Power Mismatch and Civil Conflict: An Empirical Investigation*

Massimo Morelli †   Laura Ogliari ‡   Long Hong §

This version: September 28, 2022

Abstract

This paper empirically shows that the imbalance between an ethnic group's political and military power is crucial to understand the likelihood that a group engages in a conflict. We develop a novel measure of a group's military power by combining machine learning techniques with rich data on ethnic group characteristics and outcomes of civil conflicts in Africa and the Middle East. We couple this measure with available indicators of ethnic groups' political power as well as with a novel proxy based on information about the ethnicity of cabinet members. We find that groups characterized by a higher mismatch between military and political power are approximately 30% more likely to engage in a conflict against their government. We also find that the effects of power mismatch are nonlinear, which is in agreement with the predictions of a simple model that accounts for the cost of conflict. Moreover, our results suggest that high-mismatched groups are typically involved in larger, longer, and centrist conflicts.

Keywords: Civil War, Military Power, Political Power, Mismatch, Machine Learning

---

*Morelli wishes to thank the European Research Council, ERC Advanced Grant 694583. We wish to thank Giulia Buccione, Jacopo Bellabio, Matteo Bertoli, and especially Mohamed Badaoui for indispensable and amazing research assistance and Yannick Pengl, Dominic Rohner, Austin Wright, and Ekaterina Zhuravskaya for very useful comments and suggestions. We also thank seminar and conference participants at the Political Economy of Power Relation conference, the CEPR ITRE seminar series, Zurich Political Economy Seminar Series (Z-PESS) and Sciences Po. The usual disclaimer applies.

†Department of Economics, Bocconi University, & CEPR, massimo.morelli@unibocconi.it

‡Department of Economics, University of Bergamo & CEPR, laura.ogliari@unibg.it

§Department of Economics, the University of Wisconsin - Madison, long.hong@wisc.edu

1 Introduction

When thinking about the causes of wars, the concept of *power* immediately comes to mind. Scholars, though, almost exclusively refer to military power when studying conflicts. In this paper, we provide new empirical evidence highlighting the importance of simultaneously accounting for multiple dimensions of power. Following the insights of the theory introduced by Helios et al. (2022), we document that conflicts are related to power mismatches, namely, asymmetries between the distribution of the relative military and political power of the actors involved. No systematic empirical evidence on the role of power mismatches exists, and such evidence could be crucial in order to assess the potential effectiveness of alternative policies for peace.

To progress in this direction, we build a novel measure of military power of ethnic groups, exploiting a wealth of ethnic group-level data and machine learning techniques. We combine this new measure with detailed information on civil conflict events, ethnic groups' characteristics, and measures of political power. We obtain a novel dataset, covering virtually the universe of politically relevant ethnic groups in Africa and the Middle East for the period 1992-2012. By using these new measures of the relative military and political power of ethnic groups, we can calculate the groups' power mismatch and use it to appraise, at the intra-state level, the role it plays in the decision to engage in conflicts.[1]

Using different specifications that account for many observed and unobserved confounding factors, we find that high-mismatch groups are approximately 30% more likely to engage in conflict against their government. Additionally, we observe a convex relationship between power mismatch and conflict: when the mismatch is small, a marginal increase in the mismatch does not raise the probability of conflict. On the other hand, a further increase in the mismatch raises the likelihood of conflict participation more than proportionally for high-mismatch groups. Lastly, we examine the relationship between mismatch and conflict characteristics, showing that mismatched groups have a higher chance to be involved in conflicts that are bigger (in terms of fatalities), last longer, and concern power-sharing at the central level. The latter result is also consistent with additional theoretical findings in Esteban et al. (2020): it is precisely when a dispute is about power distribution in a country (centrist dispute rather than a dispute about autonomy or potential secession) that the mismatch of powers is salient, and the bigger the stakes, the bigger the conflict.

The international relations literature debates the pros and cons of balance versus preponderance of power, focusing on military power alone. Even theorists emphasizing commitment problems as the primary cause of war always refer to the difficulty to commit not to use military power, and the comparison with political economic power is mostly ignored (see e.g., the seminal work of Fearon, 1995). A notable exception is found in the body of work by Cederman (see e.g., Cederman et al., 2013 for a recent detailed analysis of the different grievances that might lead to civil war). Cederman's research stresses the role of economic inequality and political exclusion as a trigger for civil conflict. In the same spirit, in this paper, we use the concept of "relative political Power"

[1] At the inter-state level Helios et al. (2022) find some preliminary supporting evidence, using GDP ratios as a very rudimentary measure of economic power.

to capture the advantage conferred to a player by the existing political institutions—for example, the relative control of the political bodies governing the allocation of resources in peace. This paper is directly linked to the recent theory proposed by Helios et al. (2022). Their model establishes that when bargaining fails – for example, for commitment or asymmetric information problems –, the probability of conflict depends on the *power mismatch* between the disputant groups, where the mismatch is defined as the difference between the relative military strength and the relative political-economic power of the disputant groups. If the two types of power are unequally but similarly distributed (e.g., the player that has greater military power also has greater political-economic power), no war should ensue. Conflict can instead arise when one of the players is relatively stronger in one of the two dimensions of power.[2] We contribute to the literature by providing the first empirical evidence that power mismatch matters for conflict.

Our paper also speaks to the debate within the literature on power sharing. On one side of this debate, it lies the argument that conflict settlements are more solid if they entailed institutions that granted power-sharing (see for a primary example Lijphart, 1977). On the other side, scholars held that such arrangements were morally wrong and quite precarious, and instead suggested electoral mechanisms as a way to build stable democracies (e.g., Horowitz, 1990, 2000, 2003; O'Flynn and Russell, 2005). Our results on the key role of power mismatch have the policy implication that conflict risk can be effectively reduced when both dimensions of power are simultaneously addressed. The example of the UNITA rebel group in Angola is telling. In the rhetoric of UNITA, civil war against the MPLA government in Luanda was justified by the group's exclusion from power at the time of independence in 1975. However, after signing a peace agreement in 1991 and losing the winner-take-all type of elections in 1992, UNITA returned to war until it was finally induced to sign a power-sharing agreement in 1994. This shows that elections are not a panacea unless the electoral system and proportion of voters can determine an implicit commitment to power sharing. However, even power sharing often fails since forms of credible power-sharing agreement do not necessarily reflect the desirable elimination of a mismatch: for example, a pure democracy with a proportional electoral system could guarantee a group with 30 percent of ethnic group voters 30 percent of political power, but if such a group has a probability of victory against the majority group much higher or much lower than 30 percent the mismatch is not eliminated.[3] Given that simply using elections (when they are fair) does not guarantee the credible elimination of a mismatch, democracy has to be supplemented by inventive institutional designs, for example, creating commitments in terms of public jobs, political roles, and military quotas in exchange for deposition of weapons, as in the peace treaty that led to the demilitarization of the IRA. Similarly, also in Colombia, the demilitarization of the FARC had to go hand-in-hand with the concession of a political role.

The paper is organized as follows: in section 2 we recall the baseline model that formalizes the mismatch theory of conflict; in section 3 we describe in detail the data collec-

[2]The result is shown to hold even when allowing for bargaining because the evolution of military and political powers and their future use in case of an indecisive war cannot be contracted ex-ante. This is one of the main reasons why the power mismatch giving rise also to the Russian attack incentives with Ukraine is difficult to re-balance at a negotiation table.

[3]See Spears (2000) for a comprehensive discussion on power-sharing agreements in Africa.

tion efforts on all fronts, discussing the relevant previous literature; In section 4 we zoom in particular on the description of the machine learning procedures used to create our main novel military power measure. Section 5 presents the empirical results, and section 6 concludes. The appendix contains all the technical aspects.

2 A Simple Theoretical Framework

Consider a government controlling group G and an ethnic group E that has to decide whether to rebel or not. Let $p \equiv \frac{p_E}{p_E + p_G}$ denote the relative political power of E.[4] Finally, let m denote the probability of winning of E in case of war against G and let c be the cost of war for each player.

Denote by S the divisible surplus and consider first a case in which $m > p$. If E decides not to challenge the status quo, her payoff is $U_E = pS$; on the other hand, in case of conflict, the payoff for E is, with the standard costly lottery assumptions, $mS - c_E$. Thus, conflict is initiated by E in a one-shot game iff $c_E < (m-p)S$. Given any ex ante uncertainty on c_E, represented by a distribution $F(\cdot)$ on the domain $[0, \infty)$, E rebels with probability $F((m-p)S)$, and hence the incentive to rebel is increasing in $(m-p)$, which represents the *mismatch*. It is also clearly increasing in the size of the divisible surplus.

Similarly, the government may have an incentive to start a (repression) conflict if $c_G < [(1-m) - (1-p)]S$ when $m < p$. In such a case conflict exist with the corresponding probability that c_G is less than $G((p-m)S)$, where $G(\cdot)$ denotes the cumulative probability distribution of the possible realizations of c_G.

Between the above two possible initiation incentives, it is much more likely that the former obtain, namely, the conflicts are usually initiated by ethnic groups that rebel against a status quo where they are given too little political-economic power. However, the data do not allow to distinguish initiation, and, moreover, both cases actually say the same thing in terms of the role of the mismatch, and we can certainly conclude that
Main prediction: Conflict is more likely to happen when $|m-p|$ is high.

The more general model, allowing for dynamics, bargaining and stalemates, can be found in Helios et al. (2022).

3 Data description

To test the validity of the mismatch theory for civil conflicts, we need at least three pieces of information: (i) what are the relevant groups that may be tempted to participate in conflicts; (ii) the group's political power; (iii) the group's military strength. To test whether the more mismatched groups are more likely to enter into conflicts, we need to know both military and political power, even for the groups that never participated in a conflict. To

[4] In a parliamentary system one measure of this in the status quo could be the relative number of seats in the parliament or the relative number of ministries in the government. But in our sample, the regimes and meaning of political power are quite different from country to country, and the construction of an appropriate measure of relative political power is one of the contributions of the paper.

the best of our knowledge, none of these pieces of information is readily available in existing datasets. This section provides a detailed description of how we identify the ethnic groups, the conflict variables, and the methodology used to measure group-level political and military power.

3.1 Ethnic conflict

Throughout the analysis, we focus on conflicts involving the government group (represented by at least one dominant ethnic group) and one or more ethnic opposition groups over the period 1992-2012. We restrict our attention to conflicts occurring in Africa and the Middle East to make sure that ethnicity represents a salient cleavage.[5] We construct a dataset that records ethnic groups and conflict information by linking the UCDP Georeferenced Event Dataset (UCDP-GED) and Ethnic Power Relations (EPR). In what follows, we describe in detail the procedure used to link the two databases.

UCDP-GED We use the UCDP Georeferenced Event Dataset (UCDP-GED) Global version 5.0,[6] which contains information on conflict events from 1989 to 2015. The database has three critical features that make it particularly suitable for our analysis. First, it classifies the type of violence, allowing us to identify conflicts in which one of the actors is a government. Second, it contains detailed information on the precise location of conflict events allowing us to easily assign such events to geo-referenced ethnic groups using their location. Third, the database reports all the incidents that result in at least one direct associated death, which allows us to extend the analysis to include small-scale conflicts.[7]

Ethnic Power Relation We use the list of ethnic groups provided in the Ethnic Power Relations (EPR) Core dataset (version 2018.1.1),[8] which identifies 817 politically relevant ethnic groups worldwide for the period 1946 - 2018. EPR defines ethnic groups according to the ethnic categories most salient for national politics in each country. An ethnic group is politically relevant if either at least one significant political actor claims to represent the interests of that group in the national political arena or if group members are systematically and intentionally discriminated against in the domain of public politics. This salient feature of the database assures that the groups in our analysis are politically relevant.[9] Moreover, we rely on the information contained in EPR to identify the ethnicity of the

[5]Table A.1 in Appendix A reports the list of countries in our sample.

[6]The database is introduced by Sundberg and Melander (2013). See Eck (2012) for a detailed discussion on the strengths and weaknesses of the database.

[7]The database also offers estimates of fatalities for *each* involved party, which will be crucial for us to infer the potential winner of the battle, which is not typically available in any civil conflict database but is essential for our military power measure as discussed in Section 4.

[8]The database was first introduced by Cederman et al. (2010) and further developed by Vogt et al. (2015).

[9]Alternatively, one could use the Geographic Representation of Ethnic Groups dataset (GREG) which digitally represents settlement patterns of ethnic groups worldwide coming from a version of the Atlas Narodov Mira (ANM) (Bruk and Apenchenko, 1964): a series of maps collected by Soviet ethnographers charting ethnic groups across space. However, besides being potentially outdated, the main limitation of this dataset is that it focuses exclusively on the list of ethnic groups given by the ANM authors, even if the linguistic differences on which the ANM focuses do not correspond to ethnic cleavages that are politically

group that holds political power. Using the geo-referencing feature of UCDP-GED and EPR, we explore the ethnic conflict and rebel groups' ethnicity (see below). It is worth noting that, due to the dynamic political environment of the country, both the politically relevant groups and the political ranks of each group are time-varying.

Government ethnicity EPR provides rich information on political power. Specifically, the database contains a variable that ranks each ethnic group's political power from 1 (discriminated) to 7 (monopoly).[10] In particular, if a group rules alone, the group is either "monopoly" or "dominance". If two groups share powers, the groups are either "senior partners" or "junior partners". If a group is excluded from the ruling, the group is either "powerless", or "discrimination", or "self-exclusion".

We follow a number of procedures in order to find the ethnicity of a government in a given year. First, we assign the ethnic group with the highest power rank as the government group if there is only one group. Typically, it is at the political rank of "monopoly" and "dominance". Second, if more than one group has the highest power rank in a given period—typically occurring for "senior or junior partners"—we manually examine whether they are allied using EPR Atlas, which documents in detail how the political rank is created. Appendix A provides a detailed discussion about our manual checks. If these groups are allied, we assign the government to all these groups. If they are not, we determine, when possible, which group enjoys a larger advantage using external sources. In some rare cases, it is impossible to determine the government's ethnicity, and in such situations, we exclude them from the analysis.

Rebel group ethnicity There is no direct linkage between rebel groups in UCDP-GED and EPR ethnicity. We apply a multi-step procedure to assign ethnicity to the UCDP-GED actors. First, we utilize the ACD2EPR conversion table developed by (Wucherpfennig et al., 2011; Vogt et al., 2015), which integrates UCDP/PRIO Armed Conflict Dataset with EPR.[11] In particular, the ACD2EPR covers 133 rebel groups of the 637 in our main sample.

For the remaining rebel groups, we provide a novel approach that helps to identify such linkage. First, we follow a strategy used in Michalopoulos and Papaioannou (2016) and Moscona et al. (2020), and assign each conflict event to an ethnic group's homeland by geo-match of the UCDP-GED and EPR databases.[12] In this step, we exclude all the conflicts that occur in the homeland of the government ethnicity. Second, we count the

relevant (see e.g., Posner, 2004; Cederman and Girardin, 2007; Chandra and Wilkinson, 2008; Wucherpfennig et al., 2011)

[10]There is another category called "political irrelevance", which we exclude from the analysis. In a few cases, the country is in a state of collapse. We exclude the country-year observations from our primary analysis.

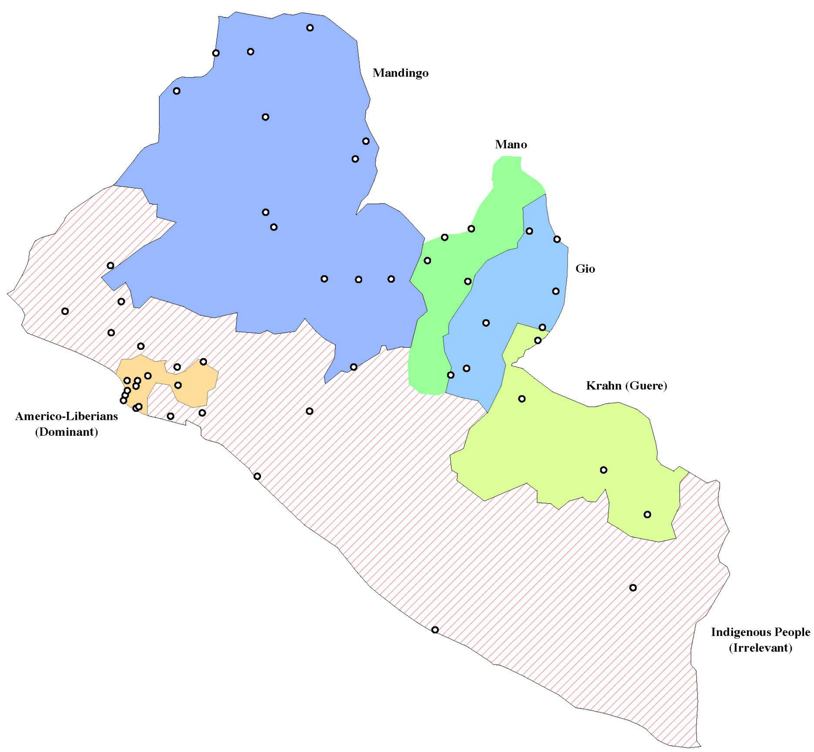
[11]UCDP/PRIO Armed Conflict Dataset, first introduced by Gleditsch et al. (2002), is an old version of the UCDP-GED. The conversion table contains a smaller set of rebel groups than that used in UCDP-GED because UCDP/PRIO only records large conflict events where the number of involved casualties is at least 25.

[12]We discard all conflict events that happened outside the boundaries of any ethnic groups. As a result, we matched 65,820 events out of the 66.396 events (99.1%). We have also done a robustness check to assign the unmatched conflicts to the nearest ethnic groups. It does not change our main results.

number of times a rebel group has a conflict event in the homeland of a particular ethnic group.[13] In cases when a conflict happens in an area where multiple ethnic groups coexist, the event is attributed to all overlapping ethnicities. Finally, we assign the ethnicity with the highest count to the rebel group. In some cases, there is a tie among the counts. We chose the homeland of the ethnic group with the highest fatalities occurred. In rare cases, the fatalities are also in a tie. We break the tie by randomly choosing one.[14]

To give a simple but concrete example, Figure 1 shows a map of Liberia with all its ethnic groups in EPR and all 47 conflict events associated with the Liberians United for Reconciliation and Democracy (LURD). We first exclude the 9 events in the dominant group's homeland and 10 associated with the irrelevant group. Using the remaining conflict events, we assign LURD to Mandingo because most of their events (13 incidences) took place in the ethnic group's homeland.

Figure 1: Rebel group coding: LURD in Liberia as an example



Notes: This is a map of Liberia, where each colored polygon represents an ethnic group listed in EPR. The dots represent all conflict events in UCDP-GED associated with the rebel group LURD.

[13]Note that we do not count the events in the homeland of the irrelevant groups since they are excluded from our main analysis.

[14]We have also conducted a manual check to identify the ethnicity of rebel groups through online resources. Among the rebel groups that we could manually identify, around 70% is in line with the results from our method.

Key dependent variables Our primary interest is the analysis of civil conflicts between incumbent governments and (ethnic) rebel groups. Using the information on involved parties in UCDP-GED, we restrict our attention to conflict events where at least one of the actors is the government. We aggregate the events at the ethnic group-year level using each rebel group's ethnicity. Hence, our primary dependent variable in our analysis is conflict *incidence*.

3.2 Relative political power

Building a continuous measure of the political power of ethnic groups with respect to the government presents challenges. The literature used relative group size and the ethnicity of the leader or relied on experts' opinions. We build two proxies of groups' political power. The first approach directly uses the discrete power rank index provided in EPR. The advantage of the measure is that it is available for all the politically relevant ethnic groups in our sample. However, since it is a discrete index (i) it does not present a lot of variation, and (ii) the index is an ordinal variable, and its value does not represent any "real" measure of power. In our second approach, we restrict our attention to a sub-sample of African countries and exploit the ethnicity of cabinet members to build a continuous measure of ethnic groups' political power. We discuss the two approaches in detail below.

Discrete measure We assign EPR's power rank as a measure of the political power of each ethnic group and define the relative political power p_{eg}^{PR} of ethnic group e to government g as the following.

$$p_{eg}^{PR} = \frac{C_e^{PR}}{C_g^{PR}} \in (0, 1)$$

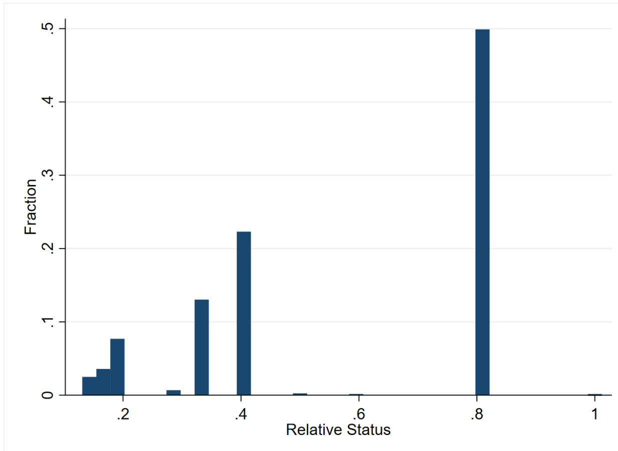
where C is the discrete power rank. Figure 2 shows the distribution of the measure. As expected, the distribution of p_{eg}^{PR} contains few points. Groups with high political power are basically those with $p_{eg}^{PR} = 0.8$, that is, groups that are considered senior partners in the government. Groups with low political power are groups that are considered either powerless or that are discriminated against.

Continuous measure We collected data about cabinet membership in 14 countries in Sub-Saharan Africa over 21 years (from 1992 to 2012).[15] We choose these fourteen countries for two reasons. First, their location is in Saharan Africa, where conflicts are most likely to be ethnic related. Second, as shown in the previous work by Francois et al. (2015), these countries are democracies and the share of cabinet members can be considered a valid measure of political power. [16]

[15] They are: Benin, Cameroon, Republic of Congo, Democratic Republic of Congo, Ivory Coast, Gabon, Ghana, Guinea, Kenya, Liberia, Nigeria, Sierra Leone, Tanzania, and Uganda.

[16] Francois et al. (2015) use a different categorization of ethnicity from EPR, which was not available at the time of their work.

Figure 2: Histogram of the relative political power p_{eg}^{PR}



To identify the ethnicity of cabinet members we follow a procedure that entails different steps. First, we obtain all the cabinet membership information from the CIA's "Chiefs of State and Cabinet Members of Foreign Governments".[17] We precisely extract all the incumbent cabinet members at the end of each calendar year. We manually check the cabinet members' names using various online sources to identify their ethnic affiliation and use that information to match the ethnic list in EPR. If direct evidence is not available, we turn to two alternatives: the ethnicity of the minister's parents and the birthplace. Specifically, using the coordinates of the birthplace and the geo-referenced EPR map, we assign the ethnic homeland of the birthplace of the cabinet member as the ethnicity. When the birthplace is not available (e.g., some minister is born in a foreign country), we use the location of his primary school or the location of the district of her first election. With this procedure, we identify 82 politically relevant ethnic groups.

The share of the cabinet seat held by an ethnicity should represent the share of the political power of the ethnic group insofar as cabinet members decide the allocation of the resources of a country. In fact, if cabinet membership contained information on political power we would expect it to monotonically decrease with the EPR Power rank measure. Table 1 confirms such a relationship: dominant groups on average hold more than 50% of the cabinet seats. On the other hand, the powerless or discriminated groups tend to have negligible shares.

We define the relative political power $p_{i,c}$ of group i in country c as the ratio of the cabinet seats of the group in a given year relative to the seats held by the government's ethnicity, g:

$$\hat{p}_{i,g,c} := \frac{\#(\text{Cabinet seats belonging to ethnicity } i)}{\#(\text{Cabinet seats belonging to ethnicity } g)}$$

Figure 3 plots the distribution of $\hat{p}_{i,g,c}$. Unlike the discrete measure, the relative political measure now is quite smooth. It also contains a mass point at zero, which indicates

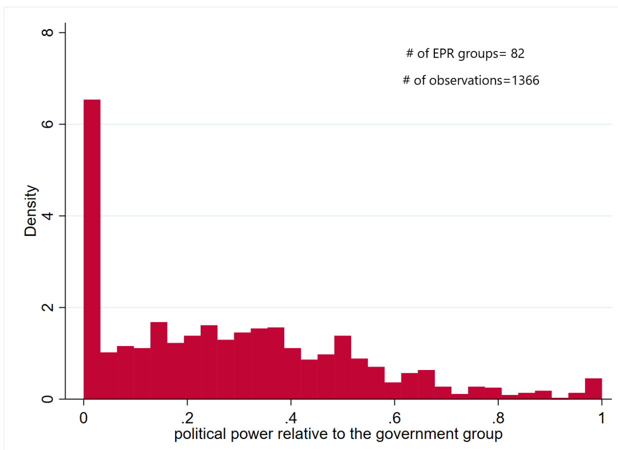
[17]For detailed information, please refer to the CIA website: https://www.cia.gov/resources/world-leaders.

Table 1: A comparison between EPR power rank and Cabinet member shares

		Political power (cabinet member shares)	
		Mean	Standard Dev.
Power Rank	Dominant	0.532	0.188
	Senior partner	0.259	0.142
	Junior partner	0.143	0.125
	Powerless	0.092	0.098
	Discriminated	0.068	0.063

that many groups are not represented in the cabinet, implying a certain degree of political power inequality in the restricted sample.

Figure 3: Histogram of relative political power



3.3 Relative military power

Finding measures of military power at the ethnic-group level is extremely challenging. Traditional information that is available at the country level—military expenditure, military personnel, trade in arms—does not exist at the group level. This is why the literature often uses the group's population size or GDP per capita as proxies of military power. In a recent paper, however, Carroll and Kenkel (2019) shows that these two variables perform no better than random guesses when used to predict the probability of winning a conflict in the context of inter-state conflict.

To be consistent with the theory described in Section 2, we define relative military strength as the probability of winning a conflict against the government. Estimating such a probability poses one difficulty: we need to compute the probability of winning a conflict for those groups that have never participated in a conflict. Moreover, even for groups that did participate in conflicts, inferring the probability of winning during the peace years requires a non-trivial technique.

We rely on some insights in Carroll and Kenkel (2019), who propose a machine-learning technique to overcome the challenges. We modify their algorithm and use an extended sample (described below) of conflicts in Asia and Africa combined with a rich set of ethnic group-level and country-level variables to infer the probability of victory for all potential conflicts between every ethnic (rebel) group and the government. The details of the machine learning procedure are described in Section 4.

3.4 Mismatch measure

With the measures of political and military power in hand, we can construct our primary independent variable: the empirical power mismatch measure $M_{e,g}$. We propose two measures of power mismatch at the group level: the Mismatch Dummy $M^D_{e,g}$, and a continuous mismatch variable. Specifically, we define $M^D_{e,g} = 1$ if the relative political power ($p_{e,g}$) is in the first quartile of the distribution and the relative military power ($m_{e,g}$) is in the fourth quartile or vice versa. In other words, $M^D_{e,g} = 1$ indicates the presence of a high imbalance between the relative political power of a group and its relative military strength. The continuous measure is defined as

$$M_{e,g} := |m_{e,g} - p_{e,g}|$$

where $m_{e,g}$ is the predicted probability of winning a conflict against the government, and $p_{e,g}$ is the relative political power measured by using the ethnicity of the cabinet members. Note that since both $m_{e,g}$ and $p_{e,g}$ change over time, our mismatch definition is also time-varying.

3.5 Control variables

In our analysis, we control for an extensive set of ethnic-level variables mainly constructed from GROWup (2019) and GRID-PRIO (v.2.0).[18]. We supplement this data with information on the location of active oil and gas fields from PETRODATA (constructed by Morelli and Rohner, 2015). Control variables are grouped into five categories, mainly following the literature on conflict. First, we include social demographic information such as population and night lights. Second, the geographic condition is captured by obtaining forest, water, climate, land area, and elevation information. Third, the location of a group is measured by controlling its distance to capital and country borders. Forth, we include the natural resources and the share of the group's homeland covered with petroleum fields. Finally, conflict history is controlled by the number of peaceful years since the last conflict event.

4 Predicting military power via machine learning

We estimate the probability of winning a conflict as a proxy of the relative military power between an ethnic group and a government. We do not directly observe the probability of

[18] GROWup is developed by Girardin et al. (2015) and GRID-PRIO is introduced by Tollefsen et al. (2012)

winning a conflict, especially when an ethnic group has never experienced a conflict. As surveyed in Carroll and Kenkel (2019), if we follow the standard approach – using a linear probability model – to predict the winning probability of a conflict, the accuracy might be as good as a random guess. We, therefore, rely on a machine learning algorithm modified from theirs and use a rich set of observed ethnic group-level variables to compute the probability of winning for all the ethnic rebel groups against their government in our sample. Below, we first describe our *training dataset* from which our machine learning algorithm is trained, then discuss the algorithm and performance.

4.1 Training set

To train the algorithm, we need to provide a training set that contains plausible conflict outcomes, i.e., which side wins the conflict. Such a variable is typically not reported in a standard database, including GED. We use a novel approach below to identify the winning side of a conflict. We leverage the idea that the winning party has a smaller fatality (relative to the population) than the losing party and thus use the fatalities ratio as a measure of whether a group wins the battle.[19]

Formally, let f_g^t, f_e^t be government g and group e's fatalities (normalized by the corresponding group's population) in year t, respectively. We define the binary outcome Y, which is equal to 1 if the government wins and 0 otherwise:

$$Y := \mathbb{I}\{F_{eg}^t = \frac{f_g^t}{f_g^t + f_e^t} \geq c\} \tag{1}$$

That is, the ethnic group wins the conflict in year t if its share of fatalities is less than a threshold c.

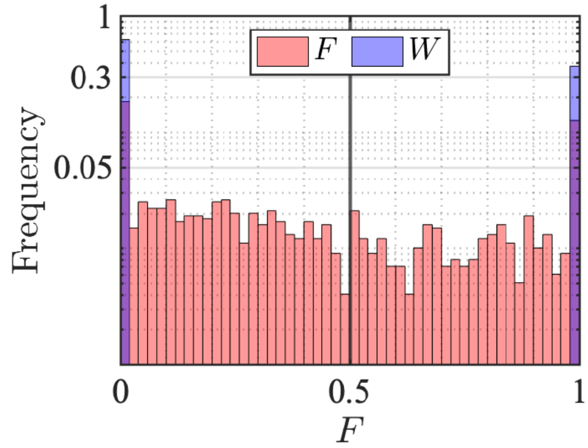
Figure 4 exhibits the distribution of F_{eg}^t in red. As it shows, there are two focal points at the two endpoints, meaning the majority of fatality ratios are concentrated at zero or one. The remaining ratios are distributed (almost) uniformly. Motivated by the data pattern, we choose the threshold $c = 0.5$. The results are also robust if we choose a different threshold.[20]

We use a rich set of predictors to train the model. In particular, we include ethnic-group level demographic, geographic, and meteorological information, as well as their external support and ethnic kinship information, extracted from the GROWup database. We include year dummies, region dummies, and the longitude and latitude of countries to capture aggregate trends. A detailed list is discussed in Appendix B.1. Instead of choosing a subset of predictors to avoid the over-fitting issue, we utilize the machine

[19] This argument is supported by the strong correlation between winning and having a smaller number of deaths (Sarkees and Wayman, 2010).

[20] An alternative dataset that documents inter-state conflict outcomes is ACLED. The battle outcome can be inferred by exploring whether the control of territory has changed. However, due to limited other information, including fatality, more than 90% of the conflict would be categorized as stalemates or unknown. Correlates of War (COW) reports both conflict outcomes, although the number of conflicts matched to our sample is very limited. Using the matched conflicts, 82% of our conflict outcomes coincide with theirs.

Figure 4: Training outcome distribution and fatalities ratio



Note: In red, the distribution of the fatality ratio F in our training set. In blue, the outcome, i.e. the *binarisation* of the ratio according to eq. 1. The frequency is represented on a logarithmic scale for better visualisation.

learning algorithm to pick the set of variables with the highest predictive power. Specifically, we adopt the advanced learning algorithm, which conducts both variable selection and cross-validation testing (more details in Appendix B.3).

Lastly, we use an extended sample of African and Asian conflicts from 1989 to 2016, with 25 or more fatalities, as our training dataset. The results remain qualitatively similar when we restrict the sample to African countries only or include small conflicts. In the final training sample, we have 726 observations with 198 predictors, where the share of outcomes where government wins is around 64%.

4.2 Algorithm and performance

We train the machine learning model to predict the conflict outcomes using the training dataset, where the algorithm involves multiple learning algorithms, as we describe in detail in Appendix B.3. In this section, we briefly summarize the training technique and discuss its performance.

Algorithm Formally, we use a binary learning model $\mathcal{M}$ based on the training data (Y, X), where X contains all the ethnic level characteristics as described in Section 4.1 and Y is a binary outcome variable defined in equation (1). The optimal trained model $\mathcal{M}^*_{(Y,X)}$ is obtained by minimizing a pre-defined loss metric $\mathcal{L}$ given our dataset (Y, X), as shown below.

$$\mathcal{M}^*_{(Y,X)} = \arg\min_{\mathcal{M}\in\mathbb{M}} \mathcal{L}\big(Y - \mathbb{I}\{\mathcal{M}(X) > 0.5\}\big). \tag{2}$$

Intuitively, we train the model to get the model with the highest prediction level given the training data. We also allow cross-validation techniques to avoid the over-fitting problem.

Using the trained model, we can predict the winning probability of an observation that does not have conflict outcomes.

Performance The performance of a binary classification model is commonly measured by the logarithmic loss metric. Given the classification model $\mathcal{M}^*_{(Y,X)}$ obtained from equation (2), the log-loss is calculated as:

$$\mathrm{L}\big(\mathcal{M}^*_{(Y,X)}\big) = -\frac{1}{N}\sum_i Y_i \log\big(\mathcal{M}^*_{(Y,X)}(X_i)\big) + (1-Y_i)\log\big(1-\mathcal{M}^*_{(Y,X)}(X_i)\big).$$

The smaller it is, the closer the likelihood of correct classification. We are interested in the cross-validated log-loss for the sake of maximal prediction accuracy. Following Carroll and Kenkel (2019), we calculate the Proportional Reduction Loss (PRL) that gives the predictive power accuracy of our model relative to a null model i.e. a classification scheme that assigns the label by majority rule (hence accurate at the 64.2%). The higher the PRL, the stronger is our prediction compared to the null model. The PRL writes:

$$\mathrm{PRL}\big(\mathcal{M}^*_{(Y,X)}\big) = \frac{\mathrm{L}_{null} - \mathrm{L}\big(\mathcal{M}^*_{(Y,X)}\big)}{\mathrm{L}_{null}}.$$

The loss L_{null} from a null model model is always higher than the trained model's loss $\mathrm{L}\big(\mathcal{M}^*_{(Y,X)}\big)$. The higher the PRL, the bigger the performance normalized difference of $\mathcal{M}^*_{(Y,X)}$ with respect to the null model, and the better is our model.

Our reported PRL is comparable with the one obtained by CK for inter-state conflict. They obtain a PRL of 23%, only slightly higher than ours. Considering that we predict the outcome of intra-state conflicts, on which the data is very limited, the performance is quite good. Among the predictors that CK use, there are important metrics such as Iron and Steel production, Military expenditure and personnel, and Primary Energy consumption. Such data is absent for ethnic groups. The variables used to build our model are much harder to calculate, and most of them rely on satellite images.

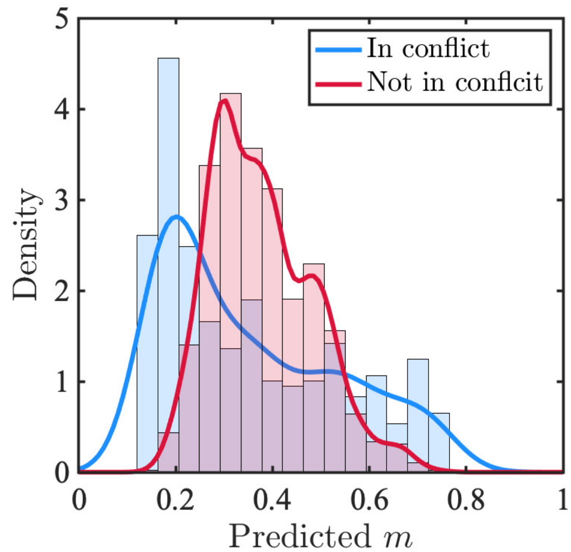
Nevertheless, our result can be critical when we compare it with the conventional military power proxies used in the literature, where population and night light have been commonly used as the main proxy for relative military power in intra-state conflicts (e.g., Esteban et al. (2012) use population). Table 2 shows that the population ratio or the night light ratio can predict the outcomes that are merely marginally better than random guessing. Our algorithm, on the other hand, performs 21% better than random guessing, which is a considerable improvement on state of the art.

Table 2: Algorithm's predictive power.

	CV Log-loss	PRL	Accuracy	Δ_{Null}(Accuracy)
Full model	0.554	15%	70.2%	9.1%
Population ratio	0.650	0.2%	64.4%	0%
Night light ratio	0.646	0.8%	65.1%	1.1%

Since groups that experience conflict are a selected sample, one concern is that the predicted military power of the groups that are in conflict (and that we use to estimate the parameters) is systematically different from that of groups that do not experience conflict. Given that in the empirical analysis we use (a transformation of) military power to predict conflict participation, we want the distribution of military power for the groups in the training sample and for the groups that do not experience conflict to have common support. Figure 5 shows the distribution of predicted military power for groups who are not in conflict (in red) and those who are in conflict (in blue). The figure shows that groups that are in conflict have a more dispersed distribution, but the two distributions have common support.

Figure 5: Distribution of the predicted military power



Finally, Table 3 reports the correlation of the predicted probability of winning a conflict, which we will use as our baseline measure of military power *m1*, with alternative predicted probabilities based on different proxies for the winning outcome. In particular, in model *m2*, we change the threshold we use to assign the outcome of the conflict based on the fatality ratio. To do so, we exploit the Correlates of War database, which is, to the best of our knowledge, the only database that reports both the outcome of a conflict and the fatalities borne by each party. We restrict the sample to intra-state conflicts against the government with a win/lose outcome, and compute the fatality ratio that maximizes the correct classification of the conflict outcome. We find this ratio to be 0.58, which yields a correct classification in 82% of the cases.[21] Model *m2* assigns a loss if the fatality ratio is greater or equal to 0.58, and the prediction of this model is extremely correlated with the baseline prediction. In model *m3*, instead, we include small conflicts (i.e., those with total fatalities lower than 25) in the training sample. Including small conflicts adds a bit of noise to the outcome variable, this reduces the performance of the model and yields a prediction of military power that is slightly less correlated with the baseline one. In model *m4*, instead, the fatality ratio is built using cumulative fatalities (fatalities over the

[21] Using our preferred threshold of 0.5 we correctly classify the outcome of the conflicts in the COW database 77% of the times.

Table 3: Correlation of military power predictions

	$m1$ Baseline	$m2$ Fatality threshold .58	$m3$ Small conflicts	$m4$ Cumulative deaths
$m1$ Baseline	1.00			
$m2$ Fatality threshold .58	0.947	1.00		
$m3$ Small conflicts	0.727	0.654	1.00	
$m4$ Cumulative deaths	0.925	0.859	0.799	1.00

entire conflict duration) instead of the yearly fatalities, keeping the threshold fixed at 0.5. Again, using cumulative fatalities yield predictions that are very similar to those of the baseline model. Appendix B.2, provides further results that show the robustness of the predictions to changes in the machine learning parameters.

4.3 Variables' relevance

We explore the question of which predictors might be the most important in predicting military power. However, due to the nature of the algorithm, it is somewhat challenging to tell how much a single predictor affects military power. We explore the importance of a particular predictor using the following algorithm (following CK). We remove a predictor of interest, rerun the entire algorithm, and track and compare the resulting PRL to the original PRL. Effectively, the larger the difference between the resulting PRL and the original one (or PRL loss), the more important the predictor is. Instead of having an extensive list of all variables, we report the PRL loss in Table 4 some of the important predictors, which are external support variables (Ext), geographic variables i.e., border distance, capital distance, and travel time (Geo), population and demographic growth variables (Pop), peace years and war history (Py Wh), affiliated ethnic groups (Tek), land characteristics (Land), country-level variables, i.e., latitude, longitude and region dummy (Country).

Table 4: Variables' predictive power.

	Ext	Geo	Pop	PyWh	Tek	Land	Country
PRL loss (%)	3.1	1.9	2.3	5.6	1	1.6	0.5

5 Empirical Analysis

This section is devoted to the exploration of the main prediction by Helios et al. (2022), namely that conflict is more likely when groups are mismatched. We start by analyzing the relationship between the government-ethnic-group power mismatch and conflict incidence using the indicator of mismatch and an extended sample of African countries. Then, we restrict the focus of our analysis to the sample of 14 African countries for which we build the continuous measure of power mismatch. Finally, we conclude the section by providing additional results on the correlation between mismatch, conflict type, and conflict duration.

5.1 Evidence using the Mismatch Dummy

We start by studying whether ethnic groups that have a high power mismatch against their government are more likely to participate in a conflict against the government. For this purpose, we use an extended sample of 46 African countries. Our main independent variable is the mismatch dummy, an indicator that equals one if the group is mismatched. We define a group as mismatched if the relative political power of the group, computed using the status-power-rank variable, is in the 1st quartile and the relative military power is in the 4th quartile or vice versa. We investigate the relationship between mismatch and conflict using three different specifications that estimate the coefficient of interest using different sources of variation. In the first specification, reported in Panel A of Table 5, we regress conflict incidence on the mismatch dummy, different sets of controls, and a full set of country$\times$year fixed effects. Including country$\times$year fixed effects controls for all time-varying country-level variables and country-level shocks that may simultaneously affect the probability of being in conflict and the military/political power of a group. Crucially, it also takes into account the fact that mismatch is always defined relative to the dominant group of a specific country.[22] Hence, the mismatch coefficient is identified by variation across ethnic groups in the same country and year.

In a second specification, reported in Panel B of Table 5, we leverage the panel dimension of the dataset and substitute country$\times$year fixed effects with ethnic group and year fixed effects, thereby using only within-group variation for identification of the mismatch coefficient. While ethnic group fixed effects absorb all time-invariant characteristics of the groups, including them reduces sample size as a group needs to have variation both in the mismatch dummy and in the incidence variable to contribute to identification.[23] [24] Finally, Panel C of Table 5 reports the results of an extremely demanding specification in which we include both country$\times$year fixed effects and ethnic group fixed effects.

Moving across the columns of Table 5, we add different sets of group-level controls to the specifications. In column (1), no control (except for the set of fixed effects) is included. In column (2) we add a variable that measures the number of years since the last conflict. While, in column 3, we include controls for the group's cross-border relationship. Specifically, using the Ethnic Power Relations Transborder Ethnic Kin Dataset (Vogt et al., 2015), we build a measure of the population of co-ethnic groups that are in power in other countries. The idea is that the size of groups of the same ethnicity in power in different countries could impact the political and military power as they may put pressure on the

[22] Indeed, both the military power and the political power measures are relative with respect to the government power.

[23] This means that groups that are always/never in conflict and groups that always have high/low mismatch are not used in the estimation of the coefficient of interest.

[24] In our sample, the definition of politically relevant ethnic groups is time-varying. Indeed one-third of the ethnic group codes in the sample are observed for less than the whole sample period. Including ethnic group fixed effect helps taking into account the fact that some groups might become irrelevant or might merge with other groups.

Table 5: Power Mismatch and Conflict Incidence

Dependent Variable:	Conflict incidence						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A – Country×Year Fixed Effects							
Mismatch Dummy	0.0722***	0.0584***	0.0705***	0.0697***	0.0526**	0.0731***	0.0689***
	(0.0207)	(0.0163)	(0.0210)	(0.0199)	(0.0230)	(0.0257)	(0.0224)
Observations	4,425	4,425	4,170	4,425	3,640	2,716	2,707
R-squared	0.459	0.600	0.497	0.485	0.523	0.562	0.690
Panel B – Rebel group & Year Fixed Effects							
Mismatch Dummy	0.138**	0.134**	0.141**	0.134**	0.138**	0.156*	0.143*
	(0.0654)	(0.0612)	(0.0689)	(0.0659)	(0.0697)	(0.0844)	(0.0778)
Observations	4,530	4,530	4,274	4,530	3,786	2,914	2,895
R-squared	0.490	0.535	0.518	0.491	0.506	0.535	0.585
Panel C – Rebel group & Country×Year Fixed Effects							
Mismatch Dummy	0.0884***	0.0599**	0.111	0.0834***	0.116***	0.123***	0.0692*
	(0.0313)	(0.0294)	(0.0837)	(0.0308)	(0.0346)	(0.0422)	(0.0382)
Observations	4,424	4,424	4,168	4,424	3,639	2,712	2,703
R-squared	0.644	0.700	0.686	0.645	0.691	0.700	0.763
Controls							
Peace years		✓					✓
Family			✓				✓
Natural resources				✓			✓
Geographic					✓		✓
Socio-ecomic						✓	✓

Note: The dependent variable is conflict incidence. Panel A reports specifications with country×year fixed effects, Panel B with ethnic group and year fixed effects, and Panel C with country×year and ethnic group fixed effects. Control variables are defined as follows. *Peace years* is the number of years since the last time the group participate in a conflict; *Family controls* are: the population of kin groups that is in power (log) and dummy variables that indicate whether groups in the family had an upgrade or a downgrade in their power rank over the previous two years. *Natural resources controls* are: dummy variables for the presence of gold, diamonds, and other precious gems; a dummy variable that indicates active oil extraction, the share of the ethnic group's homeland covered with an active oil or gas field. *Geographic controls* are: distance from the capital, distance from the closest border, area of the group homeland, and mean and standard deviation of the elevation. *Socio-economic controls* are: 1990 share of group's homeland used for agriculture, 1990 share used for pasture, 1990 share of group homeland that is urbanized, 1990 population (log), 1990 nighttime luminosity (log), and group inequality measured as $ln[(group\, nightlight\, per\, capita)/(government\, nightlight\, per\, capita)]^2$. In all specifications in Panel A and C, standard errors are clustered at the country-year level. In the specifications of panel B, standard errors are clustered at the ethnic-group level. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: One-sided Mismatch and Conflict Incidence

Dependent variable:	Conflict incidence		
	(1)	(2)	(3)
Mismatch Dummy Military	0.0769**	0.282**	0.151**
	(0.0332)	(0.138)	(0.0669)
Mismatch Dummy Political	0.0613**	0.0146	-0.0681
	(0.0279)	(0.0428)	(0.0477)
Observations	2,707	2,895	2,703
R-squared	0.690	0.588	0.771
Controls	✓	✓	✓
Country×year FE	✓		
Year FE		✓	
Rebel group FE		✓	✓

Note: The dependent variable is conflict incidence. Column (1) reports specification with country×year fixed effects, column (2) that with ethnic group and year fixed effects, and column (3) the specification with country×year and ethnic group fixed effects. All specifications include the full set of group-level controls (see notes to Table 5. Standard errors are clustered at the country-year level in columns (1) and (3) and at the ethnic-group level in column (2). *** p<0.01, ** p<0.05, * p<0.1.

government or provide logistic/material support during a conflict. In the same spirit, we add two dummy variables that signal whether groups in the family had an upgrade or a downgrade in their power rank over the previous two years.

Natural resources are arguably an important determinant of civil conflicts and could also affect the degree of political/economic power of the group sitting on the resources. To take into account these factors, we augment the specification by adding controls for the availability of natural resources at the group level (column 4). In particular, we add dummies for the presence of gold veins, diamond mines and other precious gems, and petroleum. Following Morelli and Rohner (2015), we combine georeferenced information on oil and gas fields from PETRODATA with the location of the ethnic groups to build a measure of the percentage of a group's land that is covered by an active oil or gas field.

Other factors that may impact both the decision of participating in a conflict and the allocation of power might be related to geographic characteristics (take as an example secessionist conflicts, or rebellions for regional autonomy as in Esteban et al., 2022). Hence in column (5) we repeat the exercise by adding geographic controls (log distance from the capital, log distance to the closest border, log area of the group homeland, and the mean and standard deviation of the elevation of the group homeland).

Finally, when building the mismatch variable, we primarily focus on the military and political dimensions of power. However, economic power could well be an important determinant of a group's opportunity-cost of conflict. Accordingly, in column (6), we control for economic conditions in the pre-sample period. Specifically, we add log population and log nightlight luminosity as proxies for a group GDP, and various controls for land use[25]. Moreover, we build a variable that captures the group economic inequality vis à vis the dominant group as in Cederman et al. (2010).[26] We complete the exercise by reporting a specification with all the controls (column 7).

In all specifications, the mismatch dummy is positively related to the probability of being in conflict, and—with the only exception of column (3) of Panel C—the relationship is always statistically significant. This means that, within the same country-year, high-mismatched groups have a higher probability of partaking in a conflict than groups characterized by low mismatch (Panel A). Moreover, as shown in Panel B and C, if a group becomes mismatched, the likelihood of taking part in a conflict increases.[27] The magnitude of the effect is sizable: the probability of being in conflict is 7 percentage points higher for high-mismatch groups compared to low-mismatch ones, indicating that a mismatched group is at least 35% more likely to partake in a conflict against the government than a similar group that is not mismatched.[28]

We now investigate whether the effect of power mismatch is symmetric, regardless of whether military power is larger or smaller than political power. To do so, we split the mismatch dummy into two separate indicators: *Mismatched dummy Military* equals one if military power is in the 4^th quartile of the distribution and political power is in the 1^st quartile; *Mismatched dummy Political* equals one if political power is in the 4^th quartile of the distribution and military power is in the 1^st quartile. Table 6 collects the results of this analysis. Column (1) reports the specification with country×year fixed effects, column (2) that with ethnic group and year fixed effects, and column (3) the specification with country×year and ethnic group fixed effects. In all specifications, we include the full set of group-level controls. The results suggest that, when comparing ethnic groups of the same country and year (column 1), the effect of power mismatch is symmetric. Indeed the coefficients of the two dummy variables are positive, precisely estimated, and not statistically different from each other: high-mismatch groups are more likely to experience conflict compared to low-mismatch groups irrespective of whether military power exceeds political power or vice versa. When exploiting only within-group variation (columns 2-3), though, only the dummy indicating high military power and low political power is significant. This suggests that the likelihood that a group partakes in a conflict rises when its military power increases without being balanced by an increase in political power.[29]

[25]Share of the group land devoted to agriculture, pasture, the share of urban land, all computed in 1990.

[26]The group inequality vis-à-vis the government group is defined as $[\ln(nightlight_{e,t}/nightlight_{g,t})]^2$.

[27]The number of observations changes between Panel A and B because there are a few country-year with only one ethnic group other than the government group. This means that when including country×year fixed effects, those observations drop from the sample.

[28]The average value of conflict incidence for low-mismatched groups is 0.173.

[29]The effect is not driven by the fact that there is little within-group variation in the *Mismatch Dummy Political* variable. In fact, for both dummy variables, approximately 10% of the sample switches from low to high mismatch.

5.2 Results with a continuous measure of mismatch

Thus far we used a dummy proxy of power mismatch. We now exploit the ethnicity of cabinet members for a selected sample of 14 African countries to build a continuous measure of relative political power and construct an empirical counterpart of the theoretical definition of mismatch, $M = |m - p|$. This allows us to revisit the results presented in the previous section by analyzing the impact of marginal changes in the imbalance between the two dimensions of power. Moreover, it allows us to test whether the relationship between power mismatch and conflict is linear.

We start by repeating the specification in column (7) of Table 5 replacing the *Mismatch dummy* variable with the continuous variable *Mismatch*. Table 7 reports the results. In column (1) we introduce the continuous measure of mismatch in a specification with the full set of controls and country×year fixed effects. The coefficient is positive and significant at the 5% level, and its magnitude suggests that a one standard deviation increase in mismatch (0.16) corresponds to an increase in the likelihood of participating in a conflict against the government by 2.7 percentage points (approximately 25% of the sample mean). Column (2) reports the results of the estimation of a specification with time-varying group level controls, ethnic group, and year fixed effects. The coefficient on the continuous variable of mismatch is still positive and statistically significant, and the magnitude of the effect is remarkably similar to that of the previous specification. In fact, increasing the mismatch of a group by one within-group standard deviation (0.10) raises the probability of conflict participation by 3 percentage points. In column (3), we present the result of the most demanding specification, simultaneously including both country×year and rebel group fixed effects. The results are consistent with those of the previous specifications, albeit less statistically significant.

In columns (4)-(6), we revisit the results of Table 6. Here, to gauge whether power mismatch has different effects if military power exceeds or falls short of political power, we augment the baseline specification by including an indicator that takes value one if the group's military power is greater than its political power ($\mathbb{I}\{m > p\}$) and its interaction with the Mismatch variable ($M \times \mathbb{I}\{m > p\}$). Again, we conduct this exercise using the three specifications with country×year fixed effects (column 4), ethnic group and year fixed effects (column 5), and country×year and ethnic group fixed effects (column 6). The coefficients on power mismatch are always positive and precisely estimated, while the coefficients on the interaction term are negative and generally not statistically different from zero.[30] This result seems at odds with that of Table 6 and suggests that what matters for conflict is the imbalance between different dimensions of power, not the direction of the imbalance.

Our simple theoretical framework implies that, for any realization of costs around the mean, only sufficiently high mismatches determine a rational incentive to attack. Then, a second question we can ask, taking advantage of the continuous measure of mismatch is whether the effect of mismatch on conflict is linear. To answer this question, we start by adding a quadratic term in the baseline specification, *Mismatch squared*. Results are

[30]The coefficient on the interaction term is negative and statistically significant only when exploiting cross-group variation, and the significance is mainly driven by the ethnic group "Other Kivu groups" during the Second Congo Wars (1998-2003).

Table 7: Power Mismatch - Continuous Measure

Dependent variable:	Conflict incidence					
	(1)	(2)	(3)	(4)	(5)	(6)
Mismatch	0.170**	0.306***	0.168*	0.336***	0.351***	0.303**
	(0.0791)	(0.110)	(0.0955)	(0.115)	(0.131)	(0.122)
Mismatch$\times\mathbb{I}_{\{m>p\}}$				-0.341**	-0.0985	-0.243
				(0.160)	(0.137)	(0.182)
Observations	978	985	976	978	985	976
R-squared	0.633	0.544	0.735	0.638	0.545	0.731
Controls	✓	✓	✓	✓	✓	✓
Country-year FE	✓		✓	✓		✓
Year FE		✓			✓	
Group FE		✓	✓		✓	✓

Note: The dependent variable is conflict incidence. Columns (1) and (4) report specifications with country$\times$year fixed effects, columns (2) and (5) specifications with ethnic group and year fixed effects, and columns (3) and (6) the specification with country$\times$year and ethnic group fixed effects. All specifications include the full set of group-level controls (see notes to Table 5. $\mathbb{I}\{m > p\}$ is a dummy variable that equals 1 if military power is greater than political power. Standard errors are clustered at the country-year level in columns (1), (3), (4) and(6), and at the ethnic-group level in column (2) and (5). *** p<0.01, ** p<0.05, * p<0.1.

collected in Table 8. As before, we report the estimates of the coefficients obtained using different sources of variation. The coefficients on the two variables indicate that the relationship between mismatch and conflict is convex: the coefficient on the linear term loses significance, while the coefficient on the square of the mismatch is positive and statistically significant.[31]. This finding suggests that an increase in mismatch disproportionately increases the likelihood of participating in a conflict for groups that are already sufficiently mismatched.

We further show the non-linearity of the effect using non-parametric regressions. Figure6 visualizes the results of non-parametric regressions that uses third-order B-spline as the basis. Panel (a) reports the result of a simple non-parametric regression of conflict incidence on the mismatch variable. In panel (b) we condition the regression on country$\times$year fixed effect, in panel (c), instead, we include ethnic group fixed effects. The plots clearly show that the relationship between power mismatch and conflict is non-linear and that an increase in mismatch raises the likelihood of conflict only for values of mismatch above a certain threshold. This result is immediate if one returns to the intuition of the simple theoretical framework: for any realization of costs around the mean, only sufficiently high mismatches determine a rational incentive to attack, i.e., when the mismatch is small, the costs can outweigh, in expectation, the benefits of conflict.

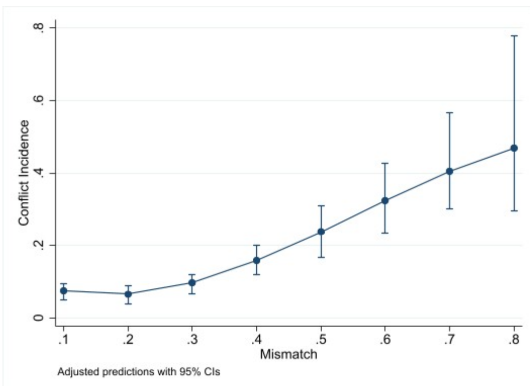
[31] In column (1) the coefficient on the square of the mismatch is only marginally insignificant, the p-value is 0.106.

Table 8: Non-linearity of the effect

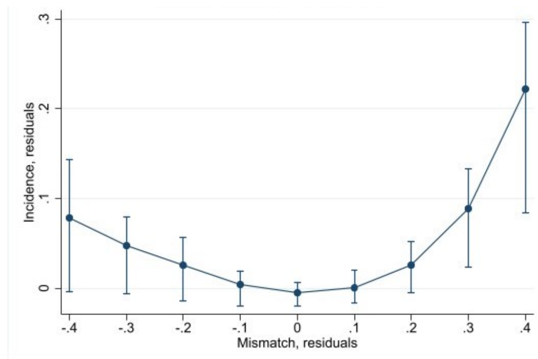
Dependent Variable:	Conflict incidence		
	(1)	(2)	(3)
Mismatch	-0.210	-0.109	-0.334
	(0.150)	(0.196)	(0.225)
Mismatch squared	0.730	0.743**	0.941**
	(0.449)	(0.298)	(0.450)
Observations	992	999	990
R-squared	0.649	0.559	0.751
Controls	✓	✓	✓
Country-year FE	✓		✓
Year FE		✓	
Group FE		✓	✓

Note: The dependent variable is conflict incidence. Column (1) reports specification with country×year fixed effects, column (2) that with ethnic group and year fixed effects, and column (3) the specification with country×year and ethnic group fixed effects. All specifications include the full set of group-level controls (see notes to Table 5. Standard errors are clustered at the country-year level in columns (1) and (3) and at the ethnic-group level in column (2). *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

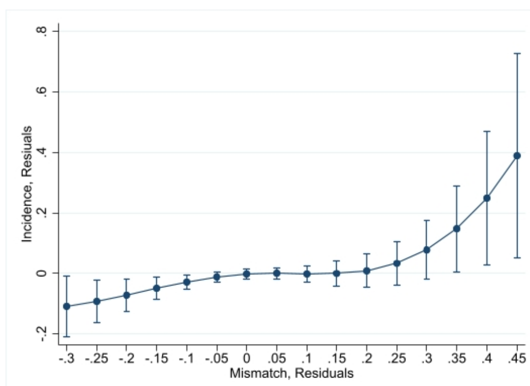
Figure 6: Non-Parametric regressions



(a) No controls



(b) Residuals after controlling for country×year fixed effects



(c) Residuals after controlling for ethnic-group fixed effects

5.3 Further results

In the previous sections, we tested the main aspect of the theory of power wars, showing a positive (and convex) relationship between power mismatch and the probability of participating in a conflict. We now explore whether power mismatch is also associated with different conflict characteristics. In particular, we focus on the size of a conflict, its duration, and on the type of grievance underlying the conflict.

First, we divide conflicts based on the type of incompatibility underlying them. The GED dataset contains a variable that categorizes civil conflicts in two types: *territorial* (i.e., the incompatibility concerns the status of a territory, secession, or autonomy) and *centrist* (i.e., the incompatibility concerns the type of political system, the replacement of the central government, or the change of its composition). We use this information to build two new measures of conflict incidence that we use as the dependent variable in the analysis. Second, we leverage the completeness of the GED dataset—not limited to civil wars, but also containing smaller conflicts— and ask whether power mismatch has different implications for major and minor conflicts. We define a conflict as *big* if the yearly average number of casualties is above 25.[32] Again, we use as dependent variable conflict incidence for big and small conflicts separately.[33]

In presenting the results, which are collected in Table 9, we replicate the structure in panels of Table 5 and report the results for the specifications with country×year fixed effects in Panel A, ethnic group and year fixed effects in Panel B, and country×year and ethnic group fixed effects in Panel C. All specifications include the full set of group-level controls used in column (7) of Table 5. The estimates in columns (1)-(4) are obtained using the extended sample of African countries and the indicator *Mismatch dummy* while columns (5)-(8) report results for the restricted sample and the continuous variable *mismatch*. The results show that power mismatch is positively correlated to conflict incidence only when considering centrist conflicts, regardless of the sample, the measure of mismatch, and the variation used to identify the coefficients. This finding is in line with the theory by Esteban et al. (2020), which highlights how dimensions such as cultural and religious group identity increase the group preferences for autonomy and may lead to territorial conflict even in the absence of a substantial power mismatch.[34] Results on the relationship between mismatch and the size of conflicts are less sharp but indicate that power mismatch is more important for big conflicts than for small conflicts. This suggests that, if the mismatch is high and cannot be bargained away, resolving the grievance will entail big and probably long conflicts.

[32]For instance, assume that a conflict lasts 3 years. In the first year, there are 10 casualties reported, in the second 20 casualties, and in the third 100 casualties, then the average number of casualties is 43.3 and the conflict is considered "big". Following the convention of UCDP/PRIO, we chose 25 as a cutoff.

[33]The sample used in the regressions for "big" conflicts is different from that used in the regressions for "small" conflicts. This is because groups that experience a big conflict are not considered in the small conflict analysis and vice versa.

[34]The results also seem reasonable when considering the dimension of political power considered in the mismatch variable. Both when using the EPR index and the continuous mismatch variable, political power is intended as participation in the decision-making process at the central level.

Table 9: Types of conflict

Conflict type:	Centrist	Territorial	Big	Small	Centrist	Territorial	Big	Small
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A – Country×Year Fixed Effects								
Mismatch Dummy	0.107***	-0.0346	0.0745***	0.00514				
	(0.0197)	(0.0228)	(0.0224)	(0.0228)				
Mismatch					0.150*	0.0571*	0.151*	0.0442
					(0.0803)	(0.0335)	(0.0824)	(0.0585)
Observations	2,525	2,309	2,531	2,337	966	879	947	903
R-squared	0.674	0.679	0.654	0.568	0.659	0.365	0.653	0.444
Panel B – Rebel group & Year Fixed Effects								
Mismatch Dummy	0.185***	-0.00958	0.101**	0.0524				
	(0.0418)	(0.0343)	(0.0394)	(0.0386)				
Mismatch					0.279**	0.0332	0.213**	0.174***
					(0.110)	(0.0338)	(0.104)	(0.0618)
Observations	2,733	2,503	2,728	2,528	973	894	955	915
R-squared	0.580	0.649	0.594	0.424	0.552	0.309	0.535	0.307
Panel C – Rebel group & Country×Year Fixed Effects								
Mismatch Dummy	0.0965**	-0.0190	0.0469	-0.0232				
	(0.0391)	(0.0390)	(0.0385)	(0.0442)				
Mismatch					0.163**	0.0232	0.104	0.0972
					(0.0807)	(0.0312)	(0.0782)	(0.0629)
Constant	-0.0846	0.135	-0.149	0.886***	0.463	1.066***	0.194	1.499***
	(0.212)	(0.176)	(0.279)	(0.234)	(0.521)	(0.303)	(0.471)	(0.407)
Observations	2,521	2,307	2,527	2,335	964	879	945	903
R-squared	0.762	0.811	0.773	0.658	0.759	0.485	0.757	0.537

Note: The dependent variable is in the column headings. Panel A reports specifications with country×year fixed effects, Panel B with ethnic group and year fixed effects, and Panel C with country×year and ethnic group fixed effects. In all specifications in Panel A and C standard errors are clustered at the country-year level. All specifications include the full set of group-level controls (see notes to Table 5. In the specifications of panel B standard errors are clustered at the ethnic-group level. *** p<0.01, ** p<0.05, * p<0.1.

Finally, we conclude this section by exploring whether conflicts against the government last longer if the groups involved are mismatched. Indeed, Helios et al. (2022) predicts that power mismatch should increase not only the likelihood of conflict participation but also the expected duration of conflicts. For this purpose, we conduct a survival analysis, where conflict termination is considered our "failure" outcome (a conflict is considered concluded if there are two or more consecutive years of peace among the parties involved). To perform this analysis, we use the sample of conflicts against the government in GED that we could attribute to the African ethnic groups in EPR. For each conflict (now our unit of observation) we know the start and end years, and the mismatch of groups involved. Figure 7 reports the Kaplan-Meier survival estimates for high-mismatch and low-mismatch groups.[35] The horizontal axis reports the duration of conflicts (in years), while the vertical axis reports the probability of "surviving"—which in our context is the probability that the conflict continues. The pattern suggests that the probability that the conflict continues for another year is higher for conflicts that involve a highly mismatched group. That is, high-mismatch groups tend to be involved in conflicts that are longer than those of low-mismatch groups. Table 10 reports the results of two tests for the equality of survival functions, both tests reject the null hypothesis that the survival functions for high- and low-mismatch groups are the same, confirming the fact that longer conflicts tend to be fought by high-mismatch groups.[36] This result is in line with that on conflict "size" and with the fact that mismatched groups tend to be involved in few and long conflicts.

Overall, the findings in this section paint a consistent picture: groups with a high power mismatch with respect to the government participate in conflicts that are bigger, last longer, and whose underlying grievance concerns the division of central political power.

[35]High/low-mismatched groups are defined as in Section 5.1.

[36]The log-rank test is constructed by giving equal weights to the contribution of each failure time to the overall test statistic. The Fleming–Harrington test reported in Table 10 gives more weight to early failures. Both the test are stratified over countries (i.e., the expected events are computed separately for each country, and those results are averaged over strata).

Figure 7: Conflict Duration: Kaplan–Meier survival estimates

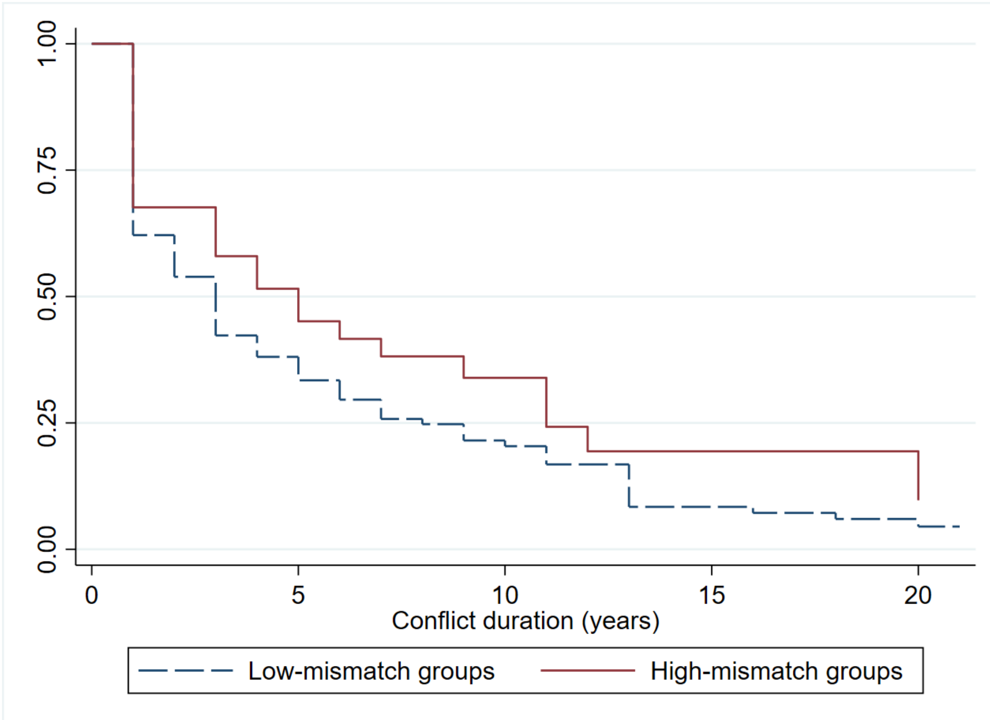


Table 10: Testing the Equality of Survival Functions

	Observed events	Expected events
Low mismatch	114	108.48
High mismatch	25	30.52
Stratified log-rank test	$\chi^2(1) = 3.59$ $Pr > \chi^2 = 0.058$	
Stratified Fleming–Harrington test (p=3, q=2)	$\chi^2(1) = 2.99$ $Pr > \chi^2 = 0.084$	

6 Conclusions

This paper makes two contributions to the literature and for future research use: it provides new measures for the study of ethnic conflict, filling an important gap, especially on the military power dimension, and it provides the first empirical analysis of the mismatch theory of power wars. About the first contribution, the use of machine learning techniques allowed us to provide a new measure of the relative military power of each ethnic group in the dyadic confrontation with the corresponding government controlling group, varying over time and context. This new measure performs well in predicting victory and allows us to improve on traditional proxies of military power used in the empirical literature on civil conflict.

As for the second contribution, theoretical work already suggested that absolute measures of military power are not enough to explain conflict participation and that researchers should focus on imbalances of different dimensions of power. In this paper, we take up the challenge to empirically show that power mismatch matters for conflict. Armed with the new measure of ethnic-group military power, we show the existence of a relationship between the likelihood of conflicts and the imbalance between relative military strength and relative political power. Exploiting data on civil (ethnic) conflicts, we find evidence that high-mismatch groups are approximately 30% more likely to take part in a conflict against their government. Moreover, these conflicts tend to be more deadly and to last longer than those where low-mismatch groups are involved.

Our findings, albeit robust to the inclusion of different sets of fixed effects and to the addition of many group-level controls, are mainly descriptive. However, they paint a picture that is extremely consistent with the theory of power wars (Helios et al., 2022). Hence, we believe that our evidence on the key role of power mismatch should encourage further research, both in the direction of precise identification and forecasting of future conflicts. In addition, our evidence bears an important policy implication: when trying to understand and prevent conflict outbreaks, one needs to pay attention to the imbalance between different dimensions of (relative) power. Focusing just on military strength or economic or political power may be misleading: militarily strong groups may not be those who start a war if they have enough political power; similarly, groups that are discriminated against may not pose a threat if they are militarily weak.

From a policy perspective, it could also be interesting to dig deeper into the origin of power mismatch and into the causes of its persistence. While a full-fledged analysis is beyond the scope of this paper, if we correlate power mismatch with country characteristics, we find that more democratic (polity 2 index), and ethnically and culturally more homogeneous countries, tend to have a larger share of high-mismatch groups. Contrary to popular belief, it seems that groups at high conflict risk reside in countries that are traditionally believed not to be at risk of conflict. This could be due to the winner-takes-all aspect of some electoral democracies, where a small difference in votes can be enough to completely exclude some (strong) groups from political power. In such cases, elections may not guarantee the credible elimination of a mismatch, and democracy may need to be supplemented by inventive institutional designs such as commitments in terms of public jobs, political roles, and military quotas, which favor power-sharing.

References

Breiman, Leo, "Stacked regressions," *Machine learning*, 1996, *24* (1), 49–64.

Bruk, Solomon I and VS Apenchenko, "Atlas narodov mira (Atlas of the World's Peoples)," *Moscow: Glavnoye Upravleniye Geodezii i Kartografii*, 1964.

Carroll, Robert J and Brenton Kenkel, "Prediction, proxies, and power," *American Journal of Political Science*, 2019, *63* (3), 577–593.

Cederman, Lars-Erik and Luc Girardin, "Beyond fractionalization: Mapping ethnicity onto nationalist insurgencies," *American Political science review*, 2007, *101* (1), 173–185.

_ , Andreas Wimmer, and Brian Min, "Why do ethnic groups rebel? New data and analysis," *World politics*, 2010, *62* (1), 87–119.

_ , Kristian Skrede Gleditsch, and Halvard Buhaug, *Inequality, grievances, and civil war*, Cambridge University Press, 2013.

Chandra, Kanchan and Steven Wilkinson, "Measuring the effect of "ethnicity"," *Comparative Political Studies*, 2008, *41* (4-5), 515–563.

der Laan, Mark J Van, Eric C Polley, and Alan E Hubbard, "Super learner," *Statistical applications in genetics and molecular biology*, 2007, *6* (1).

Eck, Kristine, "In data we trust? A comparison of UCDP GED and ACLED conflict events datasets," *Cooperation and Conflict*, 2012, *47* (1), 124–141.

Esteban, Joan, Laura Mayoral, and Debraj Ray, "Ethnicity and conflict: An empirical study," *American Economic Review*, 2012, *102* (4), 1310–42.

_ , Sabine Flamand, Massimo Morelli, and Dominic Rohner, "A dynamic theory of secessionist vs centrist conflict," 2020.

_ , _ , _ , and _ , "Stay, Split or Strike: Theory and Evidence on Secessionist vs Centrist Conflict," 2022.

Fearon, James D, "Rationalist explanations for war," *International organization*, 1995, *49* (3), 379–414.

Francois, Patrick, Ilia Rainer, and Francesco Trebbi, "How is power shared in Africa?," *Econometrica*, 2015, *83* (2), 465–503.

Girardin, Luc, Philipp Hunziker, Lars-Erik Cederman, Nils-Christian Bormann, and Manuel Vogt, "Growup-geographical research on war, unified platform," *ETH Zurich. http://growup. ethz. ch*, 2015.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Håvard Strand, "Armed conflict 1946-2001: A new dataset," *Journal of peace research*, 2002, *39* (5), 615–637.

Helios, Herrera, Massimo Morelli, and Salvatore Nunnari, "A theory of power wars," 2022.

Horowitz, Donald L, "Ethnic conflict management for policymakers," *Conflict and peace-making in multiethnic societies*, 1990, *115.*

_ , *Ethnic groups in conflict, updated edition with a new preface*, Univ of California Press, 2000.

_ , "Electoral systems and their goals," *Journal of Democracy*, 2003, *14* (4), 115–127.

Lijphart, Arend, *Democracy in plural societies: A comparative exploration*, Yale University Press, 1977.

Michalopoulos, Stelios and Elias Papaioannou, "The long-run effects of the scramble for Africa," *American Economic Review*, 2016, *106* (7), 1802–48.

Morelli, Massimo and Dominic Rohner, "Resource concentration and civil wars," *Journal of Development Economics*, 2015, *117*, 32–47.

Moscona, Jacob, Nathan Nunn, and James A Robinson, "Segmentary Lineage Organization and Conflict in Sub-Saharan Africa," *Econometrica*, 2020, *88* (5), 1999–2036.

O'Flynn, Ian and David Russell, "Power sharing: new challenges for divided societies," 2005.

Posner, Daniel N, "The political salience of cultural difference: Why Chewas and Tumbukas are allies in Zambia and adversaries in Malawi," *American Political Science Review*, 2004, *98* (4), 529–545.

Sarkees, Meredith Reid and Frank Wayman, *Resort to war: 1816-2007*, Cq Press, 2010.

Spears, Ian S, "Understanding inclusive peace agreements in Africa: the problems of sharing power," *Third World Quarterly*, 2000, *21* (1), 105–118.

Sundberg, Ralph and Erik Melander, "Introducing the UCDP georeferenced event dataset," *Journal of Peace Research*, 2013, *50* (4), 523–532.

Tibshirani, Robert and Jerome H Friedman, *The elements of statistical learning [electronic resource]: data mining, inference, and prediction: with 200 full-color illustrations*, Springer, 2001.

Tollefsen, Andreas Forø, Håvard Strand, and Halvard Buhaug, "PRIO-GRID: A unified spatial data structure," *Journal of Peace Research*, 2012, *49* (2), 363–374.

Vogt, Manuel, Nils-Christian Bormann, Seraina Rüegger, Lars-Erik Cederman, Philipp Hunziker, and Luc Girardin, "Integrating data on ethnicity, geography, and conflict: The ethnic power relations data set family," *Journal of Conflict Resolution*, 2015, *59* (7), 1327–1342.

Wucherpfennig, Julian, Nils B Weidmann, Luc Girardin, Lars-Erik Cederman, and Andreas Wimmer, "Politically relevant ethnic groups across space and time: Introducing the GeoEPR dataset," *Conflict Management and Peace Science*, 2011, *28* (5), 423–437.

A Detailed manual checks on multi-ethnic governments

This section provides the list of countries in our extended sample and the detailed manual checks on multi-ethnic governments. For each country Table A.1 reports the average number of ethnic groups in EPR over the sample period and whether the country also belongs to the restricted sample.

Table A.1: Country list

Country	Avg # of groups	Restricted sample	Country	Avg # of groups	Restricted sample
Algeria	1		Kuwait	2	
Angola	4		Liberia	4	✓
Benin	3	✓	Libya	3	
Botswana	9		Malawi	2	
Cameroon	5	✓	Mali	2	
Central African	3.69		Mauritania	2	
Chad	5		Morocco	2	
Congo	5	✓	Mozambique	2	
Congo, DRC	11.53	✓	Namibia	11	
Cote d'Ivoire	4	✓	Niger	3.96	
Djibouti	1		Nigeria	5	✓
Egypt	1.43		Saudi Arabia	3	
Equatorial Guine	4		Senegal	4	
Ethiopia	8.27		Sierra Leone	4	✓
Gabon	3.17	✓	South Africa	11.71	
Ghana	4	✓	Sudan	14.83	
Guinea	2	✓	Syria	4	
Guinea-Bissau	4		Tanzania	3.18	✓
Iran	10		Togo	1	
Iraq	4		Uganda	8	✓
Jordan	2		Zambia	4	
Kenya	6.80	✓	Zimbabwe	2	

Guinea 2009 The military coup by Capt. Camara, a Kpelle, seized control of the government after the death of the previous president. Kpelle is not a politically relevant ethnic group. EPR documentation states that "Generally, ethnicity still matters for national politics and the major political parties have easily identifiable ethnic bases". In this specific period, however, "all major ethnic groups are included in the cabinet leadership" (Malinke, Susu and Peul). It is, however, possible that the government was heavily influenced by the Malinke ethnic group since the two top positions, the military junta's second man and the prime minister, were Malinke. EPR classifies the three main ethnic groups as Senior Partners. However, further references [1] report attacks of the military against the Peul. Malinke and Peul have been historically clashing. It appears right to depart from EPR classification and assume that the Malinke were dominating the government in 2009. Further support for this is the Malinke dominating after 2009.

Liberia 1990-6: the First Liberian Civil War EPR classifies the country is in state collapse. The EPR documentation states that ”there is no central authority anymore and the state is unable to perform any of its *empirical functions* outside Monrovia”. They further state that ”there is no functioning central political power during this time and the country is ruled by different rebel groups, warlords, criminal gangs etc., the term *access to state power* becomes completely meaningless”. The EPR documentation does not describe the events. A brief account of the events can be found on Wikipedia (citing academic books and BBC articles) and is summarised below.

1. 1990: Invasion by Taylor (NPFL) overthrew president Doe. Taylor controlled most of the country (80%). NPFL fighters mainly from Gio and Mano ethnic groups of northern Liberia who were persecuted under Doe's regime.

2. NPFL efforts to capture the capital city of Monrovia were thwarted by the arrival of the Economic Community of West African States (ECOWAS) cease-fire monitoring group the Economic Community of West African States Monitoring Group (ECOMOG).

3. The NPFL instead set up in 1991 an alternative national administration away from the capital (the National Patriotic Reconstruction Assembly Government - NPRAG).

4. Johnson broke away from the NPFL and founded his own party INPFL. Johnson captured quickly the capital Monrovia. NPFL's power declined however after 1992.

5. in 1992 ECOMOG declared an Interim Government of National Unity (IGNU) with Amos Sawyer as their president, with the broad support of Johnson.

6. October 1992: Taylor launches an assault on Monrovia but was pushed back.

7. The interim government lasted until the next 1997 elections won by Taylor.

Liberia 2004-5 EPR states that all countries relevant ethnic groups are Senior Partners. In the aftermath of the Second Liberian Civil War 1999-2003, the country started its democratic transition under the National Transitional Government of Liberia until the 2005 democratic Election won by Ellen Johnson Sirleaf. EPR documentation states ”the cabinet posts and National Transitional Legislative Assembly seats were equally divided between the civil society (and other neutral political forces) and the warring factions (i.e. the corresponding ethnic groups)” and that ”Chairman of this power-sharing government was Gyude Bryant, a neutral politician”.

[1] https://www.refworld.org/docid/537db96b4.html

Sierra-Leone 1993-6 EPR states that the country is in state collapse. Military coup by Cap. Cap. Valentine Strasser in 1993 who was a Creole, a non-politically relevant ethnic group. According to EPR documentation "the constitution is suspended, and all political parties and activities are banned". There is civil war since 1991 with the major opposing rebel group being the Revolutionary United Front (RUF). Sources disagree on whether Strasser favored an ethnic group: there is information that Strasser favoured the Mende. EPR documentation doubts this based on other sources. It is also important to remark that according to many sources, RUF did not fight for a certain group or region (with claims for social justice and pan-Africanism). It seems that it would be hard to attribute a correct ethnic group to both the rebels and the government. For this reason, no ethnicity is attributed to the government following EPR.

Sierra Leone 1998-2002 EPR states that the country is in state collapse. New military coup against Kabbah (from the Mende group) and The Armed Forces Revolutionary Council (AFRC), allied with the RUF, is established as the new government. "The government even lacks an official army after 1998, being completely dependent on foreign peacekeepers and local militias". Kabbah (in power in 1997)i s reinstalled by ECOMOG forces in March 1998 but the civil war continues, with rebels reentering the capital Freetown in January 1999. EPR documentation states "The dramatic increase of warring parties and the shifting alliances completely blur the picture of who holds political power" and "the virtual loss of control by the central government and the totally nebulous situation regarding political power during this period of intensified civil war, this period is again coded as *state collapse*".

Congo DRC 2004-6 Ethnicity constantly played a role in Congo DRC politics. Joseph Kabila, the new president is more moderate than the previous one: had 4 vice-presidents each one representing a different faction. Kabila is supported by two related ethnic groups which are both coded as Senior Partners: Lunda-Yeke and Luba Shaba which have are related and form Kabila main power base. Here we consider them as allies.

Congo DRC 2007-12 Joseph Kabila is re-elected with less inclusive politics than his first term. Moreover, his power-base is unchanged and both Lunda-Yeke and Luba Shaba beeing Senior Partners and considered allies.

Kenya 2008-11 Ethnicity has been dominant in Kenyan politics: political parties are organised along ethnic identities. In December 2007, ethnic divisions turned violent following suspected fraud by the president during elections. An agreement, brokered by former U.N. Secretary Kofi Annan, was reached parties reached an agreement (brokered by former U.N. Secretary Kofi Annan) on a coalition government with equally shared cabinet reflecting Kenya's ethnic diversity. The two Senior partners are therefore considered allies.

Kenya 2012 The government did not change composition. Kenyan politics are however marked by a fight against the Somali ethnic group suspected to support the terrorist organisation Al-Shabaab.

Burundi 2002-5 The two politically relevant ethnic groups, Tutsi and Hutu, share government power "50-50" according to EPR documentation The Hutu hold a larger share in government posts which is compensated by the Tutsi's hold over the army. This equal sharing, after a historical hostility between the two ethnic groups, is the result of the Arusha Agreements in 2001 (not the same as the Arusha Agreements in Rwanda): a transitional peace treaty which brought the Burundian Civil War to an end.

Zambia 1966-2012 Zambian politics didn't involve ethnical conflicts and parties have been historically inclusive. From the EPR documentation: "leaders of Zambian parties have always attempted to appoint to significant position, members of diverse ethnic group, in the hope of increasing their share of national votes".

Madagascar 2002-12 According to the EPR documentation "from the literature surveyed there was no evidence that ethnicity is the basis for political discrimination". There is however "at least one interest group claiming to represent the interest of an ethnic group" until 2001. There is also evidence that ethnic tensions played a major role in "each of the major political transitions" until 2001. However, the political parties cannot be identified as representing a specific group. In 2002, Marc Ravalomanana is elected and his "ethnicity was eclipsed by his sense of nationalism and his call for a united Malagasy people". The two political ethnicities Côtiers and Highlanders are therefore coded as Irrelevant.

Comoros 2002-12 (included since they had conflicts) The inhabitants of each island are understood as different ethnic group. The three ethnicities corresponding to the three islands are Senior Partners in 2002. Each island has a president and they are vice-presidents in the Comorian Union government. The office of the president of the Union rotates among the three islands in 4-year terms.

B Military measure

B.1 Training set

This section describes the variables that we include in the training dataset. We collect all the ethnic level characteristics from the GrowUp Dataset (Girardin et al., 2015). First, we contain land characteristics, which include the ethnic territory's land area, the percentage area of the ethnic territory covered by mountainous, forest, barren lands, shrubland, urban area, water, grassland, and pasture land, as well as the area equipped for irrigation in the ethnic territory. We also contain a set of rich geographic characteristics, which include the spherical distances and travel time from the ethnic territory centroid to the border of the nearest land-contiguous neighboring country, and to the border of the nearest neighboring country (regardless of whether the nearest country is located across international waters), and to the border of the territorial outline of the country it belongs to, as well as to the national capital city in the corresponding country. Moreover, we contain rich socio-demographic characteristics, which include the population size of the ethnic group, population growth in the last 10 years, infant mortality rate, and prevalence of child malnutrition. Climate information is also captured by controlling the yearly total amount of precipitation and mean temperature. Besides, we include the number of affiliated ethnic groups outside the country considered and the number of affiliated ethnic groups outside the country considered in a position of power to capture kinship characteristics. Finally, we include external support information such as whether the group has ever received foreign support, whether the group has received support before the year of observation, and whether the group is currently receiving support. We also include a set of aggregate level variables, which include year, country latitude and longitude, and regional indicators.

B.2 Robustness checks on the parameters of the algorithm

We vary arbitrarily the parameters of the algorithm and show that the prediction is highly robust. We report both the relative difference in PRL and the correlation of the baseline military measure and the military measure using a different set S_i of parameters.

- S_1: the number of cross-validation folds is changed from 6 to 10;
- S_2: #(CV-folds) from 6 to 10; Random Forests' depth and span is increased; Random Forests' column sampling is also increased;
- S_3: #(CV-folds) from 6 to 10, Boosted Decision Tree's learning rate increased;
- S_4: Random Forests' column sampling is decreased;
- S_5: Random Forests' column sampling is increased; Boosted Decision Tree's learning rate increased; Boosted Decision Tree's total number of allowed trees decreased;
- S_6: in both Random Forests and Boosted Decision Tree the row sampling rate is increased.

Table B.2: Robustness check: algorithm's parameters.

Models	S_1	S_2	S_3	S_4	S_5	S_6
$\mathrm{corr}(m_{S_0}, m_{S_i})$	> 0.99	> 0.99	> 0.99	> 0.99	0.98	0.97
$\Delta_{\mathrm{PRL}}^{S_o,S_i}$	0.01	0.02	0.03	0.03	0.04	0.05

B.3 Military measure: machine learning algorithm

In this section, we describe in detail our machine learning algorithm and how to use it to predict the probability of winning for each ethnic group against the government.

We follow CK and train a machine learning algorithm to infer the dyadic probability of winning. We use a stacked ensemble learner which is a method that combines "multiple learning algorithms to obtain better predictive performance than could be obtained from any of the constituent learning algorithms"[2].

Stacking, or Super Learning, is a procedure that aims to find the optimal combination of prediction algorithms. Concretely, Stacking solves a main issue in the inference problem: when not knowing the exact form of the underlying distribution, stacking allows to combine several possible forms. Stacking uses cross-validation i.e. a random partition of the training set into n subsets of equal size. The procedure consists of generating n models, each one based on a different $\frac{n-1}{n}th$ fraction of the training set and the models are then used for prediction on the $\frac{1}{n}th$ fraction left. We can then obtain the cross-validated prediction for each of the N observations in out training set. Generally, the cross-validated error of the learner is simply the average error made on each of the N predictions. It is typically used for model selection among the considered learners when the available data is scarce enough to not be able to afford a proper testing set. The Stacking algorithm instead, uses the cross-validated prediction for each observation and for each learning algorithm.

Precisely a base learner is defined as a mapping from an N-observations dataset (Y, X) (predictors X and responses Y) to a function of the predictors. Given L base learners, we obtain by cross-validation a matrix $\mathcal{Z} \in \mathbb{M}^{N \times L}$ storing the predictions. The matrix $\mathcal{Z}$ and the original response vector $Y \in \mathbb{R}^N$ is the so-called "level-one data".

The meta-learning algorithm, then, combines the L base learning models to build a single learner. The ensemble model consists of the L base learning models and the meta-learning model. In practice, the true response vector Y is regressed on the meta-predictors $\mathcal{Z}$. The form of the meta-regression function $\Psi(\mathcal{Z}) = \mathbb{E}(Y|Z)$ is left as input by the user and the parameters of Ψ are chosen as to minimise a loss function: typically the cross-validated risk $\Psi = \arg\min_{\psi \in \mathcal{P}} \sum_i \left(Y_i - \psi(Z_i)\right)^2$ where $\mathcal{P}$ is the set of functions considered. If $\mathcal{P}$ allows a great deal of complexity, penalisation or further cross-validation can be used to avoid over-fitting on the cross-validated risk criteria[3]. A commonly used meta-regression function is a linear one with $\beta_l \geq 0$ and $\|\beta\|_1 = 1$, that is a weighted average of the predictions, the weights (to each model) being determined to minimise the cross-validated risk.

To wrap up, the stacked ensemble learner of Y for a value X is obtained by evaluating the meta-regression function Ψ on the L cross-validated predictions at X of the L base learning models. Van der Laan et al. (2007) show that the super learner will perform asymptotically as well as the best learner among all base learners. Therefore, it is clear that with an appropriate meta-regression function (eg. a simple linear function) the stacked ensemble learner outperforms each of the base learners. The functional form of the meta-

[2] h2o documentation accessible at https://docs.h2o.ai/h2o/latest-stable/h2o-docs/data-science/stacked-ensembles.html.

[3] for example, the method of the Lasso can be used if $\mathcal{P}$ is the space of multi-linear functions in X.

learner that we use is a generalised linear model, the recommended one by the h2o library. Empirical studies such as Breiman (1996) show that, generally, the more diversified are the base learners forms, the better the super learner performs. For this reason, our base models consist of three different families: *random forests*, *gradient boosting machines* and *generalised linear models* that we proceed to describe. More details on each single method can be found in Tibshirani and Friedman (2001).

B.3.1 Random Forests

A random forest is a tree based learning algorithm. A simple decision tree performs well in terms of training error with a low bias but it typically suffers from high variance i.e., it tends to over-fit the training set. Random forests consist in applying bagging, or bootstrap aggregation, to decision trees. The method relies on a simple insight from statistics: averaging a set of i.i.d. observations reduces the variance. To reduce variance and therefore increase the prediction accuracy, the algorithm generates multiple decision trees and the final prediction is an average of the predictions of all the decision trees. To effectively reduce the variance, the decision trees have to be sufficiently decorrelated. Two features of a random forest algorithm realise this task: (1) each decision tree is based on a different bootstrapped training dataset which is the core of bagging; (2) each time a split is considered in a tree (branching in the decision-tree building algorithm), only a random sample of m predictors is considered among all the predictors. In particular, (2) compromises the training error and increases the bias but greatly reduces the final variance by reducing the correlation between the trees.

It is recommended to use value of m that is close to $\sqrt{p}$ where p is the total number of predictors. A smaller value of m is used when the data set is made of a large number of correlated predictors. In our case, it is not the case and we tend to span m slightly above the recommended value. The number of trees in a Random forest algorithm is not a sensible parameter: if we allow for several trees, over-fitting does not occur. It is therefore recommended to use a sufficient number of trees such that the error rate stabilises. Another important parameter in Random forests, and as a matter of fact, in any tree based algorithm, is the depth of the tree. The depth is determined by the number of branchings allowed. Allowing a high depth naturally increases the complexity of each tree and of the resulting random forest. We do not allow for high depth to not over-fit the data since the number of observations in our training dataset is not very high.

B.3.2 Gradient Boosting Machine

A gradient boosting machine, like random forest, is a tree based learning algorithm. GBMs attempt to solve the same issue as Random Forests but in a radically different way. Instead of bagging, GBMs are based on boosting: trees are generated sequentially, each tree using the information provided by the previous ones. In that sense, GBM belongs to the forward leaning ensemble methods in that information flows only in the input-to-output direction. The boosting approach is based on the belief that learning slowly is preferable to straightforwardly fitting the data. The algorithm sequentially adds trees to the existing tree to reduce the residuals and better fit the data. Once a new tree is calculated in order to reduce the residuals, it is added in a "shrunken" version that is with an attenuating coefficient called the shrinkage parameter or learning rate. The residuals are then recalculated for the updated boosted model i.e. the shrunk sum of all trees. The smaller the shrinkage parameter is, the slower will the boosted model evolve and the lower will be the learning rate. The intuition behind the a slow learning rate is that it allows to slowly improve the boosted model in regions where it relatively mis-performs without perturbing dramatically the model. In that sense it is "safer" and generally leads to better results. A learning rate that is too small may however require many trees to converge. It is usual to set a lower maximal depth than in random forests since GBM is additive and the new tree growth takes into account the other trees that have already been built. In our case, we sacrifice computational time and use a small learning rate to avoid overfitting that we compensate with a high number of trees. The maximal depth of the trees allowed has been spanned from 1 to 5.

B.3.3 Generalised linear models

Generalised linear models are very commonly used and allow a flexible extension of linear models for responses that are not necessarily normally distributed. Given that we have a binary classification problem, we use a binomial model with a logit link function. To prevent over-fitting and increase the prediction accuracy of the statistical model, we span the Elastic net regularisation from pure Lasso to pure Ridge penalisation, which are described below.

- *Lasso* imposes a constraint on the coefficients of the form $\|\beta\|_1 \leq t$ for some t and consequently performs both regularisation and variable selection. It limits the magnitude of the coefficients multiplying the predictors such that only the most important predictors are kept in the model, the others being put to exactly 0. While the variable selection feature of the Lasso is an attracting property for interpretation, it can be a drawback in prediction: some variables having possibly an impact, albeit small, on the response are nonetheless rejected.
- *Ridge regularisation* imposes a constraint of the form $\|\beta\|_2 \leq t$ for some t and it is therefore very unlikely (with probability 0) that a coefficient is put to 0. Consequently ridge regression does not perform variable selection but only shrinkage.

Because each of the two types of constraints has advantages and drawbacks, it is interesting to use different mixtures of l^1 and l^2 penalisation. The resulting Elastic net regularisation still allows for variable selection but it is less "sharp" as the l^2 dominates the l^1 penalisation. In all cases, we let the algorithm search for the optimal magnitude of penalisation which is determined through cross-validation. The trade-off is evident: a low penalisation will cause high variance and decrease the prediction accuracy while a high penalisation will induce a high bias and decrease the training accuracy.