NORTHWESTERN UNIVERSITY

Essays in Economic History and Political Economy

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Economic History and Political Economy

By

Michael Porcellacchia

EVANSTON, ILLINOIS

June 2022

Abstract

Chapter 1 proposes and tests a model where a country aligns with a foreign power to obtain its support and reduce its geopolitical risks, which also depend on the country's exposure to the other foreign powers. We show that the country's alignment with a given foreign power is increasing in the country's *index of Indirect Exposure (IE-index)* towards it – the country's relative exposure to foreign powers it is less exposed to. Empirically, we compute every country's IE-index towards each foreign power using various measures of foreign exposure (including trade dependence) and establish its importance for economic, military, and diplomatic alignment. Additionally, we find that the theory can predict how the pattern of interstate alignment changed after the collapse of the Soviet Union and China's rise. Moving beyond alignment choices, we find that the IE-index is an important determinant of trade flows and foreign aid.

Chapter 2 studies societies where a central authority can discriminate among different social groups and the tax extraction is only constrained by the groups' violence potential. We show that, in order to maximize tax extraction, the central authority will optimally adopt a divide-and-conquer strategy that we call *political favoritism*. In particular, we show that in equilibrium the central authority optimally creates a ranking among the subjects granting better taxation treatment to some groups at the expense of others even when groups are ex-ante homogeneous. Moreover, we show that when groups are heterogeneous, the central authority optimally favors wealthier groups. These observations suggest that central authorities have an active role in creating and exacerbating social stratification. Evidence from historical societies largely corroborates our prediction.

Finally, Chapter 3 studies how internal conflicts are shaped by the political settings. Building on Chapter 2, we show that political favoritism creates competition over political rents, planting the seeds for a novel class of conflicts: *status-related conflicts*. Indeed, social groups

have incentives to destroy/appropriate each other's resources, as they anticipate that the ruler optimally assigns a higher rank (thus a higher political rent) to wealthier social groups. Additionally, we show that not even the ruler is immune from the logic of violence that its own policy generates. Indeed, we show that the ruler's adoption of political favoritism motivates social groups to support (or at least not to oppose) any claim that the wealthiest group in society might have on the rulership. This observation motivates another class of conflicts: whenever a social group is in a position to challenge the primacy of the ruling group, the two groups would start fighting over resources to come out on top of the other, thus obtaining or cementing the support to be the ruler. We support the theoretical analysis with a case study from Early Modern Europe and addressing the evidence on internal conflicts from both modern and historical settings.

Acknowledgements

I am much indebted to my committee: Sandeep Baliga, Georgy Egorov, and Joel Mokyr. Without them, my Ph.D. journey would not have been possible. I am grateful for Sandeep and Georgy's time and efforts. As teachers, they initiated me to the beauty and value of applying economic theory to political economy. As advisors, they were always clear about the realities of this type of research and how it should be done. I am very thankful for their energy, and I admire them greatly. I am deeply grateful for the dedication of my committee chair, Joel. I have been his student, his teaching assistant, and his advisee. I always learned something new with every interaction. His open-mindedness and incredible scholarly depth have always been a source of inspiration. Not only my thesis but also much of my future research will be indebted to the many conversations I have had with him. I am thankful for Joel's trust, attention, and care; he kept me focused and motivated and never allowed me to feel lost. Each of my advisors contributed to the development of this thesis in a critical way. More importantly, they contributed to shaping me as a researcher, for which I will always be in their debt.

I also acknowledge the support of many other people I met at Northwestern. Among them, I thank David Austen-Smith, Walker Hanlon, Leander Heldring, Ralph Meisenzhal, Ameet Morjaria, Harry Pei, Nancy Qian, Eric Schulz, Taco Terpstra, and Edoardo Teso for sharing their time and knowledge. I also thank Cassiano Alves, Brendon Andrews, Riccardo Bianchi Vimercati, Modibo Camara, Mario Cannella, Matteo Magnaricotte, Joris Mueller, Lorenzo Stanca, Kelly Strada, Yunus Topbas, Francesca Truffa, Adriana Troiano, and Benjamin Vatter. They offered me their friendship and all the support that comes with it. I thank my friend and coauthor Matteo Camboni, with whom I shared every up and every down of this incredible journey.

Finally, I am grateful to my family. My mother Paola and father Franco have done much

more than give me life. I am extraordinarily thankful for their wisdom, their presence, their help, their enthusiasm, their trust, and their love. I am also very grateful for all the support I received from my brother Davide. His example and advice have always made it easy to make the right decisions. I also thank my grandparents, Assunta and Dorino, that I dearly miss, and Lidia and Alberto. Talking with them or thinking about them was always a source of strength and happiness. I also acknowledge the support of the family I made along the way. The new Porcellacchias, Florentina and Milo, and the Meggersee family. My girlfriend's family, Giovanna, Fabrizio, Arianna, and the Bozzolo family, Anna, Bruno, and Trudy. They all gave me love, support, and attention. And I always feel at home with them. Finally, I thank my love Giulia. She kept me happy, balanced, and focused on the future. I am deeply thankful for her help, patience, kindness, and love.

Preface

A firm designing a compensation scheme for its workers, an entrepreneur raising funds from investors, a ruler dealing with its subjects or with the influence of the great powers. These are all examples where the actions of one individual or entity (*actor*) have effects on a plurality of other individuals or entities (*recipients*). In many situations of this kind, the success or failure of the actor's objectives depends on how the recipients coordinate. In such situations, the actor might wish to encourage or discourage the recipients' coordination on a given action. While a large literature in game theory has studied coordination problems, a less explored point of view is the *organization of coordination*: how the actor can manipulate the recipients' final payoffs to encourage or discourage coordination.

The problem of organizing coordination becomes particularly consequential when we consider settings where the actor wants to discourage or encourage the recipients from taking an action that affects the incentives of the other recipients. This would be the case, for instance, when the decision of one recipient increases the other recipients incentives to make the same decision, i.e., when there are strategic complementarities. In many cases of this type, organizing coordination effectively leads to a divide-and-conquer strategy with a very specific structure – a *ranking mechanism*. In a ranking mechanism, the actor organizes recipients in a ranking, where a higher position in the ranking corresponds to a higher final payoff. When a ranking mechanism wants to discourage the recipients' coordination on an action, the payoff of a given recipient must be high enough to induce it to abstain from the action even when all the lower-ranked recipients take it. On the other hand, when the ranking mechanism wants to induce coordination on an action, the payoff of a given recipient must be high enough to induce it to take such action, even when all the lower-ranked recipients are not.

This dissertation shows that studying the organization of coordination and the ranking

mechanisms that emerge can contribute to our understanding of the political economy of different societies and types of interactions. Specifically, this dissertation aims at studying how ranking mechanisms emerge in political economy problems, how they interact with the heterogeneity of the different political actors, how the peculiar structure of ranking mechanisms might indirectly shape behavior, and whether ranking mechanisms are empirically relevant in historical and modern settings.

The first chapter of this dissertation (joint work with Matteo Camboni) studies international relations from the perspective of a country dealing with foreign influence. Specifically, we consider (the government of) a country who attempts to increase its political stability by obtaining the support of the foreign powers. To obtain the support of a foreign power, we assume that such country can align its policies to the interest of that foreign power. However, such promises of alignment have value only if the foreign power expects such regime to survive, which depends on whether the other foreign powers will support it. Thus, foreign powers are more willing to support a regime when they expect it to be supported by the other foreign powers. In such setting, organizing coordination leads to a ranking mechanism, with the country assigning a rank to each foreign power and aligning more with higher-ranked foreign powers. But how does the country assign its ranks? What are the empirical implications of such mechanism?

For the first question, we show that, when foreign powers have a heterogeneous impact on the political stability of the country, then the country would optimally assign a higher rank to those foreign powers with a greater impact. Thus, we should expect countries to align more with those foreign powers which they are more exposed to or, conversely, that have more influence over them, a dimension of heterogeneity which we call *power*. For the second question, we show that a key empirical implication of the optimal ranking mechanism is that a country's extent of alignment with a given foreign power depends on its indirect exposure: its exposure to the other foreign powers. Specifically, we show that the country's

alignment with a foreign power is increasing in the country's relative exposure to the foreign powers it is more exposed to than the one considered – a novel directed bilateral index which we call Index of Indirect Exposure (*IE-Index*).

Finally, the Chapter explores the empirical relevance of ranking mechanisms with several tests. First, we perform two event studies to show that the theory can predict how the collapse of the Soviet Union or the rise of China had geographically heterogeneous effects on how countries adjusted their economic, military and diplomatic alignment with the US and other foreign powers. Second, we compute for each country its index of Indirect Exposure towards each foreign power using various measures of foreign influence, including trade dependence. We find that a country's IE-index towards a given foreign power is a relevant determinant of a country's alignment with that foreign power. Additionally, we find that alignment with a foreign power is not further affected by a country's exposure to it, if we condition on its Indirect Exposure index towards it. Moving beyond alignment choices, we find that the IE-index is an important determinant of the allocation of foreign aid and the pattern of trade. All in all, this empirical analysis strongly supports the theory and the role of the IE-index in shaping modern international relations.

While data limitations have restricted our focus to modern settings, one of the key contributions of our paper is that it proposes a theory that can rationalize the large set of historical case studies that suggest that the great powers have very often demonstrated that they are status-conscious, i.e., they care about their relative ranking. While the typical justification of status-consciousness is that it is an intrinsic characteristic of human preferences, this does not clarify how such preference would translate into world politics and how much it should be expected to matter. What type of status matters (relative to what type of resources)? Relative to whom? How is it linked to material interests? Our analysis answers this question, it proposes the idea that status-consciousness is universally observed because it is linked to material interests, and shows that such concept is not only useful to understand historical

cases, but also modern international relations.

The second Chapter of this dissertation (joint work with Matteo Camboni) turns its attention to domestic politics, considering how a local ruler who wishes to raise taxes deals with the possibility that its subjects resist taxation. We consider a setting where subjects can take an action (resistance) which implies that, with some probability, they manage to avoid paying the taxes. In this case, the source of the strategic complementarity is that resistance is more likely to succeed when more subjects are resisting. In such setting, it is not enough for the ruler to discourage the opposition of one of the social groups, as taxation must prevent the formation of resisting coalitions. In line with the previous discussion, we show that the optimal taxation policy is a ranking mechanism, which we call *political favoritism*: each subject is assigned to a rank and higher ranks pay lower taxes. But how are rankings assigned as a function of the distribution of resources in society? What are the consequences of this policy for inequality and social unrest? How do political institutions interact with political favoritism?

To answer the first question, we consider two potential sources of heterogeneity. First, we note that a social group that normally enjoys a higher standard of living (a wealthier group) might be less willing to resist in the first place, as it has more to lose. As a consequence, we should expect that wealthier social groups should be taxed more by the central authority. To capture this, we introduce a dimension of heterogeneity that we call *economic resources*, which measures how much a social group loses when they participate in resistance and resistance fails. We show that, when economic resources are the only form of heterogeneity, then the optimal taxation scheme is approximately proportional to wealth: the central authority demands a fraction of each subject's economic resources. However, we show that each group will face a different the tax rate, which depends on the subject's position in the ranking determined by the government, with lower-ranked groups paying a higher tax rate. Specifically, we show that, richer groups should receive a lower rank, and thus in general face

a higher tax rate.

In many settings, however, wealthier social groups should be expected to have more resources to mobilize against the central authority, when they choose to resist it. To capture this, we introduce a second dimension of heterogeneity, *military resources*, which measures the heterogeneous impact of a subject's decision to resist on the probability that the resistance succeeds. First, we show that, if groups have the same economic resources but different military resources, then stronger groups should receive a higher rank and thus pay a lower tax rate on their economic resources. However, in the general case where groups can be heterogeneous in both economic and military resources, we show that what matters for the optimal assignment of ranks is each subject's militarization ratio, i.e., the ratio of their military to economic resources. As we show, this result implies that, in the natural case where military resources are a convex and increasing function of economic resources, we should expect richer (thus stronger) groups to pay lower tax rates on their wealth. As a consequence of this, political favoritism should be expected to reinforce social stratification.

Finally, the divide-and-conquer nature of political favoritism suggests that this strategy is particularly vulnerable to political institutions that foster cross-group solidarity (the subjects' ability to act like one), for instance a representative institution. We first show that when the subjects' solidarity is sufficiently credible, the government optimally abandons its original divide-and-conquer strategy, treating the involved subjects as a single bargaining block backed by the threat of their collective action. As a result, when a subset of subjects are in solidarity they receive a better taxation treatment at the expense of the government and the excluded subjects. On the other hand, if the cross-group solidarity is not sufficiently credible, then the ruler's optimal strategy would be to treat every group as separate (employing political favoritism), and be ready to deal with resistance whenever subjects actually turn out to be in solidarity. Finally, we note that when a subject ends up resisting, then every lower-ranked subject would join. Thus, the model predicts that the ruler should often

deal with coalitional resistance, fueled by the lower-ranks of society. Additionally, it predicts that events that generate new forms of cross-group solidarity (e.g., religious reforms, participation in a common army) might generate widespread social unrest.

From an applied perspective, we discuss how our findings can justify the universal attention for relative status and rankings emerging from the historical evidence, and why government policy has typically been extremely regressive, playing a prominent role in reinforcing social inequality and even *inducing* social stratification. Moreover, as we document in our historical application, the divide-and-conquer logic at the heart of our model was well understood by rulers and politicians in Medieval and Early Modern Europe. For example, we show how Louis XI avoided to clash against a coalition of its magnates by resorting to the type of divide-and-conquer strategy that characterizes our equilibrium. Finally, we show that, as predicted, in both modern and historical societies the poorest and most discriminated groups (the lower ranks) are disproportionately more likely to participate in episodes of resistance against the government and that the rise of new identity groups has historically led to widespread political instability.

Finally, Chapter 3 (joint work with Matteo Camboni) builds on the analysis of Chapter 2 to study how political favoritism shapes the pattern of internal conflicts. Specifically, we take a general perspective where conflicts are simply a costly way to induce a change in the distribution of resources in society. We consider the case where there is only one form of heterogeneity among social groups (violence potential or power), and we consider two institutional settings which we call a United and a Hierarchic Society. In a United Society, we assume that the payoff of every subject is increasing in the combined violence potential of every other subject and decreasing in the ruler's. This reflects a setting where cross-groups solidarity is sufficiently credible and the ruler is effectively constrained by collective action. In a Hierarchic society, on the other hand, the payoff of each subject is determined by political favoritism: it is increasing in the combined violence potential of weaker subjects,

and decreasing in that of stronger subjects and the ruler.

We first study how conflicts might be motivated by the groups' ambition to become the next ruling group. To do so, we allow social groups to select the identity of the ruling group. First, note that in both a United and Hierarchic Society the power of the ruler has a negative impact on the social groups' payoffs. As a consequence, in both cases, every subject strictly prefers any ruler to be as weak as possible. In a United Society, this logic is further compounded by the fact that social groups benefit from the power of every other subject. As a consequence, not only the stronger is the ruler the worse off is each subject, but also assigning the rulership to a stronger group deprives each subject of a stronger ally. Thus, in a United Society, we show that social groups prefer to assign the rulership to the weakest possible social group. In contrast with this, in a Hierarchic Society it is the strongest group who is able to gain the largest consensus and become the ruler. Indeed, in a Hierarchic Society, each social groups is positively affected by the power of weaker groups, if the are not the ruling group. Thus, no group would support the bid for power of a weaker group. This leaves the two strongest groups as the only possible pretenders for the rulership. Specifically, if the ruler selection procedure is correlated with the level of resources of the contenders, then the strongest social group would always gain the rulership. As a consequence of this argument, in a United Society, conflicts are of limited use for a group who aims at the rulership. On the other hand, in Hierarchic Societies, the two strongest groups have a strong incentive to fight each other, as becoming the strongest group allows them to become the next ruler.

Finally to study how different institutional settings (Hierarchic vs United Society) shape the pattern of infighting, we turn our attention to the possibility that groups fight each other, even when the rulership is fixed. This can represent a setting where the ruler has not yet imposed a monopoly on the legitimate use of violence, so that groups can legitimately attempt to appropriate/destroy each other's resources. In a United Societies, the incentives

to fight another subject are straightforward. As the final payoff of each subject is increasing in the power of all the other subjects, then groups will internalize any destruction of resources that results from their attack. For this reason, a subject will only attack another subject when the value of the appropriated resources and the probability of victory are high enough to cover the costs of participating in the conflict and the indirect cost that results from the possible weakening of another subject.

The incentives to fight in Hierarchic Societies are more complex. Recall that, in such societies, the final payoff of each subject is increasing in the power of all weaker subjects and decreasing in the power of all stronger subjects. As a result, becoming stronger than another subject would induce a jump up in the subject's payoff; we call such discontinuity a *status gain*. Because gaining status has value, the payoff structure that prevails in Hierarchic Societies generates a new driver of internal conflicts that can be fulfilled even when conflicts can only destroy the resources of the opponent. Indeed, since status gains depend on relative rather than absolute changes in endowments, the objective to gain status can even rationalize conflicts where both parties always end up with less resources than before, i.e. where the cost of fighting is disproportionate to any possible direct material gain. As it turns out, this prediction is not only a theoretical possibility. As we discuss in our historical application, such disproportion is one of the most salient features of internal conflicts in Early Modern Europe. In agreement with our mechanism, historians have noted that the gap between direct gains and costs can be accounted for once we consider that, in many cases, the winners of these internal conflicts would experience more access to political rents.

In conclusion, this dissertation shows that ranking mechanisms offer a valuable and relevant perspective for many applied problems in political economics. Exploiting this perspective, Chapters 1 and 2 offered a first step towards a rational-agent account of the universal role of ranks, status, and hierarchies in international relations and domestic politics. As these Chapters show, ranking mechanism do more than just reinforcing inequality, they gen-

erate it. Highlighting the unfairness of such mechanisms while at the same time showing how natural they are, the dissertation offers a new way to look at political institutions, which should also be evaluated in terms of their ability to move us away from ranking mechanisms. Finally,the dissertation shows that the existence of ranking mechanisms may have important indirect consequences on other types of interactions. Indeed, Chapter 3 shows how the pattern of conflict in Hierarchic Societies is deeply affected by the ranking mechanism that underlies them. In this case, our analysis highlights a new mechanism whereby institutional arrangements that foster society's unity (e.g., representative institutions) can be valuable: not only they may act as checks on the government's ability to expropriate its subjects, but also they may reduce the incentives for groups' infighting, and thus the incidence of internal conflicts.

Table of Contents

List of Figures

Chapter 1

How Foreign Exposure Shapes Interstate Alignment: Theory and Evidence

1.1 Introduction

Governments often manage foreign influence by aligning their policies to the interest of foreign powers (Lake, 2011). Such alignment decisions are at the heart of international relations, affecting trade (Davis et al., 2019), international finance (Eichengreen et al., 2019; Ambrocio and Hasan, 2021), foreign aid (Rommel and Schaudt, 2020), and even conflicts (Mearsheimer, 2022). This paper proposes and tests a theory of how alignment decisions are determined by a country's exposure to the influence of foreign powers.

During the Cold War, alignment patterns were particularly stark and clearly linked to the power-projection capabilities of the great powers. With few exceptions, Eastern European countries aligned with the Soviet Union, and Latin American ones with the United States. This observation suggests that countries align more with relatively more influential foreign

powers, a pattern known as *bandwagoning* (Waltz, 1979). While bandwagoning has been observed in multiple instances throughout history, a more complex pattern emerges when we look at the modern evidence (Ross, 2006; Haacke, 2019). For instance, many Latin American countries decreased their alignment with the United States after the collapse of the Soviet Union (Levitsky and Roberts, 2011). Similarly, many countries increased their alignment with the US during China's rise (Selden, 2013). These cases suggest that alignment with a particular foreign power might be increasing in a country's exposure to the other foreign powers (*indirect exposure*), a pattern often called *balancing* or *hedging* (Lim and Cooper, 2015). But how does a country's indirect exposures matter for its alignment decisions? And how can we account for the evidence on both balancing and bandwagoning behaviors?

We propose and test a model where a country invests in its relation with (aligns with) a foreign power to prevent its ingerence, and indirectly discourage the other powers' interference. The optimal strategy, which we call *asymmetric hedging*, captures both balancing and bandwagoning behaviors. For instance, when pressured by two foreign powers, a country aligns with the one it is most exposed to, but the extent of alignment is increasing in its exposure to the other one. In general, we show that the key empirical implication of asymmetric hedging is that a country's alignment with a foreign power is increasing in the country's *Indirect Exposure Index* (*IE-index*) towards that foreign power – a novel directed bilateral index measuring a country's relative exposure to foreign powers it is less exposed to than the one considered. As a first test of the model, we perform two event studies to show that the theory can predict how the collapse of the Soviet Union or the rise of China had geographically heterogeneous effects on how countries adjusted their economic, military and diplomatic alignment with the US and other foreign powers. Second, we compute for each country its index of Indirect Exposure towards each foreign power using various measures of foreign influence, including trade dependence. We find that a country's IE-index towards a given foreign power is a relevant determinant of a country's alignment with that foreign

power. Additionally, we find that alignment with a foreign power is not further affected by a country's exposure to it, if we condition on its Indirect Exposure index towards it. Moving beyond alignment choices, we find that the IE-index is an important determinant of the allocation of foreign aid and the pattern of trade.

We derive our empirical predictions from a model where the government of a country (*policymaker*) tries to reduce its geopolitical risk (i.e., increase its chances of political survival) by obtaining the support of different foreign powers (*coercers*). The policymaker has a different *exposure* to the different coercers, a commonly known characteristic capturing how different coercers have a different impact on the policymaker's survival probability. The policymaker offers a *transfer* to each coercer to obtain its support. Transfers are paid only to supporting coercers, and if and only if the policymaker survives. We interpret each transfer as a measure of a country's alignment with a given foreign power. Upon observing the policymaker's transfer offers, coercers simultaneously choose whether to support or oppose the policymaker (i.e., accept the transfer or get an outside option). These decisions determine the probability of the policymaker's survival, which is increasing in the policymaker's exposure to all supporting coercers and decreasing in the exposure to all the opposing ones.

Our main result is that the optimal[1] strategy captures both balancing and bandwagoning behaviors. In line with balancing, the extent of alignment depends on the policymaker's indirect exposures, i.e., its exposure to the other coercers. The main force behind this is that transfers are only received when the policymaker survives, which depends on the decisions of the other coercers and the policymaker's exposure to them. For instance, if the policymaker wants to obtain the support of one coercer, it must compensate it for the risk caused by all the others, which is increasing in the policymaker's exposure to all of them. However, once the policymaker has secured one supporter in this way, it can obtain a

[1]Note that the policymaker's offer will typically be compatible with multiple subgame perfect equilibria. In line with a literature in mechanism design, we address this by assuming a robustness refinement (unique implementation) that selects (almost) one equilibrium (Segal, 1999a, 2003a; Winter, 2004a).

second supporter with a lower transfer: decreasing in its exposure to the first coercer (now contributing to the policymaker's stability) but still increasing in its exposure to all the others. As a consequence, transfers will sometimes be increasing and sometimes decreasing in the policymaker's indirect exposures. Finally, in line with bandwagoning, we show that the policymaker aligns more with coercers it is more exposed to. The intuition for this is that the policymaker's geopolitical risk is more affected by the decisions of the coercers it is more exposed to. Building on this, we show that the policymaker would prefer to obtain first the support of the coercer it is most exposed to, then of the secondmost, and so on, which results in bandawgoning behavior.

As this discussion shows, our contribution is not limited to showing that indirect exposures matter for alignment choices. Indeed, our analysis provides a full characterization of how they matter: a country's alignment with a foreign power is increasing in its exposure to the foreign powers it is less exposed to, and decreasing in its exposure to the others. In this sense, because bandwagoning results from how a country hedges against its exposure to the other foreign powers, we call our strategy asymmetric hedging. Finally, we simplify the empirical implications of asymmetric hedging showing that, under a natural restriction, alignment with a given foreign power is increasing in the country's relative exposure to all foreign powers it is less exposed to: the country's Indirect Exposure Index towards that foreign power.

To test the empirical relevance of asymmetric hedging, we must first measure alignment behavior. Our source of data on alignment choices is the GDELT dataset of international interactions, a dataset generated from news articles reporting the occurrence of a particular type of interaction for different country pairs between 1979 and 2012.[2] In our main empirical analysis we focus on a variable that we call *offers*, i.e., the number of times over a year in which a country (a policymaker) pledged/offered/promised to another country (a coercer)

[2]The GDELT dataset can be accessed from `https://www.gdeltproject.org`

to expand their cooperation in the economic, military or diplomatic domains.

Second, we need to measure foreign exposure, a crucial element of our theory. In our model, exposure captures how the decision of a specific coercer impacts the policymaker's political stability (its survival probability). Due to this consequence-based approach, the theory does not need to take a stand on which variable generates power. Consequently, in our model, countries might be employing different types of resources to exert their influence, e.g., different mixes of economic and military leverage. From an empirical perspective, however, we must measure exposure. To do so, we follow two distinct paths, both leading to the same results. First, in our baseline analysis we focus on a particular type of exposure – every policymaker's *economic exposure* to each coercer: the value of the trade flow between the two divided by the value of all goods produced or consumed in the former. According to this metric, a policymaker is more exposed to a coercer if the coercer can impose more significant economic imbalances on the policymaker.[3] In Section A.4.1, we replicate all our results with a more complex but realistic measure of exposure, the Formal Bilateral Influence Capacity (FBIC) developed by Moyer et al. (2018). Such measure goes beyond economic exposure, as it also measures, for instance, a policymaker's dependence on arms imports from each coercer.

As a first empirical exercise, we focus on how the collapse of the Soviet Union or China's rise led countries to change their alignment with the other foreign powers. In line with the theory, we show that the collapse of the Soviet Union led the average country that was more exposed to the United States than the Soviet Union to reduce its alignment with the United States relative to the average country that was less exposed to the United States than the Soviet Union. Conversely, we show that the rise of China led the average country that was more exposed to the United States than China to increase its alignment with the United

[3]Economic exposure is related to the idea of trade dependence. If a coercer decides to delay, divert or block trade flows with a policymaker, then this leads to larger demand shortages or excesses in supply the more the policymaker is economically exposed to such coercer.

States relative to the average country that was less exposed to the United States than China. In addition, we find the same pattern when we study how countries changed their alignment with any foreign power, rather than only with the United States.

As a second test, we focus on the prediction that a country's alignment behavior should respond to its Indirect Exposure Index (IE-index) towards the various foreign powers. Given our measure of exposure, a policymaker's IE-index towards a given coercer is the policymaker's economic exposure to all coercers it is less economically exposed to. In our baseline analysis we show that the lag of a policymaker's IE-index towards a given coercer is a significant determinant of the economic, military, and diplomatic offers sent from the policymaker to that coercer. A standard deviation increase in IE-index leads to an increase of around 10% for each of these variables. Additionally, we find that a policymaker's exposure to a coercer does not lead to more alignment once we condition on its IE-index towards it.

Going beyond alignment decisions, we study how the IE-index can affect the allocation of foreign aid and the pattern of trade.[4] For aid, the key idea is that a policymaker with a higher IE-index towards a given coercer should receive more aid from it, as coercers are on average more likely to send aid to more aligned regimes. For trade, a standard mechanism is that a country's alignment can reduce bilateral trade frictions, an idea which we explore using a Gravity specification from Anderson and van Wincoop (2003).[5] In both cases, we find robust confirmation of the model's predictions.[6] A one standard deviation increase in the IE-index towards a foreign power increases the receipt of military, economic, or humanitarian aid from it by more than 10% and increases bilateral imports and exports by more than 20%.

[4]Our source on trade is CEPII, covering the 1949-2019 period. It can be accessed from `http://www.cepii.fr/`. Our outcome for aid is again from the GDELT dataset; we call *aid* the number of times over a year in which a country (a coercer) sent economic, military, or humanitarian aid to another country (a policymaker).

[5]F.i., Berger et al. (2013) assumes that the effect of foreign influence acts through bilateral trade costs.

[6]For trade, we address the potential endogeneity of economic exposure with a 2SLS estimation.

1.1.1 Contribution to the literature

Our project is at the intersection of international economics and political economy. After the seminal work of Hirschman (1980), the modern literature in this field has been mainly empirical, studying how exports and imports (Yeats, 1990; Berger et al., 2013; Fuchs and Klann, 2013; Mityakov et al., 2013; Du et al., 2017; Didier and Koenig, 2019; Davis et al., 2019), economic and financial aid (Kuziemko and Werker, 2006; Dreher and Jensen, 2007; Kilby, 2009; Faye and Niehaus, 2012; Rommel and Schaudt, 2020), loans (Li and Ngo, 2018; Garmaise and Natividad, 2013; Ambrocio and Hasan, 2021), and even the extent of media coverage of human rights' violations (Qian and Yanagizawa, 2009; Qian and Yanagizawa-Drott, 2017) are all linked to political considerations and power.[7] The main theoretical paper in this literature is Antràs and Padró i Miquel (2011), studying the welfare effects of the action of one foreign power that constrains the policies of a local ruler. Our contribution to this literature is to provide a simple (and empirically relevant) way to account for the complex considerations that drive alignment behavior when a country is exposed to more than one foreign power.

Our finding that the IE-index is a determinant of bilateral trade flows also speaks to the empirical literature on trade, especially the literature that employs gravity models (for reviews, cfr. Anderson 2011 or Bergstrand and Egger 2013; for recent work, see Arkolakis et al. 2012, and Anderson and Yotov 2020). These models decompose bilateral flows as only a function of the characteristics of the two countries involved, and two multilateral resistance terms, which summarize the frictions originating from the existence of other countries. Crucially, Anderson and van Wincoop (2003) shows that these multilateral resistance terms can be decomposed into an inward and outward component, where these components are specific to a particular country, not to a particular bilateral relation. This finding allows to reduce

[7] There is also a complementary strand of this literature that studies how economic considerations can affect political outcomes, e.g. Kleinman et al. (2020).

the dimensionality of the problem considerably: for instance, it is possible to account for the effect of third parties by simply including country fixed effects (one for each side of the bilateral relation), or with a normalization (Anderson, 2011). Our model shows that once we introduce the effect of foreign influence into a gravity model of trade, each bilateral trade flow is again a function of the characteristics of third parties. While country fixed effects or normalizations cannot account for this dependence, we show that this can be done by adding the IE-index to the standard estimating equation.

Finally, our paper contributes to the vast literature in international relations that studies the determinants of interstate alignment. A key contribution to this literature is Waltz (1979), who introduced the concepts of bandwagoning and balancing: aligning with the source of the threat or against it, respectively. Since the early 2000s, scholars tried to explain the more nuanced pattern of alignment that emerged after the Cold War with the concept of hedging (Wohlforth et al., 2007; Goh, 2007; Wivel, 2008; Nexon, 2009; Tessman, 2012; Selden, 2013; Guzansky, 2015; Koga, 2018; Lim and Mukherjee, 2019; Smith, 2020; Strating, 2020; Hamdi and Salman, 2020; Ambrosio, 2021). When viewed as an alignment strategy, hedging is typically described as a strategy that falls between bandwagoning and balancing (Roy, 2005; Goh, 2006; Kuik, 2016; Koga, 2018). Lacking a formal theory, however, this literature runs into several problems: different authors have conflicting views of what these strategies are, what purposes they serve, how they are related, and how to test their relevance in world politics (for the latest review highlighting these issues see Haacke and Ciorciari, 2022; but see also Kang, 2009; Selden, 2013; Lim and Cooper, 2015; Liff, 2016; Haacke, 2019; or Jones and Jenne, 2021). Our first contribution to this literature is to introduce the concept of asymmetric hedging as the cost-minimizing alignment strategy for a country exposed to the influence of multiple foreign powers. As a second contribution, our findings clarify that bandwagoning and balancing behaviors should be expected to coexist and might actually be two sides of the same strategy: when countries are adopting asym-

metric hedging, they should be expected to align more with stronger foreign powers (as with bandwagoning), and the extent of their alignment should depend on the capabilities of the other foreign powers (as with balancing). Our final contribution to this literature is to derive the observable implications of asymmetric hedging and to show that it produces empirically relevant predictions.

1.2 The key mechanism in an example

The key mechanism that drives our results is the idea that obtaining the support of a foreign power allows a local ruler to make decisions that are unpopular with the other foreign powers. We illustrate this force with two famous episodes from the Cold War.

In 1951, the newly elected Prime Minister of Iran, Mosaddeq, nationalized the Anglo-Iranian Oil Company. This decision triggered tensions with Britain which, in 1953, successfully orchestrated a coup against the Iranian president, with the support of the United States. The decision of the US to join Britain against Mosaddeq was in part driven by the fear that the Soviet Union could be soon attempting to sponsor their own regime change in Iran (Barr 2018, pp. 160-174). In other words, the undoing of Mosaddeq was the absence of any clear alignment with any of the foreign powers.

In 1956, the Egyptian President, Nasser, nationalized the Suez Canal, another major blow to British interests. This time the United States declared their support for Nasser, as they knew that Nasser enjoyed the support of the Soviet Union and thus were not worried about a communist take-over (Barr 2018, pp. 233-247). As Nasser was backed by the two global superpowers, Britain had no choice but to accept the nationalization.

1.3 A model of Alignment Behavior

Our model studies the determinants of the relation between a policymaker p and multiple coercers $c \in C$. In the examples of Section 1.2, a coercer c would be a foreign power such as the Soviet Union, the United States or Great Britain. Instead, the policymaker p would be Nasser or Mosaddeq, i.e., the ruler of some home country that tries to enlist the support of each c. We consider a game with the following timing and characteristics:

- The policymaker publicly announces a transfer for each coercer, $\mathbf{B} = (B_c)_{c \in C}$.
- Given the offer, coercers simultaneously choose whether to support the policymaker ($c \in IN$) or oppose it ($c \in OUT$).
- Given the decisions of the coercers, with probability $Pr(FAIL)$ the policymaker fails and with probability $1 - Pr(FAIL)$ it survives.
- A coercer that supports the policymaker obtains $-\kappa < 0$ when the policymaker fails and obtains $B_c - \kappa$ when it survives.[8] A coercer that opposes the policymaker obtains an outside option of 0, regardless of whether the policymaker fails or survives.

We interpret the event $FAIL$ as the local ruler p losing power in its country (regime change or failure) and B_c as the benefit that foreign power c obtains from its relation with p. The assumptions that the benefit B_c is received only if the local ruler remains in power and that $-\kappa < 0$ imply that supporting a policymaker when the policymaker fails is worse than opposing it. This captures the idea that supporting a local ruler exposes to the risk that it fails. To focus on this feature in the simplest way, we assume that rejecting the offer yields a fixed outside option of 0.

[8] The assumption that B_c is only received when the coercer supports the policymaker could be derived from a setting where p offers transfers that are conditional on the coercers' decisions and that are constrained to be non-negative.

1.3.1 Power and the endogenous probability of failure

To model how the probability of success depends on the choices of the coercers, we introduce an element of heterogeneity which we call *power* and denote by W_c. We assume that the power of the policymaker, W_p, and the distribution of power of the coercers, $\mathbf{W} = (W_c)_{c \in C}$, is common knowledge.

Given the coercers' decisions, i.e., sets $IN \subseteq C$ and $OUT = C \setminus IN$, we define w_{OUT} as the share of total power controlled by the coercers that oppose the policymaker, i.e., we let

$$w_{OUT} = \frac{\sum_{j \in OUT} W_j}{\sum_{j \in C} W_j + W_p}$$

Given a continuous and strictly increasing function $F(\cdot)$, we make the following assumption:

$$Pr(FAIL) = F(w_{OUT})$$

This assumption justifies our interpretation of W_c as the power that c has over p. Indeed, the characteristic W_c measures how the probability of p's survival depends on c's decision. In other words, W_c measures the importance for p of inducing c's support, capturing the intuitive idea that the decisions of *stronger* coercers are more important. Note that we are modeling power by its effect on the final outcome, not as an endowment of some specific resource. This allows us to think of power as generated by different types, or even by a different mix, of coercive resources.

In addition, this assumption justifies our interpretation of accepting/rejecting the offer as a decision on whether to support/oppose the policymaker. Indeed, this specification assumes that the probability of success depends positively on the power of those that accept the offer and negatively on the power of those that reject it. This is in line with our motivating examples in Section 1.2, where opposing the regime was the same as making a military

intervention against it.[9] But it is also in line with a setting where the decision to support a regime is the same as choosing whether to encourage or discourage trade with it. Indeed, when a foreign power decides to discourage trade with another country, then this can cause a recession and possibly a regime change. On the other hand, encouraging trade with another country can benefit its economy, thus reducing the probability of regime change.

In Section A.3.2 we compare the empirical performance of this model against the prediction derived under the more general assumption that $Pr(FAIL)$ is a decreasing function of the power of all supporting coercers (and the policymaker) and an increasing function of the power of all opposing ones. The analysis shows that this more general prediction does not add much to our empirical findings, justifying our focus on the simpler specification presented here.

1.3.2 Coercers' problem: supporting or opposing a local ruler

In line with the motivating examples in Section 1.2, a crucial feature of our setting is that the incentive of a coercer to support the policymaker depends on its expectations about what others do. Given the vector of offers, $\mathbf{B}$, and the set of coercers that are expected to oppose the policymaker, $OUT \subseteq C$, the decision of coercer c on whether to support the policymaker is as follows

$$c \in IN \quad \Longleftrightarrow \quad [1 - F(w_{_{OUT}})]B_c - \kappa \geq 0$$

Naturally, this condition is more likely to be satisfied when B_c is greater. This means that B_c is a tool for the policymaker to induce c to support it. But also, this condition is more likely to hold when $w_{_{OUT}}$ is smaller. Thus, the game features strategic complementarities:

[9] To interpret the choice $f \in NO$ as a military intervention, it would be natural to assume that the value of the outside option is increasing in W_c. It is possible to show that our results hold under this assumption.

players are more willing to support the policymaker, when the policymaker is expected to be supported by a more powerful coalition of coercers.

In international relations there are at least two sources of strategic complementarities. One is the simple idea that that a country is more likely to reach its objectives when it convinces other countries to join its action, rather than to oppose it. Another more subtle source of strategic complementarity can be appreciated from the foreign policy stances of the United States and the Soviet Union during the Cold War (the so-called Truman and Brezhnev doctrines). Both powers pledged a military intervention whenever they feared that the other foreign power was about to make an intervention. In this case, the source of the strategic complementarity comes from the fear that the opponent might gain an advantage by catching the other power off guard.

In the motivating examples of Section 1.2, the first source of strategic complementarity emerges from the role of the United States: the logistic support of the United States was crucial to successfully substitute Mosaddeq; on the other hand, Great Britain immediately aborted its action against Egypt when it was clear that they did not have the support of the United States. The second source of strategic complementarity can be understood by the role of the Soviet Union: the United States decided to support Britain against Mosaddeq because it feared that the Soviet Union could make its move against Mosaddeq. Instead, it sided with Nasser because it did not fear any action of the Soviet Union.

1.3.3 Policymaker's problem: buying the foreign powers' support

Rather than restricting ourselves to a particular specification of the policymaker's payoff, we will simply make the assumption that the objective of the policymaker is to maximize its probability of survival at the minimum cost $\sum_{c \in C} B_c$.

However, because our model features strategic complementarities and coercers choose

simultaneously, then the same offer $\mathbf{B}$ will typically be consistent with multiple equilibria. Thus, whether a particular offer $\mathbf{B}$ maximizes the probability of survival will typically depend on whether the policymaker expects the coercers to coordinate on one particular equilibrium or on another one.

In line with a recent literature (Segal, 2003a; Winter, 2004a), we impose a cautiousness refinement, assuming that the policymaker's objective is to maximize its probability of survival in the worst-case equilibrium compatible with the given offer.

Definition 1 (Robust offers) *A vector of offers $\boldsymbol{B}$ is robust if there is no SPE where the policymaker offers $\boldsymbol{B}$ and one or more coercer oppose the policymaker.*

In the context of our game, this requirement is the same as imposing that the offers must be such that the only rationalizable response for each coercer is to support the ruler. Or, in terms of equilibrium, this is equivalent to restricting attention to offers that are only compatible with equilibria where all coercers accept them (unique implementation, as in Halac et al., 2021).

An important feature to highlight is that focusing on robust offers is not the same as focusing on worst-case outcomes, as robust offers are still part of an equilibrium.[10] In other words, a local ruler that makes robust offers must not only make sure that each foreign power wants to support it individually, but also that they do not end up withdrawing their support out of fear that another foreign power will do so. This assumption is an essential ingredient of our motivating examples in Section 1.2. Great Britain could sway the United States into opposing Mosaddeq by playing on the possibility of a communist take-over. On

[10]There is a large literature in International Relations that focuses on worst-case scenarios (Tang, 2008). For instance, Mearsheimer (2001) justifies this assumption as follows "Intentions are ultimately unknowable, so states worried about their survival must make worst-case assumptions about their rivals' intentions" (p. 45). Unlike this assumption, we do not subscribe to the idea that intentions are unknowable, as our equilibrium analysis is restricting the set of possible intentions using rationality and common belief in rationality. But within the outcomes that are compatible with this restriction, then we focus on worst-case scenarios.

the other hand, the United States supported Nasser as there was no doubt that Nasser was supported by the Soviet Union.

1.3.4 Simple case: the value of political alignment

To build intuition for our main result, we discuss a simple example where a policymaker p with power $W_p = 0.4$ deals with two identical coercers c and $-c$, with power $W = 0.3$.

Suppose that the policymaker offers B_2, the smallest transfer that induces c to support the policymaker when $-c$ is expected to support the policymaker as well:

$$B_2 \quad \text{such that} \quad [1 - F(0)]B_2 - \kappa = 0$$

Offering B_2 to both coercers is part of an equilibrium: if the two coercers actually expect the other to support the policymaker, then it is rational for them to accept B_2. But the offer $\mathbf{B} = (B_2, B_2)$ is not robust: there is an equilibrium where neither of the two coercers accept B_2 as they expect that the other will not support the policymaker.

An offer $\mathbf{B}$ that certainly satisfies our robustness requirement is $\mathbf{B} = (B_1, B_1)$, where $B_1 > B_2$ is the smallest transfer that induces c to support the policymaker when $-c$ is expected not to support the policymaker:

$$B_1 \quad \text{such that} \quad [1 - F(0.3)]\, B_1 - \kappa = 0$$

While by construction $\mathbf{B} = (B_1, B_1)$ is a robust offer, it is not cost-minimizing. Once the policymaker offers B_1 to c, then $-c$ will anticipate that c will always support the policymaker. As a consequence, the policymaker can reduce the offer that it makes to $-c$ from B_1 to $B_2 < B_1$, without compromising the robustness of its offer.

But, in this example, c and $-c$ are identical. So, there are two cost-minimizing robust

offers that are equally appealing for the policymaker (B_1, B_2) or (B_2, B_1). Both these offers can be thought of as the result of a two-step procedure: assign a rank to each coercer and then offer B_1 to rank 1 and B_2 to rank 2. Indeed, the next Section shows that this structure is the key feature of every robust cost-minimizing offers.

As this example shows, our analysis captures the key mechanism highlighted in the examples of Section 1.2: obtaining the support of a foreign power allows the local ruler to make decisions that are unpopular with the other foreign powers.

1.3.5 Main result: why aligning with the stronger powers

Generalizing the logic of Section 1.3.4 to the case with multiple and heterogeneous coercers, we reach our main result:

Theorem 1 *Any cost-minimizing robust offer can be constructed as follows:*

- *Assign a rank to each c so that more powerful coercers obtain a higher rank;*
- *To each coercer, offer the smallest B that induces it to support the policymaker when all and only higher-ranked coercers support the policymaker.*

As discussed in the example, the theorem is driven by the logic that making an offer that guarantees the support of a coercer allows the policymaker to reduce the offers that it makes to the other coercers, without reducing its probability of survival. This is driven by strategic complementarities: when a country obtains the support of one of the coercers, this increases its probability of survival, which allows the policymaker to offer smaller transfers to the other coercers without losing their support.

The only element of the theorem that is new relative to the discussion of Section 1.3.4 is the finding that the ranking is pinned down by the distribution of power: higher ranks are assigned to stronger coercers. To understand this, note that every equilibrium offer must

be increasing in the power of all lower-ranked coercers and decreasing in the power of all higher-ranked coercers. Thus, to minimize total costs, the policymaker prefers to assign higher ranks to stronger rather than weaker coercers.

1.3.6 Comparative statics and the Indirect Exposure Index

To lay the ground for the empirical analysis, we discuss some comparative statics on how changes in the distribution of power affect equilibrium transfers (and thus payoffs).

For notational simplicity, from now on we assume that there are no ties in the distribution of power, i.e., we assume that $W_c = W_{c'}$ if and only if $c = c'$. In this case, there is a unique cost-minimizing robust offer $\hat{\mathbf{B}}$. For each c, let $W^{tot} = \sum_{j \in C} W_j + W_p$, such offer is

$$\hat{B}_c = \frac{\kappa}{1 - F\left(\frac{\sum_{\ell : W_\ell < W_c} W_\ell}{W^{tot}}\right)} \tag{1.1}$$

The first element that emerges is that $\hat{B}_c$ depends on the whole distribution of power $\mathbf{W}$ in a non-trivial way: it is a non-monotone function of W_c and of each $W_{c'}$. To make progress in characterizing this dependence, we introduce the definition of a coercer's power rank:

Definition 2 *The **Power Rank** of $c \in C$ is a function $\rho : \mathbb{R}^N_+ \to \{1, 2, .., N\}$ such that for each $c, c' \in C$, $\rho^c(\boldsymbol{W}) < \rho^{c'}(\boldsymbol{W})$ if and only if $W_c > W_{c'}$. We say that c has higher (lower) power rank than c' if $\rho^c(\boldsymbol{W}) < \rho^{c'}(\boldsymbol{W})$.*

From Equation A.4, we can see that, once we fix the power rank, we can provide the following comparative statics:

Proposition 1 *Consider changes in W_c that do not affect $\rho^c(\boldsymbol{W})$, then*

- *for any c' with $W_{c'} > W_c$, $\hat{B}_{c'}$ is increasing in W_c,*
- *for any c' with $W_{c'} < W_c$, $\hat{B}_{c'}$ is decreasing in W_c.*

To understand the logic behind these comparative statics, recall that, from the previous theorem, every coercer receives a transfer that induce it to support the policymaker when all weaker coercers oppose the policymaker and all stronger ones support it.

With this in mind, it is natural to expect the equilibrium transfer to increase in the power of weaker coercers: a policymaker who is opposed by stronger coercers have a lower expected probability of survival and must thus promise a larger transfer to induce support. Symmetrically, we can see why the equilibrium transfer must be decreasing in the power of stronger coercers: when the policymaker is supported by stronger coercers, its probability of survival is higher, and it can thus induce others to support it with a lower transfer.

In addition, when the power rank is fixed, we can also show the counter-intuitive result that a country's equilibrium payoff is decreasing in its own power:

Proposition 2 *$\hat{B}_c$ is decreasing in any change in W_c that does not affect $\rho^c(\boldsymbol{W})$.*

The logic for this result is straightforward: when a country chooses to support the policymaker, then its power contributes to increasing the policymaker's probability of survival. As a consequence, for fixed power rank, when a coercer becomes more powerful, the policymaker can reduce its transfer while still inducing its support.

When we study changes in power that also affect the power rank, then the model features some important non-monotonicities and non-linearities. For instance, an increase in power that increases a coercer's power rank can improve its equilibrium payoff, but at the same time this positive effect can be smaller when the change is larger. As discussed in the introduction, these rich dynamics can capture some important and counter-intuitive patterns of modern and historical international relations. However, as we will shortly show, we can still provide a simple and complete characterization of how the distribution of power affects payoffs, even when the power rank is at stake. A first step to do so is to introduce the definition of the Indirect Exposure Index:

Definition 3 *The **Indirect Exposure Index (IE-index)** of c over p is defined as the combined share of total power W^{tot} of every coercer that is weaker than c, i.e. the function*

$$IE - index^c(\boldsymbol{W}) = \frac{\sum_{\ell:W_\ell<W_c} W_\ell}{W^{tot}}$$

Given this definition and Equation A.4, we can establish the following result:

Proposition 3 *The equilibrium payoff that each c obtains from p is increasing in c's IE-index over p. Moreover, the distribution of power $\boldsymbol{W}$ affects the distribution of equilibrium payoffs only via the vector of IE-indexs, i.e., $(IE - index^c(\boldsymbol{W}))_{c\in C}$.*

This proposition tells us that all the non-monotonicities of our model can be summarized by a single one-dimensional index of the distribution of power, the IE-index. In addition, as we show in Proposition 4, we can show that our model is compatible with a log-linear relation between equilibrium transfers and IE-index, implying a relatively simple and easy to test relation.

Note that, given a distribution of power $\mathbf{W}$, coercers with higher power rank have a larger Indirect Exposure Index. As a consequence, the IE-index can be thought of as formalizing the idea that the power rank matters in international relations, as assumed by a large literature in international relations (see Section 1.1.1). Moreover, the dependence of equilibrium payoffs on the Indirect Exposure Index clarifies *how* the power rank matters: it matter through the Indirect Exposure Index.

1.3.7 Towards the empirics: multiple policymakers and periods

In order to speak to the empirical analysis and historical applications, we expand the model to multiple time periods $t \in T$ and multiple policymakers $p \in P$, and make some simpifying assumptions about the parameter space. We will expand the model in the simplest possible

way by simply assuming that the interaction between every policymaker and every coercer is a repetition of the game that we analyzed above.[11] Still, to express the previous results in this more complex environment, we need to develop some notation.

Every policymaker $p \in P$ faces a distribution of power in each t that we denote by $\boldsymbol{W}_{pt} = (W_t^p, (W_t^{c\to p})_{c\in C})$, where $W_t^{c\to p}$ is the power that coercer c has over policymaker p in period t and W_t^p is the (domestic) power of country p, i.e., its ability to protect itself from external influence. We let $W_{pt}^{tot} = \sum_{j\in C} W_t^{c\to p} + W_t^p$ be the total power that is relevant for p in period t. Given a distribution of power among the coercers and W_{pt}^{tot}, we denote the IE-index that country c has over country p in period t as the following function

$$IE - index_t^{c\to p}(\mathbf{W}_{pt}) = \frac{\sum_{\ell\in C W_t^{\ell\to p} < W_t^{c\to p}} W_t^{\ell\to p}}{W_{pt}^{tot}}$$

Finally, we denote the transfer (or payoff) vector $\boldsymbol{B}_{pt} = (B_t^{p\to c})_{c\in C}$, where $B_t^{p\to c}$ is the transfer (payoff) offered by policymaker p in each period t to coercer $c \in C$.

For what concerns the model's parameters, we allow the cost $\kappa > 0$ of supporting the policymaker to be heterogeneous in each period and for each policymaker, we thus denote it by κ_{pt}.[12] In addition, we assume that the function that determines the probability of failure, i.e. $F(\cdot)$, takes the following functional form $F(w) = 1 - e^{-\lambda w}$, for any $w > 0$ and some $\lambda > 0$. Under these assumptions, we can derive from the previous analysis the following result

Proposition 4 *The equilibrium transfer (thus payoff) $\hat{B}_t^{p\to c}$ that each c receives from each*

[11] This amounts to assuming that countries are not sufficiently patient to play anything different than the repetition of the equilibrium that we discussed.

[12] In the empirical analysis, this parameter will be absorbed by the fixed effect capturing all unilateral time-varying characteristics of each policymaker.

p in every period t is such that:

$$\ln(\hat{B}_t^{p\to c}) = \ln(\kappa_{pt}) + \lambda IE - index_t^{c\to p}(\boldsymbol{W}_{pt}) \tag{1.2}$$

Note that the assumption $F(w) = 1-e^{-\lambda w}$ with $\lambda > 0$ means that λ measures the responsiveness of the probability of survival to the power of the coercers that support the policymaker. In other words, λ is the channel that introduces the strategic complementarities into our model. As we can see above, the Indirect Exposure Index matters if and only if $\lambda > 0$, clarifying that strategic complementarities are the key mechanism whereby the IE-index affects equilibrium payoffs.

1.4 Empirical analysis: how the Indirect Exposure Index shapes international cooperation

In our empirical analysis, our unit of observation is a country pair over in a given year. As we will show, both our measures of power and our outcomes are *directed variables*. For instance, the aid that country B sends to country A is directed, as it is typically different from the aid that A sends to B (see Table 1.1 for other examples).

As we deal with directed variables, we specify the role played by each country in the pair. In line with the model's terminology, when we study how the relation between country A and B is affected by the power over country B, we call country A a coercer and country B a policymaker. To emphasize this aspect of our data, we will refer to each observation as a policymaker-coercer pair.

The crucial empirical prediction that distinguishes our theory from other natural hypotheses is that the relation between a coercer A and policymaker B is not determined by A's power over B but by the power that every other country has over B. Specifically, our

key variable of interest is the Indirect Exposure Index of A over B, the share of total power over B of all countries that have less power over B than A does. As formalized in Equation 1.2, we wish to test the hypothesis that coercer A obtains larger geopolitical rents from B (e.g., more favorable policies) when it has a higher IE-index over B.

Our baseline specification allows the same country to be both a policymaker for multiple coercers and a coercer for multiple policymakers.[13] We test our model with three empirical exercises. First, in Section 1.4.3, we show that the IE-index plays a much more fundamental role than a country's own power to predict the pattern of international cooperation. Secondly, Section 1.5 shows that our model can accurately predict the effect of the fall of the Soviet Union and the rise of China (two large changes in IE-index) on the average foreign power as well as on the United States. Finally, Section 1.6 demonstrates the empirical relevance of the IE-index to understanding bilateral trade flows.

1.4.1 Measuring power at the bilateral level: Economic Power

Until now, we considered a coercer as more powerful when its decision to support/oppose the policymaker had a larger impact on the probability of survival of the policymaker (the policymaker's political stability). For the empirical analysis we must abandon this fairly general definition and measure power from observable characteristics. Our empirical analysis focuses on economic power, the ability of a country to destabilize another one by harming its economy.

A simple way to measure this is to define the power of every coercer c over a policymaker p in period t as the value of the trade flow between the two, $Trade_{cpt}$.[14] In other words,

[13]This analysis uses a global sample, where every country is both a policymaker and a coercer. The results are robust to focusing on subsamples where only foreign powers (flexibly defined) can be coercers, and they are excluded from the set of policymakers (see Section A.4.3).

[14]Our source of data is the publicly available CEPII Gravity, `http://www.cepii.fr/`. Such a dataset includes yearly observations at the country pair level on imports, exports, population, GDP, and many other variables used in Gravity models. The dataset covers the 1949-2019 period. Appendix A.2 details the

this measures power as the ability of a country to manipulate trade flows causing economic imbalances in another country.

In line with this, the ability of policymaker p to protect itself from the power of all the coercers is instead defined as $GDP_{pt} - Exports_{pt}$, the value of all finished goods and services produced *and* consumed in country p in period t (the size of p's economy in period t when isolated from the international system).[15] Based on this, we can conclude that the total economic power that is relevant for country p in period t is $GDP_{pt} + Imports_{pt}$. We can thus provide the following definition of a coercer's relative economic power over a policymaker:

Definition 4 *A coercers c's **Relative Economic Power** (REP) over p in period t is defined as the share of economic power over p that is controlled by c in period t:*

$$E_t^{c \to p} = \frac{Trade_{cpt}}{GDP_{pt} + Imports_{pt}}$$

We let $\mathbf{E}_{pt} = (E_t^{c \to p})_{c \in C}$ be the distribution of power shares over p in a particular period, i.e., the vector collecting the REP that each c has over p in t.

Note that, from the policymaker's perspective, the concept of Relative Economic Power captures its trade dependence relative to c, a standard measure of a country's vulnerability towards another one (Moyer et al., 2018). In Section A.4.1 we show that we can replicate all of our results with a more complex but realistic measure of power.

The Indirect Exposure Index of Economic Power

As we have seen in Definition 3, the Indirect Exposure Index of a coercer is simply the share of power of all weaker coercers. Because Definition 4 introduces a country's Relative

construction of trade flows.

[15] Recall that a country's GDP is the value of all goods and services produced within the country. As a consequence, it includes exports but it excludes imports. That is why we are considering $GDP_{pt} - Exports_{pt}$ rather than $GDP_{pt} - \sum_{c \in C} Trade_{cpt}$.

Economic Power as the country's power share, we can define the Indirect Exposure Index of economic power directly from the distribution of Relative Economic Powers.

Definition 5 *The **Indirect Exposure Index (IE-index)** of economic power is the share of total economic power over p controlled by coercers with less economic power over p than c*

$$IE - index_t^{c \to p}(\mathbf{E}_{pt}) = \sum_{\ell \in C : E_t^{\ell \to p} < E_t^{c \to p}} E_t^{\ell \to p}$$

Importantly, recall that we defined the Relative Economic Power $E_t^{\ell \to p}$ as the *share* of total economic power over a policymaker. Thus, the trade flow between coercer and policymaker (the coercer's economic power over a policymaker) is in the denominator of the coercer's IE-index over the policymaker. Consequently, while there is a positive correlation between the IE-index of c over p, their trade flows, and the coercer's REP over such policymaker, any increase in the economic power of c over p will decrease its IE-index over p.

In Table A.4.1, we provide some summary statistics for our measures of power, and we show that the correlation between the IE-index and the REP is 0.71. However, such correlation is reduced to 0.57 when applying the fixed effects of our baseline specification (Table A.4.2), and it goes to 0.41 when we also condition on the coercer's rank in the distribution of power (see Section 1.4.3).

1.4.2 (Indirectly) measuring transfers: International Interactions

Our model characterized a relation between the distribution of power and the aggregate transfer (payoff) that each coercer obtains from each policymaker. We will test this indirectly by focusing on the relation between power and various types of international interaction.

Our source on international interactions is the GDELT dataset: an *Event Dataset* with

a global coverage from 1979 to 2012.[16] Event Datasets like GDELT generate their variables from news articles. GDELT employs a machine-learning algorithm to infer the occurrence of a specific type of interaction between country pairs using news articles from multiple media agencies (see Leetaru and Schrodt 2013 for details). GDELT generates a separate variable for each type of interaction, where the possible types are the categories of the CAMEO taxonomy of international interactions (Gerner et al., 2009).

In our main empirical analysis, we focus on six variables, each corresponding to a specific CAMEO category: economic, military, and diplomatic *offers*, and economic, military, and humanitarian *aid*.

The first three variables focus on the decisions of each policymaker relative to each coercer. Indeed, each type of *offer* sums the number of times over a year in which a country (a policymaker) pledged/offered/promised to another country (a coercer) to expand their cooperation in some specific domain (economic, military, or diplomatic).

Studying aid, on the other hand, allows us to take the perspective of the coercer. Indeed, each type of *aid* is the sum of the number of times over a year in which a country (a coercer) sent some aid type (economic, military, or humanitarian) to another country (a policymaker).

Appendix A.2 provides a detailed discussion on the construction of such variables.[17] Table 1.1 reports more details. Finally, in Section A.4.2 we replicate our results for an alternative type of outcomes from GDELT.

[16]The GDELT dataset is publicly available, `https://www.gdeltproject.org`

[17]As discussed in Appendix A.2, we scale these variables to improve our interpretation of the coefficients, and we transform them in a way that avoids dropping the zeros when taking the log.

1.4.3 Testing the theory: how power affects International Interactions

As the transfer in our model, each type of *offer* increases the coercer's willingness to support (rather than to oppose) the policymaker. Instead, we are interested in aid because coercers who receive a larger equilibrium payoff from a policymaker should be more willing to support it in times of need.

These ideas can be shown to lead to a positive log-log relation between each type of offer that the policymaker sends to the coercer (or each type of aid that the coercer sends to the policymaker) and the aggregate transfer that the coercer receives from the policymaker. Thus, by Proposition 4, there should be a positive log-linear relation between each of our outcomes of interest and the coercer's Indirect Exposure Index (IE-index) on a policymaker.

Let $\ln(Y_{cpt})$ denote a specific type of offer or aid. We focus on the following empirical specification:

$$\ln(Y_{cpt}) = \beta_1 IE - index_{t-1}^{c\to p}(\mathbf{E}_{p,t-1}) + \beta_2 E_{t-1}^{c\to p} + \alpha \ln(Trade_{cpt}) + \alpha_{ct} + \alpha_{pt} + \alpha_{cp} + \varepsilon_{cpt} \tag{1.3}$$

The crucial element of interest in this regression is β_1. This parameter is positive if and only if $\lambda > 0$, where λ is the link between the IE-index and the equilibrium transfer, as discussed in Proposition 4. In other words, finding $\beta_1 > 0$ is a test of the relevance of the IE-index for the equilibrium transfers. Moreover, note that we are estimating a separate β_1 for each type of aid and each type of offer, implying that we can use our various types of outcomes to cross-validate the analysis.[18]

[18]We do this exercise because it is the most restrictive. We have verified that, under natural conditions, if the various types of offers are complements, then the only thing that matters for each type of offer is the overall value that the policymaker wants to send to the coercer, thus the IE-index of the coercer. Similarly, suppose the various types of aid are complements. In that case, the only thing that matters for each type of aid is the overall investment that the coercer is willing to make to increase the policymaker's probability

The inclusion of $E_{t-1}^{c\to p}$, c's Relative Economic Power (REP) over a policymaker p, allows us to distinguish our theory (predicting $\beta_1 > 0$) from the natural hypothesis that relatively stronger countries should obtain better international relations (predicting $\beta_2 > 0$).

In Table 1.2 we estimate Equation 1.3 with the restriction that $\beta_2 = 0$, in the odd Columns of Table 1.3 we estimate Equation 1.3 with $\beta_1 = 0$ and in the even Columns of Table 1.3 we estimate the unrestricted specification. Finally, in Table 1.4, we estimate Equation 1.3 controlling for the major source of correlation between IE-index and REP: the power rank. In the odd Columns, we include fixed effects for each coercer's rank in the distribution of economic power. Instead, in the even Columns, we report coefficients estimated in a subsample with no changes in rank.[19]

Discussion of the Specification

Equation 1.3 includes a restrictive set of fixed effects: α_{pt} and α_{ct} net out all time-varying unilateral characteristics of every coercer and every policymaker (e.g., GDP, Population, etc.), and α_{cp} nets out all time-invariant bilateral characteristics (e.g., distance, common language, etc.) between each policymaker and each coercer.

Consider, as an example, the case of offers. The inclusion of α_{ct} implies that the coefficient β_1 is estimated using the variation in the number of offers that coercers receive *in the same year* but from different policymakers. In addition, the inclusion of α_{pt} normalizes such variations to account for the number of offers that each policymaker has sent in that year. Finally, the inclusion of α_{cp} means that all these variations are relative to the long-run outcome of the particular coercer-policymaker pair under consideration. In Section A.3.1 we

of survival, which is an increasing function of the equilibrium transfer, thus the IE-index of the coercer. Additionally, it can be shown that each of these outcomes contains some separate information about the IE-index. Indeed, in Appendix A.3.6, we show that we can find a relation between the IE-index and each of these outcomes when we control for the others.

[19]Footnote 22 defines a coercer's rank, Footnote 23 defines rank fixed effects, and Footnote 24 provides details on the subsample of interest. The correlation between REP and IE-index is 0.41 when we add rank fixed effects.

discuss why we include these fixed effects and we show that results are robust to alternative specifications.

Furthermore, note that Equation 1.3 controls for the log of the trade flow between policymaker and coercer, $\ln(Trade_{cpt})$. Consequently, we are dropping all coercer-policymaker pairs that do not trade with each other, which is consistent with the idea that our theory should apply to coercers with some power over a given policymaker.

Finally, note that we are using the distribution of economic power of the previous period ($\mathbf{E}_{p,t-1}$) to measure a country's Relative Economic Power (REP) and to construct its IE-index. Because we focus on lags and control for the same-year trade flows, our results are unlikely to be driven by some direct effect of trade or some omitted variable affecting trade flows and international interactions in the same year. To further corroborate this, in Section A.3.3 we show that our results do not change when we exclude trade flows and/or when we use the same-year distribution of economic power to calculate the REP and/or the IE-index.

Results

Table 1.2 reports the result of estimating Equation 1.3 while fixing the coefficient of the lagged Relative Economic Power (REP) at zero, i.e., fixing $\beta_2 = 0$. Under this restriction, the Indirect Exposure Index (IE-index) lag emerges as an important determinant of every type of offer and aid. A standard deviation increase in the IE-index (a change of 0.06) leads to a change that is between 12% (for economic offers) and 4% (for military aid).

As Table 1.2 does not condition on the coercer's REP, it could simply reflect the natural hypothesis that stronger countries obtain better international relations. Indeed, there is a positive correlation between a country's IE-index and its REP. Moreover, the odd Columns of Table 1.2 show a positive effect of the REP on international interactions, when we omit the IE-index (i.e., fixing $\beta_1 = 0$ in Equation 1.3).[20]

[20] A comparison of the within R-squared between the odd Columns of Table 1.2 and Table 1.2 reveals that

To address this, in the even Columns of Table 1.3 we estimate Equation 1.3 without any restriction. We find that when we estimate the effects of both the Indirect Exposure Index and the Relative Economic Power, the IE-index remains positive and significant. In contrast, the REP turns negative (and significant). This surprising result is in line with the theoretical prediction that, for a given power rank, an increase in power can be detrimental.[21]

Finally, in Table 1.4 we compare the effect of the IE-index and the REP while controlling for the major source of correlation between the two: the coercer's rank in the distribution of REP over a given policymaker.[22] In the odd Columns of Table 1.4 we include fixed effects for the coercer's rank in the distribution of REP.[23] In the even Columns of Table 1.4 we restrict attention to subsamples where there are no changes in rank.[24] Table 1.4 confirms the result that, conditioning on the REP, the effect of the IE-index remains positive and significant for every outcome. On the other hand, conditioning on the IE-index, the effect of the REP is either indistinguishable from zero or negative.

---

a coercer's IE-index fits the data much better than its REP. In addition, while the estimated coefficient of the IE-index and REP have similar magnitudes, note that the standard deviation of the IE-index is three times as large as that of the REP, implying very different counterfactuals (see Tables A.4.1 and A.4.2).

[21] Note that in our specification, we include a policymaker-year fixed effect, α_{pt}. Consequently, when we study the effect of the Indirect Exposure Index (the share of power of all weaker coercers) and the coercer's Relative Economic Power (its share of total power), we are using as a reference category the share of power of all stronger coercers. In general, a country's share of power (its REP) should have the same effect as the share of power of all stronger coercers (the reference category). However, if there is any measurement error that leads to misclassifying some weaker coercers as stronger (and vice versa), then we might find a negative and significant effect of a country's REP for given IE-index.

[22] For each $c \in C$ this variable is defined as $\rho_t^{c\to p}(\mathbf{E}_{pt}) = 1 + |\{\ell \in C : E_t^{c\to p} > E_t^{\ell\to p}\}|$

[23] Formally, we add $\sum_{k\geq 1} \alpha_k \mathbf{1}\{\rho_{t-1}^{c\to p}(\mathbf{E}_{p,t-1}) = k\}$ to Equation 1.3, where $\rho(\cdot)$ is defined as in Footnote 22.

[24] To maximize the sample size, for each policymaker-coercer pair we drop observations where the coercer has a power rank different than the mode of its power rank over that policymaker.

1.5 Event Studies: impact of the collapse of the Soviet Union and the rise of China on other foreign powers

This section performs two event analyses for the collapse of the Soviet Union (1989-1994) and the rise of China (2002 onwards). The objective is to study how these events affected the same coercer differently relative to different policymakers.

According to our model, we expect that the collapse of the Soviet Union harmed coercers *where* they were more powerful than the Soviet Union, relative to *where* they were weaker. Similarly, we expect that countries benefited from the rise of China *where* they were stronger than China compared to *where* they were weaker. We test these predictions indirectly by studying how these events affected aid and offers relative to different policymakers.

To formalize a test, let $\tilde{c}$ be either China or the Soviet Union (Russia, after 1991). We define $\underline{T}_{\tilde{c}}$ as the first year of the event of interest and $\bar{T}_{\tilde{c}}$ as the last year.[25] Moreover, we let $\mathbf{X}_{cpt}$ be a vector of controls including the log of the trade flow between c and p in period t and the *same-year* Relative Economic Power (REP) of c over p (i.e., $E_t^{c\to p}$). Letting $\ln(Y_{cpt})$ be the log of a specific type of offers received or of aid sent, we estimate

$$\begin{aligned}\ln(Y_{cpt}) =&\beta \times \mathbf{1}\{E_{t-1}^{c\to p} > E_{t-1}^{\tilde{c}\to p}\} \times \mathbf{1}\{\underline{T}_{\tilde{c}} \leq t \leq \bar{T}_{\tilde{c}}\}+ \\ &+ \alpha_0 \mathbf{1}\{E_{t-1}^{c\to p} > E_{t-1}^{\tilde{c}\to p}\} + \alpha \mathbf{X}_{cpt} + \alpha_{ct} + \alpha_{pt} + \alpha_{cp} + \varepsilon_{cpt} \qquad (1.4)\end{aligned}$$

With this specification, β captures a Difference in Differences type of variation. The first difference is the change in c's relations where it was stronger than the USSR (or China). The second difference is the change in c's relations where it was weaker than the USSR (or

[25] We assume that $\underline{T}_{\tilde{c}}$ is 1989 for the fall of the Soviet Union and 2002 for the rise of China. Instead, $\bar{T}_{\tilde{c}}$ is assumed to be 1994 for the the Soviet Union and 2012 (the last year in our sample) for China.

China). A positive coefficient β means that, for the average coercer, the first change is larger than the second. Our model predicts $\beta < 0$ for the collapse of the Soviet Union and $\beta > 0$ for the rise of China.

Table 1.5 reports the estimated coefficients. Finally, Table 1.6 performs the same analysis in the subsample of policymaker-coercer pairs where the United States is the coercer.

Discussion of the specification

Just as in our main specification, we included a very restrictive set of fixed effects (see Section A.3.1 for a detailed discussion). The inclusion of α_{ct} and α_{pt} captures the average effect on each country within each year of the fall of the Soviet Union and the rise of China (for instance, the overall extent of aid that a policymaker needs/that a coercer is willing to send in a given year). Consequently, this specification allows us to focus on heterogeneous effects. For instance, it enables us to study how the average policymaker changes its behavior (its offers) differently with different coercers *in the same year*. Also, how these events lead the average coercer to change its behavior (its aid) differently with different policymakers *in the same year*.

While we see Equation 1.4 as the sharpest test of our predictions, Section A.3.1 shows that our results are robust to alternative fixed effects specifications. Finally, as we discuss in Section A.3.3, our estimated effects do not depend on how we control for trade flows or the Relative Economic Power, suggesting that these variables do not contribute to our results.

Results

Table 1.5 confirms the predictions of our theory, finding a large and statistically significant differential effect. Where coercers were stronger than the Soviet Union, their relations worsened compared to where they were weaker; where they were stronger than China, their relations improved compared to where they were weaker.

Moreover, the estimates of Table 1.5 imply a much larger effect than the one we estimated in the previous section.[26] This finding is consistent with the idea that changes in the power of global geopolitical players such as the Soviet Union or China might have a more considerable effect on the pattern of international relations than changes in the power of the average country.

Finally, Table 1.6 shows that our model accurately describes the heterogeneous effect of the collapse of the Soviet Union or the rise of China on the relations between the United States and the rest of the world. Again, we find a very large effect for both events. The Fall of the Soviet Union (the rise of China) hurt (benefited) the United States in those regions of the world where it was stronger than the Soviet Union (China) compared to those regions of the world where it was weaker. Additionally, comparing Tables 1.5 and 1.6 we see that the United States was affected by these two events more than the average coercer.

Discussion of the empirical mechanisms with an example

The objective of this section is to discuss the mechanism behind our empirical results. To this end, we focus on a simple case study. Consider, for instance, how the fall of the Soviet Union affected the relations between the United States and any Latin American or Eastern European country. In line with our theory, the fall of the Soviet Union had a non-monotonic effect, depending on whether the Soviet Union was the strongest country (as in Eastern Europe) or whether the United States was (as in Latin America). In Eastern Europe, the US took advantage of the collapse of the previously dominant power by extending its network of alliances (Shifrinson, 2020).[27] In Latin America, on the other hand, the collapse of the Soviet Union generated a more relaxed geopolitical environment, which led to an overall

[26]The Fall of the Soviet Union had an impact of -0.014 on the average IE-index ($+0.02$ for the rise of China).

[27]In 1999, the Czech Republic, Hungary, and Poland joined NATO. Later, NATO added Bulgaria, Estonia, Latvia, Lithuania, Romania, Slovakia and Slovenia (2004), Albania and Croatia (2009), Montenegro (2017), and North Macedonia (2020).

reduction in the extent of alignment with the US (Levitsky and Roberts, 2011). Figure 1.1 shows the diverging political trajectories of Latin America and Eastern Europe. After the fall of the Soviet Union, we observe a drastic and permanent fall in the number of leftist governments in Eastern Europe. In contrast, in Latin America, we see a steady increase.

These patterns are the result of a variety of contributing factors and different mechanisms. For instance, the global powers were ready to make *direct interventions* to secure their interests, thus directly constraining domestic politics in Eastern Europe and Latin America.[28] But sometimes regime change was only *indirectly* linked to the actions of the United States or the Soviet Union. As an example, a coup d'état against a misaligned government is certainly more likely when the perpetrator anticipates the recognition or even support of one of the global powers.[29] Finally, there is a *deterrence mechanism*, which is active even when we observe no regime change. Indeed, politicians might avoid promoting or adopting specific policies that hurt the interests of the global powers if they want to remain in power.

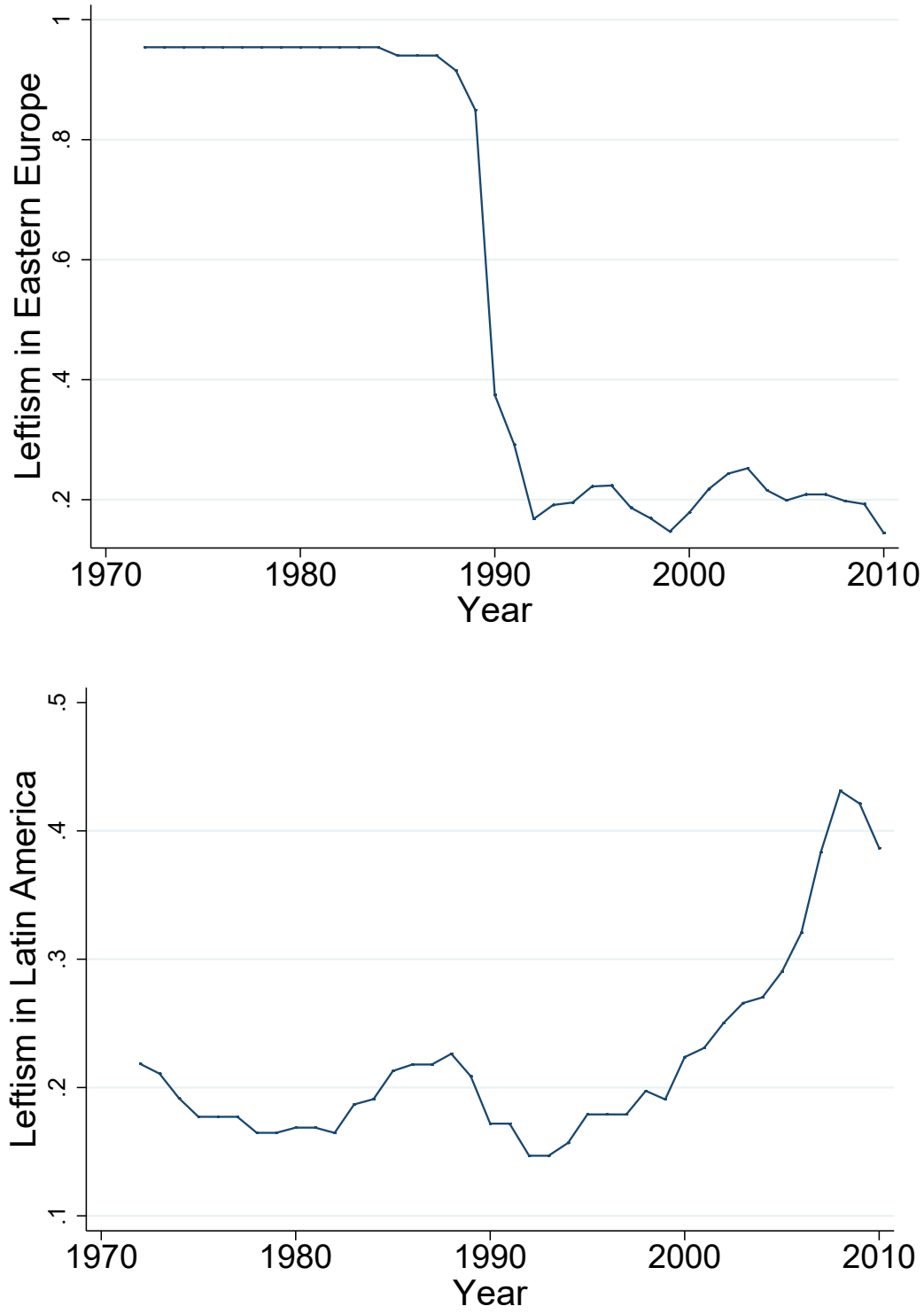
As we focused on standard equilibrium analysis, our theoretical analysis explored the deterrence effect, showing that a rational local ruler who wants to obtain the support of foreign powers at the lowest possible cost must satisfy a certain type of constraint. However, the crucial aspect of the theory is not the deterrence effect but rather how the constraint on policy-making depends on the distribution of power. Indeed, it would be easy to build an evolution-based model where the other mechanisms also play a role in keeping the equilibrium policies at the constraint identified by our theory.

The critical insight of the theory is that the constraints on domestic policy-making are tighter in periods of more intense geopolitical competition and that countries try to reduce

[28]Famous examples of direct interventions include the Soviet intervention in Czechoslovakia of 1968 or the US intervention in the Dominican Republic in 1965.

[29]During the Cold War, the United States and the Soviet Union quickly legitimized coups against governments that did not align with their interests. For instance, the United States recognized the 1973 coup d'état that ousted the Chilean president Salvador Allende in favor of general Augusto Pinochet. Similarly, the Soviet Union supported the 1978 coup in Afghanistan against its former ally Daoud Khan.

Figure 1.1: Diverging political paths for Eastern Europe and Latin America



Note: the two figures represent the time series plots of leftism in the average Latin American government (bottom) and Eastern European government (top). Leftism is variable *v2exl_legitideolcr_1* from the V-DEM dataset (accessible from `https://www.v-dem.net`). This variable determines the extent to which a country's government is using communism or socialism as its legitimizing ideology.

this pressure by aligning themselves with stronger powers. With this in mind, we see that even if the empirical mechanisms are more complex than the deterrence effect, our theory is still helpful to understand the patterns discussed above. For instance, our theory suggests that the rise of the left in Latin America was made possible by the more permissive geopolitical environment that followed the collapse of the Soviet Union. In such an environment, Latin American politicians could start running for office on an explicitly leftist platform without facing the risk of being deposed once in office. As a result, we observed a reduction in the alignment between the average Latin American government and the United States.

1.6 Additional Results: Effect of power on imports and exports

As our final empirical exercise, we study whether power affects *imports* and *exports* in a way that our theory can capture.

One possible mechanism is that each country (a policymaker) implements more favorable policies towards those countries (coercers) that have more power over them to keep their support. Another possibility is that more powerful countries can more effectively discourage the adoption of policies that would hurt their economic interests, including reducing their trade flows. In both cases, we should expect stronger countries to import and export more as they obtain more favorable trade policies and thus lower trade frictions.

We study $Flow_{d,t}^{c \to p}$, the flow of either imports or exports (depending on $d \in D$) from a coercer c to a policymaker p in period t. We estimate a standard gravity equation from Anderson and van Wincoop (2003). We assume that bilateral flows (in logs) are a function of the (log of) GDP, (the log of) Population, and the GDP per capita of coercer c and policymaker p. Moreover, we add coercer-policymaker pair fixed effects, year fixed effects,

and the lag of the bilateral flow.[30] We collect these variables in the vector $\mathbf{X}_{cpt}$. To assess the relevance of our model, we also add the (lagged) IE-index of c over p and the lagged relative economic power (REP) of c over p, i.e. $E_{t-1}^{c\to p}$.[31] Specifically, we consider the following:

$$ln(Flow_{d,t}^{c\to p}) = \beta_1 IE - index_{t-1}^{c\to p}(\mathbf{E}_{p,t-1}) + \beta_2 E_{t-1}^{c\to p} + \mathbf{\Gamma}'\mathbf{X}_{cpt} + \varepsilon_{cpt} \tag{1.5}$$

The problem with directly estimating this equation is the endogeneity. Indeed, our measure of power, economic power, is calculated from (past) trade flows. To address this, we apply a 2SLS estimation. We first identify variations in economic power that are not directly linked to trade policies. We then use such variations to assess the effect of power on imports and exports. The following section discusses this strategy in greater detail.

Empirical strategy: dealing with endogeneity

In the first stage, we use variations in the distribution of the coercers' GDP to predict changes in the distribution of the relative economic power of each coercer ($\mathbf{E}_{p,t-1}$). We let $\mathbf{GDP}_t$ denote the global distribution of GDP in period t, and let GDP_t^{tot} be the world GDP, i.e., the sum of every country's GDP.

Using a country's GDP as a measure of its economic power, we can define a country's relative economic power (REP) based on GDP as follows:

$$E_{t-1}^{c\to p}(\mathbf{GDP}_{t-1}) = \frac{GDP_{c,t-1}}{GDP_{t-1}^{tot}}$$

Based on this, we can define the GDP-based IE-index of c over p as the share of global GDP

[30] Introducing the lag of the outcome leads to Nickel bias, but as justified by Berger et al. (2013) the benefits of introducing the lagged outcome in this setting are likely to outweigh the costs. In our setting, Nickel bias will be less than -0.01, as $T \approx 70$, and the autoregressive component is close to 0.5.

[31] Variables that affect bilateral trade flows are usually introduced in the estimating equations in a log-linear fashion. See, for instance, Berger et al. (2013).

of all countries (except p) that have less GDP than c:

$$IE-index_{t-1}^{c\to p}(\mathbf{GDP}_{t-1}) = \sum_{\ell\neq p: GDP_{\ell,t-1}<GDP_{c,t-1}} E_{t-1}^{\ell\to p}(\mathbf{GDP}_{t-1})$$

Note that c's GDP is in the denominator of its GDP-based IE-index, so any change in its GDP that does not affect its GDP rank will decrease its GDP-based IE-index.

Given these definitions and Equation 1.5, we can express the two stages of our strategy:

First stages:

$$IE-index_{t-1}^{c\to p}(\mathbf{E}_{p,t-1}) = \theta_1 IE-index_{t-1}^{c\to p}(\mathbf{GDP}_{t-1}) + \theta_2 E_{t-1}^{c\to p}(\mathbf{GDP}_{t-1}) + \mathbf{\Gamma}_1'\mathbf{X}_{cpt} + \varepsilon_{cpt} \tag{1.6}$$

$$E_{t-1}^{c\to p} = \theta_3 IE-index_{t-1}^{c\to p}(\mathbf{GDP}_{t-1}) + \theta_4 E_{t-1}^{c\to p}(\mathbf{GDP}_{t-1}) + \mathbf{\Gamma}_2'\mathbf{X}_{cpt} + \varepsilon_{cpt} \tag{1.7}$$

Second stage:

$$ln(Flow_{d,t}^{c\to p}) = \beta_1 \widehat{IE-index}_{t-1}^{c\to p}(\mathbf{E}_{p,t-1}) + \beta_2 \widehat{E}_{t-1}^{c\to p} + \mathbf{\Gamma}'\mathbf{X}_{cpt} + \varepsilon_{cpt} \tag{1.8}$$

In the second stage (Equation 1.8), $\widehat{IE-index}_{t-1}^{c\to p}(\mathbf{E}_{p,t-1})$ and $\widehat{E}_{t-1}^{c\to p}$ are the fitted values from the two first stage regressions (Equations 1.6 and 1.7), and $\mathbf{X}_{cpt}$ is the vector including the standard gravity controls that we discussed earlier.

In the spirit of the analysis of Section 1.4.2, we will assess the relevance of our theory with three empirical exercises. First, we focus on the effect of the IE-index omitting the REP (i.e., fixing $\theta_2 = \beta_2 = 0$). Second, we focus on the effect of the relative economic power by omitting the IE-index (i.e., fixing $\theta_1 = \beta_1 = 0$). Finally, we carry out the entire analysis outlined above.

The crucial disadvantage of using GDP to measure economic power is its unilaterality: a coercer has the same power over every policymaker in the same period. Consequently,

we can only estimate β from variations within each coercer-policymaker pair and relative to that year's overall average of all coercer-policymaker pairs. Fortunately, however, we can rely on a well-established literature to guide our choice of control variables.

Results

Table 1.7 summarizes the results of the strategy described in Section 1.6. Columns (1) and (3) study the effect of the Indirect Exposure Index (IE-index) on imports and exports without controlling for the relative economic power (REP), columns (2) and (4) study the effect of the REP without controlling for the IE-index, and, finally, columns (3) and (5) estimate the full model described by Equations 1.7, 1.6, and 1.8.

Comparing column (2) with (3) and column (5) with (6), we find again that a country's relative economic power does not capture the effects of power on bilateral relations. Instead, we find that the Indirect Exposure Index has a large and significant impact on both imports and exports. A standard deviation increase in the lagged Indirect Exposure Index (a change of 0.06) leads to a 20% increase in imports and exports. When we also control for the coercer's REP, this number becomes larger than 30%. Additionally, note that the process for imports and exports features an autoregressive component whose coefficient is around 0.5, implying that the long-run effect of a one standard deviation increase in the IE-index is at least twice as large as the ones stated.

1.7 Robustness: definitions, specifications, inference

In Appendix A.3 we perform several robustness checks to confirm that the Indirect Exposure Index is an important empirical determinant of international interactions and trade. This section summarizes some of these exercises.

Outcome definitions. As discussed in Appendix A.2, before analyzing our outcomes

derived from GDELT, we apply a transformation to avoid dropping the observed zeros when taking the log. In Section A.3.4 we show that our results do not change when we avoid any transformation or when we make other standard transformations.

Fixed effects. Section A.3.1 shows that all the results are robust to every fixed effect specification that is less stringent than our baseline specification, but also a specification that is more stringent than our baseline. Panel D in Table A.3.3 shows that the results are robust to the inclusion of time-varying fixed effect for each pair of geopolitical regions.[32] As we discuss in Section A.3.1, this exercise shows that our results do not depend on the formation or the dissolution of supranational institutions (e.g., the European Union or the Soviet Union).

Independent Variables and Additional Controls. Section A.3.3 studies relatively minor deviations from our baseline specification. It shows that our results do not change when we omit trade flows, condition on the same-year REP, or study the same-year rather than lagged IE-index. Moreover, Section A.3.3 shows that results are robust to the inclusion of the lagged outcome.[33] Finally, Section A.3.6 shows that the results do not change when we study the number of offers received conditioning on the number of offers sent, aid sent, or both. Similarly, results do not change when we study aid sent, conditioning on the number of offers sent, received, or both. These last findings justify our focus on distinct outcomes as cross-validation of the theory.

Standard Errors. Section A.3.5 shows that the statistical significance of our estimates is robust to alternative standard error specifications. In our baseline analysis, we use two-way clustered standard errors (coercer and policymaker). This choice does not account for the possibility of dependence across country pairs of the errors, which is likely if our model

[32]The netting out of all time-varying characteristics of a coercer-policymaker pair (say c and p) that are the same across other coercer-policymaker pairs (say c' and p'), where coercers c and c' are in the same geopolitical region S, and policymakers p and p' are also in the same geopolitical region S'.

[33]Results are robust even when allowing outcome lags to have a time-varying effect, a coercer-specific effect, or a policymaker-specific effect.

is correct.[34] Tables A.3.7 and Table A.3.8 show that our inference is robust to accounting for dependence across country pairs, as we show our coefficients to be significant even when we adopt three-way clustered standard errors (adding the time dimension), or Driscoll-Kraay standard errors (Driscoll and Kraay, 1998).[35]

1.8 Validation: power, outcomes, subsamples

Appendix A.4 validates our results by replicating the analysis with an alternative measure of power, other outcomes from GDELT and restricting the analysis to some subsamples.

Alternative measure of power. To make our baseline analysis as transparent and straightforward as possible, we focused on power generated by the value of trade flows. Thus, our analysis ignores military power and the possibility that different trade compositions (e.g. exporting raw materials rather than luxury goods) with the same value can generate different power levels. To address these limitations, in a validation exercise, we replicate our analysis with an alternative measure of power: the Formal Bilateral Influence Capability (FBIC).[36] Such variable is a composite measure of power, based on arms transfers, aid dependence, membership in intergovernmental organizations, and other similar features (Moyer et al., 2018). Interestingly, the FBIC has a relatively low correlation with our measure of economic power (see Tables A.4.1 and A.4.2). Still, we successfully replicate all of our results using the FBIC, finding similar (if a bit larger) counterfactual exercises.

Alternative outcomes from GDELT. In our main analysis, we focused on offers received and aid sent to measure the value that a coercer obtains from its relation with a

[34] Recall that the key testable prediction of our model is that bilateral outcomes depend on third-parties characteristics.

[35] As discussed in Section A.3.5, unlike three-way clustering (by coercer, policymaker, and year), Driscoll-Kraay standard errors allow for a very general form of dependence across country pairs *and* auto-correlation in such component. In our robustness checks, we allow for up to seven lags of dependence across country pairs.

[36] This measure is freely accessible from `https://korbel.du.edu/fbic`.

policymaker. In Section A.4.3, we show that we can replicate our analysis with another type of outcome: *agreements*. As we discuss in Section A.4.3, agreements measure the initiation, resumption, improvement, or expansion of cooperation in a particular domain (economic, military, diplomatic, and judiciary), the ratification, signature, and finalization of an agreement or treaty, regardless of its nature (political agreement). Tables A.4.6, A.4.7, and A.4.8 replicate our analysis for these outcomes, providing further supporting evidence to the role of the IE-index in shaping the pattern of international cooperation.

Subsamples. Our final validation exercise replicates our analyses in different subsamples of interest. In line with our interpretation of the model describing how foreign powers affect domestic decision-making, we focus on subsamples where only foreign powers can be coercers, and they cannot be policymakers. Section A.4.3 shows that our results hold for various definitions of foreign powers. In the most restrictive definition of foreign powers, we consider only 8 coercers and less than 15% of our entire sample. In the least restrictive case, we study 40 coercers (and 30% of our sample).[37] Finally, Section A.4.3 replicates the analysis focusing on asymmetric relations: country pairs where the coercer has more power over the policymaker than vice versa.

1.9 Future directions and model's extensions

One of the main objectives of this paper is to provide an empirically-grounded foundation to study various applications related to geopolitical competition. We now discuss some of these applications highlighting the unique perspective offered by our model.

[37] We define as foreign powers those countries that are ever in the top 5, 10, 20, or 30 in terms of GDP. See Table A.4.9 for a list of all countries that qualify as coercers and the sample size of each subsample.

1.9.1 Great Power Wars and arms' races

The first natural extension of our model would be to endogenize the distribution of power. For instance, we could study foreign powers (coercers) who invest in their coercive capacities or can spend resources to harm the coercive capacities of other foreign powers (e.g., with a conflict).

Our model provides a unique perspective on this theme because the payoff of a country jumps up (down) when its power rank increases (decreases). As a first implication, note that arms race are very likely occurrences in our model: the incentive to invest in power crucially depends on whether other countries invest in power, as countries compete to obtain higher power ranks. Second, within our model, countries might be willing to invest resources into weakening their competitors, e.g., participating in costly conflicts that can only destroy the opponent's resources.

In addition, note that our model would make predictions on the conflict pattern that are very different from the standard wisdom that conflicts occur when strong countries find it profitable to prey on small ones. First, within our model, the source of tension is asymmetric but in the opposite direction: conflicts emerge from the desire of a country to overcome a stronger one. Thus, we should expect countries to display aggression towards stronger ones, a puzzling pattern that has spurred an entire literature in international relations (Paul, 1994). Second, foreign powers should be more likely to fight when they are more symmetric, as in that case, a small change in the distribution of power (a less costly conflict) can lead to a change in rank, thus a large change in the equilibrium payoff. Again, the idea that power parity may be conducive to war is at odds with the standard theory but not with the evidence on both historical and modern conflicts (Organski and Kugler, 2015; Lemke, 2002; Allison, 2017; Renshon, 2017a).

1.9.2 Geopolitical polarization and strategic protectionism

Another natural way to endogenize the distribution of power is to take the perspective of a local power (a policymaker) who can change its relative dependence on the various foreign powers (e.g., with its trade policies).

In this case, we can show that when the local ruler has the opportunity to redirect its trade away from one foreign power and towards another one, it would always want to increase its relative dependence on the more powerful of the two. In other words, our model introduces a force that pushes local rulers to prefer polarization to diversification. Such finding is consistent with episodes of geopolitical polarization such as the Cold War: the formation of separate blocs that barely interact with one another.

Finally, this extension would highlight that countries have an incentive to reduce their overall exposure towards foreign powers (e.g., reduce trade volumes), speaking to the existence of *strategic industries* that receive protection from foreign competition.

1.9.3 Ideological competition

In our model, we take a reduced-form perspective assuming that local rulers send transfers to foreign powers. But these relationships are more complicated in the real world. For instance, there might be a higher political cost in aligning with a foreign power relative to another one, depending on whether such foreign power is viewed with favor or disfavor by the local population. Thus, in an extension, it would be natural to assume heterogeneity in this type of cost and to assume that foreign powers can invest in reducing it (e.g., with ideological and cultural propaganda).

Importantly, this type of heterogeneity should reflect in the equilibrium transfers. Indeed, recall from Section 1.3.4 that the local ruler is indifferent on which foreign power to favor when they have the same power. Adding cost heterogeneity would break the indifference and

lead the local ruler to align itself with the least costly of the two powers. But this implies an incentive for foreign powers to invest in this dimension, which can explain why ideological competition is another important aspect of geopolitical rivalries.

1.9.4 International institutions

While we expanded the model to include multiple periods, our analysis remained effectively static, as we assumed that countries are not patient enough to play the typical cooperative strategies that might emerge in repeated games. Relaxing this assumption would open the door to various complexities. However, it would also allow us to study institutional arrangements that might reflect this type of strategies. For instance, in our model, two foreign powers could increase their payoff by credibly committing to act together when pressuring local rulers. This idea adds a geopolitical explanation as to why countries might be willing to join their forces by creating supranational organizations that can coordinate their behavior.

Moreover, suppose countries anticipate that the natural state of international relations generates payoffs that depend on the power rank (as in Section 1.3). Additionally, suppose that they expect that this aspect fuels infighting (as discussed in Section 1.9.1). In that case, they might develop institutions to change the game's rules, change the payoff structure, and thus reduce the tensions between them. In our model, the dependence on the power rank emerges from the desire of local rulers to policymaker their policies to foreign powers. Then, it is natural to expect that a more rule-based international system, a strengthening of non-discriminatory clauses on trade, and an increased focus on multilateral (rather than bilateral) approaches are all steps towards a more peaceful future.

1.10 Conclusions

This paper studies how the distribution of power shapes the international system. Our main theoretical result is that the equilibrium payoff of each foreign power is not determined by its relative power, but by the power share of all weaker foreign powers: a variable we call the *Indirect Exposure Index (IE-index)*. Specifically, greater IE-index implies higher payoff. As we discuss, this simple prediction produces rich enough dynamics to resolve seemingly unrelated international relations puzzles. Importantly, the IE-index, a one-dimensional function of the distribution of power, summarizes all these dynamics, producing straightforward empirical predictions.

Specializing the theory, we predict that when a country's IE-index increases, it should have better economic, military, and diplomatic relations, send more economic, military, and humanitarian aid, and export and import more. We confirm the empirical validity of each of these predictions by testing them with bilateral data on international interactions and trade. Additionally, we show that the IE-index plays a much more fundamental role than a country's own power. Indeed, controlling for a country's IE-index, the estimated effect of a country's power becomes insignificant or even negative. Moreover, consistently with the theory, we show that the fall of the Soviet Union hurt the United States (and the average foreign power) where it was weaker than the USSR relative to where it was stronger. Additionally, we show that the United States (and the average foreign power) has been benefiting from the rise of China where it is stronger than China relative to where it is weaker. All in all, this empirical analysis strongly supports the theory and the role of the IE-index in shaping modern international relations.

Finally, we show that our theoretical framework can be extended in various directions, promising new perspectives on a variety of themes: from great power rivalries, conflicts, geopolitical polarization, the role of ideology in great power competition, and how changes

in the institutional setting feed into each of these dimensions.

1.11 Tables

Table 1.1: description of outcomes from the GDELT dataset

Variable	Description of the interaction [CAMEO category]
Economic Offer	Offer/promise/agree to, or otherwise indicate willingness or commitment to engage in or expand economic ties [0311]. Offer/promise/agree to, or otherwise indicate willingness or commitment to engage in or expand material cooperative exchange not otherwise specified [031].
Military Offer	Offer/promise/agree to, or otherwise indicate willingness or commitment to engage in or expand military ties [0312].
Diplomatic Offer	Offer/promise/agree to, or otherwise indicate willingness or commitment to expand diplomatic ties or cooperation [032].
Economic Aid	Extend/provide monetary aid and financial guarantees, grants, gifts and credit [071]. All provisions, extension of material aid, not otherwise specified [070].
Military Aid	Extend/provide military and police assistance including arms and personnel [072].
Humanitarian Aid	Extend/provide humanitarian aid, mainly in the form of emergency assistance [073].

Table 1.2: effect of the Indirect Exposure Index (IE-index) on International Interactions

	$Offers^{p \to c}$			$Aid^{c \to p}$		
	Econ	Mil	Dip	Econ	Mil	Hum
$IE - index^{c \to p}$ (lag)	1.90***	1.11***	1.02***	1.50***	0.71***	1.15***
	[0.27]	[0.24]	[0.21]	[0.23]	[0.19]	[0.21]
Trade (ln)	0.01	0.00	0.00	-0.00	0.00	-0.00
	[0.01]	[0.01]	[0.003]	[0.004]	[0.004]	[0.005]
Observations	281,190	281,298	253,732	281,305	281,305	281,305
Adj. R-2	0.47	0.27	0.16	0.60	0.43	0.36
Within R-2	0.005	0.001	0.001	0.004	0.001	0.002

Note: the unit of observation is a coercer-policymaker (c, p) pair in a given year. The Indirect Exposure Index (IE-index) of c over p is the share of total (economic) power over p controlled by every coercer with less power over p than c. The outcomes are the log of the number of times over a year t in which country p sends Economic, Military, or Diplomatic Offers to country c, or in which country c sends Economic, Military, or Humanitarian aid to country p. All specifications control for the log of the trade flow between p and c and include coercer-policymaker pair fixed effects, as well as time-varying fixed effects for each coercer and for each policymaker. Standard errors (in parenthesis) are two-way clustered (c, p). Stars represent statistical significance of the single hypothesis test: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 1.3: how the inclusion of a coercer's Indirect Exposure Index (IE-index) changes the estimated effect of its Relative Economic Power (REP)

Table 1.3.A — $Offers^{p \to c}$

	Economic		Military		Diplomatic	
	(1)	(2)	(3)	(4)	(5)	(6)
$IE - index^{c \to p}$ (lag)		2.16***		1.18***		1.05***
		[0.32]		[0.25]		[0.25]
$REP^{c \to p}$ (lag)	1.91***	-1.44***	1.47***	-0.36	1.48***	-0.17
	[0.70]	[0.54]	[0.40]	[0.50]	[0.39]	[0.43]
Trade (ln)	0.02***	0.01	0.01	0.00	0.01**	0.00
	[0.005]	[0.01]	[0.01]	[0.01]	[0.004]	[0.003]
Observations	281,190	281,190	281,298	281,298	253,732	253,732
Adj. R-2	0.46	0.47	0.27	0.27	0.16	0.16
Within R-2	0.001	0.005	0.003	0.001	0.000	0.001

Table 1.3.B — $Aid^{c \to p}$

	Economic		Military		Humanitarian	
	(1)	(2)	(3)	(4)	(5)	(6)
$IE - index^{c \to p}$ (lag)		1.70***		0.81***		1.25***
		[0.26]		[0.27]		[0.27]
$REP^{c \to p}$ (lag)	1.53**	-1.10*	0.74**	-0.51	1.44**	-0.49
	[0.68]	[0.56]	[0.36]	[0.55]	[0.61]	[0.64]
Trade (ln)	0.01**	-0.00	0.01*	0.003	0.01	-0.001
	[0.004]	[0.004]	[0.004]	[0.004]	[0.004]	[0.005]
Observations	281,305	281,305	281,305	281,305	281,305	281,305
Adj. R-2	0.58	0.58	0.43	0.43	0.36	0.36
Within R-2	0.001	0.004	0.0002	0.001	0.0003	0.002

Note: the unit of observation is a coercer-policymaker (c, p) pair in a given year. The Indirect Exposure Index (IE-index) of c over p is the share of total economic power over p of all coercers with less power over p than c. The Relative Economic Power (REP) of c over p is c's share of total economic power over p. Outcomes are the log of the number of times in year t in which country p sent Economic, Military, or Diplomatic Offers to c (Table 1.3.A), or in which country c sent Economic, Military, or Humanitarian aid to p (Table 1.3.B). All specifications control for the log of trade between p and c, policymaker-coercer pair fixed effects, and time-varying fixed effects for each coercer and policymaker. Standard errors (in parenthesis) are two-way clustered (c, p). Stars represent statistical significance of the single hypothesis test: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 1.4: Indirect Exposure Index (IE-index) vs. Relative Economic Power (REP) including power Rank fixed effects or in subsamples with no change in rank

Table 1.4.A $Offers^{p \to c}$

	Economic		Military		Diplomatic	
	(1)	(2)	(3)	(4)	(5)	(6)
$IE-index^{c \to p}$ (lag)	2.75***	2.62***	1.47***	1.64***	1.48***	2.60***
	[0.50]	[0.53]	[0.35]	[0.49]	[0.33]	[0.70]
$REP^{c \to p}$ (lag)	-1.39***	0.18	-0.39	-0.18	-0.24	0.31
	[0.50]	[0.72]	[0.50]	[0.86]	[0.47]	[0.87]
Trade flow (ln)	0.00	-0.01	0.00	0.01	0.00	0.01
	[0.004]	[0.02]	[0.004]	[0.02]	[0.003]	[0.02]
Rank fixed effects	Yes	No	Yes	No	Yes	No
Rank = Modal Rank	No	Yes	No	Yes	No	Yes
Observations	281,179	42,164	281,287	42,189	253,716	35,549
Adj. R-2	0.47	0.53	0.27	0.28	0.16	0.10
Within R-2	0.004	0.003	0.001	0.001	0.001	0.002

Table 1.4.B $Aid^{c \to p}$

	Economic		Military		Humanitarian	
	(1)	(2)	(3)	(4)	(5)	(6)
$IE-index^{c \to p}$ (lag)	2.15***	2.18***	0.82**	0.93*	1.56***	1.81***
	[0.40]	[0.49]	[0.37]	[0.51]	[0.42]	[0.57]
$REP^{c \to p}$ (lag)	-0.82*	0.41	-0.57	0.04	-0.50	-0.05
	[0.48]	[0.50]	[0.51]	[0.96]	[0.49]	[0.83]
Trade flow (ln)	-0.00	-0.01	0.00	0.01	-0.00	-0.01
	[0.003]	[0.01]	[0.003]	[0.01]	[0.004]	[0.02]
Rank fixed effects	Yes	No	Yes	No	Yes	No
Rank = Modal Rank	No	Yes	No	Yes	No	Yes
Observations	281,292	42,228	281,292	42,228	281,292	42,228
Adj. R-2	0.59	0.65	0.43	0.44	0.36	0.38
Within R-2	0.003	0.004	0.0004	0.0003	0.001	0.001

Note: the unit of observation is a coercer-policymaker (c, p) pair in a given year. Odd columns include economic power rank fixed effect; even columns restrict the samples to cases where the economic power rank of c over p equals its mode. The IE-index of c over p is the share of total economic power over p of all coercers with less power over p than c. The REP (Relative Economic Power) of c over p is c's share of total economic power over p. Outcomes are the log of the number of times in year t in which p sent Economic, Military, or Diplomatic Offers to c (Table 1.4.A), or in which c sent Economic, Military, or Humanitarian aid to p (Table 1.4.B). All specifications control for the log of trade between p and c, (c, p) fixed effects, and time-varying fixed effects for each coercer and policymaker. Standard errors (in parenthesis) are two-way clustered (c, p). Stars represent statistical significance of the single hyp. test: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 1.5: differential impact of USSR collapse (China rise) on the average coercer comparing *where* it is stronger than USSR (China) with *where* it is weaker

	$Offers^{p \to c}$			$Aid^{c \to p}$		
	Econ	Mil	Dip	Econ	Mil	Hum
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: fall of USSR						
Stronger vs Weaker	-0.31***	-0.20***	-0.18***	-0.34***	-0.16***	-0.28***
	[0.04]	[0.07]	[0.03]	[0.04]	[0.05]	[0.05]
Observations	287,895	287,999	261,003	287,904	287,904	287,904
Adj. R-2	0.45	0.26	0.15	0.58	0.42	0.35
Within R-2	0.002	0.001	0.001	0.003	0.001	0.001
Panel B: rise of China						
Stronger vs Weaker	0.29***	0.20***	0.14***	0.32***	0.14***	0.27***
	[0.05]	[0.06]	[0.03]	[0.04]	[0.03]	[0.05]
Observations	288,829	288,936	261,594	288,824	288,824	288,824
Adj. R-2	0.45	0.27	0.15	0.58	0.43	0.35
Within R-2	0.004	0.001	0.001	0.005	0.001	0.003

Note: the unit of observation is a coercer-policymaker (c, p) pair in a given year. The variable Stronger vs Weaker estimates β of Equation 1.4. This coefficient captures the differential effect on each coercer c of the fall of the Soviet Union (Panel A) or of the rise of China (Panel B) comparing the outcomes related to policymakers over which c was stronger than the USSR (or China) with those policymaker over which it was weaker than the USSR (or China). Outcomes are the log of the number of times in year t in which country p sent Economic, Military, or Diplomatic Offers to c (Table 1.4.A), or in which country c sent Economic, Military, or Humanitarian aid to p (Table 1.4.B). All specifications control for the log of trade between p and c, the Relative Economic Power (REP) of c over p, i.e. c's share of total economic power over p, policymaker-coercer pair fixed effects, and time-varying fixed effects for each coercer and policymaker. Standard errors (in parenthesis) are two-way clustered (c, p). Stars represent statistical significance of the single hypothesis test: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 1.6: differential impact of USSR collapse (China rise) on the United States comparing *where* the USA is stronger than USSR (China) with *where* it is weaker

	$Offers^{h \to USA}$			$Aid^{USA \to h}$		
	Econ	Mil	Dip	Econ	Mil	Hum
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: fall of USSR						
Stronger vs Weaker	-0.42***	-0.80***	-0.14	-0.33***	-0.38***	-0.58***
(USA only)	[0.07]	[0.10]	[0.10]	[0.06]	[0.10]	[0.09]
Observations	5,256	5,266	3,499	5,281	5,281	5,281
Adj. R-2	0.61	0.43	0.24	0.76	0.50	0.50
Within R-2	0.01	0.02	0.002	0.02	0.005	0.01
Panel B: rise of China						
Stronger vs Weaker	0.86***	1.19***	0.25	0.79***	0.26*	0.76***
(USA only)	[0.11]	[0.19]	[0.17]	[0.09]	[0.14]	[0.14]
Observations	5,265	5,275	3,503	5,290	5,290	5,290
Adj. R-2	0.61	0.42	0.24	0.76	0.50	0.49
Within R-2	0.02	0.01	0.004	0.02	0.003	0.01

Note: the unit of observation is a US-policymaker (US, p) pair in a given year. The variable Stronger vs Weaker estimates β of Equation 1.4 restricting $c = US$. This coefficient captures the differential effect on the US of the fall of the Soviet Union (Panel A) or of the rise of China (Panel B) comparing the outcomes related to those policymakers over which the US was stronger than the USSR (or China) with those where it was weaker than the USSR (or China). Outcomes are the log of the number of times in year t in which country p sent Economic, Military, or Diplomatic Offers to the US (Table 1.4.A), or in which the US sent Economic, Military, or Humanitarian aid to p (Table 1.4.B). All specifications control for the log of trade between p and the US, the US's Relative Economic Power (REP) over p, i.e. the share of total economic power over p controlled by the US, policymaker-US fixed effects, and year fixed effects. Standard errors (in parenthesis) are robust. Stars represent statistical significance of the single hypothesis test: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 1.7: comparison between effects of Indirect Exposure Index (IE-index) and Relative Economic Power (REP) on imports and exports in Gravity Model

	c's Imports from p (ln)			c's Exports to p (ln)		
	(1)	(2)	(3)	(4)	(5)	(6)
$\widehat{IE-index}^{c\to p}$ (lag)	3.33*		5.06**	3.26**		4.43***
	[1.86]		[2.50]	[1.44]		[1.66]
$\widehat{REP}^{c\to p}$ (lag)		-0.981	-11.45		1.07	-7.93
		[9.35]	[9.79]		[8.22]	[4.92]
Observations	813,093	813,093	813,093	813,091	813,091	813,091
Adj. R-2	0.32	0.32	0.32	0.36	0.35	0.35
K-Paap F	9.08	9.53	49.12	9.06	9.84	52.74

Note: the unit of observation is a coercer-policymaker (c, p) pair in a given year. The table reports 2SLS estimates, where the second stage uses (lagged) trade flows to measure power and the first stages uses the distribution of GDP to measure power (see Equations 1.7, 1.6). The Indirect Exposure Index (IE-index) of country c (coercer) over country p (policymaker) is the share of total economic power over p of all coercers with less power over p than c. The Relative Economic Power (REP) of c over p is c's share of total economic power over p. All specifications control for coercer-policymaker fixed effects, year fixed effects, the lag of the outcome, the GDP (log), the population (log), and the GDP per capita of both c and p. Standard errors (in parenthesis) are two-way clustered (c, p). Stars represent statistical significance of the single hypothesis test: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Chapter 2

Political Favoritism and Social Unrest

2.1 Introduction

Across human history, the creation or consolidation of central authorities has often reinforced inequality and induced social stratification (Boix, 2015). A fundamental driver of this process is *political favoritism*, the practice of organizing social groups into ranks and granting privileges based on relative status. From the Inka to the Chinese Empire, from Medieval Europe to the Swazi society in South Africa, central authorities have typically resorted to political favoritism, assigning higher status and thus higher political rents to their richer and more powerful constituents. Despite the relative standing of social groups is much less transparent today, in many areas of the world rulers are still known to systematically favor some social or ethnic groups at the expense of others, with targeted public spending and taxation.

A growing literature studies the political economy of societies with no monopoly over violence, highlighting how the ruler's ability to concede political rents is crucial in maintaining social stability and keeping the ruler in power (North et al., 2009; Bueno De Mesquita et al., 2005; Acemoglu et al., 2004; Francois et al., 2015). However, little is known about

the mechanisms behind political favoritism. Addressing such gap is important not only because political favoritism is universally observed, but it is also a very consequential practice. Indeed, political favoritism might be an important determinant of inequality and social stratification. Additionally, it is often observed to be linked to social unrest. Indeed, many historical accounts see the resentment of the lower ranks of society as a driver of rebellions and, even today, discrimination and political ambitions are observed to fuel internal conflicts. We contribute to this literature arguing that political favoritism naturally emerges in societies with no monopoly on violence and deriving its consequence for the dynamics of inequality and the pattern of social unrest.

Our main result is that political favoritism arises in the shadow of violence, as the ruler's response to the subjects' ability to join forces against its policy. Indeed, in the (unique) equilibrium outcome of our game, the government undermines the subjects' ability to coordinate against taxation by assigning ranks (*status*) to each social group and asking lower taxes to higher ranked groups. In particular, the equilibrium tax on each group is sufficiently low to dissuade it not only from resisting alone but also from joining any resisting coalition that contains only lower-ranked subjects. Additionally, we show that political favoritism should be expected to increase inequality. First, we show that even identical groups should be treated differently by the central authority. Second, we show that when subjects are ex ante heterogeneous, then favoritism should be expected to be highly regressive, thus reinforcing inequality. Indeed, we show that the ruler should optimally assign higher status to richer (thus stronger) social groups. Finally, we study how political favoritism interacts with the rise of new identity groups or the creation of a representative institution. We think of these episodes as creating solidarity among previously distinct social groups, i.e., with some probability such groups act as one. We show that these events should be expected to trigger social unrest: the ruler has an incentive to treat these groups as separate, effectively testing their solidarity. Finally, we derive a clear link between political favoritism and the pattern of

social unrest: when social unrest is observed, we should find that it is fueled by lower-ranks of society.

Our model considers a society fragmented in different social groups, the decision units of our framework. Depending on the setting, we can think of such groups as ethnic or identity groups, warlords with their armed sequitur, clans, geopolitical sub-units within empires, kingdoms or principalities, or even organized interest groups such as guilds or corporations within the same geopolitical unit.[1] Each group i is heterogeneous with respect to two characteristics: economic resources (B_i), and military resources (W_i). While only economic resources directly enter the utility functions, military resources are still payoff relevant as they determine the probability of winning any dispute that may arise. The allocation of resources is common knowledge and possibly heterogeneous across different social groups.

A group (the government) publicly demands a transfer vector (taxes) from the other groups (the subjects) who sequentially choose whether to resist or comply. Resisting subjects obtain to withhold their portion of taxes in case the resistance succeeds, but lose all their resources in case the resistance fails. As we assume the success probability of resistance to be increasing in the sum of the military power of all resisting groups, subjects anticipate that their choice to resist may have a *cascade effect*: whenever a subject chooses to resist, it also encourages others to do so. Because of these cascade effects and the sequentiality of subjects' responses, the ruler always faces the largest resisting coalition consistent with its tax demands. Solving for the optimal government policy (tax vector) in a setting where subjects can jointly oppose taxation and there are cascade effects is our key innovation to

[1] We deal with potential membership in multiple groups by assuming that the central authority can target every intersection of groups: not only it can target Italians or merchants, but also the group of all Italian merchants. Limits on the government's ability to discriminate are studied in Section **??**. In Section **??** we allow for some degree of uncertainty in the composition of society, as we assume that with some probability the central authority expects two or more groups to act together as if they were a single decision unit.

the literature on policy-making in the shadow of violence.[2]

As a first step of the analysis we show that, since the government can perfectly target tax demands to individual groups, the optimal taxation must avoid any wasteful resistance on path.[3] As a result, in equilibrium, taxes must not only discourage every subject from resisting alone but also prevent the formation of resistance cascades, where multiple subjects sequentially join forces against taxation.[4] One way to avoid resistance would be to demand a tax so low that no subject would resist even when anticipating that all the other subjects would join. However, the government is able to extract more resources by adopting a divide-and-conquer strategy. In particular, the government can demand such a low tax only from one group, its favorite, and impose higher taxes on all the others, exploiting the fact that they anticipate that the favorite group would never join their resistance. Similarly, the government can single out another subject, its second-favorite, demanding a tax low enough that such group would not join a resistance formed by all the other groups but the favorite. Iterating such argument, we show that the optimal way to undermine coordination is to institute a ranking of all subjects (favorite, second-favorite, ..), and demand from each subject the maximum tax that does not induce it to start resisting, knowing that all lower-ranked but no higher-ranked subjects would join.

Such divide-and-conquer logic is at the heart of our first main result, establishing that the optimal government policy is characterized by *political favoritism*: the tax burden on each group does not only depend on its resources but also on its position (*status*) in a ranking endogenously determined by the ruler. To pin down how the optimal ranking relates to the distribution of resources, we specialize the model by assuming that the probability of successful resistance is given by the proportion of the total military power controlled by

[2]We discuss how our taxation result relates to the literature in both the next Section and Section 2.2.3. In Section 3.8 we discuss the possibility of other applications beyond the study of taxation.

[3]Note that this result is not trivial as subjects impose externalities on each other.

[4]We show that the order in which subjects are called to choose whether or not to resist does not matter for the equilibrium outcome.

resisting groups, $\frac{W_{resist}}{W_{tot}}$.[5] Under this specification, we show that the government's optimal ranking of subjects is uniquely determined by the militarization index $\frac{W}{B}$, with higher index corresponding to higher status.[6] As a corollary, once we allow for a natural positive dependence between military and economic resources, we show that taxation is regressive: richer (hence stronger) subjects are optimally granted higher status and thus a preferential taxation treatment.

The divide-and-conquer nature of political favoritism suggests that a key assumption of our analysis is that there are clearly defined social groups who are unitary but separate decision units. In many historical instances, however, the structure of society has undergone changes: the rise of a new cross-group identity or the creation of a representative institution have allowed previously separate social groups to act as one. To capture the potential fragility of such events and keep the analysis simple, we characterize how government policy responds to the possibility that some social groups are *in solidarity*, i.e. they militarily support each other, with some commonly known probability $p \in [0, 1]$.

We first show that when the subjects' solidarity is sufficiently likely, i.e. when p is sufficiently high, the government optimally abandons its original divide-and-conquer strategy, treating the involved subjects as a single bargaining block backed by the threat of their collective action. As a result, the members of such partnerships receive a better taxation treatment at the expense of the government and the excluded subjects.[7] On the other hand, if the probability of solidarity is sufficiently low, then the ruler's optimal strategy would be to treat every group as separate (employing political favoritism), and be ready to deal with

[5]Note that this specification, thoroughly discussed in Section 2.2.4, captures the idea that the success probability of a resistance is negatively affected by non-resisting subjects.

[6]Independently of the particular specification, whenever the probability of success of a resistance is increasing in the military power controlled by the resisting groups, we obtain a partial order: for equal economic (military) resources, groups with more military (economic) resources should always be ranked higher (lower).

[7]Such "unity makes strength" finding is surprisingly absent from the literature. While historically representative institutions are associated with higher taxes, note that taxes are just an interpretation: our actual focus is on the net effect on groups' payoffs of all government policies.

resistance whenever subjects actually turn out to be in solidarity.

Finally, we show that whenever social unrest is observed (for instance, because of changes in the composition of society) then we should expect a link between political favoritism and participation in social unrest. Specifically, if one group ends up resisting, then all lower-ranked subjects should join. For instance, this implies that when subjects turn out to be in solidarity and resist, then all subjects who pay a higher tax rate should also participate in their resistance.

From an applied perspective, our findings can justify the universal attention for relative status and rankings emerging from the historical evidence, and why government policy has typically been extremely regressive, playing a prominent role in reinforcing social inequality and even *inducing* social stratification. Moreover, as we document in our historical application, the divide-and-conquer logic at the heart of our model was well understood by rulers and politicians in Medieval and Early Modern Europe. For example, we show how Louis XI avoided to clash against a coalition of its magnates by resorting to the type of divide-and-conquer strategy that characterizes our equilibrium.

Turning our attention to the pattern of social unrest, we show that, as predicted, in both modern and historical societies the poorest and most discriminated groups (the lower ranks) are disproportionately more likely to participate in episodes of resistance against the government. Additionally, we argue that the analysis of how solidarity affects political favoritism might shed light on why the emergence of new identity groups has historically led to widespread political instability. For instance, the Protestant Reformation in XVI Century Europe is widely credited for the endemic conflicts and revolts of that century. Similarly, British officers attributed the Indian Mutiny of the XIX Century to an insufficient attention on part of the British in avoiding the development of nationalistic feelings between the various ethnic and social groups residing in the Indian subcontinent. Finally, the analysis of how solidarity affects government policy highlights the strong incentives of the central authority

to fight any form of inter-group solidarity. This finding might explain the various episodes of repression against new identity groups and/or the segregationist policies that have been employed universally, from Ancient Rome to China, to the European colonial empires.

2.1.1 Contribution to the Literature

Our paper speaks to the large body of literature that studies societies where the only effective constraint on the central authority is the subjects' violence potential. A review of this growing body of literature can be found in Gehlbach et al. (2016). Even if quite a few contributions have specifically addressed the distribution of political rents (Bueno De Mesquita et al., 2005; Acemoglu et al., 2004; Padró i Miquel, 2007; Acemoglu et al., 2008; North et al., 2009; Francois et al., 2015; Diermeier et al., 2017), to the best of our knowledge, this literature did not account for the practice of political favoritism yet. The most closely related paper, Acemoglu et al. (2004), focuses on a model where the ruler's optimal strategy exploits off-path threats to make sure that subjects anticipate that no other group would join in opposing the government. Such divide-and-conquer strategy implies that the subjects' ability to cooperate against the central authority does not affect the allocation of political rents, thus making them independent of the relative standing of each subject. We further discuss the relation between their result and our result in Section 2.2.3. In Francois et al. (2015), the ruler discourages the formation of revolutionary coalitions by including in its cabinet the leaders of the top ethnic groups. However, conditional on being included in (or excluded from) the cabinet, coalition-building, thus relative standing, plays no role. As a consequence, political favoritism does not arise. Finally, a paper in this literature that bears some technical similarities with our analysis is Roemer (1985), which studies how a ruler can discourage the formation of a revolutionary coalition by adopting a divide-and-conquer strategy that relies only on targeted punishments. However, assuming that the ruler can-

not distribute political rents, this paper cannot account for the role of the government in generating social stratification and the pervasiveness of political favoritism.

Moving beyond the application of our model, our taxation game relates to the literature on contracting with externalities pioneered by Segal (1999b). Within such literature, some papers have already proposed models where ranking-mechanisms, similar to political favoritism, can arise in equilibrium. For example, Winter (2004b) shows that the optimal mechanism that a principal can adopt to uniquely implement effort from a set of n identical agents is indeed a ranking mechanism that involves non-symmetric rewards.[8] Moreover, our paper is certainly related to Halac et al. (2020), which introduces agents' heterogeneity in an applied problem and highlights how the agents' position in the ranking scheme crucially depends on their relative characteristics. Our model reverses the classical focus of this literature (including Winter 2004b and Halac et al. 2020) by considering the case of a principal who wants to discourage (rather than encourage) the agents' coordination. This key difference leads to a new model where a ranking mechanism (political favoritism) arises as the unique equilibrium mechanism, even without imposing unique-implementability.[9]

2.2 Model: taxation and the threat of resistance

We consider a society made up of $n+1$ decision-makers (groups), n subjects and one ruler. Each group i is endowed with possibly heterogeneous economic resources B_i and military resources W_i, which are public information. Groups derive their utility from economic resources; military ones only determine a group's odds in its disputes with the ruler. To maintain full generality, we characterize the optimal taxation treating these resources as

[8] Segal (2003b), Eliaz and Spiegler (2015) and, in the context of persuasion, Inostroza and Pavan (2020) present similar results.

[9] The assumption that agents choose sequentially is crucial for the uniqueness of our equilibrium, but is not enough to imply the principal optimally adopts a ranking mechanism to incentivize the agents. If both principal and agents benefit from coordination (as in the literature), sequentiality would uniquely select the principal-preferred equilibrium, where no ranking mechanism is at play.

separate. This means that our characterization applies to any relation between military and economic resources. Our model features the following timing:

1. Groups receive their initial endowment of military and economic resources
2. Given the distribution of resources, The ruler demands taxes from each group aiming at maximizing total tax extraction.
3. Groups observe the vector of tax demands $\boldsymbol{T}$ and sequentially choose whether to comply or to resist taxation.
4. Compliant groups pay their taxes. Resisting groups join their forces against taxation and receive a tax break, if successful, and a a fine (e.g., lose everything) otherwise.
5. Payoffs are realized.

The assumption that subjects sequentially choose to resist taxation delivers a unique SPE outcome. As we discuss in Section 2.2.3, the outcome of our taxation game is also an outcome of a game where subjects respond simultaneously. In particular, it corresponds to the equilibrium that arises when the ruler expects the subjects to form the largest possible resisting coalition consistent with the tax demands, i.e. the one implemented by a wary ruler who selects the worst-case optimal policy (max-min).[10] Finally, in Lemma 2 we show that the order in which subjects respond is irrelevant for the equilibrium outcome. From this point of view, our SPE is protocol-free: it is robust to changes in the specification of the order of moves, in the spirit of the equilibria analyzed in Diermeier et al. (2017).

[10]Similarly, in spirit, to the equilibrium concept proposed by Dworczak and Pavan (2020), the ruler chooses its strategy under the assumption that subjects would select the ruler's worst equilibrium.

2.2.1 Payoffs: compliance and resistance

We first consider the case with a fixed composition of society, i.e., we treat each group as a separate decision maker. Resources are distributed according to a vector $(\mathbf{B}, \mathbf{W})$ about which all groups have complete information. The ruler publicly demands from each subject $i \in I$ a tax $T_i \in \mathbb{R}$. Upon observing the required tax vector $\mathbf{T}$, groups sequentially choose whether or not to resist $a_i \in \{R, NR\}$; we denote by *resist* the set of all subjects that do so,

$$resist := \{i \in I : a_i = R\},$$

and by *loyal* the set of subjects who instead choose to comply, $loyal := I \backslash resist$.[11] The choice of one group to resist is irreversible, public, and it entails a cost $c > 0$. By choosing to comply, group i obtains utility

$$u\left(B_i - T_i, NP, W_i\right) = B_i - T_i.$$

On the other hand, if group i chooses to resist then it would either end up obtaining a tax break $\beta(T_i, B_i)$ or paying $\lambda(T_i, B_i)$ on top of T_i, depending on whether the resistance succeeds or fail. To capture the idea that richer groups are less willing to resort to violent action, we make the following assumption:

Assumption 1 (Higher income discourages resisting) *The functions $\beta(T_i, B_i)$ and $-\lambda(T_i, B_i)$ are weakly decreasing in B_i and at least one is strictly so.*

A natural case in which such assumption holds is when rebels completely avoid taxation in case of success, i.e. $\beta(T_i, B_i) = T_i$, and lose everything otherwise, i.e. $\lambda(T_i, B_i) = B_i - T_i$. Consistently, we assume that when group i resists, the ruler recovers from it $0 \leq R_i \leq B_i$ if resistance fails, and 0 otherwise.

[11] As shown by Lemma 2 the actual order in which subjects are called into action is irrelevant.

The probability that resistance is successful depends on its size; in particular it is determined by a general contest success function

$$\Pr(\text{resistance success}) := f(W_{resist}, W_{ruler}, W_{tot})$$

where $W_{resist} := \sum_{j \in resist} W_j$ is the sum of the military resources controlled by the resistance, and $W_{tot} := \sum_{j \in I} W_j + W_{ruler}$. The contest success function f is strictly increasing in W_{resist}, strictly decreasing in W_{ruler} and weakly decreasing in W_{tot}. The expected utility of a resisting group i is thus

$$E[u(B_i, W_{resist}, W_{ruler}, W_{tot})] = f(W_{resist}, W_{ruler}, W_{tot}) \times B_i - c.$$

The ruler's objective is to maximize the expected sum of collected transfers.

2.2.2 Technical results: the structure of the equilibria

Since the game is dynamic and agent's information is complete, we focus on subgame perfect equilibria (SPE). In this section we establish three characteristics of any SPE.

When a tax vector is demanded, subjects sequentially choose whether or not to join the resistance. The following lemma establishes that, for any tax demand $\mathbf{T}$ there is at most one set of resisting subjects that, paired with $\mathbf{T}$, can be part of a SPE.

Lemma 1 *Any tax vector* $\mathbf{T}$ *can be at most part of one SPE outcome. Furthermore, in any SPE all subjects comply when indifferent.*

As we show in Appendix B.1, in a SPE no subject joins the resistance unless it has strict incentive to do so; otherwise the ruler could profitably deviate by demanding a slightly lower T_i.[12] Together with the complete information assumption, this observation allows us to

[12] As we show in the proof, once we allow $resist_{-i}$ to also be affected by i's decision to resist, the argument

determine, for any fixed **T**, a unique candidate SPE. In light of Lemma 1, we can show that the order in which the subjects choose whether or not to join the resistance is irrelevant.

Lemma 2 *The set of equilibrium outcomes of the taxation game does not depend on the subjects' order of play.*

Given the perfect information assumption, when subject i chooses whether or not to comply, it can perfectly anticipate what the following subjects will choose in response to its decision. Clearly this is not enough to show that the order of play does not matter. However, the results follows by noting that the decisions to resist are strategic complements (i.e. each subject's incentives to resist is increasing in the resistance size).

Finally, our last preliminary result establishes a "no resistance" result which is key for the construction of the equilibrium.

Lemma 3 *There exists no SPE of the taxation game that involves resistance.*

The interaction between the ruler and a compliant subject is zero sum. Instead, resistance is less than zero-sum, as it entails a loss (the dis-utility c) that the ruler cannot appropriate. Together with the complete information assumption and the fact that the ruler can perfectly target groups when demanding taxes T_i, this observation might seem to trivially imply the result. The subtlety lays on the fact that, in our setting, there are externalities: i's decision to resist might impact the set $resist_{-i}$. In principle, the ruler might be willing to incur in an inefficiency when dealing with i, if that enables it to avoid incurring in an even greater inefficiency when dealing with the rest of society. However, because the decisions to resist are strategic complements, discouraging i from resisting could only further discourage other groups from resisting, implying an even greater incentive for the ruler to avoid any inefficient outcome.

is strengthened. i's decision to resist might lead to an even worse payoff for the ruler, as it increases the willingness to resist of every other player, and it increases the likelihood of success of any group who might be resisting.

2.2.3 Political favoritism: the relevance of social status

In this section we show the optimality of favoritism in any SPE of the taxation game. With favoritism, we mean the act of treating subjects differently, beyond what can be explained by their intrinsic characteristics. In particular, a taxation system is favoristic if the tax on each subject does not only depend on the subject's resources but also on some order superimposed by the ruler (social status).

To gain intuition, we start by considering a ruler dealing with two identical subjects. A naïve intuition could suggest that as these subjects are identical, we should look for an equilibrium where the ruler treats them symmetrically, demanding the same tax T. By Lemma 3, we know that an optimal T must not trigger any resistance, thus it cannot be greater than $\bar{T}^{Alone}$, which satisfies the no-resistance-alone constraint:

$$B \times f\left(W, W_{Ruler}, W_{tot}\right) - c = B - \bar{T}^{Alone}$$

However, note that if the ruler were to offer $\bar{T}^{Alone}$ to both subjects, then the first subject who is called to respond would anticipate that if it were to start resisting taxation, then the other subject would join, as its expected value of resisting would increase to

$$B \times f\left(W + W, W_{Ruler}, W_{tot}\right) - c > B \times f\left(W, W_{Ruler}, W_{tot}\right) - c.$$

Anticipating that the other subject would join, the first subject would optimally resist if demanded $\bar{T}^{Alone}$. Thus, offering a tax equal to $\bar{T}^{Alone}$ to both players would lead to resistance, which by Lemma 3 cannot be optimal.

As a result of such coordination between subjects, in order not to trigger resistance and still demand the same tax to both subjects, the ruler could only demand $\tilde{T}$ equal to $\bar{T}^{Together}$,

the tax that makes subjects indifferent between complying and resisting taxation together:

$$B \times f\left(W + W, W_{Ruler}, W_{tot}\right) - c = B - \bar{T}^{Together}.$$

However, the ruler is better off by resorting to non-symmetric offers, and thus favoritism. In particular, the ruler could collect more taxes and still avoid resistance by favoring one subject over the other, demanding $T_i = \bar{T}^{Together}$ and $T_j = \bar{T}^{Alone}$ respectively. Given that subjects are identical, we can think of the difference between the tax paid by the favored group, $\bar{T}^{Together}$, and the one paid by the oppressed group, $\bar{T}^{Alone}$, as "the value of social status", where "status", a seemingly arbitrary preferential treatment, is an equilibrium object that emerges from the ruler's optimal divide-and-conquer strategy.

To see how such result generalizes to multiple subjects, note that by Lemma 3 the optimal policy must not only make each subject unwilling to resist alone, it must also prevent the formation of resistance cascades that would lead to multiple groups coordinating against taxation. This could be obtained by demanding every subject $\bar{T}^{n-1}$, a tax so low that no subject would resist even when anticipating that all other $n-1$ subjects would join (the equivalent of $\bar{T}^{Together}$ above, but for n subjects). However, the government could increase its extraction by offering such low tax only to one group, its favorite. In this way, the government could increase the tax of every other group from $\bar{T}^{n-1}$ to $\bar{T}^{n-2}$, as they would anticipate that the favorite subject would never join their resistance. Similarly, the government can single out another subject, its second-favorite, demanding a tax $\bar{T}^{n-2}$ and increase the tax on every other subject but the favorite to $\bar{T}^{n-3}$. Iterating such argument, we show that the optimal way to undermine coordination is to institute a ranking of all subjects, each corresponding to a differential taxation treatment. The next Theorem shows that such favoritism is a key feature of the optimal policy, as taxes depend on the government's ranking $O^{\star}$. Moreover, the Theorem also characterizes the optimal tax corresponding to each status, i.e. each group's

position in the government's ranking.

Theorem 2 (Optimality of Favoritism) *Tax demand* $\mathbf{T}^\star$ *is part of a subgame perfect equilibrium of the taxation game if and only if for all* i

$$T_i^\star = c + B_i \times \underbrace{\left(1 - f\left(\sum_{j:O^\star(j)\leq O^\star(i)} W_j, W_{ruler}, W_{tot}\right)\right)}_{decreasing\ in\ O^\star(i)} \tag{2.1}$$

where $O^\star : \{1, ..., n\} \longrightarrow \{1, ..., n\}$ *is a re-ordering of the subjects such that*

$$O^\star = \arg \max_{order\ O} \sum_i T_i^\star.$$

The above theorem shows that the optimal policy for the government is to set up a favoristic taxation system. The ruler establishes a ranking $O^\star$ among the subjects and demands transfers accordingly, imposing relatively higher taxes on lower-ranked subjects. In equilibrium, the ruler demands from each subject the maximum tax that does not induce it to resist when all (and only) the groups with lower status do. Differently from Acemoglu et al. (2004), in our model the ruler is unable to fully shut down every coordination among subjects and cannot induce them to believe that no one would follow if they were to start resisting.[13] Consistently with a divide-and-conquer logic, however, the government taxation policy is still able to induce subjects to (correctly) anticipate that those with higher status would not join, if they were to start resisting.

Furthermore, notice that the sequentiality in the subjects' resistance decision allows us to have a unique equilibrium outcome. If we modeled these choices as simultaneous, we would have, as usual, a multiplicity of equilibria. The equilibrium outcome, in our model, would also be an equilibrium outcome of this different game. In particular, it would correspond to

[13] In Acemoglu et al. (2004) all coordination among subjects is shut down as their timing allows the ruler to change tax demands after one group replies, but before the other does.

the equilibrium that arises when the ruler expects the subjects to form the largest possible resisting coalition consistent with the tax demands, i.e. the one implemented by a wary ruler who selects the worst-case optimal policy (max-min).[14] On the other hand, the level of extraction found in Acemoglu et al. (2004) corresponds to one where the ruler uses the max-max criterion. In that case, the ruler chooses its tax demands assuming that subjects will form the smallest coalition that is compatible with $\mathbf{T}$, which corresponds to the case where coordination between subjects is impossible.

Finally, note that, while the heterogeneity among groups guides the ruler's decision of who to favor, the favoristic structure of the taxation system has nothing to do with such heterogeneity. In fact, as shown in the example with two identical subjects, Theorem 2 prescribes to treat all subjects differently even when they are identical, assigning them to different positions in the ranking $O^\star$. The divide-and-conquer logic is the main driver of favoritism, not groups' heterogeneity.

2.2.4 Determinants of social status: the State as a motor of social stratification

Theorem 2.1 has established that the optimal taxation is favoristic: it is a function of the government's ranking. Such result does not answer the question of whether government policy should be expected to be progressive or regressive. To answer this question, in this Section we show the determinants of status: how a group's relative endowment determines its position in the optimal social order (thus its tax burden).

[14] Similarly, in spirit, to the equilibrium concept proposed by Dworczak and Pavan (2020), the ruler chooses its strategy under the assumption that subjects would select the ruler's worst equilibrium.

In light of Theorem 2, given group i's status $O_i^\star$, the optimal tax $T_i^\star$ satisfies

$$T_i^\star = B_i \times \left(1 - f\left(\sum_{\ell: O_\ell^\star \leq O_i^\star} W_\ell, W_{Ruler}, W_{tot}\right)\right) + c$$

Note that such tax is higher when group i is richer and when i and all subjects who are ranked below i are weaker from the military viewpoint. If all groups had equal military power, then the government's ranking would assign a higher status to poorer groups. Indeed, a lower status corresponds to a higher tax rate, thus it is better to assign lower status to richer groups. On the other hand, if every group had the same economic resources, then groups with more military resources would be assigned a higher rank. To see this, first note that, by the properties of the equilibrium, subjects who are ranked above a particular group benefit from that group's military resources (it reduces their equilibrium tax burden). From the ruler's perspective, this means that the amount of resources it can extract from a particular subject decreases in the military power of all lower-ranked subjects. As a consequence, by assigning a higher status to stronger groups, the ruler maximizes the sum of taxes that it ends up extracting from society.

Note that, given any specific distribution of resources $(\mathbf{B}, \mathbf{W})$ and contest success function $f(\cdot)$, it is possible to compute the optimal ranking. On the other hand, in order to derive some general properties over the link between social status and distribution of resources, we need to specialize the contest success function.

Assumption 2 *The probability that resistance against taxation of subjects $j \in \{1, 2, \ldots, i\}$ is successful is given by:*

$$f\left(\sum_{j \leq i} W, W_{ruler}, W_{tot}\right) = \frac{\sum_{j \leq i} W_j}{W_{tot}}$$

While this is a standard contest success function, it is important to notice that the odds of

a successful resistance is negatively affected by the resources of subjects who are not joining the resistance. This would be the case if neutrality works against the resistance as much as siding with the ruler, or if neutrality is not an option (at least not an optimal one).[15]

In modern contexts, the above assumption captures a situation where the generals of the State's army are drawn from the same group (e.g. ethnicity) as the dictator, but some of the lower ranks are members of other ethnic groups: an ethnic structure corresponding to many African armies today (Harkness, 2016). In such a situation, deserting the army is typically treated as treason, i.e. it is the same as directly attacking it. Similarly, in what are known as feudal societies subjects had an obligation to respond to a call to arms made by their monarchs: deserting any such call was costly and could be considered as no different than directly siding with the enemy. Finally, this assumption can also capture other forms of resistance that do not directly involve an army. For instance, if resistance takes the form of a strike, then not directly participating in it is the same as working against its success.

Given this specification, we can pin down the determinants of the government's optimal ranking for any distribution of **W** and **B**:

Theorem 3 (Determinants of Status) *Under Assumption 2, the optimal government's ranking is determined by the militarization index* $\frac{W_i}{B_i}$. *In particular, re-ordering the subjects so that* $\frac{W_1}{B_1} \leq ... \leq \frac{W_i}{B_i} \leq ... \leq \frac{W_n}{B_n}$, *the equilibrium tax would be determined by*

$$T_i = c + B_i \times \left(1 - \frac{\sum_{j \leq i} W_j}{W_{tot}}\right)$$

for all i. Thus, relatively more militarized subjects are assigned a higher status.

[15]Note that neutrality would be a dominated action for the subjects in any micro-foundation where, when resistance fails, the ruler punishes (e.g. substitute) the leadership of the various groups as if they participated in the resistance. This would be the case, for example, if we assume that some group members have the behavioral type of "loyalty", i.e. they are not strategic and always side with the ruler: in this case, whenever a leader fails to join the ruler, the ruler would optimally substitute her, appointing a member of the same group at random, which would have a higher chance of being a loyal type.

The theorem shows that with this specific contest success function we are able to find an index that determines the status of every subject for any possible distribution of **W** and **B**. In particular, every subject's position in the government's optimal ranking is determined by the relative militarization ratio $\frac{W}{B}$: subjects with a greater militarization index pay a lower tax rate.

While the militarization index is inversely related to economic resources, this does not mean that richer groups will be necessarily taxed more. In fact, typically economic and military resources are not independent. In particular, we may expect richer groups to hold a larger share of their economic resources as military resources, which would translate into $\frac{W_i}{B_i}$ being larger for richer groups. Since the investment in military power is not the focus of this paper and explicitly allowing for that would significantly complicate the model, we adopt the reduced-form approach of allowing for a relation between each B_i and W_i. In particular, for the rest of the paper we assume that there exists a function $g(\cdot)$ mapping economic resources into military resources and that such mapping reflects the idea that richer groups can commit a larger share of their economic resources to military power.

Assumption 3 *Military resources of each group i satisfy $W_i = g(B_i)$, for some continuous function $g(\cdot)$ such that $\frac{g(B_i)}{B_i}$ is increasing in B_i.*

Notice that the assumption above would be satisfied by any convex function $g(\cdot)$, reflecting the idea that with double the economic resources a group can do better than just duplicate the number of soldier it hires, a natural condition if we consider a setting in which military power can be hired (e.g. there are mercenaries).[16] Another function satisfying the assumption above is $g(B_i) = B_i - s$, reflecting the idea that, if needed, groups can transform into military resources only economic resources in excess to a certain subsistence level s, assumed to be

[16] A richer group could invest in better equipment and/or generals. Also, it is easy to imagine that if one group can hire more soldiers, then it will be cheaper for it to hire even more soldiers, as fighting for the stronger party involves less risks and thus requires a smaller payment.

the same across groups. The following corollary establish the government's optimal taxation when $\frac{g(B_i)}{B_i}$ is increasing in B_i.

Corollary 1 *Under Assumption 3, subjects with greater economic resources have higher status. In particular, the equilibrium tax on each subject is*

$$T_i(\mathbf{B}) = c + B_i \times \left(1 - \frac{\sum_{\ell:B_\ell \leq B_i} W_\ell}{W_{tot}}\right)$$

The corollary shows that when the function that maps economic into military resources satisfies Assumption 3, then we have a clear social order where status is pinned down by the relative endowments of the subjects. The analysis suggests that government policy should be expected to be highly regressive, as richer groups obtain a favorable taxation treatment.

2.2.5 Uncertain composition of society: solidarity and social unrest

A fundamental assumption of our framework is that society is exogenously divided in non-atomistic groups that constitute the decision units of our model. In this section, we aim to shed some light on how (possibly stochastic) changes in such social divisions affect the equilibrium outcomes. In particular, we consider the possibility that two or more subjects are in *solidarity*, i.e. they act together as if they were a single decision unit. In particular, we consider two types of solidarity: *common identity* (i.e. subjects sharing the same norms, goals, activities, and internalizing each others' payoffs), and *pledged support* (i.e. subjects promising to support each other).

We start our analysis from pledged support, extending our baseline model allowing for the possibility that, with some probability $p \in [0, 1]$, a subset G of subjects will uphold an alliance stipulating that they will resist whenever a fellow member wants to. As a result, with

probability p, the subjects in G are, de facto, part of the same group $\mathcal{G}$ with $B_{\mathcal{G}} = \sum_{i \in G} B_i$ and $W_{\mathcal{G}} = \sum_{i \in G} W_i$ and they either do not resist or resist together. It is easy to show that the equilibrium that arises if there is no uncertainty about the groups in G behaving as a single decision unit, i.e. $p = 1$, shares the same characteristic as our baseline model where we essentially replace the subjects in G with the united group $\mathcal{G}$. In the equilibrium with $p = 1$ the ruler would thus treat every member of G as a single group $\mathcal{G}$ and demand from all $i \in G$ a tax

$$T_i = c + \left(1 - \frac{W_{\mathcal{G}} + \sum_{j \in I \setminus G : W_j < W_{\mathcal{G}}} W_j}{W_{tot}}\right) B_i. \tag{2.2}$$

If group $\mathcal{G}$ arose from the creation of a common identity rather than pledged support we would obtain similar results. Assume that, with some probability p, a subset G of subjects have successfully created a common identity, i.e. they internalize each other's utility $u_i(\mathbf{B}) = u_j(\mathbf{B}) = \sum_{s \in G} B_s$, where $\mathbf{B}$ is the full vector of economic resources that remain at the end of the game. Also in this case, if $p = 1$, the ruler wants to steer clear of resistance and thus treat them as a single group $\mathcal{G}$, demanding from all $i \in G$ a tax T_i such that[17]

$$\sum_{i \in G} T_i = \sum_{i \in G} \left(c + \left(1 - \frac{W_{\mathcal{G}} + \sum_{j \in I \setminus G : W_j < W_{\mathcal{G}}} W_j}{W_{tot}}\right) B_i\right). \tag{2.3}$$

The comparison of the tax demands in Equations 2.2 and 2.3 with our baseline model (the case where $p = 0$) shows that when fully credible, i.e. $p = 1$, solidarity decreases the total taxation imposed on members of G, i.e. $\sum_{i \in G} T_i$. Moreover, in the case of pledged support, each member enjoys lower taxation (strictly so, with the only possible exception of the strongest subject in G).[18] Note that also the strongest member $i \in G$ is strictly better

[17] Note that, given the payoff structure with a common identity, any transfer of tax from $i \in G$ to $j \in G$ would keep both groups indifferent.

[18] All $j \in G$ strictly benefit if $j \neq i$, with $i \in argmax_{s \in G}(W_i)$. Note that in the case of common identity the comparison between the tax paid by a single member under $p = 0$ and $p = 1$ would be meaningless.

off if there exists a subject $j \in I \setminus G$ such that $W_{\mathcal{G}} > W_j > W_i$.

So far we only considered the case with $p = 1$; however, a new common identity might be frail and, certainly, not all pledges are cast in stone. It is therefore natural to consider the case in which there is uncertainty on whether the members of G will be in solidarity or not, i.e. $0 < p < 1$. To start the analysis of this case, suppose that G is the set of all subjects. Note that only two options can be optimal for the ruler: either the ruler imposes taxes as under $p = 1$, avoiding any possible resistance, or it gambles on the inability of the subjects to stick together and demands taxes as under $p = 0$.

Proposition 5 *If G includes all subjects, i.e. $G = I$, there exists $\bar{p}$ such that:*[19]

- *If $p < \bar{p}$, the ruler demands taxes as if $p = 0$ and with probability p all subjects resist.*
- *If $p > \bar{p}$, the ruler demands taxes as if $p = 1$ and avoids resistance.*

In terms of equilibrium payoffs, the proposition shows that a credible (even if not perfect) threat of collective action (i.e. p high enough but not necessarily 1) leads the subjects to reap the full benefits of being in solidarity. Note that an infinitesimal change in p from below to above $\bar{p}$ would lead to a discontinuous change in the equilibrium payoff of all subjects involved (except, at most, one) and a discontinuous reduction in the ruler's payoff. On the other hand, when p is not high enough, which could be the case when a new common identity arises or a new representative institution or alliance is formed, tax demands do not decrease and we should expect to observe social unrest, rebellions and/or conflicts against central authorities (with probability p).

Recall that the structure of the optimal taxation was such that each group was discourage to resist the government when all subjects with lower social status decided to resist, whereas

[19]From the ruler indifference condition it is easy to show that $\bar{p}$, where $\bar{p}$ is such that $\frac{\left(\frac{\sum_{i\in I} B_i \sum_{j\leq i} W_j}{W_{tot}}\right)}{(k+|G|c)} = \frac{\bar{p}}{(1-\bar{p}),}$ where k is the ruler's cost of facing a resistance and $|G|$ is the number of subjects in G.

all the others did not. As we now have social unrest in equilibrium, we can use this structure to make predictions on which social groups should be expected to participate in social unrest. The next result establishes that lower social status (higher taxation) predicts participation in revolts.

Remark 1 *If subject i resists, then every subject j such that $\frac{T_j}{B_j} > \frac{T_i}{B_i}$ also resists.*

The result tells us that correlation between government discrimination and participation in revolts should not be interpreted as evidence against the idea that government policy is constrained by the threat of social unrest. Indeed, the result shows that such correlation should be expected even when social unrest is the only relevant constraint on the government.

To conclude the analysis we now consider the general case where G might only include a subset of the subjects and the subjects in G are able to stick together with probability $p_G > 0$. In particular, we are interested in characterizing who benefits and who loses from such agreements, to see when we might expect them to emerge and who should be expected to promote or contrast them. The first result is that the ruler is always hurt from any new form of solidarity among its subjects. If a subset of subjects is in solidarity with some probability, then either this creates a new binding constraint on the ruler's maximization problem, or it creates an inefficiency (resistance) which reduces the ruler's expected payoff. As a consequence, the expected payoff of the ruler is unambiguously lower than in our baseline model (i.e. with $p = 0$).

Remark 2 *Whenever a subset of groups G are in solidarity with some probability (i.e. $p_G > 0$), the expected payoff of the ruler is strictly smaller than with $p = 0$.*

The previous result establishes that the ruler is an enemy of any form of solidarity that might emerge among its subjects. Perhaps more surprisingly, the ruler is not the only agent that can be negatively affected by a change in the composition of society. In particular,

we show that the relative impact on excluded and included groups depends crucially on the credibility of solidarity (i.e. on p).

Proposition 6 *Suppose that a subset $G \subset I$ of subjects are in solidarity with probability $p_G > 0$. There is $\bar{p}_G$ such that, with respect to the case where $p_G = 0$,*

- *If $p_G > \bar{p}_G$ then all members of G are better off, whereas every excluded subject $j \in I \backslash G$ is worse off.*

- *If $p_G < \bar{p}_G$ then every member of G can be strictly worse off (the strongest member of G is at most indifferent), whereas every excluded subject is better off.*

To better understand the first part of the result note that, when p_G is sufficiently high, the ruler treats the groups in G as a unique decision unit (i.e. as under $p = 1$) and no resistance takes place in equilibrium. As the new decision unit $\mathcal{G}$ is trivially stronger than any of its members, by the same logic of Theorem 2 and Corollary 1, in equilibrium $\mathcal{G}$ would never be ranked below any excluded subject who is weaker than the strongest subject in G, so no element of G can lose. In particular, all members of G get to be treated better than how the strongest was treated under $p = 0$. Note that, even the strongest member of G, i.e. $\arg\max_{j \in G} W_j$, obtains a status gain and strictly benefits from the alliance if there is any excluded subject s who is stronger than $\arg\max_{j \in G} W_j$ but weaker than G as a whole, i.e. $W_s \in (max_{j \in G} W_j, W_{\mathcal{G}})$. On the other hand, all excluded subjects that are stronger than the weakest member of G, i.e. $\arg\min_{j \in G} W_j$, and weaker than G as a whole, would lose from the formation of G as they incur in a status loss: some of the subjects that used to contribute, with their military power, in reducing its tax burden would now contribute in increasing it, since they are now ranked higher as part of the newly formed decision unit $\mathcal{G}$. In terms of representative institution, this result suggests that, if it fails to be fully inclusive,

excluded members may have the incentive to dismantle it.[20]

The case when p is small is more subtle. Note that, in this case, the ruler designs its offers accepting that, with probability p, the members of G will resist. By the structure of the equilibrium demands, the members of G would typically not be the only ones resisting: they would also be joined by all lower ranked subjects. But then, when p is sufficiently small, the ruler might benefit from ranking some of the subjects excluded from G above all members of G, so to limit the size of the resistance in case the members of G end up resisting (which happens with probability $p_G > 0$). As a result, excluded subjects benefit (possibly strictly so) from the formation of a fragile new alliance or common identity. Conversely, some members of G might actually prefer not to be part of an alliance whose credibility is low, i.e. strictly prefer $p = 0$ to $p_G \leq \bar{p}_G$, and thus might benefit from openly rejecting the alliance or taking actions that credibly deny their adherence to the new common identity. This rises concerns about the possibility that such a fragile $\mathcal{G}$ forms in the first place. Finally, the proposition suggests that excluded subjects benefits from undermining the alliance when it is sufficiently credible (i.e. transform $p > \bar{p}_G$ in $p < \bar{p}_G$), but they might have incentives to spread rumors about possible but unlikely anti-ruler alliances (i.e. transform $p = 0$ into $p < \bar{p}_G$) that do not include them. From this point of view, conspiracy theories might have a precise political meaning precisely because of their low credibility.

2.3 General Patterns

2.3.1 Relevance of social status

The first key pattern that we focus on is that attention for the relative standing of elite members or even social groups, in the form of rank and precedence, is a ubiquitous feature

[20] While this is beyond the scope of the current paper, such result might explain the puzzling finding that many representative institutions have been dismantled in the course of history (Stasavage, 2020).

of human societies. This fact is consistent with our prediction that the optimal policy should be a function of not only the individual characteristics of each group, but also of their position in an order endogenously determined by the ruler (Theorem 2). Such finding highlights the crucial importance of rankings as an *input* in the policies of central authority.

While the Roman Empire collapsed in Western Europe, it survived in its Eastern domains as the Byzantine Empire (330 CE - 1453). The Byzantine Emperor was the head of a court which featured a famously strict etiquette that mandated different dressings, behaviors, and prerogatives based on the title held by a particular subject. At the heart of the Emperor's power was the ability to manipulate the hierarchy of imperial titles, the so-called taxis (Bougard et al., 2009 p.363). The relation between this abstruse system and the exercise of power is expressed in the preface of the Treaty of Philotheus (X Century CE): "order is necessary so that the imperial power is upheld by the subjects. By mimicking the Godly order of the universe, imperial authority finds a way to consolidate its power" (quoted in Bougard et al. 2009, p.364; own translation).

A similar attention to ceremony and ritual can be found in Japan during the Nara (710 CE - 794), Heian (794 CE - 1185) and Kamakura (1185 CE - 1333) periods. Just as in the Byzantine Empire, "ritual and ceremony were not quaint or meaningless customs designed to occupy the time of bored courtiers, they were a visible symbol of the social order and served an important function in vitalizing and renewing the polity (..) from the very beginning, court ritual and ceremony *were* politics" (Friday 2004, p.32; emphasis in the original). Courts throughout history feature similarly complex systems of precedence that regulated access to the central authority and that induced an almost perfect stratification within the elite. For China, Chung-Mien and Herbert (1987) offers a glimpse into the intricate system of ranks of the Tang dynasty (618 CE - 907); McKnight (1985) discusses how ranks were associated with different legal and political privileges during the Song dynasty (960 CE - 1279); Duyvendak (2008) discusses how in 1794-1795, during the last Dutch Embassy to the court of the Qing

Dynasty (1644 CE -1911), dressings and manners revealed the rank of each courtier. In South-East Asia, Ghazzali (1933) provides a detailed account of how the etiquette of the XX Century Malaysian Court was intertwined with relative standing.

In a surprisingly close parallel, the courtly societies of Western Europe during the Early Modern period also featured an "extremely elaborate etiquette, a whole system of carefully orchestrated social rituals of deference, a finely gradated and strictly observed hierarchical order" (Zmora 2002, p.76). It should also be stressed that the attention for rank and precedence in Western Europe was not a novelty introduced by the rise of courtly society in the Early Modern period. Indeed, most Western European societies, at least since the Late Middle Ages, linked the political power and the fiscal/legal standing of their members to their rank within the polity. For instance, each geopolitical unit within the Holy Roman Empire (962 CE - 1806) was classified as kingdom, electorate, principality, archduchy, grand duchy, duchy, margraviate, etc. The different titles implied a particular position in the prevailing system of precedence, a different set of rights (e.g. a different extent of representation) and obligations (including a different autonomy or fiscal regime). In general, rank mattered a great deal in the politics of the time and was closely determined by the Emperor (e.g. Strauss 1978, p.301; Stollberg-Rilinger 2018, Chapter 2; or Bryce 1901).

The importance of rankings is clearly present in other civilizations of the World as well. For instance, in her study of the Swazi society in Southern Africa, Kuper writes: "The clans are graded into a rough hierarchy, and the rank of a clan is measured by its position in the national structure" (Kuper 1949, p.111-112). Such structure is not static, but reflects the current balance of power of the clans in society. In particular, she notes that "[a]n outstanding man may "elevate the clan name" and his lesser clansmen will enjoy reflected glory" (Kuper 1949, p.111-112). Importantly, such rank was crucially linked to government policy as "individuals can raise their position through winning office from the King" (Kuper

1978, p.572).[21] South America was not exempt from this pattern. In general, Earls (1969) and Jenkins (1995) reveal that the societies in the Central Andes gave key importance to the relative standing of social groups (and their leaders) before (and after) the arrival of the Spaniards in the XVI Century. For instance, every Inka Emperor ranked the relevant kin-groups; interestingly, it has been hypothesized that such ranking was not rigidly determined, but it could be changed at will by the Emperor (D'Altroy 2015, pp.263-277). After a joint analysis of political systems, myths and social organization in both Quechua and Inka civilizations, Jenkins (1995) concludes that "the apparent historical recurrence of age and rank structures in myth and social organization indicates that the structures themselves are in some way valued - a valuation, moreover, that may be significant in the organization of Andean history" (Jenkins 1995, p.92).

The attention for relative standings took a global dimension in the Modern Age, with the expansion around the world of the European colonial empires. In particular, the development in Europe of Scientific Racism in the mid XIX Century can be thought of as an attempt to construct a global system of precedence that encompasses all peoples. Indeed, intellectuals in the field of eugenics and Social Darwinism attempted to reconstruct the genealogy of all 'human races' and eventually provided a classification of every ethnicity based on their supposed descent (Dennis, 1995). This intellectual endeavour was in part the response to and in part the ideological foundation for the ethnic policies employed in the colonies.[22]

[21] While this is just a matter of definitions, in the language of our theory winning offices is not a means to raise in the ranks, but it is the outcome (or even the observable implication, if the ranking itself is not visible) of raising in the ranks.

[22] There is a large literature on how administrative classifications in the census of the colonial authorities have *created* new ethnic groups assigning them a preferential (or discriminatory) status (Brubaker, 2009; Cahill, 1994). Social Darwinism did not just affect the European elites, for a general discussion of how the adoption of Social Darwinism affected the minorities residing in China and Japan see Dikotter (1997), for Japan specifically see Weiner (2009), and for Latin America see Stepan (1991). From an ideological standpoint, it is interesting to note that Japanese scholars in the Modern period accepted a world-view that "reproduced Western racial hierarchies, privileging the West as the apotheosis of 'civilization' (bunmei), trailed by a 'semi-civilized' (hankai) Asia and 'barbaric' (yaban) Africa" (Weiner 2009, p.95). Similarly, "Latin American scientists not only imported these concepts that ranked humankind into inferior and superior races; they also produced their own raciological theories that then influenced policy making regarding

For instance, the administration of the Spanish Empire was characterized by "an unhealthy obsession with caste" whereby an individual's "character and quality were held to vary according to the colour and origins of the 'ethnic types'" (Cahill 1994, p.338-339). This approach led to an extremely detailed ranking of each ethnicity based on their relative whiteness, distinguishing a mulato (someone whose mother had African origins and a Spanish father) from a testerón (someone with a mulato father and a Spanish mother), from a quarterón (someone with a testerón father and a Spanish mother), etc. Such distinctions had legal and fiscal consequences; natives who could afford to buy their 'whiteness' in the process known as 'gracias al sacar' often did so (see f.i. Cahill 1994, Voss 2008, and Twinam 2015).

2.3.2 The State as a motor of social stratification

As a second key pattern, we document how the creation or strengthening of central authorities has typically increased social stratification at the ethnic, territorial, clan and/or other levels (Section 2.3.2). This observation speaks to our model's prediction that the optimal assignment of social status should reflect the relative power of the various social groups (Theorem 3 and Corollary 1) and thus be particularly regressive. Indeed, central authorities would optimally reinforce social inequality and even *induce* social stratification: even if social groups were ex ante identical, the optimal policy would arbitrarily favor some social groups over others.

Boix (2015) shows how the establishment of central authorities have increased the stratification of a wide array of societies, exacerbating economic inequality and compounding economic well-being with a preferential political status and legal treatment.[23]

immigration as well as access to housing, sanitation, and education by blacks, whites, Indians, and mestiços" (de Santana Pinho 2009, p.42)

[23]North et al. (2009) identifies hierarchy as a defining characteristic of "the natural state," a concept that North et al. (2013) fruitfully applies to modern developing countries.

A common feature of the various centralizing drives that followed the fall of the Roman Empire in Western Europe was the tendency to reinforce hierarchy in society (Bougard et al., 2009). For instance, the Carolingians (800 CE - 888) carefully leveraged "a system in which politics was centred on the endowment and removal of honores [i.e. public offices]" (Innes, 2000; pp.262) creating distinctions among the local leaders who obtained multiple offices (in principle delegated from the Emperor) and those who did not. Such strategy allowed the central authority to obtain the loyalty of and a share of the proceedings from those who were fortunate enough to be vested with offices, as they had to be "careful not to give rulers any reason to seize them and give them to new favourites" (Innes 2000, p.89).

After centuries of political fragmentation, following the collapse of the Carolingian Empire, the reprisal of state-building efforts in XIV Century Europe "exercised profound influence on the creation and calibration" of an elite within the elite (Zmora 2003, pp.66-67) and were an "inexorable motor of inequality" (Zmora 2002, p.59). The trend continued in the Early Modern period: while some nobles obtained little more than exemption from taxation, others obtained large pensions and covered multiple offices, until "influence and prestige became the preserve of a restricted circle of men (and sometimes women), leaving a good many hopeless and resentful" (Zmora 2002, p.97). In general, the history of Western Europe from the Middle Ages to the Modern period shows that state-building deliberately "reinforced disparities and hierarchy in society in general and among the nobility in particular" (Zmora 2002, p.97).[24]

[24]We interpret social stratification as a deliberate outcome of state-building because of the practice of assigning multiple offices to the same family (and even individual) and even assigning regressive pensions. Multiple offices were assigned everywhere in Europe. For an in-depth example of Late Medieval and Early Modern France see Zmora (2002) pp.47-48. For the Holy Roman Empire, see Zmora (2003) where it is stated that while Franconia did not feature the same degree of social stratification of Early-modern England or Castille, this was still in large part by the fact that the most prominent families were fulfilling "important functions simultaneously". For instance, in the period 1487-1528, 50% of the margravial councillors were also district governors and 75% had some other appointment (see Zmora, 2003; p.41 and references therein). While pensions could in principle compensate for any undesired stratification induced by any possible rigidity in the assignment of administrative policies, the regressive allocation of pensions by the monarchs appears to aim at deepening rather than leveling such stratification. See p.47 Zmora (2002) for an example of such

The idea that state-formation or expansion might lead to growing elite stratification also appears in Asia. For instance, in Japan, the efforts to centralize powers during the Heian period (794 CE - 1185) ended up delegating central powers to a few selected clans, creating a deep stratification within the elite as "eligibility for any given post in the bureaucratic hierarchy became progressively more circumscribed, limited to smaller and smaller numbers of houses" (Friday 2004, p.7). Even in China, where the access to administrative positions was based on an exam, it can be argued that the allocation of administrative positions induced a centrally-controlled social stratification. The enormous expenses related to the preparation of the exam made sure that the candidates would disproportionately come from members of the local elites, thus reinforcing stratification within the population. Moreover, as the success rate in the exam was governed by an heterogeneous system of local quotas, the central authority could control the relative allocation of such positions across the various territorial units, thus inducing stratification within the regional elites (Elman, 2000). Indeed, as Bai and Jia (2016) has shown, even if the selection into the administration was based on an exam, it was also fundamentally intertwined with concerns for social stability.

A different example of how central authorities have induced stratification can be found in the governance of vast territorial empires. Just before the arrival of the Europeans, the city of Cuzco expanded its authority over much of the Pacific Coast of South America. While doing so, they offered differential gradations of 'Inka citizenship' to the various communities they subdued, favoring some communities over others and creating net distinctions between local elites and local populations (Moseley, 2001; Chacon, 2009; Cahill, 1994). Indeed, during the Inka expansion "local divisions according to ethnicity and eliteness were recognized and used" (Lamberg-Karlovsky 1989, p.188) and new differentiation were created where they did not previously exist (D'Altroy 2015, Chapter 9). In a surprisingly similar fashion, when the city of Rome expanded its authority over much of the Italian peninsula in the III Century

a regressive distribution of pensions in 1480 Castille and in Late Medieval France.

BCE, it assigned different gradations of Roman citizenship to the various communities it subdued. While at the time Rome was a relatively small city, it still managed to project its authority over much of the Italian Peninsula thanks to a carefully knitted system which assigned differential privileges and obligations to the various communities (De Petra, 1866; Kendall, 2013; Sherwin-White, 1973).

The same approach can be found again in the Modern period when the various European monarchies built colonial Empires that stretched for much of the World. At least since the second half of the XIX Century, the British Empire explicitly pursued a policy modeled on the Roman example: "divide et impera was the old Roman motto and should be ours" (Stewart, 1951; p.54). The practice of employing a "highly differentiated treatment of native populations (..) [to] ensure Imperial stability" (Hausteiner, 2016; p.579) was especially elaborate in India, where it was thought that "the more nationalities and castes and religions, the more secure we shall be" (Stewart, 1951; p.53). Morrock (1973) describes how British rule appears to institute a finely grated ranking where social groups different in terms of geography, religion, occupation and traditional caste were assigned to a differential treatment. The British conduct was by no means an exception. As mentioned earlier, the administration of the Spanish Empire was based on a caste system that recognized increasing liberties and rights but at the same time less fiscal obligations based on 'whiteness.' Indeed, "a number of broad-brush racial *cum* ethnic classifications provided the basis for fiscal demands of Crown and Church alike" (Cahill 1994, p.325; emphasis in the original). Similar policies have been applied by the French Empire. For instance, during the colonial administration of Laos in the XX Century, "the French used a system of cross-racial administration which established a brutal ethnic hierarchy, and, especially in northern Laos, pitted the various ethnic groups against each other. The French used traditional racial hierarchies where they were strong, reinforced them where they were weak, and created them where they did not exist. (..) [T]he conditions of the mass of the tribesmen on the lower rungs of the hierarchy was little better

than slavery, since their traditional subordination was compounded by heavy French taxes which produced unbearable demands on their meager resources." (Adams and McCoy 1970, p.80, quoted in Morrock, 1973).

Many contemporary societies still employ discriminatory policies towards the ethnic groups that reside within their boundaries, granting varying degrees of autonomy or citizenship to particular regions or ethnic groups, targeting specific ethnic groups with taxes or public investments, and carefully balancing the inclusion/exclusion of social groups in the political system (see Kasara 2007, Burgess et al. 2018, Francois et al. 2015 for empirical evidence). Even if the relative standing of each group is not as explicit as in historical societies, the systematic nature of such policies and the experience with colonial practices suggest that it might be useful, even within these countries, to understand favoritism in its divide-and-conquer connotation, i.e. as a deliberate strategy of self-interest governments that are threatened by the possibility that their subjects form coalitions against them.[25]

2.3.3 Solidarity and social unrest

Turning our attention to the analysis in Section 2.2.5 we provide evidence on three interrelated facts. First, there is an apparent link between favoritism and social unrest: more discriminated groups are disproportionately more likely to participate in social unrest (Remark 1). Second, governments often appear averse to the creation of solidarity links among their subjects, whereas subjects often attempt to form or strengthen them (Remark 2 and Proposition 6). Third, the creation of cross-group solidarity can destabilize societies (Proposition 5).

[25]Note that the absence of an *explicit* ranking system in contemporary societies does not show that our model does not apply to them. Indeed, as long as tax demands correspond to the one determined by our model, then government policy may still conform to our optimal policy (i.e. depend on the relative status of each social group) even if no relative status is officially recognized or explicitly granted. While the explicit granting of ranks crucially depends on the culture of the time, the optimality of our tax demands depend on the absence of a monopoly over violence, which is certainly relevant in many contemporary societies as well.

Discriminatory policies and favoritism are often associated with social unrest. For instance, Boucekkine et al. (2019) notes that, in post-colonial Africa, opposition against governments is associated with the extent to which the government discriminates among the various ethnic groups. One possible mechanism is shown by Michalopoulos and Papaioannou (2016), which shows that ethnic groups that were split by the arbitrary drawing of national borders are more likely to be discriminated by the central government and are also more likely to participate in conflicts. In line with this, in-depth case studies of government oppositions and civil wars, such as Collier and Sambanis (2005a) and Collier and Sambanis (2005b), highlight that the rebel side of the conflict was often motivated by experienced or expected discriminatory policies by the relevant central authority. Similar findings carry over to historical contexts. For instance, Marín-Guzmán (1995) explains revolts in al-Andalus (modern Spain) during the IX and X Century CE as caused by the favoritism towards the Arab ethnic groups displayed by the Umayyad central government. In Western Europe, during the Late Middle Ages and the Early Modern period, it is often observed that favoritism towards particular families was a frequent trigger for armed revolts and that "favorites" are disproportionately on the loyalists side of armed rebellions.[26] Furthermore, the loyalty of favored groups was taken for granted by both rebels and monarchs alike. Indeed, in the case of Medieval England, Valente (2003) notes that before targeting the King, the violent action of rebellious magnates would typically be aimed against the King's favorites. Also, in many instances, Kings appear to be surprised when their favorites end up joining revolts (e.g. Zmora 2002, p.14). Finally, in a more recent setting, Cahill (1994) shows how the participation in rebellions of the various social groups (ethnic and territorial units) against the Spanish Empire can be explained by the extent of discrimination suffered by such groups.

[26]Typically, magnates rallied against the King's "evil councillors,' i.e. the more favored magnates who dominated the councils. See Valente (2003) for evidence on Medieval England, see Zmora (2002) for evidence on the Early Modern period (e.g. p.69 or p.97); for a general discussion of this feature during the XVII Century see Zmora (2002) p.99.

Our theoretical finding that the subjects benefit from being in solidarity captures the commonplace assertion that 'unity makes strength'. Our model shows that not only 'unity' increases the subjects' bargaining power, but it also undermines the main strength of the ruler, which is based on the ability to employ its divide-and-conquer strategy. This result runs counter to most of the results on the multilateral bargaining literature, including legislative bargaining, which often finds that there are non-increasing returns to bargaining power.[27] However, this logic seems consistent with the understanding of the workings of representative bodies by their early members. For instance, Coke (1644) reports that when asked what might be "a principal motive for them to have good success in Parliament" it was answered by both Commons and Lords alike that Parliaments protected them from the King's "Divide & impera", i.e. divide and conquer, practice.[28] Conversely, Coke (1644) says that "ancient Parliament men out of Record" have pointed to the lack of unity within and between the Houses as one of the few possible reasons for a limited success of the English Parliament.[29] While Parliaments (and in particular the English Parliament) are often viewed as exceptional institutions, a closer look at the evidence reveals that the subjects' attempts to create and institutionalize solidarity links among them is an extremely common event. For instance, the vertical alliances that are at the basis of feudal societies can be thought of attempts to create solidarity links: a subject takes an oath to support another subject, even in its

[27] The standard solution to bargaining problems with multiple agents with heterogeneous bargaining power is the Nash solution, which features constant returns to bargaining power. In the legislative bargaining literature, recently reviewed by Eraslan and Evdokimov (2019), an increase in bargaining power may be detrimental (e.g. Acemoglu et al., 2008).

[28] Quote from Coke (1644) Chapter 1, p.35. The author reports a sentence in Latin: "Eritis insuperabiles, si fueritis inseparabiles. Explosum est illud diverbium; Divide, & impera cum radix & vertex imperii in obedientium consensu rata sunt", which can be translated as follows: "You would be invincible if you were inseparable. Once the subjects' collegial agreement became the foundation and summit of authority, the proverb Divide and Rule have been defeated".

[29] "It is observed by ancient Parliament men out of Record, that Parliaments have not succeeded well in five Cases. 1. when the King has been in displeasure with his Lords, or with his Commons. 2. When any of the Great Lords were at variance between themselves. 3. When there was no good correspondence between the Lords and the Commons. 4. When there was no unity between the Commons themselves. 5. When there was no preparation for the Parliament before it began" (Coke, 1644, Chapter 1, p.35).

actions against the central authority (Valente, 2003). Another typical example is the case where multiple subjects form an horizontal alliance thus promising to stand for one another: such pact is at the basis of communes or leagues, which were very common in Europe during the Middle Ages and the Early Modern period. An example from Ancient history is the figure of the Tribune of the Plebs in Ancient Rome. Such figure 'represented' the Roman plebeians as he had the right to interfere in any action against a plebeian taken by a magistrate (i.e. a patrician), and the tribune was backed by the threat of collective action as all plebeians were supposed to act together to either protect the actions of the tribune or avenge any damage he suffered. Another example, this time from Medieval history, is Clause 61 of Magna Carta which required all barons to act together in case the King did not uphold the provisions of that charter (irrespective of who, among the barons, was hurt by such violation).

While subjects have attempted to create solidarity links, rulers were often busy trying to prevent or suppress them. This attitude can explain the various episodes of repression against new identity groups and/or the segregationist policies that have been employed universally, from Ancient Rome to China, to the European colonial empires. For instance, when Ancient Rome consolidated its power on the Italian peninsula and then in the Mediterranean basin, it regulated which community was allowed to trade or marry outside its administrative boundaries (Sherwin-White 1973, p.113). While this has certainly reduced the overall efficiency of trade (and thus Rome's tax revenues), such policies can be understood as an attempt to avoid the formation of a common identity or alliances that could unite otherwise separate subjects. A few centuries after the collapse of the Roman Empire, the Tang dynasty of China implemented similar policies in their Empire, as they imposed on the seven most prominent clans a ban on inter-clan marriages (Tackett 2006, p.67; Johnson 1977, pp.50-51). Finally, it is certainly well known that the European colonial empires adopted segregationist policies in most of the lands that fell under their control. A more interesting fact that directly links this practice to the mechanism we propose emerges from the reaction of the British

officers to the Indian mutiny of 1857: they proposed to reform the army and create separate units for each Indian ethnic and social group to avoid that national feelings could resurface in the future (Stewart 1951, p.53). Interestingly, even today we often observe countries to feature stark ethnic divisions in the organization of the armies: our analysis suggests that it might be a way to avoid the formation of cross-group solidarity.

Finally, various examples can be put forth to support the proposition that newly-formed solidarity links can destabilize society. In the context of Early Modern Europe, the Protestant Reformation of the XVI Century clearly shows how the rise of a new identity may lead to widespread social unrest (see e.g. Zmora 2002, Chapter 4). Another example is the aforementioned Indian Mutiny of 1857 against the British Colonial Empire. British officers attributed the event to an insufficient attention in avoiding the development of nationalistic feelings among the various ethnic and social groups residing in the Indian subcontinent (Stewart, 1951).

2.4 Case Study: Louis XI of France and the League of the Public Weal

We now present a detailed case study from Late Medieval and Early Modern Europe that provides direct evidence in favor of the mechanism that we propose. The example is representative of other similar episodes in many ways.[30]

Between 1461 and 1483, France was ruled by Louis XI of France, a King whose historical reputation as a Machiavellian villain persists to the current days (Bakos, 2013). Thanks to

[30]On the very same year, in Castile, we have the farce of Ávila, "a not dissimilar confrontation between king and nobles" (Zmora 2002, p.40). As Zmora (2002) remarks, the similarity of the episodes is due to the similarity of the forms of government in Castile and France. In particular, such societies "were already possessed of some of the hallmarks of 'absolutist' monarchies", including the legitimacy of royal taxation and a standing army. Yet, the magnates could still cover a position of power thanks to their territorial control which presented them as "mediators between king and populace" (Zmora 2002, p.41).

the large availability of sources and the King's particular character, his well-documented political career reveals a lot of the political workings of the societies of Late Medieval and Early Modern Europe. A lesson that was not forgotten by near contemporaries such as Charles V, Holy Roman Emperor, who would refer to Louis's earliest biography as a "textbook for Kings".[31]

Already in the XV Century, the King of France was a powerful figure, as with the exception of "limitations dictated by his own prudence or the recalcitrance of subjects, the royal prerogative remained virtually unlimited" (Kendall 1971, p.24). After ascending the throne, Louis XI set about modernizing the central administration and expanding royal prerogatives. This reforms brewed resentment that was visible even outside of France. In 1463, Cosimo de' Medici, the famous ruler of Florence and a friend of Louis, told a Milanese ambassador that "His Majesty [Louis XI of France] will have great trouble and more than he would like, because he rules his realm to suit his own ideas" (Kendall 1971, p.129). A prediction that became reality two years after. In 1465, the magnates announced that they joined against their King in the League of the Public Weal. By September 1465, the King of France was besieged into his capital by a united front of the magnates of the realm.

At the beginning of September 1465, the besieged King started negotiating with the magnates. From the very start of the negotiation it was immediately clear that his strategy would be to "deal with [the magnates] singly, a strategy which heightened their mutual suspicions and lowered their individual resistance" (Kendall 1971, p.183). During the very first week of negotiations, the magnates attempted to signal their solidarity by formally assembling and swearing "a solemn oath to sign no separate appointments with the King" (Kendall 1971, p. 178). Despite this public display, Louis was "well aware of the pressures

[31] The quote on Charles V is from Kendall (1971) p.253. Our discussion is heavily based on Kendall (1971). See Kendall (1971) p.27 for a summary of the primary sources that he used. Appendix I (pp.378-381) of Kendall (1971) offers a detailed discussion of the reliability of the Memoires of Philppe de Commynes, the earliest biography on the political career of Louis (for a recent translation, see De Commynes 1972).

straining the unity of the princes" and "he realized that he must find the means to divide them" (Kendall 1971, p. 178).

By the end of September, the King put his strategy in motion by setting himself to "amply fulfilling the expectations of one member of the League in order to separate him from the others" (Kendall 1971, p. 178). Louis considered that only two magnates had the "authority to serve his turn, the Duke of Brittany and the Count of Charolais" (Kendall 1971, p.178), the two strongest elements of the coalition of magnates. Louis decided to approach the Count of Charolais, the man whose army he fought in the bloody battle of Monthéry two months before.

Before the end of September, Louis XI of France met his subject Charles, Count of Charolais (the future Charles the Bold, Duke of Burgundy). Initially, Charles repeated "the standard demands of the princes." However, after the end of the meeting, the Count of Charolais believed that as "he was indispensable to his confederates, they must expect his own demands to be satisfied first and be willing to be guided by his wishes regarding peace." Louis' strategy was working. The Count of Charolais, the most powerful adherent of the League was now pushing the others "to come to some kind of compromise terms with the King"; only the others though, for "there was no question of compromising his own demands, the justice of which Louis XI now recognized so gratifyingly" (Kendall 1971, p.180). In other words, "the unnatural unanimity of the princes was beginning to crack (..) Not without reason, they already distrusted each other's appetites and feared the King's uncanny arts of persuasion" (Kendall 1971, p. 178).

When Louis met Charles for the second time, "[t]he Count not only accepted the King's terms on the spot but promised to secure the adherence of his confederates" (Kendall 1971, p.182). With the most powerful member of the League on his side, Louis carried on with his strategy. On October 10, the King singled-out the Duke of Brittany making an offer that the Duke gladly accepted. With Charolais and Brittany pushing for peace, the King

could now bargain with the other subjects from a position of strength. When targeting the Count of Armagnac, Louis only offered a pension and a minor command in the standing army. Caught "in a violent rage" against his "more fortunate allies", the Count of Armagnac accused: "under the pretense of assembling for the public weal of the realm, each had come (..) for his own interest". Charles, the Count of Charolais, responded that "he should be content with the offer that had been made him, which was more than he deserved" (Kendall 1971, p.184).

When the King dealt with the Duke of Nemours, he found that he was "happy to receive from the King a few crumbs of privileges" (Kendall 1971, p.184). To the magnates belonging to the house of Valois-Anjou, the King offered a subsidy paid over several years, a small help for their campaign in Southern Italy. On the other hand, Duke John, of the House of Valois-Orleans, was simply informed that the King would not accept any of his requests. Towards the end of October, with an advantageous agreement in their hands, the Burgundians and the Bretons were preparing to leave. Pushing everyone to their limits, on October 28 Louis did not present a draft of his proposals, but rather a formal document: "it was take-it-or-leave-it" (Kendall 1971, p.185). By the next day, the "unhappy magnates" accepted the offer as it was, making it the Treaty of Saint-Maur-des-Fossés (Kendall 1971, p.185). Peace was established, all actions were pardoned and the common interest had succumbed to the magnates' individual appetites, or as Commynes, our main source, remarks "the Bien Publique [the public good] had been converted into the Bien Particulère [individual interests]," (Kendall 1971, p.179). On November 3, the League was dissolved. "At the cost of satisfying the appetite of the Count of Charolais, Louis had succeeded in imposing, on princes sworn to have his skin, a treaty which dismissed their armies by concessions that mocked their demands" (Kendall 1971, p.186).

2.4.1 Discussion

The case study offers key insights into the forces that underline the political equilibrium of XV Century France and, arguably, Western Europe. While Louis XI was considered an excellent negotiator by his contemporaries, the final settlement that he forces onto his magnates is essentially a return to the status quo that prevailed under his father, before his 'reckless' reforms attempts. This suggests that the strategy used by Louis XI might be a good description of the logic behind the prevailing political equilibrium in France at the time.

In general, there is no reason to believe that Late Medieval France was special in terms of the determinants of the political balance between rulers and magnates. Indeed, observers from as diverse backgrounds as Sir Edward Coke, Immanuel Kant and James Madison have all described "divide et impera" or "divide and conquer" as a normal practice or crucial strength in how monarchs dealt with their magnates.[32]

The case study offers an opportunity to discuss how the three facts presented in Section 2.3 can all be ascribed to the divide-and-conquer nature of the policies of central authorities, just as our model proposes. The bargaining strategy of Louis shows the possible role played by *orders* (e.g. ranks, systems of precedence, etc.) in determining the political equilibrium. In spite of the magnates' oath to stand for each other's requests, Louis made sure to deal with them singly, offering separate agreements to each of them. The King granted all the requests that directly involved the strongest magnate. With the strongest (his favorite) on his side, he could then somewhat erode the requests that pertained to the second strongest (his second-favorite). With both the strongest and the second-strongest pushing for peace,

[32]See Section **??** for the relevant reference from Coke (1644). Kant includes "divide et impera," or to "destroy the unity" of the magnates and to "separate them from the people", as one of the maxims of tyranny (Kant 1983, p.130). Kant even claims that such strategy is so well known that it might not work (Kant 1983, p.131). In a letter to Jefferson on October 24, 1787, Madison famously described "divide et impera" as the "reprobated axiom of tyranny," see Madison (1787).

the King could deal with the remaining magnates from a position of increasing strength. As predicted by our model, the strategy of making offers conditional on relative status successfully dissolved the unity of the magnates, which led to a treaty that frustrated the expectations of all but the strongest elements of the coalition. The pervasiveness of rankings may simply be a way to facilitate or emphasize the adoption of similar strategies by central authorities. This idea, which is formalized by our model, is also consistent with David Hume's interpretation of the assignment of ranks and privileges, as he said that such practice could be an appropriate way to impose social order and cement the position of the ruler (Miller, 1980).

From a general perspective, it is natural to expect that more powerful groups obtain greater concessions because they have more bargaining power. However, as our model predicts, Louis assigned precedence in negotiation based on the magnates' relative powers.[33] Such strategy further aggravates any natural form of social stratification, as not only more powerful groups obtain better deals, but also more bargaining power. This is at the basis of the idea that government policies may *induce* social stratification: even if two magnates were identical, the ruler would still assign precedence to one of them, thus arbitrarily assigning him more bargaining power (thus a better deal).

The case study also offers an opportunity to discuss the link between social unrest and favoritism. First of all, note that the case study is an example of what we called gambling against solidarity in Section 2.2.5. The ruler ignored the displays of solidarity, even if this could have led to an armed confrontation. Additionally, note that, by design, the strategy of

[33]We have seen how Louis was sure that he would start his negotiations from either Charles or the Duke of Brittany, as they alone had the "authority to serve his turn", i.e. were sufficiently powerful to fit his strategy. In applications we expected the assignment of priority in negotiations to be in large part determined by a history of friendship or trust. However, this expectation is clearly subverted by Louis's strategy. Indeed, only a few months before, the King's army had clashed against Charles' army in a bloody battle, where both Louis and the Count lost many of their comrades. The Duke of Brittany was the subject whose treacherous dealings with England had risked opening a new phase of the Hundred Years' War, which was probably averted only because the English Crown was itself in a deep crisis, the devastating War of the Roses (1455-1485).

Louis made sure that if negotiation broke down at any point in time, then the King would have faced the military opposition of a coalition of only the lowest-ranked subjects, which is in line with the general pattern on revolts of Section 3.6.2. The case study also clarifies the nature of the jealousies that are often proposed as an explanation for the conflicts between magnates (see Section 3.6.2). The fact that the bargaining power of each magnate is based on the power of all the lower-ranked magnates can explain both the bitter reaction of the Duke of Armagnac against his "more fortunate allies" and the response of Charles, who noted that the Duke obtained more than he deserved.[34]

2.5 Conclusions and future directions

When lacking a monopoly over violence, central authorities ordinarily reinforce social stratification with regressive policies that disproportionately favor some social groups at the expense of others. While such practice is universally adopted, it is puzzling as many historians and political scientists have considered it a driver of political instability. As central authorities have an interest in maintaining social stability, why would they resort to such practice? To answer this question, we developed a model of a society where social groups can use their military power to resist government policies (social unrest). We found that, to discourage the formation of resisting coalitions, the central authority creates a complete (order) ranking among the subjects and adopts targeted policies that assign higher political rents to higher-ranked subjects. Additionally, we have shown that when social groups are heterogeneous, then such policies would appear to reinforce social stratification, as the government would optimally assign higher status (thus lower tax rates) to richer (thus stronger)

[34]The frustration of the Duke is natural given that the subjects that Louis singled-out before him (the higher-ranked subjects) obtained greater concessions by leveraging on the Duke's military resources as well as their own. Charles' response can be understood by noting that the Duke was not the last in the ranking and thus he also obtained "more than he deserved", in the sense that his bargaining power did not only reflect his own power, but also the military power of the magnates that did not obtain an offer from Louis before the Duke (i.e. the magnates ranked below the Duke).

groups. Finally, we expanded the analysis studying how the formation of cross-group solidarity can generate social unrest and how the pattern of social unrest appears to be linked with political favoritism.

While the taxation game analyzed in Section 2.2.1 is relatively simple, we argue that its main insight could be relevant for other applications. Indeed, from a formal point of view, the game simply captures a situation where a proposer makes an offer to a group of respondents, who choose whether to accept it or not. The fundamental innovation to this traditional setting is the existence of an option to "resist": an outside option whose value increases in the amount of respondents that take it. From a conceptual point of view, such outside option captures what Hirschman (1970) calls "voice" which, in contrast to "exit", is a largely understudied, yet pervasive, feature of economic interactions. With minor tweaks, our taxation game could be used to study problems as varied as a firm offering wages when its workers can go on strike, a government trying to regulate interest groups, a manager motivating his or her workers who can boycott the success of his or her projects, a producer setting quality when consumers can voice their complaints to a regulating board or to other potential consumers (one of the many examples in Hirschman, 1970).[35]

[35]Our primitives, W_i and B_i, can be interpreted in flexible ways. For instance, W_i is just a general "power" variable (power of a lobby, number of votes a group is able to sway, credibility in the eyes of other consumers, etc.). Similarly, B_i is a general "stakes" variable: it allows for heterogeneity in how much a group has to lose from opposing the proposer. Applications that consider redundant the presence of both W_i and B_i can easily shut down one of the two by assuming that it is equal for every decision maker.

Chapter 3

Internal Conflicts and the Political System

3.1 Introduction

Conflicts among members of the same society permeate human history and, even today, represent the vast majority of armed conflicts.[1] Depending on the parties involved, the size of the events, and how violence is expressed, internal conflicts are categorized in various ways: civil wars, non-state conflicts, genocides, domestic terrorism, riots, communal violence, etc. Abstracting from their qualitative and quantitative differences, all these events induce a change the distribution of resources in the society, via resource appropriation or, more often, destruction. This paper studies internal conflicts as the decisions of social groups to use their violence potential to force a change in the distribution of resources.

Because internal conflicts have a deep interrelation with a country's development, its social fabric and political system, the study of conflicts is an inter-disciplinary effort. While

[1] From 1945 to 1997 there were 108 civil wars, with 11.4 million deaths and an average duration of 1.665 days (Levy and Thompson, 2010). In each year between 1000 and 1800 Europe experienced internal conflicts (Kokkonen and Sundell, 2020).

economists study conflicts as motivated by "greed", a cost-benefit analysis of the value of conflicts, political scientists often stress the importance of "grievances". A major distinction between the two approaches is the role played by the political system, in particular government policy. Most models of greed typically abstract away from it, on the other hand government policy is often identified as a leading source of grievances. In particular, a recurring theme in the grievances literature is that conflicts are motivated by government's favoritism: the preferential treatment of particular social groups at the expense of others.[2]

By overlooking the role of government policy, the greed literature may be missing an important part of the story, which is often highlighted by in-depth narrative accounts of specific conflicts (Collier and Sambanis, 2005b): what is the impact of government policy on the incentives for in-fighting? On the other hand, the absence of a formal model in the grievances literature creates ambiguities on the mechanism linking government policy, in particular discrimination, to conflicts: when are conflicts a good strategy to obtain a more favorable treatment from the central authority? Our paper aims at bridging the gap between the greed and grievances literatures by studying conflict in a model where payoffs are shaped by government policy.

We reflect the findings of Chapter 2 by considering two institutional settings which we call a United and a Hierarchic society. In a United Society, the payoff of every subject is increasing in the violence potential of every other subject and decreasing in the ruler's. In a Hierarchic Society, the payoff of each subject is increasing in the violence potential of weaker subjects, and decreasing in that of stronger subjects and the ruler. We first study how conflicts might be motivated by the groups' ambition to become the next ruling group. To do so, we allow social groups to select the identity of the ruling group. While in a United

[2]For reviews of the theoretical greed literature see, for instance, Baliga and Sjöström (2013) or Jackson and Morelli (2009). Reviews that discuss both the greed and grievances literature include Blattman and Miguel (2010), Collier and Sambanis (2005b) and Sambanis (2004). Note that not all models of "greed" abstract away from government policy and not all of the grievances literature focuses on government policy. We discuss our relation to the literature more precisely in the next section.

Society we find that the weakest group should be expected to become the ruler, we show that in a Hierarchic Society groups would optimally select the strongest as their ruler. As a consequence, in a United Society, conflicts are of limited use for a group who aims at the rulership. On the other hand, in Hierarchic Societies, the two strongest groups have a strong incentive to fight each other, as becoming the strongest group allows them to become the next ruler. We then turn to study the motivations that can drive conflicts between subjects and how they differ in a United and Hierarchic Society. In a United Society, conflicts can only be motivated by resources appropriation. As a consequence, conflicts can only occur when some groups are too (relatively) weak to effectively defend themselves, and when conflicts are sufficiently profitable. On the other hand, in Hierarchic Societies, we find that conflicts among subject can occur even when they are not profitable, and we show that they are particularly likely to occur when two subjects have similar levels of resources. We discuss a case study from Early Modern Europe and address the evidence on both modern and historical settings to show that the concept of Hierarchic Society is useful to make sense of the observed patter of internal conflicts.

Our model considers a society fragmented in different social groups, the decision units of our framework.[3] Each group i is characterized by an endowment of resources B_i. This resource has two functions. First, it can be used for consumption, and it thus contributes directly to the group's final payoff. Second, it determines the social group's violence potential or *power*, i.e., its ability to win a conflict. We assume that a groups violence potential is $g(B_i)$, a continuous, increasing and convex function of its endowment B_i. The initial allocation of resources is common knowledge and possibly heterogeneous across different social groups. Social groups coexist within a society that can be of two types: Hierarchic or United. In

[3]Depending on the setting, we can think of such groups as ethnic or identity groups, warlords with their armed sequitur, clans, geopolitical sub-units within empires, kingdoms or principalities, or even organized interest groups such as guilds or corporations within the same geopolitical unit. See Chapter 2 for a more detailed discussion.

a Hierarchic Society, the final payoff of each subject (non-ruling group) is decreasing in the sum of power of the ruler and all stronger subjects, and it is increasing in the sum of power of all weaker subjects. In a United Society, the final payoff is increasing in the sum of power of every subject and decreasing in the ruler's power. These different types of societies capture different institutional arrangements, as discussed in Chapter 2. We assume that the type of society where social groups coexist is common knowledge and fixed.

Our baseline model unfolds in two phases: a first phase of conflict, where groups can fight each other to change the distribution of resources, and a second phase where social groups select the ruling group. After the ruling group is selected, payoffs are realized. This time structure allows us to capture how a different institutional arrangement (Hierarchic vs United Societies) affects the link between distribution of resources and ruler identity, and thus how conflicts can be used to become the next ruling group.

We start by studying the ruler selection phase. First, note that in both a United and Hierarchic Society the power of the ruler has a negative impact on the social groups' payoffs. As a consequence, in both cases, every subject strictly prefers any ruler to be as weak as possible. In a United Society, this logic is further compounded by the fact that social groups benefit from the power of every other subject. As a consequence, not only the stronger is the ruler the worse off is a group, but also assigning the rulership to a stronger subject deprives each subject of a stronger ally. Thus, in a United Society, we show that social groups prefer to assign the rulership to the weakest possible social group. In contrast with this, in a Hierarchic Society it is the strongest group who is able to gain the largest consensus and become the ruler. Indeed, in a Hierarchic Society, each social groups is positively affected by the power of weaker groups. Thus, no group would ever support the bid for power of a weaker group. This leaves the two strongest groups as the only possible pretenders for the rulership. Specifically, if the ruler selection procedure is correlated with the level of resources of the contenders, then the strongest social group would always gain the rulership.

We then turn to the conflict phase of our model, studying how the incentives for groups to appropriate/destroy each other's resources depend on the result of the rulership phase. We make two main assumptions. First, conflicts are unilateral decisions of the attacker. Second, the probability that the attacker wins a conflict is increasing in its power and decreasing in the power of the defender. The finding that the weakest group becomes the ruler in a United Society suggests that conflicts are an unlikely strategy to gain the rulership. Thus, in a United Society, conflicts against the current ruling group can only be motivated by trying to making it even weaker than it is, and possibly appropriate valuable resources from it. On the other hand, as the strongest social groups becomes the ruler of a Hierarchic Society, we should expect conflicts to be particularly likely to occur when the two strongest groups have similar levels of resources. Indeed, we show that when the initial endowments of the two strongest group is close enough, then the weaker of the two would always want to fight the strongest, in an attempt to gain the rulership by becoming the strongest group (hegemonic conflicts). As we discuss, this analysis speaks to the evidence on civil wars from post-colonial Africa. First of all, it can explain why countries with representative institutions (the bedrock of United Societies) are usually less prone to civil wars. Second, the analysis offers a novel rationalization as to why a condition of relative parity between the two largest groups in society (polarization of resources) should be expected to correlate with civil wars.

To study how different institutional settings (Hierarchic vs United Society) shape the pattern of infighting, we turn our attention to the possibility that groups fight each other, even when the rulership is fixed. This can represent a setting where the ruler has not yet imposed a monopoly on the legitimate use of violence, so that groups can legitimately attempt to appropriate/destroy each other's resources. In a United Societies, the incentives to fight another subject are straightforward. As the final payoff of each subject is increasing in the power of all the other subjects, then groups will internalize any destruction of resources that results from their attack. For this reason, a subject will only attack another subject

when the value of the appropriated resources and the probability of victory are high enough to cover the costs of participating in the conflict, and the indirect cost that results from the possible weakening of another subject. As we show, this argument implies that conflicts are driven by the probability of victory, and thus conflicts should only occur when resource appropriation is sufficiently profitable and when social groups are more unequal.

The incentives to fight in Hierarchic Societies are much more complex. Recall that, in such societies, the final payoff of each subject is increasing in the power of all weaker subjects and decreasing in the power of all stronger subjects. As a result, becoming stronger than another subject would induce a jump up in the subject's payoff. In line with the terminology adopted in Chapter 2, we call such discontinuity a *status gain*. Because gaining status has value, the payoff structure that prevails in Hierarchic Societies generates a new driver of internal conflicts that can be fulfilled even when conflicts can only destroy the resources of the opponent. Indeed, since status gains depend on relative rather than absolute changes in endowments, the objective to gain status can even rationalize conflicts where both parties always end up with less resources than before, i.e. where the cost of fighting is disproportionate to any possible direct material gain. As it turns out, this prediction is not only a theoretical possibility. Studying the endemic infighting among the nobility of Early Modern Germany, Zmora identifies such disproportion as one of their most salient features. In agreement with our mechanism, the same historian notes that the gap between direct gains and costs can be accounted for once we consider that "many feuders were appointed to high princely offices during or immediately following their feuds," concluding that such conflicts "brought them to the fore and then closer to the prince (..) [and] [t]heir status rose concordantly" (Zmora 2003, pp.116-117). To the best of our knowledge, our theory provides the first formal and rational-agent treatment of the notion of status-related conflicts which is commonly used in fields such sociology, criminology, anthropology, biology and international

relations.[4]

Finally, our analysis unveils a novel empirical predictions by combining the peculiar nature of status conflicts with the standard assumption that the probability that a group wins a conflict is increasing in its relative military power. In line with the evidence from contemporary societies, we find a U-shaped relationship between the similarity of the relative endowments of two groups and their likelihood of fighting. On the one hand, when conflicts can lead to gain resources, the incentive to fight is always greater against a weaker subject (an easier prey); thus, as it is standard in the conflict literature, inequality of powers should motivate infighting. On the other hand, when two groups are sufficiently close in terms of resources, the weaker of the two has an extra-incentive to fight: even a small change in their relative endowments might lead to a status gain, thus a large change in its final payoff.

To conclude, note that, just as hegemonic conflicts, status-related conflicts would only occur in Hierarchic Societies. Thus, our analysis highlights a new mechanism whereby institutional arrangements that foster society's unity (e.g., representative institutions) can be valuable: not only they may act as checks on the government's ability to expropriate its subjects, but also they may reduce the incentives for groups' infighting, and thus the incidence of internal conflicts.

3.1.1 Contribution to the Literature

Conflicts attracted the attention of scholars across different disciplines. In this section, we review the most closely related literature and discuss our contribution to the understanding

[4]There is "a strong consensus that status is a critically important human motivation" (Renshon 2017b, p.6), and status has been described as a "universal feature of social groups" (Cheng et al. 2013, p.104). The concept is used by criminologists, anthropologists and biologists to describe conflicts within street gangs, youth gangs or social groups of animals, see e.g. Wrangham and Wilson (2004), Papachristos (2009), Franck (1985), Barkow (1989), or Cheng et al. (2013). It is also widely used by sociologists, international relations scholars and historians to discuss internal conflicts and both the pattern of interstate wars or specific episodes such as the First World War, see f.i. Renshon (2017b), or Chapter 8 in Tortella (2010). Philosophers and political scientists who have discussed these types of conflicts ascribed them to (intrinsic) preferences for relative status, "will to power" in Nietzsche (1967) and "animus dominandi" in Morgenthau (1948).

of internal conflicts.

To the best of our knowledge, our theory provides the first formal treatment of the notions of status-related and hegemonic conflicts, i.e. conflicts aiming at becoming stronger than the opponent or the strongest group in society.[5] Philosophers and political scientists who have discussed these types of conflicts ascribed them to preferences for relative status, in Nietzchean terms "will to power" (Nietzsche, 1967) or "animus dominandi" according to Morgenthau (1948). Our theory highlights the crucial link between status-related conflicts and Hierarchic Societies. On top of such theoretical contribution, status-related conflicts produce novel empirical predictions on how conflicts may be triggered by income differences (or similarity), which Ray and Esteban (2017) identifies as a priority for future research. In particular, to our knowledge, we are the first to highlight that in Hierarchic Societies there is a robust *policy-induced* reason as to why "conflict is more about the change in the relative income status of two similar groups rather than the overall level of inequality" (Ray and Esteban, 2017, p. 280).[6]

A large literature has focused on internal conflicts, both from a theoretical and empirical perspective. Most of the theoretical contributions focus on conflicts as a way to change the identity of the ruler (Besley and Persson, 2011; Morelli and Rohner, 2015; Esteban and Ray, 2011). Unlike this literature, we view conflicts only as a way to alter the distribution of resources in the society: conflicts can at most lead to the destruction or appropriation of resources. This approach, which is in line with the "conflict motives" literature reviewed in Baliga and Sjöström (2013),[7] allows us to consider, in a single framework, also conflicts not

[5]The concept is used by anthropologists and biologists to describe conflicts within youth gangs or social groups of animals (Wrangham and Wilson, 2004). It is also widely used by sociologists and historians to discuss internal and inter-state conflicts, such as WWI (e.g. Chapter 8 in Tortella, 2010).

[6]We explicitly address the empirical evidence of Esteban et al. (2015), Montalvo and Reynal-Querol (2005), and Mitra and Ray (2014, 2019) in Section 3.6.2. Our model explains some of the evidence as a consequence of government policy (favoritism). This is complementary to the current literature that only considers how changes in income affect benefits and (opportunity) costs of conflicts (Bates et al., 2002; Dal Bó and Dal Bó, 2011) or investments in defensive equipment (Mitra and Ray, 2014).

[7]Such literature, also reviewed in Blattman and Miguel (2010) and Jackson and Morelli (2009), mostly

directly aimed at a rulership change, which are often observed but did not yet receive much attention (Sundberg et al., 2012). Such conflicts, for instance the Hindu-Muslim conflicts studied in both Mitra and Ray (2014) and Mitra and Ray (2019), are less likely to make the headlines than outright civil wars. Nonetheless, they are potentially very harmful for a country's social fabric and its local development.

While we model conflicts as only a way to change the distribution of resources, the pattern of conflict emerging from our model is crucially shaped by the political framework. From this point of view, our model contributes to understanding the interrelation between political institutions and internal conflicts. While the literature that studies the impact of (potential) conflicts on political institutions and policy is well developed,[8] our paper is one of the first to focus on the opposite direction: how political institutions (and thus policy) shapes the pattern of internal conflicts. Other papers taking a similar approach are Besley and Persson (2011) and Esteban et al. (2015). Besley and Persson (2011) studies how the incidence of coups and repression depends on the ruler's decision between public good provision and pork barrel spending. Esteban et al. (2015) studies how institutional constraints to unfairness in surplus sharing may create the incentives for the ruling group to engage in mass-killings. Our main results are entirely complementary to these findings, as we highlight a novel mechanism linking political institutions to internal conflicts: hierarchies. With respect to Besley and Persson (2011) and Esteban et al. (2015), which confine their attention to societies made up of two groups (government and opposition), our model allows for multiple, possibly heterogeneous groups. This enables us to express our conflict predictions as a function of the whole distribution of resources in society, a research direction advocated by Alesina et al.

studies why groups fight rather than bargain and does not specifically address internal conflicts.

[8] A relatively small literature studies how policy-making is affected by the risk of regime changes (Acemoglu et al., 2004; Francois et al., 2015; Padró i Miquel, 2007). A larger literature, entirely complementary to our work, studies how political institutions emerge as a consequence of (potential) conflicts (Weingast, 1997; Greif, 1998; Acemoglu and Robinson, 2000; Bates et al., 2002; Myerson, 2008; Besley and Persson, 2009, 2010; Ticchi and Vindigni, 2008; Fearon, 2011; Dincecco et al., 2011; Gennaioli and Voth, 2015).

(2016).

In the study of internal conflict, behavioral explanations have received particular attention. One possible reason is that many modern internal conflicts oppose distinct ethnic groups. Since ethnicity is a non-economic marker, this suggests that such conflicts are driven by hatred, not interest (Horowitz, 1985). Models with behavioral economics features have connected internal conflicts and polarization of resources, finding strong support in the data (Esteban and Ray, 1999; Montalvo and Reynal-Querol, 2005; Esteban et al., 2012a). We derive a similar prediction, without deviating from the standard rational-agent framework. By linking ethnic conflicts to the government policy of Hierarchic Societies, our model can rationalize the evidence that polarization of resources only matters when interacted with institutional features (Esteban et al., 2012b) and that political institutions may effectively curtail ethnic violence (Reynal-Querol, 2002).[9]

Another behavioral literature, especially popular among political scientists, motivates conflicts as based on grievances (Collier and Sambanis, 2005b). This literature is connected to our work as it often sees government's favoritism (in particular ethnic-biased policies) as an important source of grievances (e.g. Boucekkine et al., 2019).[10] Just as this literature, we show that favoritism does indeed create the incentives to start conflicts. However, we show that the motivation to fight may come from the expectation to gain a more favorable ethnic-biased policy in the future, at the expense of the defeated ethnic group. Unlike grievances based on past policies, such forward-looking motive is linked to expectations about future policies and institutions. As a result, our theory enables us to show how a society can move beyond a period of ethnic-biased policies without internal conflicts, or go through internal

[9]Critics of the primordialist view, as Esteban et al. (2012b), still often make the reduced-form assumption that ethnic groups have somewhat incompatible preferences. Our contribution is to get rid of all preference-related considerations as we link conflicts among ethnic groups to the political logic of ethnically-divided Hierarchic Societies.

[10]A classic example in the context of post-colonial Africa is that the Rwandan genocide was caused by hatred created by colonial policies that escalated tensions between Hutu and Tutsi (King, 2013). Collier and Sambanis (2005b) makes a similar statement for Nigeria and Cyprus.

conflicts without an history of ethnic-biased policies.[11]

3.2 Conflicts and Government Policy

Our model considers $n+1$ groups, each characterized by economic resources B_i and military resources $W_i = g(B_i)$, where $g(\cdot)$ is a continuous, increasing, and convex function.[12]

3.2.1 Conflict Technology

A conflict occurs when a group, the attacker A, unilaterally chooses to attack another group, the defender, D.[13] To simplify the analysis, we abstract away from the complexity generated by multiple and possibly interrelated conflicts. In particular, we assume that, when choosing whether or not to start a conflict, each group does not internalize the possibility of other conflicts. Such assumption would be consistent with groups being naïve in their conflict decisions as well as with a model in which at most one group has the opportunity to start a conflict.

We model conflicts as a costly way to induce a change in the distribution of resources $\mathbf{B}$, i.e., conflicts are a costly way for a group to appropriate or destroy the resources of another group. We allow for the attacker to choose the stakes (size) of the conflict: a real number b which determines the maximum amount of resources that can be affected by the conflict.

[11] For instance, our theory may help explain why ethnic conflicts broke out after the collapse of the USSR even if ethnicity was not an important dimensions of Soviet policy-making, as Marxist ideology emphasizes class struggle.

[12] See Chapter 2 for a discussion of this assumption.

[13] While studying conflict as the unilateral decision of the attacker is in line with Bates et al. (2002) and Mitra and Ray (2014), note that this assumption is a deviation from the literature that studies internal conflicts with a general-equilibrium approach. Such literature, which is pioneered by Grossman (1991) and also includes Dal Bó and Dal Bó (2011), models a society where productive activities coexist with a resource-appropriation sector, and interpret the equilibrium allocation of labor to the predatory sector as a measure of the incidence of internal conflicts. While this approach offers rich empirical predictions on how economy-wide conditions make it easier or harder to recruit soldiers, it is not ideal for our focus on how the inter-group distribution of resources may trigger internal conflicts. A combination of the two approaches might provide valuable insights to bring to the data.

We assume that participating in a conflict implies a cost in terms of utility that is a function of the conflict's size, $\kappa(b)$, where $\kappa(0) = 0$ and $\kappa(\cdot)$ is a continuous, differentiable, increasing and convex function for $b \geq 0$. As we will discuss, our results are not driven by the possibility of very cheap small conflicts, so that our main results would also hold for $\kappa(0) > 0$.

As conflicts change the distribution of resources, we must make assumptions on what we call the *conflict technology*, i.e., how resources change after a conflict. To simplify the exposition, we assume that a conflict has two possible outcomes: either the attacker wins or loses. With this, a conflict technology can be interpreted as the answer to two questions: what is the attacker's probability of victory? How do resources change when the attacker wins and when it loses?

We assume that the probability that the attacker wins is a function increasing in the attacker's power and decreasing in the defender's power, $\Pr(A \text{ Wins}) = p(W_A, W_D)$. For the second question, we assume that every conflict induces a (possibly inefficient) transfer of resources from the losing to the winning side that is increasing in the stakes of the conflict, b.

Simple Example

Before introducing the general class of conflict technologies that we consider, we start from a simple example.

Suppose that when the attacker wins a conflict of size b then it manages to appropriate b resources from the defender while suffering a destruction of γb resources (where $\gamma \geq 0$). On the other hand, the defender loses b and suffers a destruction of δb resources (where $\delta \geq 0$). Note that we also need to specify how resources change when the defender wins. In this simple example we can just assume that, when the defender wins, there is no appropriation of resources and there is no destruction, i.e, resources are unchanged for both attacker and defender.

Letting ΔB_i represent how resources change for group i, we can summarize how the attacker's resource change with the following expression:

$$\Delta B_A = \begin{cases} (1-\gamma)\bar{b} & \text{w.p. } p(W_A, W_D) \\ 0 & \text{w.p. } 1 - p(W_A, W_D) \end{cases}$$

While from the perspective of the defender the consequences of a conflict are represented by the following stochastic change in resources:

$$\Delta B_D = \begin{cases} -(1+\delta)\bar{b} & \text{w.p. } p(W_A, W_D) \\ 0 & \text{w.p. } 1 - p(W_A, W_D) \end{cases}$$

First, note that when $\gamma > 1$, then the attacker will end up with less resources than it had before the conflict, even after winning. This is a class of conflicts that we will call non-profitable conflicts. One of our main results will be to show under what conditions we can expect non-profitable conflicts to take place.

Second, note that, when δ increases, conflicts become more destructive in a way that does not affect the attacker. We will study how the incentive of the attacker are affected by changes in δ.

Finally, note that p is increasing in W_A and decreasing in W_D. This suggests that conflicts should be more likely when the relative power of the attacker increases. While this correctly suggests that inequality triggers conflict, we will also show that, under some conditions, equality can be a trigger for conflicts.

General Case

In the general case, the gain/loss in resources for a group $\ell \in \{A, D\}$ when group $s \in \{A, D\}$ wins the conflict is given by $B_\ell^s - B_\ell = \Delta B_\ell^s(W_A, W_D, b)$, where we assume that ΔB_ℓ^s is increasing in W_ℓ and decreasing $W_{-\ell}$ and continuous in all arguments. The size of the conflict b only affects the magnitude of such change in resources, i.e., $|\Delta B_\ell^s(W_A, W_D, b)|$ is increasing and concave in b and is 0 when $b = 0$.[14] We call conflict technology C the couple $\left(p^C, \Delta^C B\right)$ of winning probability and change in resources that is available to the attacker.[15]

Note that these specifications allow for an offender or defender advantage to reflect in the probability of victory and in the function that determines the gains/losses in resources. For instance, we may have that a defender has a higher chance of winning, but when it loses it loses more resources. Also, while we are assuming that a stronger group has a greater chance of winning, the specification allows that, when the weaker group wins, then more resources of the defeated party are lost, or vice versa.

We assume that fighting cannot create new resources, i.e. for any winner $s \in \{A, D\}$, $B_A^s + B_D^s \leq B_A + B_D$. In the previous example, this requirement would be satisfied by assuming that $\gamma \geq 0$ and $\delta \geq 0$. Note that we do not require conflicts to be necessarily destructive, they can be purely redistributive. This would be the case, in the previous example, with $\delta = \gamma = 0$. Finally, we impose restrictions on how gains and losses are distributed among the winner and loser, and how winning and losing impacts a group's resources:

Assumption 4 *The gains/losses in resources are such that:*

- *The conflict favors the winner, $\Delta B_s^s \geq \Delta B_{-s}^s$, strictly so when $s = A$.*
- *Winning and not fighting are better than losing, i.e.* $\min\{\Delta B_s^s, 0\} \geq \Delta B_s^{-s}$.

[14] We also assume that a group can never lose more than its endowment of resources. Thus the changes in resources are always bounded by the actual resources available to the fighting parties.

[15] Figure **??** in the Appendix depicts the reduced-tree corresponding to the conflict phase.

Note that the previous assumption does not imply that the winner ends up with more resources than the loser, nor that it ends up with more resources than it previously owned. In terms of the previous example, this requirement is simply assuming that $\gamma > 2 + \delta$, i.e., the destruction suffered by the attacker is not high enough to make the defender better off than the attacker, when the defender loses. As it will become clear in the next sections, conflicts will arise in equilibrium, even when non-profitable, i.e. even when the winner is not able to appropriate enough of the loser's resources to cover his own losses (e.g., $\gamma > 1$ in the previous example). Finally, we can define the following class of conflict technologies:

Definition 6 *A conflict technology C is non-profitable if $B_s^s \leq B_s$, otherwise it is profitable.*

Arguably, this class of conflicts captures most modern and historical internal conflicts.[16] Examples of such conflicts include also those that do not involve resource appropriation at all, such as bombings, sabotages, killings, terrorist attacks.

3.2.2 Government Policy

To focus on how the political framework contributes to the pattern of conflict, we assume that, when choosing whether or not to fight, social groups fully anticipate the effect that any change in the distribution of resources would have on government policy, and thus on their equilibrium payoff.

Letting $\hat{T}_i(\mathbf{B}, \mathbf{S})$ be the optimal taxation prevailing under political system S when the distribution of resources is $\mathbf{B}$, we assume that a group i in political system S who is deciding

[16] Collier and Sambanis (2005b) concludes that resource appropriation is at most a mechanism to finance a conflict, but not a motive. Berman et al. (2017) uses the commodity super-cycle to show that natural resources can explain between 14% and 24% of conflicts in modern Africa. Finally, note that civil conflicts typically last long and entail high costs, estimated by Collier and Sambanis (2005b) at 105% of GDP. Even in historical contexts, scholars "firmly believe that medieval warfare was not an economically profitable activity" (Grillo and Settia, 2018, p. 222).

whether to start a conflict anticipates the following final payoff:

$$V_i(\mathbf{B}, S) = B_i - \hat{T}_i(\mathbf{B}, \mathbf{S})$$

Building on the analysis in Chapter 2, we restrict our attention to two possible political systems:

Definition 7 *Let I be the set of non-ruling groups and let $W_{tot} = \sum_{j \in I} W_j + W_{ruler}$. We consider two political systems, defined by a specific government policy:*

- *We call United Society one where each subject i obtains $V_i(\boldsymbol{B}) = B_i \times \frac{\sum_{\ell \in I} W_\ell}{W_{tot}}$ and the ruling group obtains $V_{ruler}(\boldsymbol{B}) = B_{ruler} + \sum_{j \in I} \left(1 - \frac{\sum_{\ell \in I} W_\ell}{W_{tot}}\right) B_j$*
- *We call Hierarchic Society one where each subject i obtains $V_i(\boldsymbol{B}) = B_i \times \frac{\sum_{\ell \in I: B_\ell \leq B_i} W_\ell}{W_{tot}}$ and the ruling group obtains $V_{ruler}(\boldsymbol{B}) = B_{ruler} + \sum_{j \in I} \left(1 - \frac{\sum_{\ell \in I: B_\ell < B_j} W_\ell}{W_{tot}}\right) B_j$*

While the set-up of payoffs that we attributed to Hierarchic Societies looks quite specific, in Chapter 2 we have shown that we can derive that formula from a general model where the ruler employs divide-and-conquer tactics to maximize its tax extraction. Also, the focus on Hierarchic Societies is justified by North et al. (2009), who show that "social hierarchies" are characteristics of what they call a limited access order or natural state, the most common arrangement of human societies (North et al., 2009, p. 12).

Note that the results in Chapter 2 suggest that a society is more likely to be United when it features a representative system of government. We will thus interpret differences in the pattern of conflict that prevails in Hierarchic and United Societies as induced by the political institutions of the latter type of society, and especially by representative institutions.

3.3 Benchmark: Conflicts in a Stateless Society

As our benchmark, we consider the case of a Stateless Society. Such terminology, used by Acemoglu and Robinson (2019), corresponds to a set of social groups that interact with each other but do not have a central authority that affects their payoffs (e.g., via taxes). Note that this case can also be interpreted as one where the decisions to start a conflict do not take into account how conflicts affect the political equilibrium. To study conflict, we consider a model with the following timing:

1. Groups receive their initial endowment of resources $\mathbf{B}^0$.
2. *Conflict phase*: one group A can attack another group D to attempt to change the distribution of resources. The resulting distribution is called $\mathbf{B}^1$.
3. *Stateless Society:* Given $\mathbf{B}^1$, the payoff to each group i is $V_i(\mathbf{B}) = B_i$.

Definition 8 *The attacker's expected resource gain is defined as follows:*

$$\mathbb{E}[G \mid W_A, W_D, b] = p(W_A, W_i)\Delta B_A^A(W_A, W_i, b) + (1 - p(W_A, W_i))\Delta B_A^i(W_A, W_i, b)$$

In this benchmark, it is easy to see that conflicts are started only if the expected value of appropriated resources compensates for the disutility of fighting and the attacker's expected losses. As a consequence of this, if the attacker anticipates that it will end up with less resources then it started with (as in a non-profitable conflict), then it will not start a conflict. Naturally, this implies that non-profitable conflicts cannot be rationalized in a stateless society.

Remark 3 *Group A will start a conflict against i if and only if there exists b such that* $\kappa(b) \leq \mathbb{E}[G \mid W_A, W_D, b]$. *No non-profitable conflicts can occur.*

This result implies that, in our benchmark, the only aspect of the conflict technology that matters is how the attacker is expected to be affected by the conflict. Specifically, attackers only care about the amount of resources that they can win and that they can lose, they do not care about the effects on the defendant. For a similar reason, in this setting, the other groups are not affected by the occurrence of the conflict or by its outcome. As these results will contrast with what we find in the next sections, we summarize these observations in the following remark:

Remark 4 *The following observations are true:*

- *The decision to start a conflict is only affected by the consequences accruing to the attacker.*
- *Only attacker and defender are affected by the conflict.*

Finally, because the probability of winning a conflict is higher when the attacker becomes stronger or the defender becomes weaker, we can show that inequality triggers conflict.

Remark 5 (Inequality triggers conflict) *Suppose that* $(\Delta B_A^A, \Delta B_A^D)$ *does not depend on the power of attackers and defenders. If group* A *wants to attack* i*, then it would also want to attack* j *such that* $W_j < W_i$*. If group* A *wants to attack* i*, then also group* B *would want to attack* i *if* $W_B > W_A$

Clearly, under some conditions, the above remark would also be true if $(\Delta B_A^A, \Delta B_A^D)$ depended on the balance of power between attacker and defender. Indeed, inequality would be an even greater trigger of conflicts if a stronger attacker can be expected to gain more resources when it wins, and lose less when it loses. Recall that $\kappa(b)$ is convex, $\kappa(0) = 0$ and $\mathbb{E}[G \mid W_A, W_D, b]$ is linear in b. The only role played by these assumptions is that, without one of them, we might have that a group who wants to attack group i does not want to

attack a weaker group because it cannot select a high enough b to recover the costs of the conflict.

3.4 Monopoly on the Legitimate Use of Violence

We first consider a society where there is an effective monopoly on the legitimate use of violence. In such society, conflicts can only occur when a group challenges the legitimacy of the central authority, i.e., when one social group proposes itself as a replacement for the current ruling group.

As we model conflicts as only a way to change the distribution of resources, we must first understand why a conflict can help a social group to become the next ruler. To this end, we consider a model where, after the conflicts have occurred, groups decide their next ruler. This allows us to characterize how the support for the rulership depends on the distribution of resources in the society, and thus how conflicts can lead social groups to become the next ruler. We consider a model with the following timing:

1. Groups receive their initial endowment of resources $\mathbf{B}^0$. A group is randomly selected to be the incumbent ruler and the society is commonly known to be either Hierarchic or United.

2. *Conflict phase*: one group A can attack the ruling group to attempt to change the distribution of resources. The resulting distribution is $\mathbf{B}^1$.

3. *Ruler Selection*: given $\mathbf{B}^1$,

   3a) Each group can propose itself as the next ruler;

   3b) Groups choose whether to support the incumbent, the challenger, or stay neutral.[17]

[17]We focus on the equilibrium in which groups which are indifferent between incumbent and challenger

3c) The group with the largest support becomes (or remains) the ruler.

4. Given $\mathbf{B}^1$, the identity of the ruling group, and whether the society is Hierarchic or United, the payoffs to each group are realized according to Definition 7.

This particular time structure allows us to capture, in a simple way, how government policy (i.e., whether social groups live in a Hierarchic or United Society) affects the selection of the ruling group, and how this affects the pattern of internal conflict.

3.4.1 Ruler Selection

In order to become the ruler, the challenger needs to collect enough support relative to the current ruler. In particular, given a challenger and an incumbent, we assume that each group can either stay neutral or side with one of the candidates for a negligible but positive cost.[18] Given these decisions, the challenger becomes the ruler if:

$$\sum_{j \in \, C\text{-}support} W_j > \sum_{j \in \, I\text{-}support} W_j,$$

where *C-support* and *I-support* are the set of groups that choose to support the challenger or the incumbent, respectively.

As a first step of the analysis, in the following lemma we characterize subjects' preferences over the identity of the ruler as a function of whether they are in a United or Hierarchic Society.

Lemma 4 *In a Hierarchic Society, each subject i strictly prefers a ruler h which is stronger $W_h > W_i$ rather than a ruler ℓ which is weaker $W_\ell < W_i$. Furthermore, subject i is indifferent*

stay neutral.

[18] This assumption, implying abstention of indifferent voters, allows us to focus on changes in rulership strictly preferred by a more powerful coalition than the one that strictly prefers the status quo.

among all stronger subjects. In a United Society, every group strictly prefers the weakest possible ruler.

The first part of this result appears to run counter to the intuition that groups should benefit from a weak ruler, given its purely extractive role. While it is true that subjects benefit from a weakening of the ruling group, as tax demands are increasing in the military power of the ruler, W_{ruler}, Lemma 4 implies that every group i would actually prefer the strongest group over any weaker group ℓ, $W_\ell < W_i$, as the ruler.

The logic of the result rests on the characteristics of the taxation that we associate with Hierarchic Societies. As it is clear from Definition 7, in a Hierarchic Society, the taxation demands that the ruler imposes on i is backed by the military power of the ruler and that of all groups who are richer (and stronger) than i. This follows from the fact that whenever (in a deviation) resistance occurs, the ruler deals with it using its weapons and those of all subjects who have more resources than the resisting ones. This means that, for any i, the military power of any subject h with more resources would be employed to extract taxes from i, independently of the ruler's identity. Thus, in order to minimize their tax burden, subjects would strictly prefer to have a ruler with more resources than them.

On the other hand, in a United Society, the resources of all subjects are used to keep the tax burden low for every other subject. As a consequence, assigning the rulership to a strong group would be counterproductive: not only it would lead to a stronger ruler, capable of extracting higher taxes, but it would lead to the loss of a formidable ally in the subjects' objective to limit the ruler's ability to extract resources from society.

With this in mind, we can characterize the endogenous choice of a ruler in the following theorem which identifies the strongest group as the natural ruler for a Hierarchic Society, and the weakest ruler as the natural ruler for a United Society.

Theorem 4 *In a Hierarchic Society, no group can successfully challenge the ruler when the*

rulership is held by the strongest group. Also, the strongest group can successfully challenge any ruler. In a United Society, every group can successfully challenge a stronger group, unless such group has more than half of the total resources in society. Also, the weakest group can successfully challenge any ruler, unless the ruler has more than half of the total resources in society.

for a United Society, the result shows that the only case in which the strongest group can become ruler is when its power is greater than that of all other groups combined, that is when the support decision of the other groups does not matter for ruler selection.

3.4.2 Hegemonic Conflicts

In this section we explore how the ruler selection phase affects the pattern of conflict. In a United Society, the weakest group becomes the ruler. As a consequence, a conflict is not a good strategy to become the next ruler. On the other hand, in a Hierarchic Society, becoming the strongest group leads a social group to become the next ruler. As a consequence, groups have a high incentive to invest in changing the distribution of resources (i.e., to fight) when such change can make them the strongest social group.

To be more precise, note that ruler selection considerations introduce a discontinuity in the incentives to attack: the expected payoff of a conflict is significantly higher when it enables the attacker to become the strongest group. This discontinuity is the key ingredient to highlight how similarity triggers conflicts.

To gain intuition, consider a country where a resource-rich region is divided between the two strongest ethnic groups in society. Each group can attack the other to gain control of a gold mine: depending on the odds of success and the cost of fighting, the attack may or may not be deemed profitable.[19] If the two groups live in a Hierarchic Society, the two

[19] As the odds of winning depend on the balance of power, we would expect conflicts to be associated with a more unequal distribution of power. The result that inequality of power leads to conflict is a robust result

groups anticipate that the strongest of the two groups will be able to become the next ruler. Thus, when gaining control of a gold mine makes a group stronger than the other, then the value of launching an attack jumps up. As this discontinuity is only present when groups are sufficiently similar, then similarity between groups is a trigger for conflicts.

The next Theorem formalizes the idea that, in Hierarchic Societies, under a very general class of conflict technologies, similarity between the two strongest groups triggers conflicts.

Theorem 5 (Hegemonic Conflicts) *In a Hierarchic Society, for every distribution of resources* $\mathbf{B}$ *with no ties the second-strongest group* n *would want to start a conflict whenever it is sufficiently close to the strongest group* $n+1$*, i.e., there exists* $H > 0$ *such that* n *starts a conflict whenever* $B_{n+1} - B_n < H$*.*

Intuitively, the above mentioned discontinuity in the payoffs provides strong incentives for a subject to start a conflict that allows it to become the strongest social group in case of success and not become too weak in case of failure. If the resources of the strongest group get closer, the probability of winning a conflict against it increases and a smaller conflict size is needed to overcome it. A conflict with smaller size b is associated with both a smaller cost $k(b)$ and smaller resource losses, which in turn implies a smaller risk of becoming too weak in case of failure. Thus, we should expect conflicts to occur when the resources of the strongest and second strongest groups move sufficiently close to each other.

Note that the result is not driven by the possibility of very cheap conflicts. Indeed, a similar result could be established even for $\kappa(0) > 0$. Note that the value of the rulership will in general depend on the structure of society. Moreover, because tax revenues depend on the whole distribution of resources in society, then our model shows that civil wars should not only depend on the characteristics of the fighting parties, but also on a measure of the whole distribution of resources in society. This is another finding that justifies the approach

of the conflict literature that focuses on "greed" (Jackson and Morelli, 2009).

of the empirical literature on ethnic conflict (e.g. Arbatli et al., 2020), which shows that the incidence of conflicts can be predicted by statistics of the country-level distribution of resources (e.g. fractionalization or polarization). Finally, note that our result does not require the conflict technology to be profitable: engaging in a non-profitable conflict can still make a group the strongest, as long as the extent of destruction on the other party is worse than on itself. Thus, similarity can trigger conflicts where even the winner loses resources. Going back to the earlier example, our result says that, when the two groups are sufficiently similar, the weaker of the two would be willing to start a conflict even if the best possible outcome were to destroy the other group's gold mine.

3.5 No Monopoly on the Legitimate Use of Violence

As a second case, we consider a setting with no monopoly on the legitimate use of violence. This means that we allow subjects to fight each other. To avoid overlaps with the previous section, we assume that the identity of the ruling group is fixed. We consider a model with the following timing:

1. Groups receive their initial endowment of resources $\mathbf{B}^0$, one group is randomly selected to be the ruler, and the society is commonly known to be either Hierarchic or United.

2. One group A can attack another group D to attempt to change the distribution of resources. The resulting distribution is $\mathbf{B}^1$.

3. Given $\mathbf{B}^1$, the identity of the ruler, and whether the society is Hierarchic or United, payoffs are realized according to Definition 7.

3.5.1 United Society

Just as in the Stateless Society benchmark, conflicts in a United Society are only started if the expected value of appropriated resources is high enough. However, the incentive to fight against another subject will in general be lower than in a Stateless Society.

Remark 6 *The incentive to attack another subject in a United Society is strictly lower than the incentive to attack the same group in a Stateless Society.*

Note that the only benefit from a conflict against another subject is the gain in resources that might be obtained. Consider first the case where a conflict is simply a transfer of resources, with no destruction. In this case, the resources gained have less value than in a Stateless Society, as the attacker would have to pay taxes on the additional resources it has appropriated. In addition to this, if the conflict causes some destruction, then this would lead to a higher tax rate, which further reduces a subject's incentive to fight against another subject, compared to the Stateless Society benchmark.

Note that this implies that an attacker's incentive to fight is not only determined by the consequences accruing to itself, but also by how the defendant is affected. While destroying the resources of another subject has a negative effect (via a higher tax rate), destruction can be good when the defendant is the ruling group. As a consequence of this, unlike the Stateless Society benchmark, we find that non-profitable conflicts might occur, and when they do they are always against the ruling group. More in general, as taxation generates interdependence among all subjects, we also find that conflicts have an effect on every other subject. We can thus establish the following:

Remark 7 *The following observations are true:*

- *Consider two conflict technologies C and C' whose only difference is that C' imposes higher losses on the defendant when it loses. Then, a change from C to C' would make a*

conflict against another subject less attractive, whereas it would make a conflict against the ruling group more attractive.

- *Non-profitable conflicts can happen, but only against the ruling group.*
- *In a conflict between subjects, every other subject prefers the conflict outcome that causes the less overall destruction*

Finally, we can show that the pattern of conflict among subjects is the same as in the Stateless Society benchmark. Indeed, the incentive to fight still comes from the desire to appropriate the resources of another group, as a consequence inequality will still be a driver of conflicts..

Remark 8 (Inequality Triggers Conflicts) *Suppose that* $(\Delta B_A^A, \Delta B_A^D)$ *does not depend on the power of attackers and defenders. If subject* A *wants to attack subject* i*, then it would also want to attack subject* j *such that* $W_j < W_i$*. If subject* A *wants to attack subject* i*, then also subject* B *would want to attack* i *if* $W_B > W_A$*.*

3.5.2 Hierarchic Society: Status-Related Conflicts

In a Hierarchic Society, subjects pay higher taxes than in a United one. This feature reduces the value of appropriating resources and thus it reduces the value of starting conflicts. However, it would be incorrect to say that the incentives are lower than in a United Society. Indeed, in a Hierarchic Society, becoming stronger (weaker) than another subject leads to a discontinuous jump up (down) in a subject's payoff. In line with the terminology of Chapter 2, we call such discontinuity a change (gain or loss) in *status*. This feature increases the incentives to fight compared to a Stateless or United Society, as a small change in resources can trigger a status gain and thus be particularly valuable (by the same logic discussed for hegemonic conflicts). The second element that shapes the incentives to fight in a Hierarchic

Society is that groups benefit from the resources of weaker groups (such resources reduce their tax burden), whereas they are hurt from the resources of stronger ones and the ruling group. As a consequence, conflict incentives will depend on whether the defender is stronger or weaker than the attacker. The following result summarizes this dependence.

Remark 9 *The following observations are true:*

- *Consider two conflict technologies C and C' whose only difference is that C' imposes higher losses on the defendant when it loses. Then, a change from C to C' would make a conflict against a weaker subject less attractive, whereas it can increase the attractiveness of a conflict against a stronger subject or the ruler.*
- *Non-profitable conflicts can happen, but only against stronger groups or the ruler.*
- *Every subject who is stronger (weaker) than both fighting parties prefers the conflict outcome that causes the lower (higher) overall destruction.*

A final and key difference between Hierarchic Societies and the other types is that inequality is no longer a key driver of conflict. Indeed, the finding that non-profitable conflicts can only happen against stronger subjects already shows that we cannot establish a result analogous to what we established for United Societies. More interestingly, note that it is also not true that if group i wants to attack a weaker group j then this would also be true for any group stronger than i. For instance, suppose that $\kappa(b) = b$ and suppose that the conflict technology is such that we have a transfer of b from the defender to the attacker when the attacker wins and there is no change in resources when it loses. With such conflict technology, no group would want to attack a weaker group, unless the attack allows them to gain status (become stronger than another subject). As a consequence, we might find that a group i wants to attack a weaker group j even when a group stronger than i does not. While this discussion clarifies that inequality is not a trigger for conflicts, we find that equality is.

Indeed, in line with the analysis of Hegemonic Conflicts, we can show that equality triggers conflicts:

Theorem 6 (Status-Related Conflicts) *In a Hierarchic Society, for every distribution of resources* $\mathbf{B}$ *with no ties a subject* i *would want to start a conflict whenever it is sufficiently close to a stronger group* $i+1$*, i.e., there exists* $\Delta > 0$ *such that* i *starts a conflict whenever* $B_{i+1} - B_i < \Delta$*.*

The intuition behind this result is the same as Theorem 5. However, while in Theorem 5 the discontinuity came from the prospect of becoming the next ruler, in this case the discontinuity is embedded in the fact that status has value: becoming stronger than another subject leads to a jump up in a group's payoff. While fighting against a stronger subject is probably the most intuitive way to gain status, if the conflict technology is profitable, a subject can also become stronger than a given subject by appropriating the resources of another (e.g., weaker) subject. Thus, also the resource level of third parties might be relevant when studying the incentive to attack. This feature has methodological implications for the empirical study of internal conflicts. While counter-intuitive, our model shows that the nation-wide distribution of resources may be an important determinant of internal conflicts, even if the observed conflicts are regionally clustered. This finding supports the mainstream empirical approach against Berman et al. (2017)'s recent criticism.[20]

In our historical case study we will build on this analysis to show how status-related conflicts can help explain some puzzling observations. First, note that status-related conflicts can be non-profitable, i.e., they are consistent with social groups being willing to participate in highly destructive conflicts with each other over relatively minor stakes (i.e. for relatively minor resource gains). Indeed, as these conflicts are motivated by the relative level of

[20]In light of our model, we agree with Berman et al. (2017) that using the country-year level as the sole unit of analysis may be too aggregate. However, our analysis emphasizes that information at the country-level may still be relevant even to understand regionally clustered conflicts.

resources of the two fighting parties, groups can engage in conflicts that make them weaker than they were before, as long as this makes them stronger than their opponent. Second, note that the driver of status-related conflicts should be to establish a group's superiority over another group in the eyes of the central authority. As a consequence, we should expect that the conduct of conflict and the various rituals of reconciliation should be aimed at *publicly* sanctioning the superiority of one group over the other one. Finally, we should observe these conflicts to be followed by a social group's advancement (e.g. its leaders obtaining important administrative positions), as becoming stronger than another group should lead to a fiscal gain (i.e. lower taxes or higher political rents).

3.6 General Patterns

3.6.1 Inequality or Equality as a Trigger for Conflicts

Our first result (Theorem 5) speaks to the evidence on civil wars, which, in the main datasets on conflicts (e.g. COW), are defined as conflicts involving the government that are not one-sided (i.e. there is no side that is negligible compared to the other). Indeed, the objective of a hegemonic conflict is to obtain the rulership by becoming stronger than the current hegemonic power which should be the current ruling group. Thus, we should expect hegemonic conflicts to be coded as civil wars. From an empirical perspective, our model suggests that civil wars should be more likely when the balance of power between the two strongest groups approaches parity. In effect, this observation may explain why polarization is associated with the incidence of civil wars even when controlling for other measures, such as fractionalization (Montalvo and Reynal-Querol, 2005). Indeed, polarization is meant to capture a situation in which there are two main blocs with power close to parity, an idea dating back to Horowitz (1985)'s observation that conflicts arise in societies where a large ethnic minority faces an

ethnic majority.[21] Additionally, our finding that hegemonic conflicts should only occur in a Hierarchic Society suggest that institutions pushing towards a United Society should have the positive side-effect of reducing the incidence of civil wars. As discussed in Chapter 2, United Societies should be expected to have representative institutions. While the mere existence of a de jure representative institution does not imply its de facto relevance, it is encouraging to observe that existence of de jure representative political systems correlates with a lower probability of civil wars in post-colonial Africa (Reynal-Querol, 2002).

Moving beyond civil wars, our theoretical results also speak to the empirical evidence on conflicts that are not aimed at the rulership, in particular Esteban et al. (2015), Mitra and Ray (2014) and Mitra and Ray (2019). Esteban et al. (2015) studies conflicts perpetrated against minorities at risk. These minority groups "suffer from threats or discrimination" (Esteban et al., 2015, p.1116); in terms of our model we can think of these groups as weaker with respect to the majority groups threatening them. According to our model, the incentives to start a conflict against such minorities is provided only by the value of the resources that can be appropriated from them. Consistently with this, Esteban et al. (2015) finds that smaller minorities (lower W) are at higher risk. Moreover, Esteban et al. (2015) highlights how the presence of mineral resources (oil and diamonds) is another important risk factor.[22] We can rationalize this finding as a consequence of the fact that mineral resources are arguably easier to appropriate and are thus associated with a more profitable conflict technology.[23] Finally, note that Esteban et al. (2015) specifically focuses on genocides which can be seen in our model as a degeneration of conflicts against weaker subjects. Indeed, the incentive to fight a minority is driven by the favorable odds of winning. As each victory makes the stronger

[21] An example of a country with high polarization and low fractionalization is Guatemala, where the two main groups include 97% of the country with the dominant one at 55% (Montalvo and Reynal-Querol, 2005).

[22] These results are in Table 4 of Esteban et al. (2015); note that the regression also controls for a group's satellite light intensity, a proxy for local economic resources.

[23] Recall that, according to our model, conflicts against weaker groups require a profitable conflict technology, which we expect to be the case when there are mineral resources at stake.

group even stronger, then we should expect the conflict to start again at the next available chance, until the weaker group is completely exterminated or it has nothing else of value to steal.

On the other hand, our result that equality is a trigger for conflicts speaks particularly neatly to the evidence of Mitra and Ray (2019). The authors find that conflicts are more (less) likely when the poorer (richer) group experiences a positive economic shock. Our model speaks to such finding as it predicts status-related conflicts to emerge exactly when two groups become more similar. Since status-related conflicts can occur also when the conflict technology is non-profitable, it should not strike as a puzzle if such conflicts were to appear to only cause destruction. Finally, note that status-related conflicts could easily degenerate in a long-lasting sequence of conflicts between two groups that struggle for local supremacy: until the two groups are sufficiently far apart in term of resources, there will be a conflict at the next available chance. This description is very similar to the reality of Hindu-Muslim relations of post-independence India studied by Mitra and Ray (2014, 2019) and many other "communal" conflicts in the developing world; furthermore, it is strikingly reminiscent of the conflicts during the Middle Ages among small and large lineages alike (see the historical case study). While such pattern may be easily interpreted as a self-reinforcing dynamics between internal conflicts and ethnic or partisan hate, our model suggests that the political logic of a Hierarchic Society may add to any other consideration a robust *policy-induced* economic interest for the occurrence of such conflicts, thus fueling them.

3.6.2 Status-Related Conflicts

The idea that status in hierarchic systems may fuel infighting is very widespread. For instance, Earls (1969) notes that a typical trait of Andean societies and mythology is the idea that the contiguous ranks of society should be particularly likely to fight each other.

Similarly, elaborating on the hierarchy-reinforcing policies of the French Empire in Laos, Adams and McCoy note that the "system of control without direct contact had another advantage for the French. By deflecting peoples' anger at the race above or below them on the scale, the French administration was able to exploit extremely independent and volatile groups without ever incurring any direct hostilities" (Adams and McCoy 1970, p.80).

In Medieval and Early Modern Europe, the idea that the distribution of privileges could generate the incentives for subjects' infighting was well understood by contemporaries, who often speak of 'jealousy' as underlining the prevalence of internal conflicts. For instance, Dante Alighieri mentions jealousy as one of the three 'vices' that led to internal conflicts in Florence (Alighieri 1314, Canto 6). The connection between conflict-inducing jealousies and government policy is brought to its extreme logical consequences by an advisor of Charles VII of France (r. 1422-1461): "when you cease to pay the said pensions, the lords will be more contented than they are at present; for it's all these jealousies that produce covert hatreds, and each one thinks that he deserves to have more than another." (quoted in Vale 1974, p. 226). Note that all these hatreds were only an indirect threat to the central authorities, as they triggered subjects' infighting rather than revolts. Indeed, Kendall (1971) notes that the policies of Charles VII of France led magnates to aim their resentment at each other, as in the decade-long warfare that opposed the House of Armagnac and the House of Comminges in Southern France (Kendall 1971, p. 47). While such conflict, as the ones discussed in our historical case study, involved relatively minor political players, the idea that political favoritism may trigger infighting has been applied to larger-scale wars. For instance, Kaminsky (2002) discusses how the modern historical literature has shown that concerns over the relative distribution of privileges was a key motivation for the nobles who fought each other during the Hundred Years' War (1337-1453). Similarly, the Thirty Years War (1618-1648), one of the most destructive conflicts of the Modern Age, was in part fueled by the rivalry between Bavaria and the Palatinate, rooted in the desire of the former to be

elevated to the rank of prince elector (Gutmann, 1988). Interestingly, the rivalry between Bavaria and the Palatinate persisted into the XVIII Century, this time being "over their relative ranking within the electoral college" (Wilson 2014, p.14).

3.7 Case Study: Seigneurial Wars in Late Medieval and Early Modern Europe

One of the most consistent features of Medieval society is the pervasiveness of internal conflicts. Indeed, the "medieval European nobility spent much of their time fighting each other, in what modern historians call private wars, guerres privèes, Privatkriege." (Kaminsky, 2002, p.55) While the conflicts "of the great noble families" are the most likely to be recorded by the surviving sources, we know that conflicts between whole kindreds and their armed supporters were common "at every level of society", in urban centers and rural districts alike (Bloch, 1961; p.127).

Among all forms of private wars, our focus will be on what historians call seigneurial wars or the noble feud, which can be defined as "large-scale, organized, politically-motivated violence waged among non-royal parties" (Firnhaber-Baker 2010a, p.37). Unlike 'feuds' in the sense of "cyclical, vindicatory violence waged by kin groups", the basis of seigneurial wars was not "wounded honour or anger" (Firnhaber-Baker 2010b, p.91) but contested lordship rights (i.e. the distribution of political rents). Moreover, we will focus on seigneurial wars that took place in Late Medieval and Early Modern Germany (i.e. the Germanic domains of the Holy Roman Empire), basing our work on Zmora (2003, 2015, 2020), which studies the almost 300 seigneurial wars that took place between 1440 and 1570 in Franconia. This geographical and temporal focus is necessary to have a sharper view of the pattern of conflict behavior, however seigneurial wars were a widespread European phenomenon.[24] At the European level,

[24]Conflicts between kin-groups are a constant feature of most documented pre-modern European history,

seigneurial wars interacted with other conflicts, including important inter-state wars such as the Hundred Years' War.[25] Even when they did not interact with inter-state wars and only involved a limited number of military units, seigneurial wars caused "extensive devastation" and an "array of social problems and economic disruptions" (Zmora, 2003).

3.7.1 Characteristics of Seigneurial Wars in the German lands

Approaching seigneurial wars from a rational-agent perspective rises three key puzzles. First, fighting does not appear to be motivated by a positive balance between expected material gains and costs. Second, there is a strange ritualistic and behavioral focus on the public humiliation of the opponent. Third, while central authorities tried to limit the practice of seigneurial wars through mediation attempts or outright prohibitions, ex post they often appear to provide the rewards that may motivate the infighting in the first place. As discussed, our model rationalizes these findings, as they are natural expressions of what we called status-related conflicts.

The first feature is noted by Zmora that mentions "the apparently considerable disproportion between the comparatively little economic value of the disputed objects and the intransigence that the feuders displayed" as a "salient characteristic" of seigneurial wars

at least since the first city-states in Ancient Greece (Zorzi 2009, pp.8-9, 17). Besides the case of Franconia, these conflicts are also well documented for Italian city-states in the Late Middle Ages. A general view is expressed in Martines (1972), Vigueur (2004) and Zorzi (2009). For studies on specific Italian communes see Lansing (1991) for Florence, Di Santo (2016) for Rome, Gentile (2009) for Parma, and Zorzi (2009) for Florence, Parma, Mantova, Lucca and Pisa. We focus on Germany as Zmora distinguishes between seigneurial wars and blood-feuds. Importantly, seigneurial wars did not only appear in places with a relatively weak central authority. The Kingdom of France was also afflicted from them, despite the regulations against feuding: more than fifty thousand royal pardons can be related to seigneurial wars in the period 1302-1568 (Kaminsky, 2002). In Southern France alone, we can count somewhere between 59 and 72 cases during the XIV Century (Firnhaber-Baker 2010a, p.91). Even in England, despite the relatively strong position of the English Crown vis-à-vis the aristocracy and the early royal prohibitions against private wars, seigneurial wars were "endemic from the fourteenth to the seventeenth centuries" (Kaminsky 2002, p.75).

[25] While the Hundred Years' war is typically characterized as one of the first conflict between nation states (France and England), Kaminsky (2002) discusses how recent scholarship on the surviving documentary evidence presents a picture where local rivalries and "the maintenance of individual power and privilege" were the key driver of the choice to side with the French or English Crown (Little 1984, p.176; quoted in Kaminsky 2002, p.72).

(Zmora 2020, p.221). This disproportion was evident even to contemporary observers. When reporting that one of the most devastating seigneurial wars of early-modern Franconia was over the rights to use a worthless sheep-run, the narrator himself commented in bewilderment "this was the beautiful Helen over which the two princes (...) went to a veritable Trojan War" (Fries 1713; quoted in Zmora 2003, p.87). Interestingly, this pattern does not only emerge from an ex-post reading of the evidence. Indeed, the actors immediately involved seemed perfectly aware of such disproportion, but willing to fight nonetheless. For instance, when the margrave of Brandenburg-Ansbach went to war with another Franconian prince over tolls, he noted "with the help of God we would not let 4,000 Gulden be taken away from the lordship even if it should cost us three times the worth of the tolls" (quoted in Zmora 2020, p.221). In general, any theory of seigneurial conflicts must come to terms with the fact that when fighting over lordship rights, "[a]ristocrats, from lesser nobles to princes and kings, accepted the disproportionately high costs of defending these rights" (Zmora 2020, p.221).

Both Firnhaber-Baker and Zmora discuss the role of public displays and acts of dominance as integral part of the conduct of fighting parties. "Wars almost invariably involved acts of violence and domination that produced no direct material benefit for the attackers, but which humiliated their enemy, resulting in the loss of prestige for the opposition." Moreover "[s]uch acts often involved public symbols of status such as banners and other regalia" (Firnhaber-Baker 2010b, p.97). This feature was shared by the seigneurial wars analyzed by Zmora as well, and was understood by contemporaries. For instance, "a commonplace maxim" that circulated among fighters in Franconia was "he who sustains the damage must often suffer the scorn as well" (quoted in Zmora 2020, p.222). In general, there is a general attention among the fighting parties to publicize their actions, especially towards the relevant overlord who were asked to be present at decisive events, such as duels or pacification ceremonies.

Finally, the evidence from Early Modern Franconia directly links seigneurial wars to the

practice of political favoritism. Indeed, Zmora notes as striking that "many feuders were appointed to high princely offices during or immediately following their feuds," concluding that "feuds brought them to the fore and then closer to the prince (..) [and] [t]heir status rose concordantly" (Zmora 2003, pp.116-117) Importantly, because "lordships were in large part controlled or distributed by princes", princely favors could not be simply extorted with brute force from another family. Yet, "[u]nderlying all [conflict motives] was the concomitance of the competitions over accumulation and concentration of land-lordship and territorial lordship" (Zmora 2003, pp.117-118). This competition was an indirect effect of the princely policies: "making choices, by favouring some fighter over others, princes influenced the process of social stratification among the nobility" (Zmora 2003, p.112). So that "[seigneurial wars] helped to create winners and losers. In association with other strategies, feuds could result in accumulation and concentration of lordship" (Zmora 2003, p.118). Zmora concludes that "one of the most salient traits of the feud in this period of intense princely territorialisation" is that many "German noblemen owed their rise to prominence to a timely prosecution of feuds," as seigneurial wars provided the "political moment that generated the 'critical mass' necessary for [their] further ascent." It is only after taking these political outcomes into account that a seigneurial war can appear as a "a well-thought-out enterprise" (Zmora 2003, pp.25-26).

3.8 Conclusions and Future directions

This paper studies how conflicts can be indirectly affected by political institutions, showing how the pattern of internal conflicts differ in a United and a Hierarchic society. While we have shown that our model can be helpful to analyze both modern and historical evidence, we did not directly test our model's predictions. This is the main direction that we wish to pursue in future work.

In terms of the theoretical framework, we presented a purposefully stark setting to capture the two conflict motives that emerge in Hierarchic Societies: hegemonic, and status-related conflicts. The framework can be expanded in various directions. As we wanted to focus on the relation between social groups, our model was silent about the internal organization of each group. In principle, we could embed within our framework a model that captures within-group considerations, like the idea in Padró i Miquel (2007) or the model proposed by Esteban and Ray (2011). This would certainly highlight novel insights and new predictions to bring to the data. Furthermore, as we wanted to stress the robustness of the incentives to fight identified by our model, we worked under a general conflict technology. By specifying a sharper conflict technology, we would certainly be able to derive additional predictions on the pattern of conflict. This extension could be guided by previous research that have analyzed the incentives to invest in military equipment (Mitra and Ray, 2014) or that highlights how the costs and benefits of conflicts may depend on conditions of the local economy (Dal Bó and Dal Bó, 2011). Another promising avenue for future research is to allow for participation of other groups in conflicts, a topic we only touched upon when studying how the other social groups are affected by conflicts. Finally, an obviously important extension would be to explore the logic of our model within a repeated game. Such dynamic extension would enable to identify long-run patterns of conflicts.

Finally, while in our model it is taken for granted, we might be interested in studying how the very existence of a society, i.e. the fact that different groups are subject to the same authority, impacts the pattern of conflicts. The analysis suggests that weaker groups would be less frequently attacked when inside a society, as stronger subjects internalize any damage they may inflict when attacking weaker social groups. If greed, i.e. the desire to appropriate other groups resources, is the main driver of conflicts, then the creation of a society can reduce the incidence of conflicts even without considering any direct role of the government to prevent them. Pushing this logic to the extreme, it would be interesting

to study the incentives of social groups to organize themselves in a society with a central authority. For instance, starting from state-less societies as those discussed in Acemoglu and Robinson (2019) and study when we should expect state-building to take place. Especially the the result that, in a Hierarchic Society, social groups would select the strongest of them as their ruler appears to have the most connections with the discussion and evidence in Acemoglu and Robinson (2019).

Bibliography

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

Acemoglu, D. and J. A. . Robinson (2000). Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective. *Quarterly Journal of Economics 115*(4), 1167–1199.

Acemoglu, D. and J. A. Robinson (2019). *The Narrow Corridor: States, Societies, and the Fate of Liberty*. Penguin Press.

Acemoglu, D., J. A. Robinson, and T. Verdier (2004). Kleptocracy and Divide-and-Rule: A Model of Personal Rule. *Journal of the European Economic Association 2*(2-3), 162–192.

Adams, N. S. and A. W. McCoy (Eds.) (1970). *Laos: War and Revolution*. HarperCollins Publishers.

Alesina, A., S. Michalopoulos, and E. Papaioannou (2016). Ethnic inequality. *Journal of Political Economy 124*(2), 428–488.

Alighieri, D. (1314). *La Divina Commedia. Inferno*.

Allison, G. (2017). *Destined for war. Can America and China escape Thucydides's trap?* Houghton Mifflin Co.

Ambrocio, G. and I. Hasan (2021). Quid pro quo? political ties and sovereign borrowing. *Journal of International Economics 133*, 103523.

Ambrosio, T. (2021). Security hedging after nazarbayev? the future of kazakhstan's alignment with russia. In *Understanding Kazakhstan's 2019 Political Transition*, pp. 29–51. Springer.

Anderson, J. E. (2011). The gravity model. *Annual Review of Economics 3*(1), 133–160.

Anderson, J. E. and E. van Wincoop (2003). Gravity with gravitas: A solution to the border puzzle. *American economic review 93*(1), 170–192.

Anderson, J. E. and Y. V. Yotov (2020). Short run gravity. *Journal of International Economics 126*, 103341.

Antràs, P. and G. Padró i Miquel (2011). Foreign influence and welfare. *Journal of International Economics 84*(2), 135–148.

Arbatli, C. E., Q. H. Ashraf, O. Galor, and M. Klemp (2020). Diversity and Conflict. *Econometrica 88*(2), 727–797.

Arkolakis, C., A. Costinot, and A. Rodríguez-Clare (2012, 02). New trade models, same old gains? *American Economic Review 102*(1), 94–130.

Bai, Y. and R. Jia (2016). Elite Recruitment and Political Stability: The Impact of the Abolition of China's Civil Service Exam. *Econometrica 84*(2), 677–733.

Bakos, A. E. (2013). *Images of Kingship in Early Modern France: Louis XI in Political Thought, 1560-1789*. Routledge.

Baliga, S. and T. Sjöström (2013). Bargaining and war: A review of some formal models. *Korean Economic Review 29*(2), 235–266.

Barkow, J. H. (1989). *Darwin, sex and status: Biological approaches to mind and culture*. University of Toronto Press.

Barr, J. (2018). *Lords of the Desert: Britain's Struggle with America to Dominate the Middle East*. Simon and Schuster.

Bates, R., A. Greif, and S. Singh (2002). Organizing violence. *Journal of Conflict Resolution 46*(5), 599–628.

Berger, D., W. Easterly, N. Nunn, and S. Satyanath (2013). Commercial imperialism? political influence and trade during the cold war. *American Economic Review 103*(2), 863–96.

Bergstrand, J. H. and P. Egger (2013). Gravity equations and economic frictions in the world economy. In *Palgrave handbook of international trade*, pp. 532–570. Springer.

Berman, N., M. Couttenier, D. Rohner, and M. Thoenig (2017, 6). This mine is mine! How minerals fuel conflicts in Africa. *American Economic Review 107*(6), 1564–1610.

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

Besley, T. and T. Persson (2010). State Capacity, Conflict, and Development. *Econometrica 78*(1), 1–34.

Besley, T. and T. Persson (2011). The logic of political violence. *Quarterly Journal of Economics 126*(3), 1411–1445.

Blattman, C. and E. Miguel (2010). Civil war. *Journal of Economic Literature 48*(1), 3–57.

Bloch, M. (1961). *Feudal Society, Volume 1: The Growth of Ties of Dependence* (Taylor & F ed.). London and New York: Routledge & Kegan Paul Ltd.

Boix, C. (2015). *Political Order and Inequality*. Cambridge University Press.

Boucekkine, R., R. Desbordes, and P. Melindi-ghidi (2019). Social Divisiveness and Conflicts: Grievances Matter!

Bougard, F., D. Iogna-Prat, and R. Le Jan (2009). *Hiérarchie et stratification sociale dans l'Occident médiéval (400-1100)*. Brepols.

Brubaker, R. (2009). Ethnicity, race, and nationalism. *Annual Review of Sociology 35*, 21–42.

Bryce, J. (1901). The Holy Roman Empire.

Bueno De Mesquita, B., A. Smith, R. M. Siverson, and J. D. Morrow (2005). *The Logic of Political Survival*. The MIT Press.

Burgess, R., R. Jedwab, E. Miguel, A. Morjaria, and G. Padró i Miquel (2018). The Value of Democracy : Evidence from Road Building in Kenya. *American Economic Review 105*(6), 1817–1851.

Cahill, D. (1994). Colour by Numbers: Racial and Ethnic Categories in the Viceroyalty of Peru, 1532—1824. *Journal of Latin American Studies 26*(2), 325–346.

Chacon, Y. (2009). An inquiry into inka structures of domination.

Cheng, J. T., J. L. Tracy, T. Foulsham, A. Kingstone, and J. Henrich (2013). Two ways to the top: Evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *Journal of Personality and Social Psychology 104*(1), 103–125.

Chung-Mien, T. and P. A. Herbert (1987). The T'ang System of Bureaucratic Titles and Grades. *Tang Studies* (5), 25–31.

Coke, E. (1644). *The Fourth Part of the Institutes of the Laws of England: concerning the jurisdiction of Courts*. London: Printed by Miles Flesher for William Lee and Daniel Pakeman.

Collier, P. and N. Sambanis (Eds.) (2005a). *Understanding Civil War: Evidence and Analysis, Volume 1. Africa*. World Bank Publications.

Collier, P. and N. Sambanis (Eds.) (2005b, 8). *Understanding Civil War: Evidence and Analysis, Volume 2. Europe, Central Asia, and Other Regions*. World Bank Publications.

Dal Bó, E. and P. Dal Bó (2011, 8). Workers, warriors, and criminals: Social conflict in general equilibrium. *Journal of the European Economic Association 9*(4), 646–677.

D'Altroy, T. N. (2015). *The Incas* (2nd ed.). Wiley-Blackwell.

Davis, C. L., A. Fuchs, and K. Johnson (2019). State control and the effects of foreign relations on bilateral trade. *Journal of Conflict Resolution 63*(2), 405–438.

De Commynes, P. (1972). *Memoirs: The Reign of Louis XI, 1461-83*. Penguin Books.

De Petra, G. (1866). *Sulle condizioni delle città italiche dopo la guerra sociale: con applicazioni alle colonie di Pompei e Pozzuoli*. Stamperia della R. Univ.

de Santana Pinho, P. (2009). White but Not Quite: Tones and Overtones of Whiteness in Brazil. *Small Axe 13*(2), 39–56.

Dennis, R. M. (1995). Social Darwinism, Scientific Racism, and the Metaphysics of Race. *The Journal of Negro Education 64*(3), 243.

Di Santo, A. (2016). *Guerre di torri: violenza e conflitto a Roma tra 1200 e 1500*. Roma: Viella.

Didier, L. and P. Koenig (2019). Has china replaced colonial trade? *Review of World Economics 155*(2), 199–226.

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

Dikotter, F. (Ed.) (1997). *The Construction of Racial Identities in China and Japan*. Hong Kong University Press.

Dincecco, M., G. Federico, and A. Vindigni (2011). Warfare, Taxation, and Political Change: Evidence from the Italian Risorgimento. *Journal of Economic History 71*(4), 887–914.

Dreher, A. and N. M. Jensen (2007). Independent actor or agent? an empirical analysis of the impact of us interests on international monetary fund conditions. *The Journal of Law and Economics 50*(1), 105–124.

Driscoll, J. C. and A. C. Kraay (1998). Consistent covariance matrix estimation with spatially dependent panel data. *Review of economics and statistics 80*(4), 549–560.

Du, Y., J. Ju, C. D. Ramirez, and X. Yao (2017). Bilateral trade and shocks in political relations: Evidence from china and some of its major trading partners, 1990–2013. *Journal of International Economics 108*, 211–225.

Duyvendak, J. (2008). The Last Dutch Embassy To the Chinese Court (1794-1795). *T'oung Pao 34*(1), 1–137.

Dworczak, P. and A. Pavan (2020). Preparing for the Worst but Hoping for the Best: Robust (Bayesian) Persuasion.

Earls, J. (1969). The Structure of Power in Quechua Mythology. *Journal of the Steward Anthropological Society 1*, 63–82.

Eichengreen, B., A. Mehl, and L. Chiţu (2019). Mars or mercury? the geopolitics of international currency choice. *Economic Policy 34*(98), 315–363.

Eliaz, K. and R. Spiegler (2015). X-games. *Games and Economic Behavior 89*, 93–100.

Elman, B. A. (2000). *A cultural history of civil examinations in late imperial China*. University of California Press.

Eraslan, H. and K. S. Evdokimov (2019). Legislative and Multilateral Bargaining. *Annual Review of Economics 11*(1), 443–472.

Esteban, J., L. Mayoral, and D. Ray (2012a). Ethnicity and conflict: An empirical study. *American Economic Review 102*(4), 1310–1342.

Esteban, J., L. Mayoral, and D. Ray (2012b). Ethnicity and conflict: Theory and facts. *Science 336*(6083), 858–865.

Esteban, J., M. Morelli, and D. Rohner (2015). Strategic mass killings. *Journal of Political Economy 123*(5), 1087–1132.

Esteban, J. and D. Ray (1999). Conflict and distribution. *Journal of Economic Theory 87*(2), 379–415.

Esteban, J. and D. Ray (2011). A model of ethnic conflict. *Journal of the European Economic Association 9*(3), 496–521.

Faye, M. and P. Niehaus (2012). Political aid cycles. *American Economic Review 102*(7), 3516–30.

Fearon, J. D. (2011). Self-enforcing democracy. *Quarterly Journal of Economics 126*(4), 1661–1708.

Firnhaber-Baker, J. (2010a). Seigneurial War and Royal Power in Later Medieval Southern France. *Past & Present* (208).

Firnhaber-Baker, J. (2010b). Techniques of seigneurial war in the fourteenth century Techniques of seigneurial war in the fourteenth century. *Journal of Medieval History 36*(1), 90–103.

Franck, R. H. (1985). *Choosing the right pond: Human Behavior and the quest for status*. Oxford University Press.

Francois, P., I. Rainer, and F. Trebbi (2015). How Is Power Shared in Africa? *Econometrica 83*(2), 465–503.

Friday, K. F. (2004). *Samurai, Warfare and the State in Early Medieval Japan*. Taylor & Francis e-Library.

Fries, L. (1713). Historie, Nahmen, Geschlecht, Wesen, Thaten, gantz Leben und Sterben des gewesenen Bischoffen zu Wirtzburg und Hertzogen zu Francken [1544]. In J. P. Ludewig (Ed.), *Geschicht-Schreiber von dem Bischoffthum Wirtzbürg*, pp. 838–45. Frankfurt am Main.

Fuchs, A. and N.-H. Klann (2013). Paying a visit: The dalai lama effect on international trade. *Journal of International Economics 91*(1), 164–177.

Garmaise, M. J. and G. Natividad (2013). Cheap credit, lending operations, and international politics: The case of global microfinance. *The journal of finance 68*(4), 1551–1576.

Gehlbach, S., K. Sonin, and M. W. Svolik (2016). Formal Models of Nondemocratic Politics. *Annual Review of Political Science 19*(1), 565–584.

Gennaioli, N. and H. J. Voth (2015). State Capacity and Military Conflict. *Review of Economic Studies 82*(4), 1409–1448.

Gentile, M. (2009). *Fazioni al governo: Politica e società a Parma nel Quattrocento*. Roma, Italia: Viella.

Gerner, D. J., P. A. Schrodt, and Ömüur Yilmaz (2009). Conflict and mediation event observations (cameo) codebook.

Ghazzali, M. (1933). Court Language and Etiquette of the Malays. *Journal of the Malayan Branch of the Royal Asiatic Society 11*(2), 273–287.

Goh, E. (2006). Understanding "hedging" in asia-pacific security. *PacNet 43*, 31.

Goh, E. (2007). Great powers and hierarchical order in southeast asia: Analyzing regional security strategies. *International Security 32*(3), 113–157.

Greif, A. (1998). Self-Enforcing Political Systems and Economic Growth: Late Medieval Genoa. In R. H. Bates, A. Greif, M. Levi, J.-L. Rosenthal, and B. R. Weingast (Eds.), *Analytic Narratives*, Chapter 1, pp. 23–63. Princeton, NJ: Princeton University Press.

Grillo, P. and A. A. Settia (2018). *Guerre ed eserciti nel Medioevo*. Bologna: il Mulino.

Grossman, H. I. (1991). A General Equilibrium Model of Insurrections. *American Economic Review 81*(4), 912–921.

Gutmann, M. P. (1988). The Origins of the Thirty Years' War. *The Journal of Interdisciplinary History 18*(4), 749–770.

Guzansky, Y. (2015). Strategic hedging by non-great powers in the persian gulf. In *Great Powers and Geopolitics*, pp. 231–252. Springer.

Haacke, J. (2019). The concept of hedging and its application to southeast asia: a critique and a proposal for a modified conceptual and methodological framework. *International Relations of the Asia-Pacific 19*(3), 375–417.

Haacke, J. and J. D. Ciorciari (2022). Hedging as risk management: Insights from works on alignment, riskification, and strategy.

Halac, M., I. Kremer, and E. Winter (2020). Raising Capital from Heterogeneous Investors. *American Economic Review 110*(3), 889–921.

Halac, M., I. Kremer, and E. Winter (2021). Monitoring teams.

Hamdi, S. and M. Salman (2020). The hedging strategy of small arab gulf states. *Asian Politics & Policy 12*(2), 127–152.

Harkness, K. A. (2016). The Ethnic Army and the State: Explaining Coup Traps and the Difficulties of Democratization in Africa. *Journal of Conflict Resolution 60*(4), 587–616.

Hausteiner, E. M. (2016). Managing the world: Conceptions of imperial rule between republicanism and technocracy. *History of European Ideas 42*(4), 570–584.

Hirschman, A. O. (1970). *Exit, voice, and loyalty: Responses to decline in firms, organizations, and states*. Harvard university press.

Hirschman, A. O. (1980). *National power and the structure of foreign trade*, Volume 105. University of California Press.

Horowitz, D. L. (1985). *Ethnic Groups in Conflict*. Berkeley: University of California Press.

Innes, M. (2000). *State and Society in the Early Middle Ages: The Middle Rhine Valley, 400-1000*. Cambridge University Press.

Inostroza, N. and A. Pavan (2020). Persuasion in Global Games with Application to Stress Testing.

Jackson, M. O. and M. Morelli (2009). Reasons for War: an Updated Survey. In C. Coyne (Ed.), *Handbook on the Political Economy of War*.

Jenkins, D. (1995). Temporal Processes and Synchronic Relations: Age and Rank in the Central Andes. *Ethnohistory 42*(1), 91.

Johnson, D. G. (1977). *The Medieval Chinese Oligarchy*. Boulder: Westview Press.

Jones, D. M. and N. Jenne (2021). Hedging and grand strategy in southeast asian foreign policy. *International Relations of the Asia-Pacific*.

Kaminsky, H. (2002). The Noble Feud in the Later Middle Ages. *Past & Present* (177), 55–83.

Kang, D. C. (2009). Between balancing and bandwagoning: South korea's response to china. *Journal of East Asian Studies 9*(1), 1–28.

Kant, I. (1983). *Perpetual peace and other essays*. Hacket Publishing.

Kasara, K. (2007). Tax Me If You Can: Ethnic Geography, Democracy, and the Taxation of Agriculture in Africa. *American Political Science Review 101*(1), 159–172.

Kendall, P. M. (1971). *Louis XI*. London, UK: George Allen and Unwin Ltd.

Kendall, S. (2013). *The Struggle for Roman Citizenship: Romans, Allies, and the Wars of 91-77 BCE*. Gorgias Press.

Kilby, C. (2009). The political economy of conditionality: An empirical analysis of world bank loan disbursements. *Journal of Development Economics 89*(1), 51–61.

King, E. (2013). *From classrooms to conflict in Rwanda*. Cambridge University Press.

Kleinman, B., E. Liu, and S. J. Redding (2020). International friends and enemies. Technical report, National Bureau of Economic Research.

Koga, K. (2018). The concept of "hedging" revisited: the case of japan's foreign policy strategy in east asia's power shift. *International Studies Review 20*(4), 633–660.

Kokkonen, A. and A. Sundell (2020). Leader Succession and Civil War. *Comparative Political Studies 53*(4), 434–468.

Kuik, C.-C. (2016). How do weaker states hedge? unpacking asean states' alignment behavior towards china. *Journal of Contemporary China 25*(100), 500–514.

Kuper, A. (1978). Rank and Preferential Marriage in Southern Africa: The Swazi. *Royal Anthropological Institute of Great Britain and Ireland 13*(4), 567–579.

Kuper, H. (1949). *An African aristocracy*. London: Oxford University Press.

Kuziemko, I. and E. Werker (2006). How much is a seat on the security council worth? foreign aid and bribery at the united nations. *Journal of political economy 114*(5), 905–930.

Lake, D. A. (2011). *Hierarchy in international relations*. Cornell University Press.

Lamberg-Karlovsky, C. C. (Ed.) (1989). *Archaeological thought in America*. Cambridge University Press.

Lansing, C. (1991). *The Florentine Magnates: lineage and faction in a medieval commune*. Princeton, NJ: Princeton University Press.

Leetaru, K. and P. A. Schrodt (2013). Gdelt: Global data on events, location and tone, 1979-2012. *ISA Annual Convention*.

Lemke, D. (2002). *Regions of war and peace*, Volume 80. Cambridge University Press.

Levitsky, S. and K. Roberts (2011).

Levy, J. S. and W. R. Thompson (2010, 6). Balancing on land and at sea: Do states ally against the leading global power? *International Security 35*(1), 7–43.

Li, K. and P. T. Ngo (2018). Geopolitics and international bank flows.

Liff, A. P. (2016). Whither the balancers? the case for a methodological reset. *Security Studies 25*(3), 420–459.

Lim, D. J. and Z. Cooper (2015). Reassessing hedging: The logic of alignment in east asia. *Security Studies 24*(4), 696–727.

Lim, D. J. and R. Mukherjee (2019). Hedging in south asia: balancing economic and security interests amid sino-indian competition. *International Relations of the Asia-Pacific 19*(3), 493–522.

Little, R. G. (1984). *The Parlement of Poitiers: war, government and politics in France, 1418-1436*. Royal Historical Society.

Madison, J. (1787). The Papers of James Madison, vol. 10, 27 May 1787-3 March 1788.

Marín-Guzmán, R. (1995). The Causes of the Revolt of Umar ibn Hafsun in al-Andalus (880-928): A Study in Medieval Islamic Social History. *Arabica 42*(2), 180–221.

Martines, L. (Ed.) (1972). *Violence and Civil Disorders in Italian Cities, 1200-1500*. Los Angeles: University of California Press.

McKnight, B. E. (1985). Song Legal Privileges. *Journal of the American Oriental Society 105*(1), 95–106.

Mearsheimer, J. J. (2001). *The tragedy of great power politics*. WW Norton & Company.

Mearsheimer, J. J. (2022). John mearsheimer on why the west is principally responsible for the ukrainian crisis. *The Economist March 19th 2022*.

Michalopoulos, S. and E. Papaioannou (2016). The long-run effects of the scramble for Africa. *American Economic Review 106*(7), 1802 – 1848.

Miller, D. (1980). Hume and Possessive Individualism. *History of Political Thought 1*(2), 261–278.

Mitra, A. and D. Ray (2014). Implications of an economic theory of conflict: Hindu-Muslim violence in India. *Journal of Political Economy 122*(4), 719–765.

Mitra, A. and D. Ray (2019). Hindu-Muslim Violence in India: A Postscript from the Twenty-First Century. *Advances in the Economics of Religion* (March), 229–248.

Mityakov, S., H. Tang, and K. K. Tsui (2013). International politics and import diversification. *The Journal of Law and Economics 56*(4), 1091–1121.

Montalvo, J. G. and M. Reynal-Querol (2005, 6). Ethnic polarization, potential conflict, and civil wars. *American Economic Review 95*(3), 796–816.

Morelli, M. and D. Rohner (2015). Resource concentration and civil wars. *Journal of Development Economics 117*, 32–47.

Morgenthau, H. J. (1948). *Politics Among Nations: The Struggle for Power and Peace*. New York: Alfred A. Knopf.

Morrock, R. (1973). Heritage of Strife: The Effects of Colonialist "Divide and Rule" Strategy upon the Colonized Peoples. *Science & Society 37*(2), 128–151.

Moseley, M. (2001). *The Incas and Their Ancestors: The Archaeology of Peru (Revised Edition)*. London, UK: Thames & Hudson.

Moyer, J. D., T. Sweijs, M. J. Burrows, and H. Van Manen (2018). Power and influence in a globalized world. *The Atlantic Council*.

Myerson, R. B. (2008). The autocrat's credibility problem and foundations of the constitutional state. *American Political Science Review 102*(1), 125–139.

Nexon, D. H. (2009). The balance of power in the balance. *World Politics 61*(2), 330–359.

Nietzsche, F. W. (1967). *The Will to Power* (Vintage ed.).

North, D. C., J. J. Wallis, S. B. Webb, and B. R. Weingast (Eds.) (2013). *In the shadow of violence: Politics, economics, and the problems of development*. Cambridge University Press.

North, D. C., J. J. Wallis, and B. R. Weingast (2009). *Violence and social orders: A conceptual framework for interpreting recorded human history*. Cambridge University Press.

Organski, A. F. and J. Kugler (2015). *The war ledger*. University of Chicago Press.

Padró i Miquel, G. (2007). The control of politicians in divided societies: The politics of fear. *Review of Economic Studies 74*(4), 1259–1274.

Papachristos, A. V. (2009). Murder by Structure: Dominance Relations and the Social Structure of Gang Homicide. *American Journal of Sociology 115*(1), 74–128.

Paul, T. V. (1994). *Asymmetric conflicts: war initiation by weaker powers*. Number 33. Cambridge University Press.

Qian, N. and D. Yanagizawa (2009). The strategic determinants of us human rights reporting: Evidence from the cold war. *Journal of the European Economic Association 7*(2-3), 446–457.

Qian, N. and D. Yanagizawa-Drott (2017). Government distortion in independently owned media: Evidence from us news coverage of human rights. *Journal of the European Economic Association 15*(2), 463–499.

Ray, D. and J. Esteban (2017, 8). Conflict and Development. *Annual Review of Economics 9*(1), 263–293.

Renshon, J. (2017a). *Fighting for status*. Princeton University Press.

Renshon, J. (2017b). *Fighting For Status: Hierarchy and Conflict in World Politics*. Princeton University Press.

Reynal-Querol, M. (2002). Ethnicity, political systems, and civil wars. *Journal of Conflict Resolution 46*(1), 29–54.

Roemer, J. E. (1985). Rationalizing Revolutionary Ideology. *Econometrica 53*(1), 85–108.

Rommel, T. and P. Schaudt (2020). First impressions: How leader changes affect bilateral aid. *Journal of Public Economics 185*, 104107.

Ross, R. S. (2006). Balance of power politics and the rise of china: Accommodation and balancing in east asia. *Security Studies 15*(3), 355–395.

Roy, D. (2005). Southeast asia and china: balancing or bandwagoning? *Contemporary Southeast Asia: A Journal of International and Strategic Affairs 27*(2), 305–322.

Sambanis, N. (2004). Using case studies to expand economic models of civil war. *Perspectives on Politics 2*(2), 259–279.

Segal, I. (1999a). Contracting with externalities. *The Quarterly Journal of Economics 114*(2), 337–388.

Segal, I. (1999b). Contracting with Externalities. *The Quarterly Journal of Economics 114*(2), 337–388.

Segal, I. (2003a). Coordination and discrimination in contracting with externalities: Divide and conquer? *Journal of Economic Theory 113*(2), 147–181.

Segal, I. (2003b). Coordination and discrimination in contracting with externalities: Divide and conquer? *Journal of Economic Theory 113*(2), 147–181.

Selden, Z. (2013). Balancing against or balancing with? the spectrum of alignment and the endurance of american hegemony. *Security studies 22*(2), 330–364.

Sherwin-White, A. N. (1973). *The Roman Citizenship* (2nd ed.). Oxford University Press.

Shifrinson, J. R. I. (2020). Eastbound and down: The united states, nato enlargement, and suppressing the soviet and western european alternatives, 1990-1992. *Journal of Strategic Studies 43*(6-7), 816–846.

Smith, N. R. (2020). When hedging goes wrong: Lessons from ukraine's failed hedge of the eu and russia. *Global Policy 11*(5), 588–597.

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

Stepan, N. L. (1991). *The Hour of Eugenics: Race, Gender and Nation in Latin America*. Ithaca (NY): Cornell University Press.

Stewart, N. (1951). Divide and Rule: British Policy in Indian History. *Science & Society 15*(1), 49–57.

Stollberg-Rilinger, B. (2018). *The Holy Roman Empire: A Short History*. Princeton University Press.

Strating, R. (2020). Small power hedging in an era of great-power politics: Southeast asian responses to china's pursuit of energy security. *Asian Studies Review 44*(1), 97–116.

Strauss, G. (1978). The Holy Roman Empire Revisited. *Central European History 11*(3), 290–301.

Sundberg, R., K. Eck, and J. Kreutz (2012). Introducing the UCDP Non-State Conflict Dataset. *Journal of Peace Research 49*(2), 351–362.

Tackett, N. O. (2006). *The Transformation of Medieval Chinese Elites (850-1000 C.E.)*. Ph. D. thesis, Columbia University.

Tang, S. (2008). Fear in international politics: two positions. *International Studies Review 10*(3), 451–471.

Tessman, B. F. (2012). System structure and state strategy: Adding hedging to the menu. *Security Studies 21*(2), 192–231.

Ticchi, D. and A. Vindigni (2008). War and Endogenous Democracy.

Tortella, G. (2010). *The Origins of the Twenty-first Century*. New York, NY: Routledge.

Twinam, A. (2015). *Purchasing whiteness: Pardos, mulattos, and the quest for social mobility in the Spanish Indies*. Stanford University Press.

Vale, M. G. A. (1974). *Charles VII*. London: Eyre Methuen.

Valente, C. (2003). *The Theory and Practice of Revolt in Medieval England*. Aldershot, England: Ashgate Publishing Limited.

Vigueur, J.-C. M. (2004). *Cavalieri e Cittadini: guerra, conflitti e società nell'Italia comunale*. Bologna: il Mulino.

Voss, B. L. (2008). *The archaeology of ethnogenesis: Race and sexuality in colonial San Francisco*. University of California Press.

Waltz, K. N. (1979). *Theory of international politics*. McGraw-Hill.

Weiner, M. (Ed.) (2009). *Japan's Minorities: The Illusion of Homogeneity.*, Volume 53. Routledge.

Weingast, B. R. (1997). The Political Foundations of Democracy and the Rule of Law. *American Political Science Review 91*(2), 245–263.

Wilson, P. H. (2014). Prussia and the Holy Roman Empire 1700-40. *German Historical Institute London Bullettin 36*(1), 3–48.

Winter, E. (2004a). Incentives and discrimination. *American Economic Review 94*(3), 764–773.

Winter, E. (2004b). Incentives and Discrimination. *American Economic Review 94*(3), 764–773.

Wivel, A. (2008). Balancing against threats or bandwagoning with power? europe and the transatlantic relationship after the cold war. *Cambridge Review of International Affairs 21*(3), 289–305.

Wohlforth, W. C., R. Little, S. J. Kaufman, D. Kang, C. A. Jones, V. Tin-Bor Hui, A. Eckstein, D. Deudney, and W. L. Brenner (2007). Testing balance-of-power theory in world history. *European Journal of International Relations 13*(2), 155–185.

Wrangham, R. W. and M. L. Wilson (2004). Collective violence: Comparisons between youths and chimpanzees. *Annals of the New York Academy of Sciences 1036*(1), 233–256.

Yeats, A. J. (1990). Do african countries pay more for imports? yes. *The World Bank Economic Review 4*(1), 1–20.

Zmora, H. (2002). *Monarchy, Aristocracy and the State in Europe 1300-1800*. London and New York: Routledge.

Zmora, H. (2003). *State and nobility in early modern Germany: the knightly feud in Franconia, 1440-1567*. Cambridge University Press.

Zmora, H. (2015). *The Feud in Early Modern Germany*. Cambridge University Press.

Zmora, H. (2020). Reputation as a common source of cooperation and violent conflict: The case of the noble feud in early modern Germany. *Evolution and Human Behavior 41*(3), 218–225.

Zorzi, A. (Ed.) (2009). *Conflitti, paci e vendette nell'Italia comunale*. Firenze, Italia: Firenze University Press.

Appendix A

Appendix to Chapter 1

A.1 Proofs

Proof of Theorem 1. Consider a robust cost-minimizing offer $\hat{\mathbf{B}}$. Without loss of generality, relabel the coercers so that $i > j$ if $\hat{B}_i > \hat{B}_j$ (arbitrarily if $\hat{B}_i = \hat{B}_j$). Let $n = |C|$ be the label of the coercer that received the highest offer, let 1 be the label of the coercer with the smallest offer, and let $n+1$ be the label of the policymaker. Note that a robust cost-minimizing offer must be such that the following condition holds:

$$\left[1 - F\left(\frac{\sum_{j=1}^{n-1} W_{n-1}}{\sum_{j=1}^{n+1} W_j}\right)\right] \hat{B}_n = \kappa \tag{A.1}$$

Any smaller $\hat{B}_n$ would give up robustness, as any such $\hat{\mathbf{B}}$ would be consistent with an SPE in which every coercer oppose the policymaker. Any larger $\hat{\mathbf{B}}$ would not be cost-minimizing. Now, given that any cost-minimizing robust offer must satisfy the above condition, we know that any robust cost-minimizing offer must also satisfy the following requirement:

$$\left[1 - F\left(\frac{\sum_{j=1}^{n-2} W_j}{\sum_{j=1}^{n+1} W_j}\right)\right] \hat{B}_{n-1} = \kappa \tag{A.2}$$

Any smaller $\hat{B}_{n-1}$ would give up robustness, as any such $\hat{\mathbf{B}}$ would be consistent with an SPE in which every coercer except n oppose the policymaker. However, if $\hat{\mathbf{B}}$ satisfies Equation A.1 (which must be true as $\hat{\mathbf{B}}$ is robust and cost-minimizing), then any greater $\hat{B}_{n-1}$ would not be cost-minimizing.

Iterating this argument, we see that any robust cost-minimizing offer must be such that $[1 - F(0)]\hat{B}_1 = \kappa$, and for every $c \in C$ we have:

$$\left[1 - F\left(\frac{\sum_{j=1}^{c-1} W_c}{\sum_{j=1}^{n+1} W_c}\right)\right] \hat{B}_c = \kappa \tag{A.3}$$

The argument so far showed that any cost-minimizing robust offer can be constructed by relabeling coercers (assigning ranks) and making the smallest offer that induces each coercer to support the policymaker when all and only policymakers with a greater label support the policymaker. We are left with showing that that $B_1 > B_2$ if $W_1 > W_2$, i.e. that higher labels are assigned to stronger coercers. For notational convenience, let $w_j = \frac{W_j}{\sum_{j=1}^{n+1} W_c}$. Given the above characterization, we can derive the total cost associated with any offer $\hat{B}$ that satisfies the conditions above. Such total cost takes the following form:

$$\sum_{j=2}^{n} \hat{B}_c = \frac{\kappa}{[1 - F(0)]} + \frac{\kappa}{[1 - F(w_1)]} + \frac{\kappa}{[1 - F(w_1 + w_2)]} + ... + \frac{\kappa}{[1 - F(w_1 + w_2 + .. + w_n)]}$$

Because $F(\cdot)$ is strictly increasing, this expression implies that if there are two coercers i and j such that $i < j$ but $w_i > w_j$ then the policymaker could reduce its total costs by switching the offers that it makes to them. By construction, such deviation would not have effect on the robustness of the offers, as this total cost structure already reflects the necessary conditions summarized by Equation A.3. This leads to the conclusion that if $W_i > W_j$ then it must be that $B_i > B_j$.

■

Proof of Propositions 1, 2, 3, and 4. Restrict attention to distributions of power where there are no ties, i.e., with $W_c = W_{c'}$ if and only if $c = c'$. From Theorem 1 we know that this restriction implies that there is a unique robust cost-minimizing offer $\hat{\mathbf{B}}$. From Equation A.3 we know that, in such equilibrium, each c receives the following offer

$$\hat{B}_c = \frac{\kappa}{1 - F\left(\frac{\sum_{\ell:W_\ell<W_c} W_\ell}{\sum_{j\in C} W_j + W_p}\right)} \tag{A.4}$$

First note that the distribution of power only enters each $\hat{B}_c$ via the argument of $F(\cdot)$. The first part of Proposition 3 thus follows from the fact that we defined the IE-index of a country as exactly the argument of $F(\cdot)$, i.e. $\frac{\sum_{\ell:W_\ell<W_c} W_\ell}{\sum_{j\in C} W_j + W_p}$. The other part of Proposition 3 follows form the observation that $\hat{B}_c$ is increasing in $F\left(\frac{\sum_{\ell:W_\ell<W_c} W_\ell}{\sum_{j\in C} W_j + W_p}\right)$ and $F(\cdot)$ is increasing, thus $\hat{B}_c$ is increasing in the IE-index.

With this in mind, to establish Propositions 1 and 2 it is enough to show that the stated comparative statics hold relative to the IE-index. Proposition 1 follows immediately from the IE-index being increasing in the power of weaker countries and decreasing in the power of stronger ones. Propositions 2 follows from the observation that a country's own power is at the denominator of its IE-index.

Finally, for what concerns Proposition 4, note that once we assume that $F(w) = 1 - e^{-\lambda w}$, then the equilibrium offer is $\hat{B}_c = \kappa e^{\lambda \frac{\sum_{\ell:W_\ell<W_c} W_\ell}{\sum_{j\in C} W_j + W_p}}$. Taking the natural logarithm and adjusting the notation leads to the result stated in Propositions 4.

■

A.2 Data appendix: description and construction

This Section provides details on how we constructed our variables of interest. Appendix A.3.4 shows that our analysis is robust to alternative transformations.

A.2.1 Trade flows

Trade flows are constructed from the CEPII Gravity dataset. The dataset reports trade flows from multiple sources (IMF, UN, BACI). These sources do not always agree with each other and sometimes one source features missing values when another one does not. Additionally, IMF and UN data distinguishes between the trade value as reported by the origin vs. the destination country. For our baseline analysis we integrated the various sources to obtain the largest dataset possible. Importantly, all these trade flows are highly correlated and results are robust to alternative choices.

When building our measure of power over a given policymaker from trade flows, we prioritized information on the value of trade flows as reported by the policymaker. Specifically, we started from IMF data on policymaker-reported trade flows. If missing, we added the UN Comtrade policymaker-reported value of the trade flow. If missing, we added the coercer-reported IMF figure. If missing, we added the coercer-reported observation from the UN dataset. And finally we integrated missing observations using the BACI dataset.

We used a similar procedure when building flows of imports and exports to use as our outcomes. As in this case our focus is on the effect on the coercer, we followed the precedure outline above but prioritizing coercer-reported figures.

A.2.2 GDELT data

As mentioned, the GDELT dataset contains information on different types of international interactions (each defined as a distinct CAMEO category) between 1978 to 2012 at the

country pair level and with a daily frequency. To derive a dataset with a yearly frequency, we sum up all the interactions of the same type taking place between any two countries in distinct days but in the same year. In addition to such aggregation, every variable from GDELT is transformed in two ways: a normalization to compare the estimated coefficients and a substitution to avoid drooping the zeros. The following provides details and discusses the logic of each of these exercises.

Aggregation. Let $\mathbf{y}_{d,fh,\tau(t)}$ be the vector where $y_{d,fh,\tau(t)} = 1$ when we observe an interaction of type d (e.g., one of those described in Table 1.1) between any country c and any country p in any day τ of any year t, and $y_{d,fh,\tau(t)} = 0$ when we do not. We construct our yearly variables with the following procedure. We aggregate over the given year t for every country c, country p and event of type d, i.e. we compute $\tilde{Y}_{cpt} = \sum_{\tau:\tau^{-1}(\tilde{t})=t} y_{d,fh,\tau(t)}$. In an unreported robustness check, we show that our results are robust to alternative aggregation rules. For instance, we show that we can replicate our results when constructing outcomes by summing up, for each coercer-policymaker pair, the interactions of the same kind happening in the same year but in distinct half-weeks, weeks, half-months, months or trimesters.

Substitution. The second transformation replaces every observed 0 with 0.5, a transformation that avoids dropping the zeros when taking the log, while still preserving the shape of the log. The interpretation for this transformation is that any observed zeros does not really correspond to a situation with no interactions, but rather to a situation with very few of them. Importantly, our analysis also holds when we simply take the log (thus dropping all zeros) or when we employ the more standard transformation of adding 1 to each variable before taking the log. Finally, note that in our baseline specification we control for $ln(Trade_{cpt})$, we are thus dropping all observations with a small treatment (i.e. any coercer-policymaker pair where the coercer has zero economic power over the policymaker). Define $\underline{Y}_{cpt}$ so that $\underline{Y}_{cpt} = \tilde{Y}_{cpt}$ if $\tilde{Y}_{cpt} > 0$, and set $\underline{Y}_{cpt} = 0.5$ if $\tilde{Y}_{cpt} = 0$.

Normalization. The final transformation is a normalization such that the standard

deviation of the log of each variable is fixed to 1. Define the standard deviation of each $\ln(\underline{Y}_{cpt})$ as $\sigma_{Y,d}$. Our final variable is defined as $Y_{cpt} = \underline{Y}_{cpt}^{\sigma_{Y,d}^{-1}}$. This is the variable that we analyze in Section 1.4.2.

A.3 Robustness exercises

All Tables related to our robustness exercises can be found after Appendix A.4. To express our results compactly, we focused on reporting only the coefficients of interest and their standard errors. More details are available upon request.

A.3.1 Fixed effects

Our first robustness exercise shows that our results hold across different fixed effects specifications. We now discuss the interpretation and logic behind checking alternative fixed effect specifications, how we can interpret the differences between these exercises, and how these findings speak to our main results and, more in general, our theory. Recall that, in our baseline analysis, we included α_{pc}, α_{pt}, and α_{ct} as fixed effects. This means that we could not include the standard gravity controls such as (log of) GDP, GDP per capita, and (log of) Population for both countries. While results can be shown not to depend on whether we do so, the coefficients reported in Table A.3.3 are estimated by including the largest possible vector of Gravity controls that is compatible with the fixed effect specification we consider.

Year fixed effects. Panel A of Table A.3.3 only includes α_{pc} (and gravity controls), without year fixed effects. The estimated coefficients could be simply driven by the upward trend of both the Indirect Exposure Index and our measures of international cooperation. While both variables do feature an upward trend over time (and so do most of the Gravity controls), when we compare Panel A and Panel B of Table A.3.3 (where we include the year fixed effect) we see that the inclusion of a year fixed effect has almost no impact on the

coefficients, suggesting that the upward trend was not important for the estimates in Panel A.[1]

Why Policymaker×Year fixed effects? By including α_t, Panel B of Table A.3.3 controls for any global process or time trend affecting countries homogeneously. However, relative to our baseline specification, it omits both α_{pt} and α_{ct}. The omission of α_{pt} implies that the results in Panel B might be driven by some time-varying feature of the policymakers that might affect how (and with whom) they trade and also their international interactions. For instance, if smaller and poorer countries were more likely to have fewer international interactions (as recorded in the media) and they also were more likely to be highly dependent on one particular country for their trade, then in Panel B we would find a positively biased estimate of the relation between the Indirect Exposure Index and the overall extent of bilateral international interactions. On the other hand, if countries who have less international interactions (or are less likely to be reported in the news) were on average more likely to trade with more countries and not to depend too much on any of them, then this would induce a negative bias in the estimate reported in Panel B for how the Indirect Exposure Index affects our various bilateral outcomes. Importantly, once we include α_{pt}, then any such effect would be muted.

Effect of including Policymaker×Year fixed effects. Comparing Panel B with Panel D of Table A.3.3, where we include α_{pt}, we find that the results do not change by much, suggesting that neither of the alternative explanations discussed above induces an important bias on the IE-index. If any effect can be discerned, we see that adding α_{pt} increases the estimated coefficient, suggesting that the second hypothesis is more important for the average coercer-policymaker pair that we observe. Note that the coefficients reported control for the policymaker's GDP (log), GDP per capita and Population (log), so the effect

[1]Column (1) in Tables A.3.1 and A.3.2 show that the IE-index is important for Imports and Exports also when we only consider coercer-policymaker pair fixed effects α_{pc} without year fixed effects α_t.

of size or wealth should be captured by these variables (if the functional form is correct). But we verified that our conclusion about the comparison between Panel B and D does not change when we omit these controls.

Why Coercer×Year fixed effects? Compared to our baseline specification, another limitation of Panels A, B and D of Table A.3.3 is that they omit α_{ct}. This can be important as we wish to be robust relative to the possibility that our results are driven by differences in coercers in particular periods. For instance, consider two large countries as coercers who only interact and trade with their separate geopolitical regions. One is positioned in a thriving region where there are no other big countries, and the average country trades a lot with a lot of other small ones (but mainly with the large one). The other coercer is also the only large country in its geopolitical region, but such region is poor or more ill-connected and the average country only trades with the big one. The first coercer would have a larger average IE-index than the other one IE-index. Unfortunately, it is also easy to see how such scenarios would reflect in differences in outcomes. For instance, we might expect that in more dynamic geopolitical regions there is more potential for misunderstandings, which can translate into a larger international role as mediator (or rule-setter) for the large country than the role it would have in the poorer region. Alternatively, we can expect large countries in poorer regions to cannibalize all international interactions, possibly because they are the only country that matters in that specific geopolitical environment. All these are examples of conditions that might affect both our outcomes and the IE-index, potentially biasing our coefficients. Moreover, local conditions like the ones we described are likely to change over time, thus casting doubts on the possibility to capture these effects by simply including α_{pc}. On the other hand, our baseline analysis takes care of these issues by including α_{ct}, thus focusing on the deviations in a coercer's outcomes relative to its *yearly* average outcome.

Effect of including Coercer×Year fixed effects. Comparing Panel C of Table A.3.3 and the panels that do not include α_{ct} (Panels A, B and D of Table A.3.3), we see that

the inclusion of α_{ct} reduces our estimated coefficients. While this can be interpreted as a by-product of one of the stories that we highlighted above, another possibility, which is in line with our theory, is that the reduced coefficient comes from the fact that these fixed effects absorb also unilateral measures of c's power, e.g. the size of its military. As we can expect countries with a stronger military to be able to more effectively use their economic leverage, then the inclusion of α_{ct} implies that we are likely to be left with a type of variation that understates the effect of economic power on international interactions. Even if this is somewhat undesirable, we still prefer to include α_{ct} in our baseline specification to be as robust as possible to alternative explanations. With the fixed effects introduced in our baseline, even a time-dependent, non-linear and non-monotone effect of size, wealth or any other unilateral characteristics would be captured by α_{ct} and α_{pt}, and cannot thus affect the estimates of our baseline specifications.

Regional-pairs×Year fixed effects. Finally, Panel E of Table A.3.3 shows that our results are robust to an even more demanding specification than our baseline one. On top of α_{pc}, α_{pt}, and α_{ct}, we add time-varying fixed effects for each pair of geopolitical regions. Letting $S(j)$ is the geopolitical region of country j, these fixed effects capture *time-varying* characteristic that can be linked to the particular geopolitical region of a coercer and of a policymaker $\alpha_{S(c)S(p)t}$.[2] A potential reason to include $\alpha_{S(c)S(p)t}$ is that we wish to be robust to possible events that affect the international interactions between any two geopolitical regions in a time-varying way and that can in principle affect trade flows (thus our measures of power). For instance, the formation of the European Union and the process of integration that followed such event is likely to have had a common effect (certainly time varying, at

[2]We did this robustness check with both definitions of geopolitical regions included in the CEPII dataset. In Panel E we report the coefficients obtained from a division of the world in 10 areas: Eastern Europe and post Soviet Union, The Caribbean, Latin America, North Africa and the Middle East, Sub-Saharan Africa, Western Europe and North America, East Asia, South-East Asia, South Asia, The Pacific. In another exercise (unreported) we focus on a division of the world in 6 regions: Africa, North America, Latin America, Asia, Europe, Middle East.

least within our sample period) on all countries involved, both in their pattern of cooperation with other regions of the world *and* their pattern of trade with them. Another example is the collapse of the Soviet Union. Clearly, this event affected Eastern Europe and the ex members of the Soviet Union. It affected both their relations with the other regions of the world and their pattern of trade. Note that both the effects of the collapse of the Soviet Union and the formation of the European Union on any single coercer or policymaker would be captured by α_{ct} and/or α_{pt}, *unless* such effects have a systematic time-dependent variation that is linked to the country pair considered. The inclusion of $\alpha_{S(c)S(p)t}$ accounts for this, when such systematic variation is not specific to each specific country pair but rather to the pair of geopolitical region in which such countries reside. In Panel E of Table A.3.3 we find that the addition of this type of fixed effect to our baseline specification (reported in Table 1.3) has only a mild impact on the magnitude of the estimated coefficients, and does not affect their statistical significance. Thus, our results are unlikely to be driven by one of the many large supranational geopolitical shocks that happened in the sample period under consideration.

Finally, while we do not report the related tables, we have verified that the results in Tables 1.5, 1.6, and 1.7 are robust to an appropriate set of alternative fixed effect specifications.

A.3.2 Alternative prediction: Splitting the IE-Index

In Section 1.3.1 we assumed that the probability of political survival of the policymaker was simply a decreasing function of the power share of all countries that opposed the policymaker. A more general specification would be to allow for such probability to be increasing in the power of all supporting coercers and decreasing in the power of all coercers that oppose the regime. Applying the same argument of the proof of Theorem 1 to this more general model would yield a more general prediction: every coercer's equilibrium transfer is increasing in

the power of all weaker coercers and decreasing in the power of every stronger one.

There are two reasons why, in the main text, we ignored this more general prediction. First, we wanted to express our results in terms of power shares, as this enables us to compare our prediction with the natural prediction that relatively more powerful countries (countries with a larger power share) obtain better outcomes. Second, for robustness concerns, we wanted to include α_{pt}, a policymaker-year fixed effects. Clearly, once we include such fixed effects and we express our results in terms of power shares, then we cannot estimate the effect of the power share of stronger coercers *and* the effect of the power share of weaker ones, as the two become perfectly collinear. However, as in Section A.3.1 we showed that excluding α_{pt} from our specification does not alter the results too much, we are now in a position to discuss the merits of this generalized prediction.

In Table A.3.4 we replicate the analysis of Panel A, B, and D of Table A.3.3, but this time we add the Stronger Powers Index, i.e., the share of total power of all stronger coercers. Note that this variable is not collinear with the Indirect Exposure Index: in the absence of α_{pt}, we are left with some variation that is due to the share of total power over the policymaker that is controlled by the policymaker, i.e., the policymaker's domestic power (its ability to resist external pressure). The reason why we still express these findings in terms of shares is that we wish to make our results comparable with our baseline findings.

Table A.3.4 shows that once we control for the Indirect Exposure Index, the Stronger Powers Index is negative, albeit only marginally significant. The fact that the Indirect Exposure Index remains positive and significant and that the Stronger Powers Index is instead close to zero justifies our focus on the simpler specification of the model, where the probability of survival is a one-dimensional function. If we wanted to take the negative coefficients of the Stronger Powers Index at face value, then this would suggest that any generalization of the model would require to add an extra stabilizing impact of the power of supporting countries than what can be captured by the Indirect Exposure Index.

A.3.3 Estimating equations

The objective of this Section is to show that results are robust to a variety of deviations from the estimating equations that we focused on in Sections 1.4.3, 1.5, and 1.6.

The estimating equations of Sections 1.4.3 and 1.5 are based on the finding that we should expect a positive log-linear relation between our outcomes measuring international interactions and the Indirect Exposure Index. We thus focused on a log-linear relation, adding a restrictive set of fixed effects (see Section A.3.1 for a discussion and for alternative fixed effect specifications). On top of that, we introduced two controls. One is a coercer's Relative Economic Power (REP) over the policymaker, which enables us to distinguish our theory against the natural alternative hypothesis that stronger countries obtain better relations. The other is the trade flow between the policymaker and the coercer, as it is natural to expect that countries who trade more should have friendlier relations (and also larger IE-index).

To be as conservative as possible, in our baseline specification we controlled for the *contemporaneous* trade flow and we used the *lagged* distribution of power to compute the IE-index and the REP. In Table A.3.5 we present the coefficients under different specifications. We show that the Indirect Exposure Index remains an important determinant of international cooperation regardless of whether we include trade flows, whether we include the *contemporaneous* trade flows *and* the *contemporaneous* REP, and whether we focus on the *contemporaneous* IE-index, rather than the lagged one. A comparison between our baseline estimates (the even Columns of Table 1.2) and Panel A of Table A.3.5 (where we do not include the contemporaneous trade flow) reveals that the estimated coefficients of the IE-index are not affected by such change of specification. Panel B and C of Table A.3.5 show that when we substitute the lagged REP with the contemporaneous one, our coefficient for the lagged IE-index is slightly reduced, but remains significant (regardless of whether we

introduce trade flows as in Panel B or we do not as in Panel C). In Panel D of Table A.3.5 we show that focusing on the contemporaneous rather than lagged IE-index does not alter our results. Finally, Panel E of Table A.3.5 shows that results are robust to the inclusion of the lag of the outcome. This exercise makes sure that the estimated effects do not change too much when we allow for a dynamic specification of the empirical model. We find that this addition reduces the magnitude of the estimated impact of the IE-index, but it remains significant at the 1% level with respect to all our outcomes. In an unreported robustness check we also verified that the effects are still significant when we allow the lag of the outcome variable to have a time-varying, a policymaker-specific or a coercer-specific coefficient. Also, while we do not report the Tables for it, the same robustness exercises can be shown to hold for our event studies (Tables 1.5 and 1.6).

For the case of Imports and Exports (Section 1.6), we considered estimating Equations based on the large literature employing Gravity models to estimate bilateral trade flows. We simply added the lagged IE-index and the lagged REP to the standard estimating equations. Tables A.3.1 and A.3.2 show that deviations from our estimating equations do not change the results.

A.3.4 Outcome definitions

Section A.2 presented and justified the transformations that we performed before analyzing our outcome variables. The first of these transformations, the normalization, is simply a rescaling of the coefficient, so it does not play any role in the statistical significance of the parameters of interest. As a second transformation, we substituted any observed 0 with 0.5 to avoid dropping the zeros when taking the log. As justified, we are interested in this transformation as we wish to preserve the shape of the log as much as possible, as this is part of the prediction of our model. Table A.3.6 shows that the result are robust to alternative

transformations. Indeed, our results are found to hold even when we simply take the log (dropping all the zeros) or when we adopt the more standard approach of analyzing $\ln(1+Y)$, i.e. of adding 1 to each variable before taking the log. The same robustness exercises can be shown to hold for Tables 1.5 and 1.6.

A.3.5 Model-based standard errors

Table A.3.7 and A.3.8 show that the way we calculate standard errors does not matter for the statistical significance of our coefficients of interest. Note that in our baseline specification we use two-way clustered standard error, where we consider as clusters each coercer and each policymaker. While this allows for a general form of dependence, it does not allow for correlation across country pairs. This restriction might be at odds with the very alternative hypothesis that we are trying to test (the importance of the IE-index), as the key prediction of our model is that the bilateral outcomes of each coercer-policymaker pair should depend on the characteristics of other coercers.

As a first approach, we calculate three-way clustered standard errors (labeled "3w-Clusters c, p, t" in Table A.3.7), where we also cluster with respect to time (each year). Adding time as a clustering dimension yields standard errors that are robust to a general form of contemporaneous correlation across country pairs. As we can see, the standard errors increase by as much as 30%, but all coefficients remain significant with a p-value below 1%. A limitation of this approach is that it does not allow for the correlation across country pairs to have an auto-correlated component. To calculate standard errors that are robust to this type of error structure, we calculate the standard errors proposed in Driscoll and Kraay (1998). We calculate Driscoll-Kraay standard errors allowing for various degrees of persistence in the autocorrelation component: either four lags (labeled "D-K 4 lags" in the Table) or seven lags (labeled "D-K 7 lags" in the Table). Tables A.3.7 and A.3.8 show that

these standard errors can be twice as large as the ones we computed in the baseline analysis. However, coefficients remain statistically significant at the 1% level (except for Humanitarian Aid, whose significance is now between the 1% and the 5% level).

A.3.6 Conditioning on the other outcomes

While we interpret the analysis of different outcomes as a form of cross-validation of the theory, it is still true that these outcomes are likely to be related to each other independently of their link to the Indirect Exposure Index. In this Section we show that the IE-index has a direct effect on each of our outcomes, even when we condition on various combination of the other outcomes, or other related variables (Table A.3.9). Column (6) of Tables A.3.1 and A.3.2 does the same exercise but for imports and exports.

Panel A of Table A.3.9 controls for *all types* of offers sent by the coercer to the policymaker. This exercise allows for the possibility that countries are more likely to express their intentions to cooperate (which would be categorized as an offer) as a response to the other country expressing such intentions. While we do find a positive effect of the number of offers sent on the number of offers received (and on aid sent), Panel A shows that the IE-index remains a significant determinant of the offers received and of aid sent, suggesting that the IE-index has a direct effect on explaining the number of offers received (or aid sent) conditional on the numbers of offers sent.

Panel B of Table A.3.9 controls for all types of offers sent by the coercer *and* all types of aid sent. This exercise allows for the possibility that a policymaker is more likely to make an offer to a coercer (i.e., be reported as declaring some intention to improve cooperation with the coercer) when such coercer sends aid to the policymaker. While it can be shown that such mechanism is relevant, Panel B shows that the IE-index remains a significant determinant of the number of offers sent, conditional on the number of offers received and

on aid sent.

Finally, Panel C of Table A.3.9 controls for all types of offers sent and all types of offers received. This allows for the possibility that a country who receives more promises of future cooperation (and who sends more) is more likely to send aid. Again, we show that this mechanism is an important determinant of aid sent, but we also find that the IE-index still has a significant and positive effect on aid sent, conditional on the offers sent or received.

We interpret these exercises as revealing that the Indirect Exposure Index captures some variation that is intrinsic to each of these variables, and not only a subset of them. From this perspective, Table A.3.9 lends credence to the idea that our focus on different outcomes is a form of cross-validation.

A.4 Validation exercises

This Section shows that the analysis can be replicated with an alternative measure of power, with alternative outcomes, and restricting the attention to some subsamples of interest. To express our results compactly, we focused on reporting only the coefficients of interest and their standard errors. More details are available upon request.

A.4.1 Replication with alternative measure of power: the Formal Bilateral Influence Capacity (FBIC)

In this section we show that we can replicate all of our analysis with an alternative measure of power, the Formal Bilateral Influence Capability (FBIC).[3] This variable is a composite measure of bilateral power that measures economic power but also other dimensions of power, as it uses information on arms transfers, aid dependence, formal role in international institutions, and other measures (Moyer et al., 2018).

[3]This measure can be accessed from `https://korbel.du.edu/fbic`.

Tables A.4.1 and A.4.2 present statistics on the relation between the FBIC (denoted $I_t^{c \to p}$) and economic power, the measure we analyzed in our baseline specification (denoted $E_t^{c \to p}$). As we should expect, Table A.4.1 shows that there is a strong positive correlation between the two measures of power under consideration (0.6 between the actual variables, and 0.73 for the IE-index computed from the two variables). However, Table A.4.2 shows that this correlation is drastically reduced when we look at the variation that remains after we net out all the fixed effects that we include in our baseline specification. Specifically, when we look at deviations from α_{pc}, α_{pt}, and α_{ct}, the two variables only exhibit a 0.21 correlation and the two IE-index that can be computed from these variables have a correlation of 0.36. These statistics suggest that the two variables are effectively different, making the replication of the analysis all the more important.

Tables A.4.3 and A.4.4 replicate the results of Tables 1.3 and 1.7 substituting our measure of power with this alternative one. Note that the estimated coefficients are around half of those that we estimated in our baseline analysis (one third for the case of imports and exports). However, the standard deviation of the FBIC (and the IE-index related to the FBIC) is around three times that of economic power (see Tables A.4.1 and A.4.2). As a consequence, we can show that the impact of increasing the FBIC-based IE-index by one standard deviation is greater than the comparable counterfactual we proposed in our baseline analysis.

A.4.2 Replication with alternative outcomes: Agreements

As a further validation check, we show that we can replicate all of our analysis focusing on another set of outcomes from the GDELT datasets: *agreements*, measuring for each country pair and in each year the initiation, resumption, improvement or expansion of cooperation in a particular domain (economic, military, diplomatic, and judiciary). In addition, we also

focus on what we call *political agreements* which captures the ratification, signature and finalization of an agreement or treaty, regardless of its nature. A more precise description of these variables are reported in Table A.4.5.[4]

Similar to the offers studied in our baseline specification, agreements measure the extent of cooperation in each coercer-policymaker pair. However, unlike offers, these variables capture the actual conclusion of deals and agreements between the two parties. In other words, agreements are the realization of the offers. An important drawback of focusing on agreements is that they are undirected: the agreements that country A concludes with country B are the same that country B concludes with A. To avoid including the same outcome twice, we thus restrict attention to *asymmetric country pairs*: the subsample of coercer-policymaker pairs where the coercer has more power over the policymaker than viceversa.[5]

Tables A.4.6, A.4.7, and A.4.8 replicate our baseline analysis, the event study, and the event study restricted to the United States. All of our results are shown to hold, providing a further demonstration of the importance of the IE-index (thus our model) to understanding the pattern of international cooperation.

A.4.3 Replication in different subsamples: foreign powers as coercers and Asymmetric Relations

Our final validation exercise replicates our analyses in different subsamples of interest. In line with our interpretation of the model as describing how foreign powers affect domestic decision-making, we focus on subsamples where only foreign powers can be coercers, and

[4]We perform the same transformations that we did for all the other variables from GDELT (see Appendix A.2 for details).

[5]Results are robust to focusing on the full sample. Moreover, in Section A.4.3 we show that all of our results can be replicated when focusing on asymmetric country pairs.

they are excluded from the set of policymakers.[6] We define as *foreign powers* those countries that have been at the top of the distribution of GDP in at least a year in our sample. Table A.4.9 lists all coercers that we define as foreign powers for different definitions that we adopt (top 5, 10, 20, or 30). Table A.4.10 successfully replicates our results in these subsamples, confirming that our theory is a useful description of the relation between foreign powers and non-foreign powers (Panels A, B, C, and D of Table A.4.10). Finally, Panel E of Table A.4.10 replicates the analysis for *asymmetric country pairs*, the subsample of coercer-policymaker pairs where the coercer has more power over the policymaker than viceversa. Again, this analysis suggests that our results are important to understand asymmetric relations, where power (thus our theory) is likely to be more important.

While we do not report the relevant table, we also manage to replicate our results on imports and exports (Table 1.7) for the case of asymmetric country pairs. On the other hand, when focusing on the relations between foreign powers and non-foreign powers, we only start finding statistically significant effects when defining foreign powers as countries that have ever been in the top 40 of GDP. The problem with limiting our attention to a smaller set of coercers when analyzing exports and imports is likely due to our empirical strategy, which adopts unilateral measures of power (in our first stage estimates). As a consequence of this, we do not have enough variation when we restrict attention to a few coercers, as we cannot exploit variations *within* coercers in the same year. In line with this, we find that standard errors are very high when we focus on subsamples with few coercers.

[6] In principle, the model can also reflect other interactions that justify our focus on a global dataset. For instance, the model is compatible with a policymaker being a foreign power who wants to implement an infrastructural project (e.g., a pipeline) whose completion requires the consent or participation of multiple local rulers.

A.5 Figures and Tables of Appendix A

Table A.3.1: robustness of results in Table 1.7 to alternative specifications

	c's Imports from p (ln)					
	(1)	(2)	(3)	(4)	(5)	(6)
$\widehat{IE-index}^{c\to p}$ (lag)	6.38**	12.99**	3.92	4.80*	12.45***	5.08**
	[2.82]	[5.35]	[2.53]	[2.50]	[3.62]	[2.55]
$\widehat{REP}^{c\to p}$ (lag)	-3.42	-24.29	-7.99	-10.09	-1.87	-3.01
	[11.36]	[23.03]	[10.04]	[10.17]	[17.81]	[9.77]
Year FE	No	Yes	Yes	Yes	Yes	Yes
Outcome lag	Yes	No	Yes	Yes	Yes	Yes
c and p's GDP p.c.	Yes	Yes	No	No	No	Yes
c and p's Pop (ln)	Yes	Yes	Yes	No	No	Yes
c and p's GDP (ln)	Yes	Yes	Yes	Yes	No	Yes
Exports (ln)	No	No	No	No	No	Yes
Observations	813,093	813,099	813,093	818,462	823,485	784,286
Adj. R-2	0.53	0.06	0.32	0.32	0.29	0.53
K-Paap F	41.777	46.085	30.306	26.226	41.673	46.260

Note: the unit of observation is a coercer-policymaker (c, p) pair in a given year. The table reports 2SLS estimates, where the second stage uses (lagged) trade flows to measure power and the first stages uses the distribution of GDP to measure power (see Equations 1.7, 1.6). $IE-index^{c\to p}$ *(lag)* is the Indirect Exposure Index of c over p in period $t-1$, i.e., the share of total economic power over p controlled by coercers weaker than c in period $t-1$; $REP^{c\to p}$ *(lag)* is the Relative Economic Power of c over p in year $t-1$, i.e., the share of total economic power over p controlled by c in year $t-1$. All specifications control for coercer-policymaker fixed effects. Standard errors (in parenthesis) are two-way clustered (one cluster for each c and one for each p). Stars represent statistical significance of the single hypothesis test: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A.3.2: robustness of results in Table 1.7 to alternative specifications

	c's Exports from p (ln)					
	(1)	(2)	(3)	(4)	(5)	(6)
$\widehat{IE-index}^{c\to p}$ (lag)	5.83***	13.14***	2.42	1.51	9.32***	5.01***
	[2.00]	[3.56]	[1.76]	[1.66]	[2.71]	[1.79]
$\widehat{REP}^{c\to p}$ (lag)	-0.51	-19.64	-2.20	0.16	9.67	0.04
	[6.52]	[12.46]	[6.28]	[6.03]	[12.71]	[5.26]
Year FE	No	Yes	Yes	Yes	Yes	Yes
Outcome lag	Yes	No	Yes	Yes	Yes	Yes
c and p's GDP p.c.	Yes	Yes	No	No	No	Yes
c and p's Pop (ln)	Yes	Yes	Yes	No	No	Yes
c and p's GDP (ln)	Yes	Yes	Yes	Yes	No	Yes
Imports (ln)	No	No	No	No	No	Yes
Observations	813,091	813,098	813,091	818,460	823,483	784,284
Adj. R-2	0.57	0.08	0.35	0.35	0.32	0.58
K-Paap F	46.989	48.862	30.841	25.382	39.573	46.257

Note: the unit of observation is a coercer-policymaker (c, p) pair in a given year. The table reports 2SLS estimates, where the second stage uses (lagged) trade flows to measure power and the first stages uses the distribution of GDP to measure power (see Equations 1.7, 1.6). $IE-index^{c\to p}$ *(lag)* is the Indirect Exposure Index of c over p in period $t-1$, i.e., the share of total economic power over p controlled by coercers weaker than c in period $t-1$; $REP^{c\to p}$ *(lag)* is the Relative Economic Power of c over p in year $t-1$, i.e., the share of total economic power over p controlled by c in year $t-1$. All specifications control for coercer-policymaker fixed effects. Standard errors (in parenthesis) are two-way clustered (one cluster for each c and one for each p). Stars represent statistical significance of the single hypothesis test: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A.3.3: robustness of results in Table 1.3 to alternative fixed effects (with Gravity controls)

	$Offers^{p\to c}$			$Aid^{c\to p}$		
	Econ	Mil	Dip	Econ	Mil	Hum
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: α_{pc}, $\mathbf{X}_{ct}$, and $\mathbf{X}_{pt}$						
$IE-index^{c\to p}$ (lag)	2.90***	1.45***	1.37***	2.46***	1.03***	1.70***
	[0.57]	[0.37]	[0.36]	[0.47]	[0.36]	[0.39]
$REP^{c\to p}$ (lag)	-2.45***	-1.53***	-0.60	-2.33***	-1.08	-1.64*
	[0.81]	[0.49]	[0.47]	[0.82]	[0.66]	[0.88]
Panel B: α_{pc}, α_t, $\mathbf{X}_{ct}$, and $\mathbf{X}_{pt}$						
$IE-index^{c\to p}$ (lag)	2.80***	1.32***	1.24***	2.27***	0.88**	1.58***
	[0.58]	[0.36]	[0.36]	[0.49]	[0.36]	[0.40]
$REP^{c\to p}$ (lag)	-2.20***	-1.23***	-0.33	-1.87**	-0.79	-1.25
	[0.79]	[0.44]	[0.42]	[0.82]	[0.69]	[0.87]
Panel C: α_{pc}, α_{ct}, and $\mathbf{X}_{pt}$						
$IE-index^{c\to p}$ (lag)	1.88***	0.89***	0.84***	1.27***	0.41	0.88***
	[0.39]	[0.27]	[0.26]	[0.30]	[0.31]	[0.29]
$REP^{c\to p}$ (lag)	-0.96	-0.00	0.08	-0.33	0.10	0.12
	[0.58]	[0.52]	[0.41]	[0.54]	[0.57]	[0.56]
Panel D: α_{pc} and α_{pt}, and $\mathbf{X}_{ct}$						
$IE-index^{c\to p}$ (lag)	3.42***	1.73***	1.59***	2.98***	1.35***	2.15***
	[0.63]	[0.42]	[0.40]	[0.53]	[0.38]	[0.45]
$REP^{c\to p}$ (lag)	-3.13***	-1.76***	-0.80	-2.97***	-1.52**	-2.14**
	[0.77]	[0.53]	[0.51]	[0.83]	[0.69]	[0.93]
Panel E: α_{pc}, α_{pt}, α_{ct}, and $\alpha_{S(c)S(p)t}$, where $S(j)$ is j's geopolitical region						
$IE-index^{c\to p}$ (lag)	1.29***	0.70***	0.56***	0.79***	0.48*	0.50**
	[0.23]	[0.20]	[0.20]	[0.18]	[0.25]	[0.22]
$REP^{c\to p}$ (lag)	-1.16***	-0.45	-0.04	-0.68	-0.54	-0.26
	[0.41]	[0.44]	[0.34]	[0.42]	[0.51]	[0.54]

Note: the unit of observation is a coercer-policymaker (c,p) pair in a given year. $IE-index^{c\to p}$ *(lag)* is the IE-index, i.e., the share of economic power over p controlled by coercers weaker than c in $t-1$; $REP^{c\to p}$ *(lag)* is the REP, i.e., the share of economic power over p controlled by c in $t-1$. Outcomes are the log of the number of times in t in which p sent Economic, Military, or Diplomatic Offers to c, or c sent Economic, Military, or Humanitarian aid to p. Panel A and B control for policymaker-coercer pair fixed effects and Gravity controls for both c and p (log of GDP, log of Population, and the GDP per capita). Panel B also includes year fixed effects. Panel C adds time-varying fixed effects for c (dropping c's controls). Panel D adds time-varying fixed effects for p (dropping p's controls). Panel E modifies our baseline specification adding time-varying fixed effects for each pair of geopolitical regions as discussed in Section A.3.1. Standard errors (in parenthesis) are two-way clustered (c,p). Stars represent statistical significance of the single hypothesis test: $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

Table A.3.4: robustness of results in Table 1.3 to alternative fixed effects (with Gravity controls) and the Stronger Powers Index (SPI)

	$Offers^{p\to c}$			$Aid^{c\to p}$		
	Econ	Mil	Dip	Econ	Mil	Hum
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: α_{pc}, $\mathbf{X}_{ct}$, and $\mathbf{X}_{pt}$						
$IE - index^{c\to p}$ (lag)	2.90***	1.46***	1.40***	2.48***	1.05***	1.70***
	[0.57]	[0.37]	[0.36]	[0.47]	[0.36]	[0.39]
$REP^{c\to p}$ (lag)	-2.43***	-1.52***	-0.56	-2.29***	-1.05	-1.64*
	[0.78]	[0.48]	[0.46]	[0.79]	[0.66]	[0.86]
$SPI^{c\to p}$ (lag)	0.02	0.01	0.07	0.05	0.05	-0.01
	[0.11]	[0.06]	[0.06]	[0.10]	[0.06]	[0.08]
Panel B: α_{pc}, α_t, $\mathbf{X}_{ct}$, and $\mathbf{X}_{pt}$						
$IE - index^{c\to p}$ (lag)	2.76***	1.26***	1.21***	2.16***	0.82**	1.47***
	[0.59]	[0.35]	[0.35]	[0.48]	[0.36]	[0.39]
$REP^{c\to p}$ (lag)	-2.25***	-1.31***	-0.36	-1.99**	-0.86	-1.37
	[0.76]	[0.43]	[0.42]	[0.79]	[0.68]	[0.84]
$SPI^{c\to p}$ (lag)	-0.09	-0.14*	-0.05	-0.22*	-0.13	-0.22**
	[0.13]	[0.08]	[0.07]	[0.13]	[0.08]	[0.11]
Panel C: α_{pc} and α_{ct}, and $\mathbf{X}_{pt}$						
$IE - index^{c\to p}$ (lag)	1.87***	0.84***	0.82***	1.21***	0.36	0.80***
	[0.42]	[0.28]	[0.27]	[0.33]	[0.32]	[0.31]
$REP^{c\to p}$ (lag)	-0.97*	-0.08	0.06	-0.42	0.03	0.02
	[0.56]	[0.52]	[0.41]	[0.53]	[0.56]	[0.54]
$SPI^{c\to p}$ (lag)	-0.03	-0.14*	-0.04	-0.15	-0.12	-0.19*
	[0.13]	[0.08]	[0.08]	[0.13]	[0.08]	[0.10]

Note: the unit of observation is a coercer-policymaker (c, p) pair in a given year. $IE - index^{c\to p}$ *(lag)* is the Indirect Exposure Index of c over p in period $t-1$, i.e., the share of total economic power over p controlled by coercers weaker than c in period $t-1$; $REP^{c\to p}$ *(lag)* is the Relative Economic Power of c over p in year $t-1$, i.e., the share of total economic power over p controlled by c in year $t-1$; $SPI^{c\to p}$ *(lag)* is the Stronger Powers Index of c over p in period $t-1$, i.e., the share of total economic power over p controlled by coercers stronger than c in period $t-1$. Outcomes are the log of the number of times in year t in which country p sent Economic, Military, or Diplomatic Offers to c, or in which country c sent Economic, Military, or Humanitarian aid to p. Panel A and B control for policymaker-coercer pair fixed effects and standard Gravity controls for both coercer and policymaker (log of GDP, log of Population, and the GDP per capita). Panel B also includes year fixed effects. Panel C adds time-varying fixed effects for each coercer (thus dropping the coercer's Gravity controls). Standard errors (in parenthesis) are two-way clustered (c, p). Stars represent statistical significance of the single hypothesis test: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A.3.5: robustness of results in Table 1.3 to alternative specifications of the estimating equation

	$Offers^{p\to c}$			$Aid^{c\to p}$		
	Econ	Mil	Dip	Econ	Mil	Hum
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Trade (ln) **not** included, lag REP, lag IE-index						
$IE-index^{c\to p}$ (lag)	2.20***	1.17***	1.07***	1.69***	0.83***	1.25***
	[0.31]	[0.26]	[0.25]	[0.25]	[0.26]	[0.26]
$REP^{c\to p}$ (lag)	-1.43***	-0.36	-0.21	-1.09*	-0.50	-0.50
	[0.54]	[0.50]	[0.44]	[0.55]	[0.54]	[0.62]
Panel B: Trade (ln) **not** included, **current** REP, lag IE-index						
$IE-index^{c\to p}$ (lag)	1.80***	1.04***	0.89***	1.34***	0.69***	1.03***
	[0.27]	[0.25]	[0.21]	[0.22]	[0.23]	[0.21]
$REP^{c\to p}$	-0.09	-0.22	0.24	0.13	-0.09	0.10
	[0.42]	[0.35]	[0.25]	[0.54]	[0.46]	[0.54]
Panel C: Trade (ln) included, **current** REP, lag IE-index						
$IE-index^{c\to p}$ (lag)	1.76***	1.03***	0.86***	1.33***	0.67***	1.04***
	[0.27]	[0.24]	[0.21]	[0.22]	[0.24]	[0.22]
$REP^{c\to p}$	-0.18	-0.23	0.18	0.12	-0.14	0.11
	[0.41]	[0.35]	[0.26]	[0.54]	[0.44]	[0.55]
Panel D: Trade (ln) included, **current** REP, **current** IE-index						
$IE-index^{c\to p}$	1.92***	0.99***	1.05***	1.49***	0.58**	1.03***
	[0.31]	[0.22]	[0.25]	[0.25]	[0.28]	[0.25]
$REP^{c\to p}$	-0.95**	-0.53	-0.36	-0.47	-0.20	-0.27
	[0.47]	[0.42]	[0.35]	[0.57]	[0.52]	[0.59]
Panel E: Trade (ln) included, lag REP, lag IE-index, **add** outcome lag						
$IE-index^{c\to p}$ (lag)	1.85***	1.09***	0.71***	1.30***	0.72***	1.08***
	[0.27]	[0.23]	[0.23]	[0.20]	[0.22]	[0.25]
$REP^{c\to p}$ (lag)	-1.30**	-0.21	0.59	-1.03**	-0.52	-0.54
	[0.51]	[0.50]	[0.39]	[0.44]	[0.49]	[0.59]

Note: the unit of observation is a coercer-policymaker (c,p) pair in a given year. $IE-index^{c\to p}$ *(lag)* is the Indirect Exposure Index of c over p in period $t-1$, i.e., the share of total economic power over p controlled by coercers weaker than c in period $t-1$; $REP^{c\to p}$ *(lag)* is the Relative Economic Power of c over p in year $t-1$, i.e., the share of total economic power over p controlled by c in year $t-1$. When *(lag)* is omitted, as in Panel B, C and D, we are considering the relevant year t variable. Outcomes are the log of the number of times in year t in which country p sent Economic, Military, or Diplomatic Offers to c, or in which country c sent Economic, Military, or Humanitarian aid to p. All specifications control for policymaker-coercer pair fixed effects, and time-varying fixed effects for each coercer and policymaker. Panels C, D and E control for the log of trade between p and c. Panel E also controls for the relevant outcome in year $t-1$. Standard errors (in parenthesis) are two-way clustered (c,p). Stars represent statistical significance of the single hypothesis test: $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

Table A.3.6: robustness of results in Table 1.3 to alternative transformations of the outcome variables

	$Offers^{p\to c}$			$Aid^{c\to p}$		
	Econ	Mil	Dip	Econ	Mil	Hum
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: $ln(\tilde{Y})$ (std. dev. still normalized to 1)						
$IE-index^{c\to p}$ (lag)	1.29***	1.08*	2.41**	1.08***	0.51	1.10**
	[0.29]	[0.63]	[0.99]	[0.25]	[0.38]	[0.53]
$REP^{c\to p}$ (lag)	-1.01	-0.93	2.08	-1.17	1.35	-0.85
	[0.84]	[2.18]	[3.73]	[0.76]	[1.10]	[1.18]
Trade flow (ln)	-0.00	0.05	0.08	-0.02***	-0.00	-0.00
	[0.01]	[0.06]	[0.07]	[0.01]	[0.01]	[0.02]
Panel B: $ln(1+\tilde{Y})$ (std. dev. still normalized to 1)						
$IE-index^{c\to p}$ (lag)	2.13***	1.19***	0.98***	1.68***	0.81***	1.24***
	[0.31]	[0.25]	[0.24]	[0.25]	[0.26]	[0.27]
$REP^{c\to p}$ (lag)	-1.41***	-0.43	-0.12	-1.03*	-0.51	-0.47
	[0.53]	[0.49]	[0.42]	[0.56]	[0.54]	[0.62]
Trade flow (ln)	0.01	0.00	0.00	0.00	0.00	-0.00
	[0.00]	[0.01]	[0.00]	[0.00]	[0.00]	[0.00]

Note: the unit of observation is a coercer-policymaker (c, p) pair in a given year. $IE-index^{c\to p}$ *(lag)* is the Indirect Exposure Index of c over p in period $t-1$, i.e., the share of total economic power over p controlled by coercers weaker than c in period $t-1$; $REP^{c\to p}$ *(lag)* is the Relative Economic Power of c over p in year $t-1$, i.e., the share of total economic power over p controlled by c in year $t-1$. Outcomes are the log of the number of times in year t in which country p sent Economic, Military, or Diplomatic Offers to c, or in which country c sent Economic, Military, or Humanitarian aid to p. Rather than adopting the transformations discussed in Appendix A.2, Panel A only takes the log of the variable and then normalizes the standard deviation of the log to 1; Panel B adds 1 to every variable before taking the log, and then normalizes the standard deviation of the log to 1. All specifications control for the log of trade between p and c, policymaker-coercer pair fixed effects, and time-varying fixed effects for each coercer and policymaker. Standard errors (in parenthesis) are two-way clustered (c, p). Stars represent statistical significance of the single hypothesis test: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A.3.7: robustness of results in Table 1.3 to alternative standard errors

	$Offers^{p\to c}$			$Aid^{c\to p}$		
	Econ	Mil	Dip	Econ	Mil	Hum
	(1)	(2)	(3)	(4)	(5)	(6)
$IE-index^{c\to p}$ (lag)	2.16***	1.18***	1.05***	1.69***	0.81***	1.25**
3w-Cl c,p,t	[0.482]	[0.303]	[0.300]	[0.413]	[0.285]	[0.386]
D-K 4 lags	[0.664]	[0.327]	[0.316]	[0.575]	[0.200]	[0.507]
D-K 7 lags	[0.715]	[0.358]	[0.322]	[0.607]	[0.201]	[0.543]
$REP^{c\to p}$ (lag)	-1.44*	-0.36	-0.17	-1.09	-0.51	-0.49
3w-Cl c,p,t	[0.636]	[0.503]	[0.442]	[0.654]	[0.568]	[0.647]
D-K 4 lags	[0.707]	[0.324]	[0.281]	[0.636]	[0.475]	[0.426]
D-K 7 lags	[0.716]	[0.326]	[0.207]	[0.641]	[0.398]	[0.382]
Trade flow (ln)	0.01	0.00	0.00	-0.00	0.00	-0.00
3w-Cl c,p,t	[0.006]	[0.005]	[0.004]	[0.005]	[0.004]	[0.005]
D-K 4 lags	[0.005]	[0.003]	[0.004]	[0.005]	[0.003]	[0.004]
D-K 7 lags	[0.006]	[0.003]	[0.004]	[0.005]	[0.003]	[0.005]
Observations	281,190	281,298	253,732	281,305	281,305	281,305
Adj. R-2	0.47	0.27	0.16	0.58	0.43	0.36
Within R-2	0.005	0.001	0.001	0.004	0.001	0.002

Note: the unit of observation is a coercer-policymaker (c, p) pair in a given year. $IE-index^{c\to p}$ *(lag)* is the Indirect Exposure Index of c over p in period $t-1$, i.e., the share of total economic power over p controlled by coercers weaker than c in period $t-1$; $REP^{c\to p}$ *(lag)* is the Relative Economic Power of c over p in year $t-1$, i.e., the share of total economic power over p controlled by c in year $t-1$. Outcomes are the log of the number of times in year t in which country p sent Economic, Military, or Diplomatic Offers to c, or in which country c sent Economic, Military, or Humanitarian aid to p. All specifications control for the log of trade between p and c, policymaker-coercer pair fixed effects, and time-varying fixed effects for each coercer and policymaker. *3w-Cl* c, p, t indicates three-way clustered standard errors (c, p, t). *D-K 4 lags* and *D-K 7 lags* indicate Driscoll-Kraay standard errors allowing for up to 4 lags or 7 lags of auto-correlation in the correlation of errors across country pairs, respectively. Stars represent statistical significance of the single hypothesis test under the highest of the three standard errors reported in parenthesis: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A.3.8: robustness of results in Table 1.7 to alternative standard errors

	c's Imports (ln) Econ (1)	c's Exports (ln) Mil (2)
$\widehat{IE-index}^{c\to p}$ (lag)	5.06***	4.43***
D-K 4 lags	[1.118]	[0.701]
D-K 7 lags	[1.200]	[0.731]
$\widehat{REP}^{c\to p}$ (lag)	-11.45	-7.93
D-K 4 lags	[5.963]	[4.184]
D-K 7 lags	[7.180]	[5.041]
Observations	813,093	813,091
Adj. R-2	0.30	0.33
K-Paap F (*D-K 4 lags*)	15.548	16.043
K-Paap F (*D-K 7 lags*)	10.457	10.778

Note: the unit of observation is a coercer-policymaker (c, p) pair in a given year. The table reports 2SLS estimates, where the second stage uses (lagged) trade flows to measure power and the first stages uses the distribution of GDP to measure power (see Equations 1.7, 1.6). $IE-index^{c\to p}$ *(lag)* is the Indirect Exposure Index of c over p in period $t-1$, i.e., the share of total economic power over p controlled by coercers weaker than c in period $t-1$; $REP^{c\to p}$ *(lag)* is the Relative Economic Power of c over p in year $t-1$, i.e., the share of total economic power over p controlled by c in year $t-1$. All specifications control for coercer-policymaker fixed effects, year fixed effects, the lag of the outcome, the GDP (log), the population (log), and the GDP per capita of both c and p. *D-K 4 lags* and *D-K 7 lags* indicate Driscoll-Kraay standard errors allowing for up to 4 lags or 7 lags of auto-correlation in the correlation structure of the errors across country-pairs, respectively. Stars represent statistical significance of the single hypothesis test under the highest of the three standard errors reported in parenthesis: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A.3.9: robustness of results in Table 1.3 to controls based on the different types of outcomes

	$Offers^{p\to c}$			$Aid^{c\to p}$		
	Econ	Mil	Dip	Econ	Mil	Hum
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: control for all types of Offers sent						
$IE-index^{c\to p}$ (lag)	0.89***	0.37***	0.73***	0.98***	0.34*	0.46***
	[0.16]	[0.11]	[0.20]	[0.18]	[0.18]	[0.16]
$REP^{c\to p}$ (lag)	-0.62***	-0.46**	-0.22	-0.73*	-0.44	-0.14
	[0.22]	[0.23]	[0.39]	[0.43]	[0.39]	[0.35]
Panel B: control for all types of Offers sent and of Aid sent						
$IE-index^{c\to p}$ (lag)	0.76***	0.30***	0.59***			
	[0.15]	[0.11]	[0.19]			
$REP^{c\to p}$ (lag)	-0.53**	-0.40*	-0.10			
	[0.21]	[0.23]	[0.39]			
Panel C: control for all types of Offers sent and of Offers received						
$IE-index^{c\to p}$ (lag)				0.80***	0.27**	0.28*
				[0.15]	[0.12]	[0.15]
$REP^{c\to p}$ (lag)				-0.61	-0.43	-0.13
				[0.41]	[0.29]	[0.28]

Note: the unit of observation is a coercer-policymaker (c, p) pair in a given year. $IE-index^{c\to p}$ *(lag)* is the Indirect Exposure Index of c over p in period $t-1$, i.e., the share of total economic power over p controlled by coercers weaker than c in period $t-1$; $REP^{c\to p}$ *(lag)* is the Relative Economic Power of c over p in year $t-1$, i.e., the share of total economic power over p controlled by c in year $t-1$. Outcomes are the log of the number of times in year t in which country p sent Economic, Military, or Diplomatic Offers to c, or in which country c sent Economic, Military, or Humanitarian aid to p. All specifications control for the log of trade between p and c, policymaker-coercer pair fixed effects, and time-varying fixed effects for each coercer and policymaker. Panel A controls for all types of offers sent by c to p in t. Panel B controls for all types of offers sent by c to p in year t and all types of aid sent by country c to p in t. Panel C controls for all types of offers sent by c to p in t and all types of offers sent by p to c in t. Standard errors (in parenthesis) are two-way clustered (c, p). Stars represent statistical significance of the single hypothesis test: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A.4.1: correlations and descriptive statistics of our measures of power

	$E_t^{c\to p}$	$IE-index_t^{c\to p}(\mathbf{E}_{pt})$	$I_t^{c\to p}$	$IE-index_t^{c\to p}(\mathbf{I}_{pt})$
$E_t^{c\to p}$	1			
$IE-index_t^{c\to p}(\mathbf{E}_{pt})$	0.71	1		
$I_t^{c\to p}$	0.59	0.68	1	
$IE-index_t^{c\to p}(\mathbf{I}_{pt})$	0.44	0.73	0.86	1
Mean	0.004	0.027	0.007	0.057
Std. Dev.	0.02	0.06	0.03	0.20
Std. Dev. (top 5)	0.05	0.12	0.08	0.46
Std. Dev. (top 10)	0.04	0.11	0.09	0.43
Std. Dev. (top 20)	0.03	0.10	0.06	0.33
Std. Dev. (top 30)	0.03	0.09	0.05	0.27
Std. Dev. (Asym.)	0.03	0.08	0.04	0.21

Note: the unit of observation is a coercer-policymaker (c, p) pair in a given year. $E_t^{c\to p}$ is the economic power of c on p that we used in our baseline specification and discussed in Section 1.4.1. $I_t^{c\to p}$ is the Formal Bilateral Influence Capacity of c on p as discussed in Appendix A.4.1. Std. Dev. (top N) is the standard deviation calculated in the subsample where countries can only be either coercers or policymakers, and countries are considered coercers if and only if there is some year within our sample in which that country is in the top N of GDP. Std. Dev. (Asym.) is the standard deviation in the subsample where we drop country pairs where the policymaker has more power over the coercer than viceversa.

Table A.4.2: correlations and descriptive statistics of our measures of power after netting out baseline fixed effects

	$E_t^{c\to p}$	$IE-index_t^{c\to p}(\mathbf{E}_{pt})$	$I_t^{c\to p}$	$IE-index_t^{c\to p}(\mathbf{I}_{pt})$
$E_t^{c\to p}$	1			
$IE-index_t^{c\to p}(\mathbf{E}_{pt})$	0.57	1		
$I_t^{c\to p}$	0.21	0.28	1	
$IE-index_t^{c\to p}(\mathbf{I}_{pt})$	0.14	0.36	0.74	1
Mean	0	0	0	0
Std. Dev.	0.01	0.03	0.01	0.07
Std. Dev. (top 5)	0.02	0.05	0.04	0.21
Std. Dev. (top 10)	0.02	0.05	0.03	0.17
Std. Dev. (top 20)	0.02	0.04	0.02	0.14
Std. Dev. (top 30)	0.02	0.04	0.02	0.12
Std. Dev. (Asym.)	0.01	0.03	0.01	0.08

Note: the unit of observation is a coercer-policymaker (c, p) pair in a given year. $E_t^{c\to p}$ is the economic power of c on p that we used in our baseline specification and discussed in Section 1.4.1. $I_t^{c\to p}$ is the Formal Bilateral Influence Capacity of c on p as discussed in Appendix A.4.1. Std. Dev. (top N) is the standard deviation calculated in the subsample where countries can only be either coercers or policymakers, and countries are considered coercers if and only if there is some year within our sample in which it is in the top N of GDP. Std. Dev. (Asym.) is the standard deviation in the subsample where we drop country pairs where the policymaker has more power over the coercer than viceversa. The reported statistics are for the variables after netting out the fixed effects of our main baseline specification: α_{pc} α_{ct} and α_{pt}.

Table A.4.3: replication of results in Table 1.3 with alternative measure of power

	$Offers^{p\to c}$			$Aid^{c\to p}$		
	Econ	Mil	Dip	Econ	Mil	Hum
	(1)	(2)	(3)	(4)	(5)	(6)
$IE-index^{c\to p}(\mathbf{I}_{pt})$ (lag)	0.91***	0.52***	0.55***	0.88***	0.24***	0.69***
	[0.11]	[0.11]	[0.09]	[0.11]	[0.08]	[0.11]
$REP^{c\to p}(\mathbf{I}_{pt})$ (lag)	-1.68***	-1.31**	-0.80	-1.34*	0.30	-1.15*
	[0.46]	[0.53]	[0.55]	[0.71]	[0.56]	[0.60]
Trade flow (ln)	0.01	0.00	0.00	-0.00	0.00	-0.00
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Observations	299,354	299,463	270,822	299,472	299,472	299,472
Adj. R-2	0.47	0.27	0.16	0.59	0.44	0.36
Within R-2	0.014	0.003	0.003	0.019	0.002	0.007

Note: the unit of observation is a coercer-policymaker (c, p) pair in a given year. $\mathbf{I}_{pt}$ is the distribution of Formal Bilateral Influence Capacities (FBIC) over p, as discussed in Appendix A.4.1. The variable $IE-index^{c\to p}(\mathbf{I}_{pt})$ *(lag)* is the share of FBIC over p of all coercers with less FBIC over p than c in period $t-1$. $REP^{c\to p}(\mathbf{I}_{pt})$ *(lag)* is c's FBIC over p in period $t-1$. Outcomes are the log of the number of times in year t in which p sent Economic, Military, or Diplomatic Offers to c or in which c sent Economic, Military, or Humanitarian aid to p. All specifications control for the log of trade between p and c, policymaker-coercer pair fixed effects, and time-varying fixed effects for each coercer and policymaker. Standard errors (in parenthesis) are two-way clustered (c, p). Stars represent statistical significance of the single hypothesis test: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A.4.4: replication of results in Table 1.7 with alternative measure of power

	c's Imports (1)	c's Exports (2)
$\widehat{IE-index}^{c\to p}(\mathbf{I}_{pt})$ (lag)	1.15*	0.71*
	[0.65]	[0.37]
$\widehat{REP}^{c\to p}(\mathbf{I}_{pt})$ (lag)	-1.83	1.12
	[6.58]	[3.65]
Observations	864,140	864,148
Adj. R-2	0.30	0.33
K-Paap F	14.93	16.74

Note: the unit of observation is a coercer-policymaker (c, p) pair in a given year. $\mathbf{I}_{pt}$ is the distribution of Formal Bilateral Influence Capacities (FBIC) over p, as discussed in Appendix A.4.1. The table reports 2SLS estimates, where the second stage uses (lagged) FBIC to measure power and the first stages uses the distribution of GDP to measure power (see Equations 1.7, 1.6). The Indirect Exposure Index (IE-index) of country c (coercer) over country p (policymaker) is the share of total FBIC over p of all coercers with less power over p than c. The Relative Economic Power (REP) of c over p is simply the FBIC of c over p. All specifications control for coercer-policymaker fixed effects, year fixed effects, the lag of the outcome, the GDP (log), the population (log), and the GDP per capita of both c and p. Standard errors (in parenthesis) are two-way clustered (c, p). Stars represent statistical significance of the single hypothesis test: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A.4.5: description of alternative outcomes

Variable	Description of the interaction [CAMEO category]
Economic Agreement	Initiate, resume, improve, or expand economic exchange or cooperation. (Trade relations and other economic exchanges that are reciprocal in nature – even if the particular event in question cannot be coded as reciprocal – should be coded here) [061].
Military Agreement	Initiate, resume, improve, or expand military exchange or cooperation (Military exchanges such as joint military games and maneuvers should be coded here) [062]. Provide, share, or exchange intelligence or information [064].
Diplomatic Agreement	Initiate, resume, improve, or expand diplomatic, non-material cooperation or exchange not otherwise specified [050].
Judiciary Agreement	Initiate, resume, improve, or expand judicial cooperation [063].
Political Agreement	Ratify, sign, finalize an agreement, treaty. (Events should be coded under this category only when agreements are reportedly finalized or signed. This event code is typically reciprocal. Even when the agreement in question implies a formal commitment to boost material cooperation, provide aid, or yield in some way, the event of signing the agreement or treaty is still coded here since signing of an agreement or treaty represents diplomatic cooperation but does not guarantee implementation – whatever its terms) [057].

Table A.4.6: replication of results in Table 1.3 with alternative outcomes

	$Agreements^{c \longleftrightarrow p}$				
	Eco	Mil	Dip	Jud	Pol
	(1)	(2)	(3)	(4)	(5)
$IE - index^{c \to p}$ (lag)	1.22***	0.67**	0.83***	0.43**	0.83***
	[0.23]	[0.26]	[0.20]	[0.20]	[0.20]
$REP^{c \to p}$ (lag)	-0.18	0.55	-0.44	0.88	-0.38
	[0.57]	[0.65]	[0.52]	[0.67]	[0.53]
Trade flow (ln)	0.01	-0.01	0.00	-0.01*	0.00
	[0.01]	[0.00]	[0.01]	[0.00]	[0.01]
Observations	123,895	123,895	123,895	123,895	123,895
Adj. R-2	0.66	0.41	0.64	0.17	0.64
Within R-2	0.005	0.001	0.002	0.001	0.002

Note: the unit of observation is a coercer-policymaker (c, p) pair in a given year. $IE - index^{c \to p}$ *(lag)* is the Indirect Exposure Index of c over p in period $t-1$, i.e., the share of total economic power over p controlled by coercers weaker than c in period $t-1$; $REP^{c \to p}$ *(lag)* is the Relative Economic Power of c over p in year $t-1$, i.e., the share of total economic power over p controlled by c in year $t-1$. Outcomes are the log of the number of times in year t in which country p and the US concluded an Economic, Military, Diplomatic, Judiciary, or Political agreement. All specifications control for the log of trade between p and c, policymaker-coercer pair fixed effects, and time-varying fixed effects for each coercer and policymaker. Standard errors (in parenthesis) are two-way clustered (one cluster for each c and one for each p). Stars represent statistical significance of the single hypothesis test: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors (in parenthesis) are two-way clustered (c, p). Stars represent statistical significance of the single hypothesis test: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A.4.7: replication of Table 1.5 with alternative outcomes

	$Agreements^{c \longleftrightarrow p}$				
	Eco	Mil	Dip	Jud	Pol
	(1)	(2)	(3)	(4)	(5)
Panel A: fall of USSR					
Stronger vs Weaker	-0.34***	-0.19***	-0.32***	-0.18***	-0.33***
	[0.03]	[0.04]	[0.03]	[0.04]	[0.04]
Observations	121,120	121,120	121,120	121,120	121,120
Adj. R-2	0.65	0.40	0.62	0.15	0.62
Within R-2	0.004	0.001	0.002	0.001	0.002
Panel B: rise of China					
Stronger vs Weaker	0.43***	0.18***	0.31***	0.25***	0.32***
	[0.04]	[0.04]	[0.04]	[0.07]	[0.04]
Observations	120,126	120,126	120,126	120,126	120,126
Adj. R-2	0.66	0.41	0.63	0.17	0.63
Within R-2	0.012	0.002	0.005	0.002	0.006

Note: the unit of observation is a coercer-policymaker (c, p) pair in a given year. The variable Stronger vs Weaker estimates β of Equation 1.4. This coefficient captures the differential effect on each coercer c of the fall of the Soviet Union (Panel A) or of the rise of China (Panel B) comparing the outcomes related to policymakers over which c was stronger than the USSR (or China) with those policymaker over which it was weaker than the USSR (or China). Outcomes are the log of the number of times in year t in which country p and c concluded an Economic, Military, Diplomatic, Judiciary, or Political agreement. All specifications control for the log of trade between p and c, the Relative Economic Power (REP) of c over p, i.e. c's share of total economic power over p, policymaker-coercer pair fixed effects, and time-varying fixed effects for each coercer and policymaker. Standard errors (in parenthesis) are two-way clustered (c, p). Stars represent statistical significance of the single hypothesis test: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A.4.8: replication of Table 1.6 with alternative outcomes

	$Agreements^{US \longleftrightarrow h}$				
	Eco	Mil	Dip	Jud	Pol
	(1)	(2)	(3)	(4)	(5)
Panel A: fall of USSR					
Stronger vs Weaker	-0.37***	-0.44***	-0.27***	-0.33***	-0.27***
(USA only)	[0.08]	[0.13]	[0.08]	[0.12]	[0.08]
Observations	5,270	5,270	5,270	5,270	5,270
Adj. R-2	0.82	0.51	0.73	0.26	0.73
Within R-2	0.013	0.009	0.018	0.009	0.018
Panel B: rise of China					
Stronger vs Weaker	0.64***	0.48***	0.51***	1.11***	0.53***
(USA only)	[0.06]	[0.15]	[0.07]	[0.17]	[0.07]
Observations	5,279	5,279	5,279	5,279	5,279
Adj. R-2	0.82	0.50	0.73	0.26	0.73
Within R-2	0.025	0.005	0.019	0.013	0.019

Note: the unit of observation is a US-policymaker (US, p) pair in a given year. The variable Stronger vs Weaker estimates β of Equation 1.4 restricting $c = US$. This coefficient captures the differential effect on the US of the fall of the Soviet Union (Panel A) or of the rise of China (Panel B) comparing the outcomes related to those policymakers over which the US was stronger than the USSR (or China) with those where it was weaker than the USSR (or China). Outcomes are the log of the number of times in year t in which country p and the US concluded an Economic, Military, Diplomatic, Judiciary, or Political agreement. All specifications control for the log of trade between p and the US, the US's Relative Economic Power (REP) over p, i.e. the share of total economic power over p controlled by the US, policymaker-US fixed effects, and year fixed effects. Standard errors (in parenthesis) are robust. Stars represent statistical significance of the single hypothesis test: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A.4.9: list of coercers and sample size of subsamples of interest

Subsample	List of coercers	Sample Size
Top 5	China, Germany (or West Germany), France, Great Britain, Italy, Japan, Russia (or Soviet Union), United States	37512
Top 10	Brazil, Canada, China, Germany (or West Germany), Spain, France, Great Britain, India, Italy, Japan, Republic of Korea, Mexico, Russia (or Soviet Union), United States	52978
Top 20	Argentina, Australia, Austria, Belgium, Brazil, Canada, Switzerland, China, Germany (or West Germany), Spain, France, Great Britain, Indonesia, India, Iran, Iraq, Italy, Japan, Republic of Korea, Mexico, Netherlands, Poland, Russia (or Soviet Union), Saudi Arabia, Sweden, Turkey, Taiwan, United States	82016
Top 30	Argentina, Australia, Austria, Belgium, Brazil, Canada, Switzerland, China, Germany (or West Germany), Denmark Spain, Finland, France, Great Britain, Greece, Indonesia, India, Iran, Iraq, Israel, Italy, Japan, Republic of Korea, Mexico, Nigeria, Netherlands, Norway, Poland, Portugal, Romania, Russia (or Soviet Union), Saudi Arabia, Sweden, Thailand, Turkey, Taiwan, United States, Venezuela, Vietnam, South Africa	94938

Note: Top N is the subsample where each country c can only be either a coercer or a policymaker, and country c is a coercer if and only if there is some year within our sample when c is in the top N of that year's GDP distribution. To improve the readability of the table, we underlined countries when they appear for the first time.

Table A.4.10: replication of results in Table 1.3 focusing on subsamples of interest

	$Offers^{p\to c}$			$Aid^{c\to p}$		
	Econ	Mil	Dip	Econ	Mil	Hum
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: c in top 5 of GDP in some t and p never						
$IE-index^{c\to p}$ (lag)	1.42***	1.29**	0.93**	0.92***	1.09**	1.18***
	[0.24]	[0.39]	[0.28]	[0.25]	[0.41]	[0.26]
$REP^{c\to p}$ (lag)	-1.43	-0.45	0.03	-1.02	-1.16	-1.41
	[0.83]	[1.00]	[0.70]	[0.96]	[0.83]	[1.05]
Panel B: c in top 10 of GDP in some t and p never						
$IE-index^{c\to p}$ (lag)	1.28***	1.05***	0.81**	0.92***	0.91**	0.93***
	[0.26]	[0.28]	[0.31]	[0.25]	[0.39]	[0.27]
$REP^{c\to p}$ (lag)	-0.96	-0.48	0.40	-0.72	-1.04	-1.05
	[0.67]	[0.61]	[0.60]	[0.81]	[0.84]	[0.85]
Panel C: c in top 20 of GDP in some t and p never						
$IE-index^{c\to p}$ (lag)	1.22***	0.78***	0.51*	0.81***	0.65**	0.80***
	[0.27]	[0.19]	[0.27]	[0.24]	[0.28]	[0.26]
$REP^{c\to p}$ (lag)	-1.38***	-1.01***	0.38	-0.96	-1.17*	-1.10
	[0.48]	[0.32]	[0.56]	[0.65]	[0.66]	[0.76]
Panel D: c in top 30 of GDP in some t and p never						
$IE-index^{c\to p}$ (lag)	1.02***	0.69***	0.57**	0.80***	0.53**	0.60**
	[0.23]	[0.22]	[0.24]	[0.26]	[0.25]	[0.25]
$REP^{c\to p}$ (lag)	-0.76	-0.72*	0.25	-0.68	-0.91	-0.59
	[0.58]	[0.39]	[0.49]	[0.68]	[0.72]	[0.81]
Panel E: c has greater lag IE-index over p than p over c						
$IE-index^{c\to p}$ (lag)	1.31***	0.89***	0.67***	0.94***	0.69**	0.78***
	[0.23]	[0.21]	[0.23]	[0.24]	[0.27]	[0.23]
$REP^{c\to p}$ (lag)	-0.71	-0.57	0.44	-0.72	-0.54	-0.51
	[0.62]	[0.73]	[0.52]	[0.78]	[0.65]	[0.83]

Note: the unit of observation is a coercer-policymaker (c,p) pair in a given year. $IE-index^{c\to p}$ *(lag)* is the Indirect Exposure Index of c over p in period $t-1$, i.e., the share of total economic power over p controlled by coercers weaker than c in period $t-1$; $REP^{c\to p}$ *(lag)* is the Relative Economic Power of c over p in year $t-1$, i.e., the share of total economic power over p controlled by c in year $t-1$. Outcomes are log of times in year t in which p sent Econ, Military, or Diplomatic Offers to c, or in which c sent Econ, Military, or Humanitarian aid to p. All specifications control for the log of trade between c and p, policymaker-coercer pair fixed effects, and time-varying fixed effects for each coercer and policymaker. The Panels exhibit coefficients computed from different subsamples. Panels A, B, C, and D focuses on subsamples where countries can only be either coercers or policymakers, and countries are considered coercers if and only if there is some year within our sample in which that country is in the top 5, 10, 20, or 30 of GDP (respectively). Table A.4.9 provides the list of coercers in each of these subsamples. Panel E focuses on the subsample where we dropped country pairs where the policymaker has more power over the coercer than viceversa. Standard errors (in parenthesis) are two-way clustered (c,p). Stars represent statistical significance of the single hypothesis test: $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

Appendix B

Appendix to Chapter 2

B.1 Proofs

PROOF OF LEMMA 1: Fix tax demands $\mathbf{T}$. Consider any SPE in which, when subjects are indifferent they choose not to resist. Proceeding by backward induction, we know that the last agent to choose, call it n, observing the set of agents already resisting, $resist_n \subseteq I$, would join the resistance if and only if

$$B_n \times f\left(W_{resist_n} + W_n, W_{Ruler}, W_{tot}\right) - c < B_n - T_n$$

Knowing this, if group $n-1$ observes $resist_{n-1} \subseteq I$, the set of subjects resisting before its choice and n's choice is made, it will join the resistance if and only if

$$B_{n-1} \times f\left(W_{resist_{n-1}} + W_{n-1} + W_n \mathbb{J}, W_{Ruler}, W_{tot}\right) - c < B_{n-1} - T_{n-1},$$

$$\mathbb{J} := \mathbb{I}\left(B_n \times f\left(W_{resist_{n-1}} + W_{n-1} + W_n, W_{Ruler}, W_{tot}\right) - c < B_n - T_n\right)$$

Iterating the argument, we find similar conditions for every subject. Clearly, given the tie-breaking assumption and the common knowledge of the vectors $(\mathbf{B}, \mathbf{W})$ and the tax vector $\mathbf{T}$ demanded by the ruler, group i's choice about whether or not to join the resistance is uniquely determined by the choice of $j < i$, the player who moves before i. Further, note that if i resists when $j < i$ does not, then i will resist also when j resists. So, under this tie-breaking ruler, for any fixed tax vector $\mathbf{T}$ and order in which subjects are called to resist, there exists a unique SPE of the continuation of the taxation game. Finally, note that the tie breaking rule is without loss of generality. Indeed, if i joins the resistance when indifferent, the ruler would be better off by decreasing T_i by an infinitesimal amount so to discourage it from resisting. As a consequence, there is no equilibrium where an indifferent subject chooses to resist.[1] ■

PROOF OF LEMMA 2: Fix any tax demands $\mathbf{T}$. Consider two successive groups i and $i+1$ in a particular order $(1....n)$. Because players move sequentially and information is complete, it is easy to show that if $resist\,(\mathbf{T})$ is the set of subjects that choose to resist in an SPE when the order is $(1, i, i+1...n)$ then $resist\,(\mathbf{T})$ would also be the set of subjects that choose to resist in an SPE when the order is $(1..i-1, i+1, i, i+2..n)$. Lemma 1, establishing uniqueness of the SPE for any fixed order completes the proof: tax vector T generates resistance $resist\,(\mathbf{T})$ independently on the order in which subjects are called to choose. ■

PROOF OF LEMMA 3: In Lemma 1 we showed that the subgame following the tax demand $\mathbf{T}$ has a unique SPE. Given the complete information, the ruler can anticipate the set of subjects that choose to resist in the possible SPE $resist\,(\mathbf{T})$. Suppose that in an SPE we have $i \in resist\,(\mathbf{T})$. Since discouraging i from joining the resistance cannot trigger a resistance of greater size (reducing the size of the resistance reduces its likelihood of success and thus

[1] If $T_i = 0$ then i strictly prefers not to resist.

the value of joining), then it must be the case that

$$\underbrace{B_i\left(1-f\left(W_{resist(\mathbf{T})},W_{Ruler},W_{tot}\right)\right)}_{\substack{\text{expected resources extracted}\\ \text{from i when resistance set is resist(T)}}} > \underbrace{c+B_i\times\left(1-f\left(W_{resist(\mathbf{T})},W_{Ruler},W_{tot}\right)\right)}_{\substack{\text{a tax able to discourage}\\ \text{i from joining the resistence}}}$$

i.e. $c<0$, which cannot hold as we assumed $c>0$. ■

PROOF OF THEOREM 2: The proof evolves through different steps:

1. By Lemma 2 and 1, for any given $\mathbf{T}$ the SPE outcome is unique and order independent.
2. By Lemma 3 resistance cannot be part of a SPE. Call $\mathbf{T}^\star$ the SPE taxation demanded by the ruler and $resist_{\mathbf{T}^\star}$ the only possible set of resisting subjects that can be associated with $\mathbf{T}^\star$ in an SPE. By definition the ruler would choose $\mathbf{T}=\mathbf{T}^\star$ to maximize its utility, i.e. to maximize the expected sum of tax extraction.
3. In order to avoid resistance $\mathbf{T}^\star$ should be such that:
   - No subject needs to find optimal to resist alone: $\nexists\ i$ such that
   $$T_i^\star > c+B_i\left(1-f\left(W_i,W_{Ruler},W_{tot}\right)\right)$$
   - No two subjects find optimal to resist together: $\nexists\ s,j$ such that for all $i\in\{s,j\}$
   $$T_i^\star > c+B_i\times\left(1-f\left(W_s+W_j,W_{Ruler},W_{tot}\right)\right)$$
   - No three subjects find optimal to resist together: $\nexists\ s,j,l$ such that for all $i\in\{s,j,l\}$
   $$T_i^\star > c+B_i\left(1-f\left(W_s+W_j+W_l,W_{Ruler},W_{tot}\right)\right)$$
   - etc.

4. As by Lemma 3 the above constraints must be satisfied, it is easy to show that for any SPE which does not involve resistance, the tax vector $\mathbf{T}^{\star}$, should be such that given a re-ordering of subjects $(1, ..., n)$ (the order here does not need to coincide with the order in which subject can join the resistance), then

$$T_i^{\star} = c + B_i \left(1 - f\left(\sum_1^i W_j, W_{ruler}, W_{tot}\right)\right)$$

5. Thus it is a matter to maximize the total possible transfer with respect to such "order"

$$\max_{order\ O} \sum_i \left(c + B_i \left(1 - f\left(\sum_{j:O(j)<O(i)} W_j, W_{ruler}, W_{tot}\right)\right)\right)$$

6. Note also that the ruler would want assign a lower position in the order, $O(i) \leq O(j)$, subjects i with lower W and higher B. ■

PROOF OF THEOREM 3: From theorem 1 we know that in any SPE, the ruler would set up a ranking so to maximize

$$\max_{order\ O} \sum_i \left(c + B_i \left(1 - f\left(\sum_{j:O(j)<O(i)} W_j, W_{Ruler}, W_{tot}\right)\right)\right)$$

sub

$$T_i^{\star} = c + B_i \left(1 - f\left(\sum_1^i W_j, W_{Ruler}, W_{tot}\right)\right)$$

Note that if $f(\cdot)$ is linear in the first argument, the following is true

$$\max_{order\ O} \sum_i \left(c + B_i \left(1 - f \left(\sum_{j:O(j)<O(i)} W_j, W_{Ruler}, W_{tot} \right) \right) \right)$$
$$= nc + \sum_i B_i - \max_{order\ O} \left(-\sum_i B_i \sum_{j:O(j)<O(i)} f\left(W_j, W_{Ruler}, W_{tot}\right) \right)$$

Solving the above program is thus equivalent to the following

$$\min_{Order\ O} \left(\sum_1^n f\left(W_i, W_{Ruler}, W_{tot}\right) \sum_{O(j)<O(i)} B_j \right)$$

Consider exchanging the order between i and $i+1$ keeping everyone else fixed. Note that nothing changes for the others, so the incentive to switch i with $i+1$ is just given by the following

$$\left[f\left(W_i, W_{Ruler}, W_{tot}\right) \left(\sum_{j<i} B_j + B_{i+1} \right) + f\left(W_{i+1}, W_{Ruler}, W_{tot}\right) \left(\sum_{j<i} B_j \right) \right] +$$
$$- \left[f\left(W_{i+1}, W_{Ruler}, W_{tot}\right) \left(\sum_{j<i} B_j + B_i \right) + f\left(W_i, W_{Ruler}, W_{tot}\right) \left(\sum_{j<i} B_j \right) \right]$$
$$= \left[f\left(W_i, W_{Ruler}, W_{tot}\right) B_{i+1} \right] - \left[f\left(W_{i+1}, W_{Ruler}, W_{tot}\right) B_i \right]$$

Finally, note that this is negative if and only if

$$\frac{f\left(W_i, W_{Ruler}, W_{tot}\right)}{B_i} < \frac{f\left(W_{i+1}, W_{Ruler}, W_{tot}\right)}{B_{i+1}}$$

Thus, to maximize the total extraction $\sum_i T_i^\star$, the ruler would optimally order subjects in terms of $\frac{f(W_i, W_{Ruler}, W_{tot})}{B_i}$ with higher values corresponding to a higher position in the ranking from Theorem 2.1. Using this result and plugging in the specific $f(\cdot)$ from Assumption 2,

we establish Theorem 3. ■

PROOF OF PROPOSITION 5: Suppose that G is the set of all subjects. Note that if the ruler wants to be sure to avoid resistance in equilibrium, it would optimally demand the same taxes as under $p = 1$, call this vector $\mathbf{T}^1$. However, especially when the probability that agents would actually stick together is relatively small, the ruler might prefer to gamble and try to extract more taxes, accepting the risk of triggering some resistance. In this case, the optimal tax vector would be the one prevailing under $p = 0$, which we will call $\mathbf{T}^0$. When G includes all subjects, no other tax vector would be optimally demanded in equilibrium. Indeed any other tax vector would either generate lower income for the ruler than $\mathbf{T}^0$ or generate resistance from all subjects whenever they stick together, in which case $\mathbf{T}^1$ would be better. From the ruler indifference condition it is easy to show that $\bar{p}$, where $\bar{p}$ is such that $\frac{\left(\frac{\sum_{i\in I} B_i \sum_{j\leq i} W_j}{W_{tot}}\right)}{(k+|G|c)} = \frac{\bar{p}}{(1-\bar{p}),}$ where k is the ruler's cost of facing a resistance and $|G|$ is the number of subjects in G.■

Lemma 5 *Suppose $G \subset I$. If $p > \bar{p_G}$ then, in equilibrium, the ruler demands T_1 and avoids resistance. If $p < \bar{p}$ then the ruler triggers the resistance of G.*

PROOF OF LEMMA 5: If $p_G = 1$ then by the same logic as theorem 1, the ruler would optimally avoid rebellions by optimally demanding a tax vector T_1 as shown in the main text. Since no tax vector performs better than T_1 for the ruler in case the alliance in G actually holds we can conclude that if T_1 is optimal at $p_G = \bar{p_G}$ then it is optimal also at $p_G \geq \bar{p_G}$. Similarly, denote by T_0 the optimal tax vector chosen by the ruler if $p_G = 0$ (our baseline model). Note that under $p_G = 0$ the ruler strictly prefers T_0 to T_1 and that under $p_G > 0$ T_0 triggers protests with probability p_G. As the ruler's payoff is continuously decreasing in p_G for all T, the ruler strictly prefers not to induce any resistance when $p_G = 1$ and strictly prefers T_0 to T_1 when $p_G = 0$, then $0 < \bar{p_G} < 1$. ■

PROOF OF PROPOSITION 6: As the previous lemma establish that when $p_G > \bar{p_G}$ the ruler optimally adopts T_1 then the the case of $p_G > \bar{p_G}$ immediately follows from the same logic as under $p = 1$, as the previous lemma. Consider the case in which $p_G < \bar{p_G}$. By the same logic of Lemma 3 no subject would resist both when the G holds and when it does not. So, even if $p_G < \bar{p_G}$,by the same logic as Theorem 2 the ruler would create a ranking both within G and within $I \setminus G$, assigning stronger subjects are assigned higher status (by the same logic as Corollary 1). Further note that every excluded subject that would join every resistance joined by a member of G (WLOG the strongest) when G does not hold, would also resist when G holds. So tax should be demanded according to a ranking: if ruler does not want to trigger resistance of i when G holds, then it should rank i higher than the strongest member of G (assigning T_i taking into account that all members of G would join i's resistance); if the ruler accepts to have i resisting when G holds, then it should rank i as under T_0 demanding exactly T_i^0. Thus it is easy to conclude that all excluded subjects will be demanded at most the same taxes as under T^0 and all members of G will be demanded at least the same taxes as under T^0. So a possible candidate for the optimal taxation would obviously be T^0, but the ruler might be better off by deviating from such taxation scheme. For example consider a coalition of two subjects: $G = \{s, j\}$ such that $W_s = W_j$. If there exists $i \in I \setminus G$ such that $W_i = W_j - \varepsilon$, with $\varepsilon > 0$ infinitesimally small, the ruler would prefer to rank i above both s and j. This would result in an infinitesimal loss in case T^0 were to be accepted by G, but guarantees that i would not join the resistance when the members in G are able to stick together and choose to resist. Therefore, in this case both s and j would be strictly worse off with respect to the case $p = 0$ (assuming ranking between s and j is preserved). As a result we showed that some excluded subjects can be strictly better off and some members of G strictly worse off. ■

Appendix C

Appendix to Chapter 3

C.1 Proofs

PROOF OF LEMMA 4: The result on Hierarchic Societies follows from the payoff structure. For instance, looking at the tax, we see that

$$T_i = B_i \times \underbrace{\left(\frac{W_{ruler} + \sum_{h:B_h>B_i} W_h}{W_{tot}}\right)}_{\substack{\text{equilibrium tax rate if} \\ \text{ruler is weaker than } i}} > B_i \times \underbrace{\left(\frac{\sum_{h:B_h>B_i} W_h}{W_{tot}}\right)}_{\substack{\text{equilibrium tax rate if} \\ \text{ruler is stronger than } i}},$$

Note that the equilibrium tax imposed on i when the ruler is stronger than i does not depend on which of the stronger subject is the ruler but just on the sum of the military resources of all stronger groups. Thus each subject is indifferent on which stronger group is the ruler.

On the other hand, the part of the result pertaining United Societies follows immediately from the payoff structure which is determined by the following tax:

$$T_i = B_i \left(\frac{W_{ruler}}{W_{tot}}\right)$$

■

PROOF OF THEOREM 4: Suppose the rulership is held by the strongest group ($W_{strongest}$). Suppose that any other group challenges the ruler (say $W_{challenger}$), then by Lemma 4 every group who is weaker than the challenger would be indifferent on who gets to be ruler and every group who is stronger than the challenger would strictly prefer the strongest to remain ruler. Moreover, because we assumed that there is a negligible but positive cost in picking a side, groups would remain neutral when indifferent. As a consequence, the strongest group would remain ruler as $W_{I-support} \geq W_{strongest} > W_{C-support} = W_{challenger}$. On the other hand, if the incumbent ($W_{incumbent}$) is not the strongest group and the strongest group challenges it, then by the same argument we have $W_{I-support} = W_{incumbent} < W_{strongest} \leq W_{C-support}$. Turning our attention to United Societies, Lemma 4 tells us that every subject would always support the weakest contender for the rulership, unless they are themselves the challenger. In equilibrium, it cannot occur that the weakest social group does not obtain the rulership, unless there is one group who has enough power to make itself the ruler, i.e., it has more than half of society's resources. ■

PROOF OF THEOREM 5 and 6: Both Theorem 6 and 5 are based on the fact that there is a discontinuity that can be otaining by becoming richer than an opponent. For this reason, we will prove the two theorems jointly. Let $h_s^\ell = \Delta B_\ell^s(W_A, W_D, b)$, $\pi = p(W_A, W_D)$, $L = \sum_{j \in I \setminus \{D\}: W_j < W_A} W_j$, and $W_{tot \setminus \{A,D\}} = W_{tot} - W_A - W_D$. Group A's gain from overcoming D with a conflict of size b is:

$$
\begin{aligned}
&(B_A + h_A^A)\frac{L + g\left(B_A + h_A^A\right) + g\left(B_D + h_D^A\right)}{W_{tot\setminus\{A,D\}} + g\left(B_A + h_A^A\right) + g\left(B_D + h_D^A\right)}\pi + \\
&+ (B_A + h_A^D)\frac{L + g\left(B_A + h_A^D\right)}{W_{tot\setminus\{A,D\}} + g\left(B_A + h_A^D\right) + g\left(B_D + h_D^D\right)}(1 - \pi) + \\
&- \kappa(b) - B_A\frac{L + g\left(B_A\right)}{W_{tot}},
\end{aligned}
$$

Set $b = b(B_A, B_D)$ such that $b(B_A, B_D)$ is just enough for A to overcome D in case of success, i.e. $h_A^A(B_A, B_D, b(B_A, B_D)) - h_D^A(B_A, B_D, b(B_A, B_D)) = [B_A - B_D]^+$. Recall h_l^s is assumed to be continuous in all its arguments. As always we use $W_i = g(B_i)$. Thus for $B_A \to B_D$ we get $h_A^A(W_A, W_D, b(B_A, B_D)) - h_D^A(W_A, W_D, b(B_A, B_D)) \to 0^+$, which is true for $b(B_A, B_D) \to 0^+$, in which case $h_\ell^s \to 0$ for both $\ell, s \in \{A, D\}$. Thus for $B_A \to B_D$ and $b \to 0^+$, the previous expression goes to:

$$
\begin{aligned}
& p(W_i, W_i) \frac{L + W_i + W_j}{W_{tot}} B_i + (1 - p(W_i, W_i)) \frac{L + W_i}{W_{tot}} B_i - \kappa\left(0^+\right) - \frac{L + W_i}{W_{tot}} B_i \\
&= p(W_i, W_i) \frac{W_j}{W_{tot}} B_i - \kappa\left(0^+\right) \\
&= p(W_i, W_i) \frac{W_i}{W_{tot}} B_i - \kappa\left(0^+\right)
\end{aligned}
$$

Since by assumption $p(W_i, W_i) \frac{W_i}{W_{tot}} B_i > \kappa(0) = 0$, and both $\kappa(\cdot)$ and $h_\ell^s(\cdot)$ are continuous we have that there exists an $M > 0$ such that i would always attack another group (possibly j) for some stake b whenever $0 < W_j - W_i < M$. ■