolis: Bobbs-Merrill.
- MAILATH, G., AND L. SAMUELSON (2006): *Repeated Games and Reputations: Long Run Relationships*. Oxford: Oxford University Press.
- MALCOMSON, J. M. (2013): "Relational Incentive Contracts," in *The Handbook of Organizational Economics*, ed. by R. Gibbons, and J. Roberts, pp. 1014–1065. Princeton, NJ: Princeton University Press.
- MALINOWSKI, T. (1974): "An Amber Trading-Post in Early Iron Age Poland," *Archaeology*, 27(3), 195–200.
- MAYSHAR, J., O. MOAV, AND L. PASCALI (2022): "The Origin of the State: Land Productivity or Appropriability?," *Journal of Political Economy*, 130, 1091–1144.
- MCC. ADAMS, R. *et al.* (1974): "Anthropological Perspectives on Ancient Trade [and Comments and Replies]," *Current Anthropology*, 15(3), 239–258.
- NOZICK, R. (1974): *Anarchy, State, and Utopia*. New York, NY: Basic Books Inc.
- PICCIONE, M., AND A. RUBINSTEIN (2004): "The Curse of Wealth and Power," *Journal of Economic Theory*, 117, 119–123.
- (2007): "Equilibrium in the Jungle," *The Economic Journal*, 117(522), 883–896.
- PLATO (1943): *Plato's The Republic, ca. 375 BCE*. New York: Books Inc.
- ROUSSEAU, J.-J. (2004): *The Social Contract, 1762*. Harlow, England: Penguin Great Ideas.

- RUBINSTEIN, A., AND K. YILDIZ (2022): “Equilibrium in a Civilized Jungle,” *Theoretical Economics*, 17, 943–953.
- SCOTT, J. C. (2017): *Against the Grain: A Deep History of the Earliest States*. Yale University Press.
- SHAPLEY, L., AND M. SHUBIK (1977): “Trade Using One Commodity as a Means of Payment,” *Journal of Political Economy*, 85, 937–68.
- SKINNER, B. F. (1957): *Verbal Behavior*. Acton, MA: Copley Publishing Group.
- STARR, R. (2008): “Commodity money equilibrium in a convex trading post economy with transaction costs,” *Journal of Mathematical Economics*, 44, 1413–1427.
- TURCK, T. J., AND D. L. L. TURCK (1992): “Trading Posts along the Yukon River: Noochuloghoyet Trading Post in Historical Context,” *Arctic*, 45(1), 51–61.
- WALRAS, L. (1874): *Elements of Pure Economics*. Homewood, IL: Irwin. *Éléments d’Économie Politique Pure, ou Théorie de la Richesse Sociale*, W. Jaffe translation (1954).
- WICKSELL, K. (1898): *Interest and Prices: A Study of the Causes Regulating the Value of Money*. London: Macmillan. *Geldzins und Güterpreise*, R.F. Kahn translation (1936).