Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Power Mismatch and Civil Conflict: An Empirical Investigation

Massimo Morelli, Laura Ogliari, Long Hong

Working Paper n. 703

This Version: October 2023

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
http://www.igier.unibocconi.it

Power Mismatch and Civil Conflict: An Empirical Investigation*

Massimo Morelli † Laura Ogliari ‡ Long Hong §

October 10, 2023

Abstract

This paper empirically shows that the imbalance between an ethnic group's political and military power is crucial to understanding the likelihood that such 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 an ethnic group's 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 between 30% and 50% more likely to engage in a conflict against their government depending on the specification used. 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 and centrist conflicts. The policy implication is that power-sharing recommendations and institutional design policies for peace should consider primarily the reduction of power mismatches between relevant groups, rather than focusing exclusively on equalizing political power in isolation.

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

*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. We also wish to thank Matieu Cotonieur, Yannick Pengl, Dominic Rohner, Austin Wright, Ekaterina Zhuravskaya, the editor and anonymous referees 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.

†Bocconi University, IGIER, Baffi and CEPR; massimo.morelli@unibocconi.it

‡University of Bergamo & CEPR; laura.ogliari@unibg.it

§Ohio State University hong.1134@osu.edu

1 Introduction

When thinking about the causes of wars, the concept of *power* should come to mind. International relations scholars, however, almost exclusively refer to military power when studying conflicts, ignoring the interplay with economic and political power. By contrast, the policy debate on the role of power sharing for peace objectives focuses almost exclusively on balancing political power, but without attention to military power imbalances. In this paper, we provide new empirical evidence highlighting the importance of simultaneously accounting for multiple dimensions of power. Herrera et al. (2022) establish theoretically 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 the relative military power and relative political power of the main disputants are not aligned.[1] We contribute to the literature by providing the first empirical evidence that power mismatch matters for conflict. 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.

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 intrastate level, the role it plays in the decision to engage in conflicts.[2]

Using different specifications that account for many observed and unobserved confounding factors, we find that within the same country-year, high-mismatched groups have a much higher probability of partaking in a conflict than groups characterized by low mismatch.[3] The results are extremely robust, and the use of country$\times$year and ethnic group fixed effects provide a good baseline identification. Nonetheless, there are factors that do not allow us to interpret the result in a causal way. We describe the potential issues and provide evidence that they may not be empirically relevant in our context.

After providing the main results with appropriate robustness checks, we delve into a number of heterogeneity analyses, aimed at clarifying the mechanism and hence provid-

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

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

[3] As shown in section 5, a mismatched group is 30% more likely to partake in a conflict against the government than a similar group that is not mismatched.

ing indirect confidence in the results. When using a continuous measure of the mismatch, we observe a convex relationship between power mismatch and conflict: when the mismatch is small, a marginal increase in the mismatch is not associated with 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. Second, we show that mismatched groups have a higher chance of being involved in conflicts that are bigger (in terms of fatalities) and that concern power-sharing at the central level (centrist conflicts). The latter result is also consistent with additional theoretical findings in Esteban et al. (2022): 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. Finally we explore which country characteristics interact with the mismatch variable. We find suggestive evidence that an increase in the mismatch has greater effect in countries that are poorer, more unequal, more dependent on natural resources, and that are not full autocracies.

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 of committing 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" to capture the advantage conferred on a player by the existing political institutions—for example, the relative control of the political bodies governing the allocation of resources in peace.

Our paper speaks to the debate within the literature on power sharing. As discussed, for example, in Hartzell and Hoddie (2003) and Reilly (2012), the two opposed theories on how power sharing should be advocated to resolve or avoid conflicts are consociationalism and centripetalism: the former relates to proportional access to political power by all relevant groups, as in Lebanon; the latter refers to attempts to create multi-ethnic parties competing for power, as in Kenya, Indonesia, and Nigeria. Both theories, in any case, focus on the pros and cons of different types of distributions of political power alone. The whole debate must shift, we argue, in the direction of considering simultaneously all the relevant dimensions of power, not just political power. The most successful case of a power sharing agreement, the Good Friday Agreement of 1998, which was proposed to end conflict in Northern Ireland, was a case in which the power mismatch was indeed addressed, since deposition of weapons and access to political power and public sector jobs were part of the multidimensional deal – see e.g., O'Leary (2001). As far as the parallel literature on the usefulness of proportional representation as a peace-inducing electoral mechanism is concerned, – see e.g., Horowitz (1990, 2000, 2003, 2005) – 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 the proportion of voters can determine an implicit commitment to power sharing. However, even power sharing often fails since the form of a credible power-sharing agreement does not necessarily reflect the desirable elimination of a mismatch: for example, a pure democracy with a proportional electoral system could guarantee that a group with 30 percent of ethnic group voters gains 30 percent of political power, but if such a group has a probability of victory against the majority group that is much higher or much lower than 30 percent the mismatch is not eliminated.[4] 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, by 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 collection efforts on all fronts, discussing the relevant previous literature; In section 4 we zoom in 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 discusses the policy implications. Section 7 concludes. The appendix contains all the technical aspects and additional figures and tables.

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.[5] Let $p \equiv \frac{p_E}{p_E+p_G}$ denote the relative political power of E.[6] Finally, let m denote the probability of winning of E in the event 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, the payoff is $U_E = pS$; on the other hand, in the 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

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

[5] In Herrera et al. (2022), the focus is on showing that for any two players involved in a bilateral dispute are, the mismatch matters for war and duration. But most of the disputes that lead to a conflict involve governments (or at least one group holding power). Given that in the empirical analysis of this paper, we focus on the bilateral conflicts involving a government group, and given that these types of disputes are those ending in conflict with incomparably higher frequency, we limit attention to such pairs also in the brief sketch of the theory that we present in this section.

[6] 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.

with probability $F((m-p)S)$, and hence incentive to rebel increases with $(m-p)$, which represents the *mismatch*. It also clearly increases with the size of the divisible surplus.

Whenever players make decisions on the basis of an expected cost (rather than knowing their own cost of war), then the mismatches below such an expected cost do not lead to war, while the ones above the expected cost do. For this reason, we expect to find in the data an intuitive form of non linearity: if a given increase in mismatch happens starting from a status quo with low mismatch, the impact on risk of war should be low, because such a marginal increase makes the new mismatch still likely to remain below the expected cost. On the other hand, when the same marginal increase in the mismatch happens in a status quo with an already significant mismatch, the likelihood that the new mismatch is considered higher than the expected cost is higher. Hence the impact of a marginal increase in the mismatch should be expected to be higher in situations with an already high level of mismatch.

In the rare cases in which $p > m$, i.e., when an ethnic group has political power but is very weak militarily, the government may have an incentive to start a (repression) conflict if $c_G < [(1-m)-(1-p)]S$.[7] Conflicts 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.

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 us to distinguish initiation, and, moreover, both cases actually say the same thing in terms of the role of the mismatch. In sum:
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 Herrera 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 more mismatched groups are more likely to enter into conflicts, we need to know both military and political power, even for the groups that have never participated in a conflict. To 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 in conflict and how we measure group-level political and military power.

[7] Powell (2012,2013) and Debs and Monteiro (2014) argue that an additional reason for these types of wars could be the fear of future power shifts that would make the group that is currently weaker militarily a stronger one to repress in the future. We do not need to invoke these considerations on expectations for the future for our simple goal to establish the relevance of the mismatch as a reason for war.

3.1 Ethnic conflicts

Throughout the analysis, we focus on conflicts involving a government group (represented by at least one dominant ethnic group) and one or more ethnic opposition groups over the period 1992-2012. We restrict attention to conflicts involving a government-related group because, for such conflicts, we can build meaningful measures of relative political and military power for the two opposing sides.[8] We restrict our attention to conflicts occurring in Africa and the Middle East to make sure that ethnicity represents a salient cleavage.[9] We construct a dataset that records ethnic groups and conflict information by linking the UCDP Geo-referenced 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,[10] which contains information on conflict events from 1989 to 2015. The database has four key features that make it particularly suitable for our purposes. First, it classifies the type of violence, allowing us to identify conflicts in which one of the actors is a government, and the other is an organized rebel group. It also has separate coding for civil vs. interstate conflicts – where one of the actors belongs to a different country. Second, it provides detailed information on the precise location of conflict events, which will help us link actors from different databases. Third, UCDP-GED reports all the incidents that result in at least one direct associated death, which allows us to extend the analysis to include also small-scale conflicts. Finally, the database also provides estimates of fatalities separately for civilians and each side involved in the conflict. This feature—not available to the best of our knowledge in any other civil conflict database—is essential to estimate the military power measure, as discussed in Section 4.

Ethnic Power Relation We use the list of ethnic groups provided in the Ethnic Power Relations (EPR) Core dataset (version 2018.1.1),[11] 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 considered politically relevant if one or more significant political actors claim to represent their interests in the national political arena, or if the group's members are subjected to systematic and intentional discrimination in the realm of public politics. This feature of the database ensures that the groups analyzed are politically relevant and likely to represent organized actors who could make the decision to engage in a conflict.[12] It is worth

[8]On the contrary, in conflicts against civilians, one of the sides is not determined. In conflicts between two rebel groups, the relevant measure of political power is unclear.

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

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

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

[12]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

noting that, due to the dynamic political environment of a country, the politically relevant groups 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 to 7.[13] In particular, if a group rules alone, the group is either "monopolist" (rank 7) or "dominant" (rank 6). If groups share powers, they could be either "senior partners" (rank 5) or "junior partners" (rank 4). If a group is excluded from power, the group is either "self-excluded" (rank 3), "powerless" (rank 2), or "discriminated" (rank 1). Table B.2 shows the share of our observations in each power rank category.

We follow different steps to find the ethnicity of the government group in a given year. First, if a group rules alone, we label it a government group. Second, if more than one group has the highest power rank in a given year—this occurs for "senior partners"—we consult the EPR Atlas, which contains details of how the political rank was created, to determine whether the groups are allied. Appendix B provides a detailed discussion about our manual checks. If they are allied, we consider them both government groups and treat them as a single entity. If they are not allied, we determine who enjoys a larger advantage using external sources. If unsuccessful, we exclude them from the analysis (3.9% of the sample).

Rebel group ethnicity Our objective is to assign conflict against the government to ethnic groups in the EPR dataset. This requires linking actors of the GED dataset to Ethnic groups in EPR. There is no direct correspondence between actors in UCDP-GED and ethnic groups in EPR so, we apply a multi-step procedure to assign ethnicity to the UCDP-GED rebel groups. As a first step, we use the ACD2EPR (Version 2021) conversion table developed by Wucherpfennig et al. (2011) and Vogt et al. (2015), which integrates UCDP/PRIO Armed Conflict Dataset (Version 17.1) with EPR.[14] Out of the 369 rebel groups in our sample of countries, 78 can be matched directly through ACD2EPR (this amounts to 30.2% of the conflicts in our sample).

For the remaining rebel groups, we exploit the location of the conflicts they have been involved in to build a link with ethnic groups. We take all the conflict events (both against the government and against other rebel groups) which involve one rebel group.

We first build a link between rebel and ethnic groups, then, we assign the conflicts against the government that the rebel group has been involved in to the corresponding

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

[13]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.

[14]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.

ethnicity. As a first step, we overlay conflict events (UCDP-GED) on the geoEPR ethnic group's homeland polygons. Appendix Figure A.2 shows an illustration of this step. We start with 27,117 events in our main country list, reported in Table B.1. We specifically exclude all the conflicts where the exact location of the event is not known or coded (45.8%), drop all conflicts that are against civilians and foreign governments (25.3%), and discard all the conflict events outside the boundaries of any ethnic group (2.1%). We further exclude the events that happen in the homeland of government and irrelevant groups (6.3%) as they do not contain useful information to locate the homelands of rebel groups.[15] The sample restrictions leave us with 2,975 events. This leaves us with approximately 5,000 ethnicity-event observations which we use to match rebel groups to ethnic groups.

Second, we count the number of times a rebel group has a conflict event in the homeland of a particular ethnic group. Finally, we assign the ethnicity with the highest count to the rebel group. In some cases, there is a tie among the counts. We choose the homeland of the ethnic group with the highest number of fatalities. In rare cases, the fatalities are also tied. We break the tie by randomly choosing one.[16] Finally, having matched rebel groups to ethnic groups, we assign all the conflicts a rebel group participated in—including those occurring in the government homeland and in the homeland of irrelevant groups—to the corresponding ethnic group.[17]

To have a sense of the performance of this procedure, for the conflicts that are assigned directly through ACD2EPR, we can compare the correct rebel-to-ethnic-group match with that generated by the geo-matching procedure. If we do so, geo-matching can identify correctly 59 of the 78 matches (76%). Zooming in on the "bad matches", the procedure fails when there are few conflict events on which to base the geo-match. For this reason, we exclude from the sample matches that are based on fewer than 3 events. We further provide robustness checks where results are obtained using a sample that excludes cases where ethnicities are identified by fewer than 5 events, as well as results based only on the ACD2EPR matches.[18]

To give a simple but concrete example, Figure 1 shows a map of Liberia with all of its ethnic groups in EPR, which is geo-matched with all conflict events associated with the Liberians United for Reconciliation and Democracy (LURD) and Independent National Patriotic Front of Liberia (INPFL), respectively. Take LURD as an illustration of our geo-matching procedures. We first exclude all the events in the dominant group's homeland and the events associated with irrelevant groups. Using the remaining conflict events, we assign LURD to Mandingo because most of their events took place in that ethnic group's homeland.[19] In the case of INPFL, most of the conflict events happened in the dominant

[15]In EPR all groups that are considered irrelevant do not have information on the ethnic homeland polygons. Similarly, conflicts falling in the government homeland do not give useful information on the ethnicity of the rebel group, as we cannot assign to it the government ethnicity.

[16]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% are in line with the results from our method.

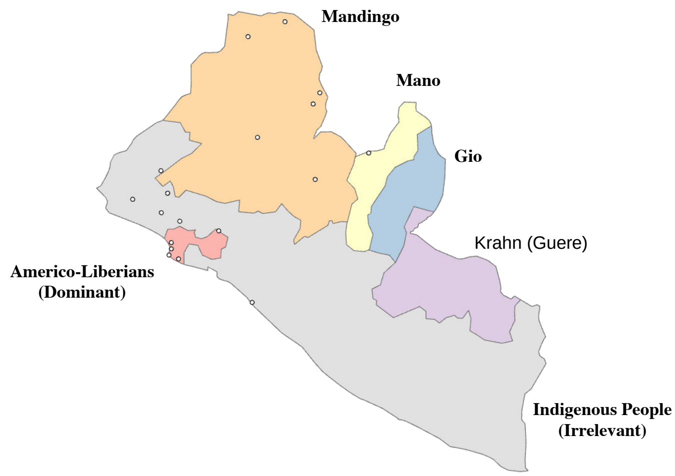
[17]Michalopoulos and Papaioannou (2016) and Moscona et al. (2020) use direct matching of battles with the ethnic group whose homeland contains the battle. Employing this direct match of battles to ethnicities in our event dataset would lead to inconsistent matches.

[18]As shown in Table 8 below, the results are close to our baseline results.

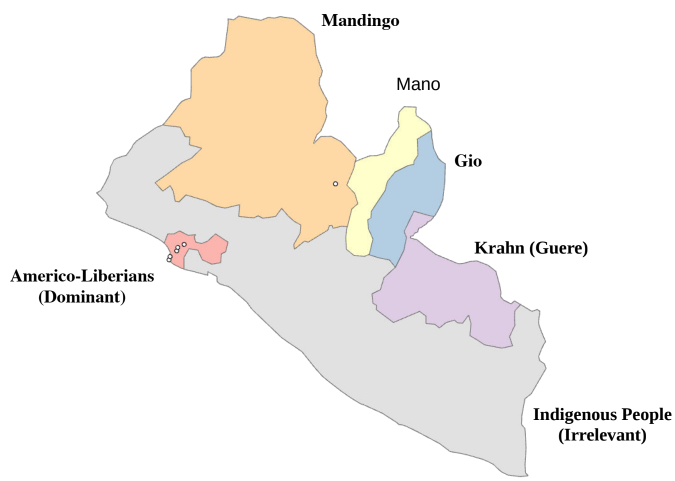
[19]We also include a similar map of the National Patriotic Front of Liberia (NPFL) in Appendix Figure A.3.

group's homeland. Using only one event for identification, we would incorrectly infer its ethnic group to be Mandingo, while ACD2EPR identified it as Gio.

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



(a) LURD conflict events



(b) INPFL conflict events

Notes: The figure represent the map of Liberia, where each colored polygon represents an ethnic group listed in EPR. The gray polygon is represents the area of groups coded as irrelevant. The dots represent all conflict events in UCDP-GED associated with the rebel group LURD or INPFL. Sources: GEO-EPR and UCDP-GED.

Key dependent variables Our primary goal is the analysis of civil conflicts between incumbent governments and (ethnic) rebel groups. Using the information on involved

We have correctly identified the rebel's ethnic group as Gio.

parties in UCDP-GED, we restrict our attention to conflict events where one of the actors is the government and the other is an organized rebel group. We aggregate the events at the ethnic group-year level using each rebel group's ethnicity. We use conflict *incidence* as the dependent variable, an indicator that equals one if the ethnic group is involved in a (at least one) conflict event in a given year.[20]

3.2 Relative political power

Building a measure of the political power of an ethnic group with respect to the government presents challenges. The literature used relative group size and the ethnicity of the leader or relied on expert opinions. We build two proxies of a group's 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" quantification 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 an ethnic group's 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 an ethnic group e to the government g as follows:

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

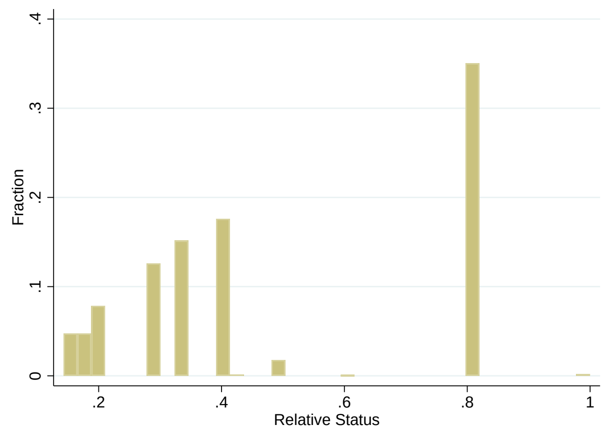
where C is the discrete power rank. Figure 2 shows the distribution of the measure. As expected, the distribution of p_{eg}^{PR} contains few points. Groups with high political power are basically those with $p_{eg}^{PR} = 0.8$, that is, groups that are considered junior partners in the government (i.e. their rank is 4 and the government group has rank 5). Groups with low political power are groups that are considered either powerless or that are discriminated against.

Continuous measure We collect data about cabinet membership in 14 countries in Sub-Saharan Africa over 21 years (from 1992 to 2012).[21] 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 works by Francois et al. (2015); Rainer and Trebbi (2014), these countries, while not democracies, are organized with some form

[20] We use all the conflict events – regardless of the location precision and whether it happened in the homeland of the government – but conflicts against civilians and foreign governments are excluded from our analysis.

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

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



Notes: The graph plots the frequencies of the relative power measure defined as the EPR power rank of the ethnic group over the EPR power rank of the group in power. Source: EPR Core Dataset.

of power sharing configuration, and the share of cabinet members can be considered as a valid measure of how political power is distributed among ethnic groups.[22]

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".[23] 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., a 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.[24]

The share of the cabinet seats 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

[22]Indeed, if you look the EPR power rank index for the government group of these countries, only 4 of them (Guinea, Nigeria, Sierra Leone, and Tanzania) have periods where the government group ruled alone (rank 6 - dominant), in all other cases the government group has rank 5 (senior partner). This is in line with Francois et al. (2015) claim that in these countries some form of power sharing is in place. Francois et al. (2015) and Rainer and Trebbi (2014) use a different, much finer, categorization of ethnicity from EPR.

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

[24]There are 2,557 raw events in the 14 countries of interest from 1992 to 2012. Among these relevant groups, there are 1,866 conflicts.

the cabinet seats. On the other hand, the powerless or discriminated groups tend to have negligible shares.

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

Notes: The table reports the average and the standard deviation of share of cabinet members by values of the EPR power rank index. Source: EPR Core Dataset and own data collection for the share of cabinet members.

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 ethnic group, g:

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

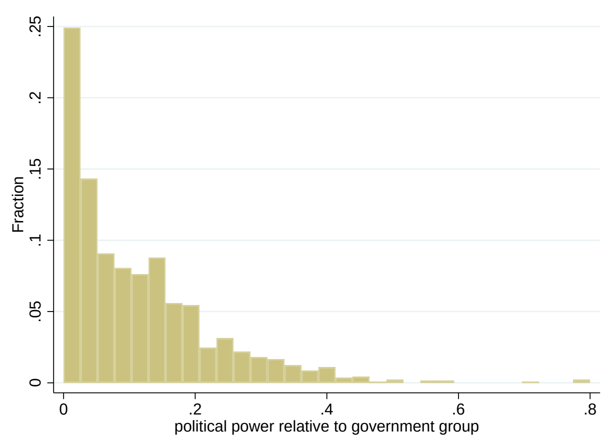
Figure 3 plots the distribution of $\hat{p}_{i,g,c}$. Unlike the discrete measure, the relative political measure now is quite smooth. It also contains a mass point at zero, which indicates that many groups are not represented in the cabinet, implying a certain degree of political power inequality in the restricted sample.

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 a 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 years of peace requires a non-trivial technique.

Figure 3: Histogram of relative political power



Notes: The graph plots the frequencies of the relative power measure defined as the number of cabinet members of ethnic group over the number of cabinet members of the group in power. Source: own data collection.

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.

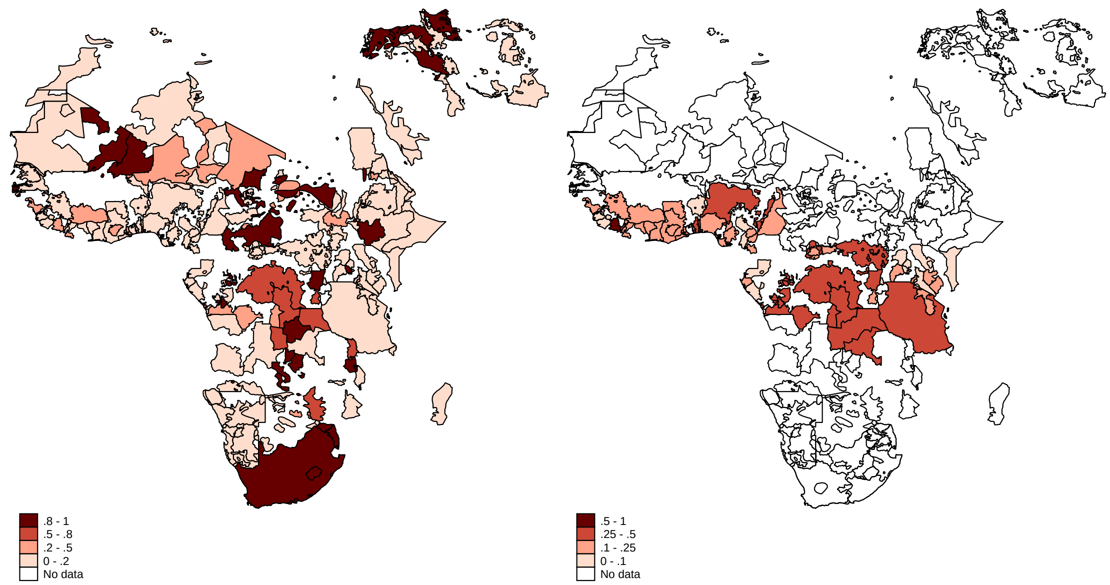
3.4 Mismatch measure

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. The mismatch dummy is an indicator that takes value one if political power is low and military power is high or vice versa. Specifically, we define the mismatch dummy as follows:

$$M^D_{e,g} = \begin{cases} 1 \;\; if \;\; (p^{PR}_{eg} \leq \bar{p}^{PR}_{p50} \wedge m_{eg} > \bar{m}_{p75}) \vee (p^{PR}_{eg} > \bar{p}^{PR}_{p50} \wedge m_{eg} \leq \bar{m}_{p25}) \\ 0 \;\; otherwise \end{cases}, \qquad (1)$$

Where $\bar{p}^{PR}_{p50}$ is the median of the distribution of relative political power computed using the EPR index, and $\bar{m}_{p25}$ and $\bar{m}_{p75}$ are the values of the 25$^{\text{th}}$ and 75$^{\text{th}}$ percentiles of the military power distribution. In other words, $M^D_{e,g}$ captures the presence of a high imbalance between the relative political power of a group and its relative military strength. Figure 4 panel (a) shows the share of years in the sample an ethnic group is mismatched according to this indicator.

Figure 4: Mismatch Measures



(a) Mismatch Dummy
Share of years with Mismatch Dummy=1

(b) Continous measure of Mismatch
Average value of the ethnic group

Notes: The figure plot the geographical distribution of the mismatched groups, darker colors mean higher mismatch. Panel (a) summarizes the mismatch dummy using the fraction of years in the sample in which an ethnic group is mismatched. Panel (b) summarizes the continuous measure of mismatch for the restricted sample using the average mismatch of the ethnic group. Sources: Polygons for the homeland of the ethnic group are sourced from the GEO-EPR dataset; data on the Mismatch Dummy are from authors' computation based on the EPR-Core Dataset; data on the continuous measure of mismatch are based on authors' data collection.

The continuous measure is defined as

$$M_{e,g} := |m_{e,g} - p_{e,g}| \tag{2}$$

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. Figure 4 panel (b) shows the spatial distribution of the average of the continuous measure of mismatch for the restricted sample. Figure A.1 in the appendix plots the within group standard deviation for the two variables.

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).[25] Control variables are grouped into five categories. These should account for possible determinants of war highlighted

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

by the literature on conflict. First, since groups that have experienced a recent conflict are more likely to participate in another conflict, we include information on the number of peace years and number of years in conflict of each ethnic group. Second, as the fates of co-ethnic groups may affect both conflict likelihood and power mismatch, we control for information about kinship relationship. Third, we include geographic characteristic such as land area, elevation, distance to capital and country borders. Forth, as natural resources may be linked both to conflict events and the political power of ethnic groups, we include information on gold and diamond veins, diamonds and gem mines and active oil production. Finally, in some specifications we include socio-economic controls: share of land devoted to agriculture, share of land devoted to pasture, population, urban population, nightlight density, and group inequality.

4 Predicting military power via machine learning

Our goal is to estimate the probability of winning for each ethnic group, against their corresponding government. However, when an ethnic group has never experienced a conflict, traditional methods cannot be used for estimation. According to Carroll and Kenkel (2019), the standard approach of using a linear probability model to predict the winning probability of a conflict may result in accuracy no better than a random guess. Therefore, we use a machine learning algorithm, adapted from theirs and incorporating a rich set of observed ethnic group-level variables, to compute the probability of winning for all ethnic rebel groups against their government in our sample. In the following sections, we describe our *training dataset* used to train our machine learning algorithm and discuss the algorithm and its performance.

4.1 Training set

In order to train our algorithm, we require a training set that includes plausible conflict outcomes – which side wins the conflict. However, this variable is typically not reported in standard databases such as GED. Therefore, we use a novel approach to determine the winning side of a conflict. We leverage the idea that the winning party should have a lower number of fatalities (relative to the population) than the losing party, and thus use the fatalities ratio as a measure of whether a group has won the battle. This approach is supported by the strong correlation between winning and having a smaller number of deaths (Sarkees and Wayman, 2010).

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.[26] We define the binary outcome Y, which is equal to 1 if the government wins and 0 otherwise:

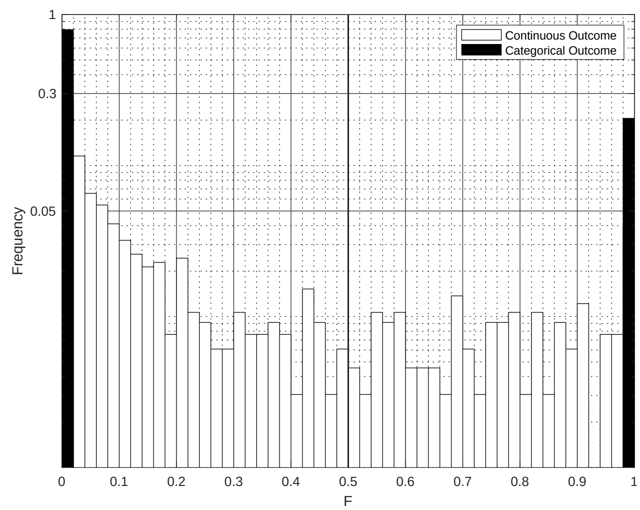
$$Y := \mathbb{I}\{F_{eg}^t = \frac{f_g^t}{f_g^t + f_e^t} \geq c\} \tag{3}$$

[26]We also included rebel-to-rebel conflicts in the training data, although it accounts for less than 5% of the observations.

That is, the ethnic group wins the conßict at time t if its share of fatalities is less than a threshold c.[27]

Figure 5 displays the distribution of F^t_{eg} in red. The graph shows two focal points at the two endpoints, indicating that the majority of fatality ratios are concentrated at zero or one. The remaining ratios are distributed (almost) uniformly. Based on this data pattern, we choose the threshold c = 0.5. However, the results remain robust even if we choose a different threshold.[28]

Figure 5: 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 binarization of the ratio according to eq. 3. The frequency is represented on a logarithmic scale for better visualization. Source: fatalities used to predict conßictsÕoutcomes are from UCDP-GED.

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 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 C.1. We adopt an advanced learning algorithm that conducts variables selection and cross-validation testing (more details in Appendix C.3).

Since machine learning methods are powerful when data points are dense, to train the algorithm, we use an augmented sample of African and Asian conßicts from 1989 to 2016. Moreover, since we want the fatality ratio to be sensible, we only consider conßicts

[27] We do not consider civilian deaths in the computation of the fatality ratio.

[28] An alternative dataset for documenting inter-state conßict outcomes is ACLED, where battle outcomes can be inferred by exploring whether control of territory has changed. However, using this method, we would categorize as stalemates more than 90% of the conßicts. Correlates of War (COW) also reports conßict outcomes, although the number of conßicts that match our sample is very limited. Using the matched conßicts, if we attribute victory to the side with fewer fatalities, we correctly guess 82% of the conßict outcomes.

with 25 or more fatalities in our training dataset.[29] In the final training sample, we have 657 observations with 198 predictors, where the share of outcomes where the government wins is around 79.5%.

Algorithm Formally, we use a binary learning model 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 (3). The optimal trained model $M^*_{(Y,X)}$ is obtained by minimizing a pre-defined loss metric L given our dataset (Y, X), as shown below.

$$M^*_{(Y,X)} = \arg\min_{M \in \mathcal{M}} L\left(Y - I\{M(X) > 0.5\}\right). \tag{4}$$

Intuitively, we look for 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 in all situations.

It is important to point out that what we label military power is the predicted probability that a group wins a conflict against the government. This implies: (i) the military power estimate is a "dyadic" measure, so it is best to compare it within a country; (ii) given the predictors we feed to the ML algorithm, what we are estimating is "potential" or structural military power, which differs from actual military personnel, military expenditure, quantity of armaments, etc.

Performance The performance of a binary classification model is commonly measured by the logarithmic loss metric. Given the classification model $M^*_{(Y,X)}$ obtained from equation (4), the log-loss is calculated as:

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

The smaller it is, the closer the likelihood of correct classification. We are interested in the cross-validated log-loss for 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 our prediction is compared to the null model. The PRL writes:

$$PRL\left(M^*_{(Y,X)}\right) = \frac{L_{null} - L\left(M^*_{(Y,X)}\right)}{L_{null}}.$$

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

Our reported PRL is comparable with the one obtained by CK for inter-state conflict. They obtain a PRL of 23%, similar to ours. Considering that we predict the outcome of

[29] Results remain qualitatively similar when we restrict the sample to countries in the extended sample or if we also consider small conflicts, but the predictive power of the algorithm is reduced.

intra-state conflicts, on which the data is very limited, the performance is quite good. Among the predictors that CK uses, important metrics include 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.

Compared with the conventional military power proxies used in the literature, population and night light have commonly been used as the main proxies for relative military power in intra-state conflicts (e.g., Esteban et al. (2012) use population). Table 2 shows that the population or night light ratio can predict the outcomes only marginally better than random guessing. Our algorithm, on the other hand, performs 27% better than random guessing based on the PRL, which is a considerable improvement on state of the art.[30]

Table 2: Algorithm's predictive power.

	Full model	Population ratio	Night light ratio
PRL	26.6%	0.5%	2.3%

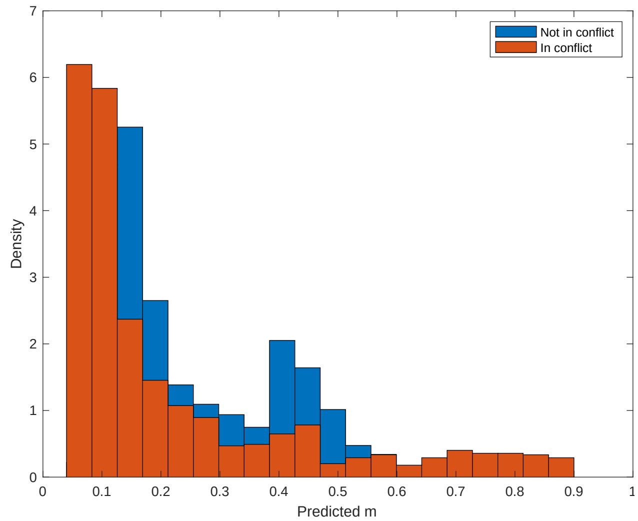
Notes: The table reports the performance of the machine learning procedure for the model with all the controls (Full model) and two simplified models that use as a predictor either the ratio between the population of the ethnic group and the population of the government group or the ratio between nightlight luminosity of the ethnic group and the nighlight luminosity of the government group. PRL is the proportional reduction in loss as described above. Sources: for conflict outcomes, UCDP-GED; for the predictors used in the machine learning algorithm, EPR Dataset Family and UCDP External Support Dataset.

Since groups that experience conflict are a selected sample, one concern is that the predicted military power of the groups in conflict (and we use it 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 6 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.

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

[30]Notice that even if the algorithm does not correctly predict who is more likely to win, by looking at fatalities the exercise might still be able to approximate how threatening the rebel group is for the government. The explanatory power of the variable could be even higher than what the validation would suggest.

Figure 6: Distribution of the predicted military power



Note: The graph presents the distributions of the predicted military power measure (i.e., the predicted probability of winning a conflict) for groups that experienced at least one conflict (in red) and for groups that are never in conflict (in blue). Source: for conflict outcomes, UCDP-GED; for the predictors used in the machine learning algorithm, Ethnic Power Relations (EPR) Dataset Family and UCDP External Support Dataset.

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.[31] Model *m2* assigns a loss if the fatality ratio is greater or equal to 0.58, and the prediction of this model is well correlated with the baseline prediction. In model *m3*, 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 model's performance and yields a prediction of military power that is slightly less well correlated with the baseline one. In model *m4*, the fatality ratio is built using cumulative fatalities (fatalities over the entire conflict duration) instead of the yearly fatalities, keeping the threshold fixed at 0.5. Again, cumulative fatalities yield predictions that are very similar to the baseline model. Appendix C.2 provides further results that show the robustness of the predictions to changes in the machine learning parameters.

4.2 Variable relevance and summary statistics

We can 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 challeng-

[31] 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	0.953	1.00		
$m3$ Small conflicts	0.824	0.715	1.00	
$m4$ Cumulative deaths	0.909	0.867	0.818	1.00

Notes: the table reports correlations among the predictions of military power obtained using different models. Specifically, model *m1* is the baseline model, model *m2* define classifies the outcome of conflict as a loss if the fatality threshold is above 0.58, moel *m3* includes in the training sample also conflicts with less than 25 fatalities, model *m4* uses the cumulative fatalities over the entire duration of the conflict as an outcome.Source: for conflict outcomes, UCDP-GED; for the predictors used in the machine learning algorithm, Ethnic Power Relations (EPR) Dataset Family and UCDP External Support Dataset.

ing to tell how much a single predictor affects military power. We explore the importance of a particular set of predictors using the following algorithm (following CK). We remove a predictor of interest, rerun the entire algorithm, 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 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). Overall, excluding one category of the predictors does not significantly change the PRL loss because the algorithm adjusts the parameters to other predictors for model fits. On the other hand, we show that external support and the geographic condition play a slightly more important role.

Table 4: Variables' predictive power

	Ext	Geo	Pop	PyWh	Tek	Land	Country
PRL loss (%)	0.94	0.68	0.23	0.23	0.08	3.31	0.56

Notes: The table reports percentage decrease in PRL we obtain eliminating groups of predictors from the model. Column headings identify the group of predictors we exclude from the model. *Ext* is the group of external support variables, *Geo* includes geographical variables, i.e., border distance, capital distance, and travel time, *Pop* includes population and demographic growth variables, *Py Wh* are peace years and war history, *Tek* variable pertaining to affiliated ethnic groups, *Land* includes land characteristics, and *Country* includes country-level variables, i.e., latitude, longitude and regional dummies. Source: UCDP-GED; Ethnic Power Relations (EPR) Dataset Family; UCDP External Support Dataset.

Finally, Table 5 provides summary statistics for the measure of the estimated military power. Observing the number, we see that the majority of the variation of the variable comes from differences across different ethnic groups of the same country year, while the within-group variation is quite low. For this reason, we always report in the analysis results exploiting both the within-group variation and variation within country-year.

5 Empirical Analysis

This section is devoted to the exploration of the main prediction by Herrera 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 in an extended sample of countries in Africa and the Middle East. [32] We combine these results with evidence obtained using a restricted sample of African countries for which we can construct a continuous measure of power mismatch. Finally, we conclude the section by providing additional results on the heterogeneous effects of power mismatch, focusing on conflict and country characteristics.

The main independent variable is the power mismatch as defined in eq. (1) for the extended sample and in eq. (2) for the restricted sample. We investigate the relationship between mismatch and conflict by estimating through Ordinary Least Squares the following model:

$$\text{Conflict incidence}_{ect} = \alpha_{ct} + \beta\,\text{Mismatch}_{ect} + X_{ect}\gamma + (\delta_e) + \varepsilon_{ect}, \tag{5}$$

where Conflict incidence is an indicator that takes value 1 if ethnic group e in country c in year t is taking part in a conflict; X_{ect} is a matrix of ethnic-group-level controls which have been highlighted by the literature as important determinants of conflicts; α_{ct} is a full set of country $\times$ year fixed effects and δ_e are ethnic group fixed effects. The inclusion of country $\times$ year fixed effects absorbs time-varying country characteristics—including the identity of the government group in power, the level of democracy, the level of ethnic fractionalization, GDP growth, inequality, dependency on natural resources, colonial origin, etc.—and ensures that the identification is achieved by exploiting variation across ethnic groups in the same country-year. The inclusion of group-level controls reduces omitted-variable-bias concerns, guaranteeing that the major confounders are accounted for. Moreover, the addition of ethnic-group fixed effects cleans for time-invariant unobserved characteristics

[32] The list of countries in the sample and the number of ethnic groups for each country is reported in Table B.1 in the Appendix.

Table 5: Military power

	Obs.	N. of groups	Mean	Median	Max	Min	sd overall	sd within country-year	sd within group
military power	8090	438	0.253	0.182	0.884	0.075	0.167	0.121	0.045

Notes: the table reports summary statistics for the military power measure computed, through the machine learning procedure, as the probability of winning of the rebel group against its government group. The statistic reported is described in the column headings. Sources: UCDP-GED; Ethnic Power Relations (EPR) Dataset Family; UCDP External Support Dataset.

at the ethnic-group level, effectively making the estimation strategy akin to a difference in difference, where we exploit variation across groups within a country-year and temporal variation within ethnic groups. Since the bulk of the variation of the estimated military power measure is across groups (78% in the extended sample and 60% in the restricted sample) in the paper we always report results of speciÞcations with and without ethnic-group Þxed effects.

The large set of Þxed effects and control variables account for many observed and unobserved characteristics that may generate a spurious correlation between mismatch and conßict incidence. At this point it is worth mentioning the remaining factors that may prevent a causal interpretation of our estimates. A reverse causality concern could be that groups that are excluded from power enter into conßict and, precisely because they are experiencing conßict, develop higher military power than groups that are not. A second concern is related to the forward-looking behavior of groups: groups with low political power may increase their military capabilities precisely because they want to enter into a conßict against their government in order to gain political power. A Þnal worry could arise concerning political power. A recent body of work (e.g., Deiwiks et al., 2012; Wimmer et al., 2009) emphasizes the role of political exclusion as a trigger for conßict. In our setting, this might be problematic if the mismatch variable is just capturing exclusion from political power. In Section 5.3 we discuss these concerns in more detail and provide evidence suggesting that these factors are not empirically relevant in our context.

5.1 Baseline results

We start off this section by presenting the results of the estimation of model (5) using the extended sample of 44 countries in Africa and the Middle East and the indicator of power mismatch Mismatch Dummy. The results are collected in Table 6. Moving across the columns of Table 6, we add different sets of group-level controls to the speciÞcations. In column (1), we report the result of a simple OLS regression of Conßict incidence on the mismatch dummy only controlling for country! year Þxed effects. This set of Þxed effects absorbs all time-varying country-level variables and country-level shocks that may simultaneously affect the probability of being in conßict and the military/ political power of a group. Crucially, it also takes into account the fact that mismatch is always deÞned relative to the dominant group of a speciÞc country.[33] Hence, the mismatch-dummy coefÞcient is identiÞed only by variation across ethnic groups in the same country and year. In column (2) we add a variable that measures the number of years since the last conßict and a variable that measures the number of years of conßict experienced by the ethnic group. In column 3, we include controls for the groupÕs cross-border relationships. SpeciÞcally, 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 government or provide logistic/ material support during a conßict. In the same spirit, we add

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

Table 6: Power Mismatch and Conflict Incidence

Dep. Var.: Conflict incidence	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mismatch dummy	0.0557***	0.0682***	0.0592***	0.0570***	0.0549***	0.0509**	0.0443**
	(0.0186)	(0.0132)	(0.0129)	(0.0127)	(0.0157)	(0.0231)	(0.0225)
Controls							
Peace years & war hist.		!	!	!	!	!	!
Family			!	!	!	!	!
Natural resources				!	!	!	!
Geographic					!	!	
Socio-ecomic						!	
Fixed Effects							
Country×year	!	!	!	!	!	!	!
Ethnic group							!
Mean dep. var.	0.114	0.114	0.114	0.114	0.131	0.143	0.114
R^2	0.333	0.479	0.665	0.669	0.690	0.716	0.767
Within R^2	0.00469	0.223	0.500	0.507	0.528	0.519	0.353
Obs.	4,220	4,220	4,220	4,220	3,572	2,645	4,220

Note: The dependent variable is conflict incidence. Control variables are defined as follows. *Peace years & war history* are the number of years since the last time the group participated in a conflict and the number of years of conflict for the group, respectively; *Family controls* are: a dummy variable indicating whether the family is politically relevant (EPR), the population of kin groups that are in power (log), dummy variables that indicate whether groups in the family had an upgrade or a downgrade in their power rank over the previous two/ten years, and an indicator of whether other kin groups in the family are currently participating in a conflict. *Natural resources controls* are: dummy variables for the presence of gold, diamonds, and other precious gems; dummy variables that indicates active oil or copper extraction. *Geographic controls* are: (log) distance from the capital, (log) distance from the closest border, (log) area of the group homeland, and (log) mean and standard deviation of the elevation. *Socio-economic controls* are: 1990 share of the group's homeland used for agriculture, 1990 share used for pasture, 1990 share of the group's 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 standard errors are clustered at the country-year level. *** p<0.01, ** p<0.05, * p<0.1. Sources: For conflict incidence peace years and war history, UCDP-GED; for the other control variables,Ethnic Power Relations (EPR) Dataset Family.

dummy variables that signal whether groups in the family had an upgrade or a downgrade in their power rank over the previous ten years or in the previous 2 years and an indicator that equals one if one of the kin groups in another country is participating in a conflict.[34]

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 of 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 productive gold veins, diamond mines and other precious gems, and petroleum. Other factors that may impact both the decision to start a conflict and the allocation of power might be related to geographic characteristics (for 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). 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 group GDP, and various controls for land use.[35] Moreover, we build a time-varying variable that captures the group economic inequality vis à vis the dominant group as in Cederman et al. (2010).[36] Even if economic inequality is a "bad control" (Angrist and Pischke, 2009, 2014), adding it to the specification is important as it shows that power mismatch has an effect on conflict that is independent from that of economic inequality. These controls are not available for all the observations in our sample, consequently their inclusion reduces the sample size by roughly one-third.

In all specifications, the mismatch dummy is positively related to the probability of being in conflict, and 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 against the government than groups characterized by low mismatch. A comparison of the mismatch-dummy coefficients across column (1)-(6) suggests that the magnitude of the effect is sizable: the probability of being in conflict is approximately 5 percentage points higher for high-mismatch groups compared to low-mismatch ones, indicating that a mismatched group is roughly 30-50% more likely to partake in a conflict against the government than a similar group that is not mismatched.[37]

In column (8) of Table 6, we leverage the panel dimension of the dataset and add ethnic-group fixed effects. Including this set of fixed effects reduces the effective number of observations used to identify the mismatch-dummy coefficient, as a group needs to

[34]We control for changes in the political power of kin groups in different temporal spans as it may take time for a group that seized power to settle and provide help and resources to a kin group member in a different country. Changing the temporal span does not have any impact on our results.

[35]these include: share of the group land devoted to agriculture, pasture, the share of urban land, all computed in 1990.

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

[37]The average value of conflict incidence for low-mismatched groups is 0.11.

have variation both in the mismatch dummy and in the incidence variable to contribute to identiÞcation.[38,39]. This speciÞcation is akin to a difference in differences model where we use for identiÞcation deviations of the mismatch variable both from the country-year average and from the individual average. The mismatch-dummy coefÞcient remains positive and signiÞcantly different from zero. The size of the coefÞcient indicates that a group that becomes mismatched has a probability of entering a conßict that is 4.4 percentage point higher (approximately 38%) than a group in its country and year whose mismatch does not change.

Thus far we have 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 deÞnition of mismatch, $M = |m \# p|$. This allows us to revisit the results presented above by analyzing the impact of marginal changes in the imbalance between the two dimensions of power.

For this analysis we make two changes: (i) the main explanatory variable is differentÑ and now is continuousÑ, (ii) the sample we use for the estimation is different. To understand whether sample selection is driving the results, for each model we report both a speciÞcation where Mismatch dummy is our main explanatory variable, and a speciÞcation where we replace the indicator with the continuous variable Mismatch cont.

Table 7 reports the results. Moving across columns we increase the number of control variables included in the model. SpeciÞcally, columns (1)-(2) only control for country! year Þxed effects; in columns (3) and (4) we introduce controls for peace years and war history, ethnic relations and natural resources; in column (4) and (5) we add controls for geographic and socio-economic characteristics, columns (7) and (8) also include ethnic-group Þxed effects. Looking at the mismatch dummy coefÞcients (odd-numbered columns), the magnitude of the effect is bigger than the one we found using the extended sample. This suggests that countries that are in the restricted sample behave similarly to those in the extended sample but the effect of power mismatch seems to be stronger in the restricted sample. Findings in the even-numbered columns, which exploit the continuous measure of mismatch, conÞrm the positive relationship between mismatch and conßict participation. Using the most demanding speciÞcation, column (8), results indicate that a one-standard-deviation increase in conßict mismatch (0.16) is associated with an increase of the probability of participating in a conßict by 2.8 percentage points (an increase that corresponds to approximately 35% of the sample mean).

While the correlation between mismatch dummy and conßict incidence is robust to the inclusion of different sets of controls and Þxed effects, the credibility of the estimates also hinges on the robustness of the matching procedure between ethnic and rebel groups. Table 8 collects the results of the estimation the baseline model (eq. 5) using samples where we eliminate potential Óbad matchesÓ. Columns (1)-(3) report the coefÞcients, standard

[38] This means that groups that are always/ never in conßict and groups that always have high/ low mismatch are not used in the estimation of the coefÞcient of interest.

[39] In our sample, the deÞnition 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 Þxed effects helps to take into account the fact that some groups might become irrelevant or might merge with other groups during the sample period.

Table 7: Power Mismatch - Continuous Measure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mismatch dummy	0.158***		0.138***		0.0695		0.112**	
	(0.0352)		(0.0343)		(0.0430)		(0.0456)	
Mismatch cont.		0.117		0.248***		0.292***		0.181**
		(0.0866)		(0.0711)		(0.0928)		(0.0739)
Controls								
Peace years & war hist.			!	!	!	!	!	!
Family			!	!	!	!	!	!
Natural resources			!	!	!	!	!	!
Geographic					!	!		
Socio-ecomic					!	!		
Fixed Effects								
Country×year	!	!	!	!	!	!	!	!
Ethnic group							!	!
Mean dep. var.	0.082	0.0818	0.082	0.082	0.082	0.082	0.0818	0.082
R^2	0.249	0.225	0.568	0.559	0.655	0.660	0.672	0.669
Within R^2	0.035	0.0032	0.444	0.433	0.545	0.552	0.312	0.307
Obs.	1,247	1,247	1,247	1,247	902	902	1,247	1,247

Note: The dependent variable is conflict incidence. See notes to Table 6 for the definitions of the groups of controls. Standard errors are clustered at the country-year level. *** p<0.01, ** p<0.05, * p<0.1. Sources: for conflict incidence, peace years and war history UCDP-GED; for the continuous measure of mismatch, authors' data collection on cabinet members of ethnic groups and estimates of military power from the machine learning procedure; for the other controls, Ethnic Power Relations (EPR) Dataset Family.

errors, R-squared, and number of observation for the model estimated in column (4) of Table 6, which includes country × year fixed effects, peace years as well as family and natural resources controls. In columns (4)-(6) are collected the results of the model in column (8) of Table 6, which adds to the specification ethnic-group fixed effects. We perform this sensitivity test for both the indicator (Panel A) and the continuous measures of mismatch (Panel B). The conditions that define a "bad match" are reported in the row headings. In general, we think of a match as potentially problematic if it is determined by few events. Specifically, in row (a) we exclude all the ethnic groups that have at least one of the matches identified by 3 or fewer events. In row (b) we define a match bad if it is based on 5 or fewer events. So we exclude all those groups which contain at least one of the matches identified by 5 or fewer events. In row (c) we keep only ethnic groups that are matched to rebel groups through the ACD2EPR conversion table, thus doing away with the geomatching procedure altogether. In row (d), we directly use as a dependent variable the measure of conflict incidence provided in the Ethnic Power relations Core Dataset. This measure, based on PRIO conflicts, not only does not use the geomatching procedure, but also includes only conflicts with more than 25 fatalities.[40] Finally, since

[40]It is not possible to run this robustness check in the restricted sample because only 2 groups are matched trough the ACD2EPR crosswalk table.

Table 8: Robustness checks on the geomatching procedure

Dependent variable: conflict incidence	" (se)	R^2	Obs.	" (se)	R^2	Obs.
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A – Mismatch Dummy						
(a) No matches identifyed by 3 (or less) events	0.0528*** (0.0129)	0.742	3,439	0.0425* (0.0199)	0.811	3,439
(b) No matches identifyed by 5 (or less) events	0.0384*** (0.0133)	0.727	3,327	0.0465** (0.0191)	0.798	3,327
(c) No geomatching (ACD2EPR only)	0.0274*** (0.00926)	0.746	3,633	0.0392** (0.0154)	0.819	3,633
(d) No geomatching - EPR incidence	0.0262*** (0.00744)	0.607	3,453	0.0363** (0.0170)	0.754	3,453
(e) No groups in coalitions	0.0571*** (0.0128)	0.655	4,010	0.0519** (0.0246)	0.761	4,010
Panel B – Mismatch Continuous						
(a) No matches identifyed by 3 (or less) events	0.171** (0.076)	0.596	841	0.143** (0.056)	0.653	841
(b) No matches identifyed by 5 (or less) events	0.0174** (0.076)	0.618	826	0.158** (0.063)	0.680	826
(c) No geomatching (ACD2EPR only)	0.156** (0.065)	0.477	1,020	0.117** (0.0459)	0.571	1,020
(e) No groups in coalitions	0.156** (0.0676)	0.555	1,195	0.120* (0.0715)	0.671	1,195
Model fixed effects	Country × Year			Ethnic group & Country ×Year		

Note: The dependent variable is conflict incidence. Column (1)-(3) report specifications with country ! year fixed effects, column (4)-(6) those with country ! year and ethnic group fixed effects. All specifications include *peace years & war history*, *family*, *natural resources* controls (see notes to Table 6). Standard errors are clustered at the country-year level in column. *** p<0.01, ** p<0.05, * p<0.1. Sources: for conflict incidence, peace years and war history UCDP-GED; for the continuous measure of mismatch, authors' data collection on cabinet members of ethnic groups and estimates of military power from the machine learning procedure; for the other controls, Ethnic Power Relations (EPR) Dataset Family.

our military power measure might be biased if ethnic groups form a coalition against the government, in row (e) we exclude those groups that according to ACD2EPR had ever fought in coalition with other groups. The coefficients of the mismatch variables are always positive and statistically significant. The size of the coefficient seems to decrease in samples built without the geomatching procedure (row c-d), especially when using the mismatch dummy in the extended sample. However, the quantification of the effect is remarkably similar to that of the baseline sample. In fact, when we do not employ geomatching, the number of groups in conflict is lower and the average value of conflict incidence is 0.04 in the extended sample in row (d) and 0.035 for the sample in row (e); this implies that, within a country-year, groups that are mismatched are about 50% more likely to participate in conflict than their non mismatch counterparts. An effect that is extremely close in magnitude to that found in Table 6.

5.2 Heterogeneity of the effect

In this section, we enrich the baseline results by examining various dimensions of heterogeneity. Firstly, we explore whether groups with high military power and low political power behave differently with respect to groups with the same absolute value of mismatch but with reversed order. To investigate the symmetry of the effect of power mismatch, we split the mismatch measures into two separate variables. Specifically, we create two dummy variables for the indicator of mismatch: *Mismatched dummy* $m > p$ equals one if political power is below the median of the distribution, and military power is in the fourth quartile; *Mismatched dummy* $p > m$ equals one if political power above the median of the distribution, and military power is in the first quartile. Similarly, *Mismatch cont.* $m > p$ contains the value of the continuous measure of mismatch if military power is greater than political power, while *Mismatch cont.* $p > m$ contains the value of the continuous measure of mismatch if military power is smaller than political power.

Table 9 presents the results of this analysis. Odd columns display the specifications with country×year fixed effects, and even columns include both country×year and ethnic-group fixed effects. All specifications include the usual ethnic group level controls. In Columns (1) and (2), we present the results for the extended sample using the indicator variables of power mismatch. In Columns (3) and (4), we narrow the sample down to the 14 Sub-Saharan countries while still using the mismatch indicators as our main explanatory variable. Finally, in Columns (5) and (6), we use the continuous measure of mismatch for the restricted sample. The results remain consistent across samples and measures of power mismatch, but they depend on the source of variation used for identification. In particular, when we exploit variation across groups within the same country, the effect of power mismatch appears symmetric: both mismatch variables have significant coefficients, and moreover, the two coefficients are not statistically different. However, when ethnic group-fixed effects are included, the effect of the political mismatch (i.e., when political power exceeds military power) becomes small and not statistically significant, and only the military mismatch (i.e. when military power is high) is positively and significantly correlated with conflict. This finding suggests that there may be some unobservable group-level characteristics that explain the distribution of mismatch across groups within countries, and once they are accounted for with the fixed effects, the likelihood that a group partakes in a conflict against the government rises when its military power increases without being balanced by an increase in political power.[41]

Table 9: Signed Mismatch

Dep. Var.: Conflict Incidence	Extended Sample		Restricted Sample		Restricted Sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Mismatch dummy, $m > p$	0.0400***	0.0782***	0.156***	0.177***		
	(0.0146)	(0.0294)	(0.0437)	(0.0503)		
Mismatch dummy, $p > m$	0.0809***	-0.0143	0.108**	0.00504		
	(0.0186)	(0.0343)	(0.0526)	(0.0699)		
Mismatch cont., $m > p$					0.234***	0.315***
					(0.0882)	(0.0974)
Mismatch cont., $p > m$					0.267***	-0.0187
					(0.0828)	(0.108)
Fixed Effects						
Country×year	!	!	!	!	!	!
Ethnic group		!		!		!
Mean dep. var.	0.114	0.114	0.0818	0.0818	0.0818	0.0818
R^2	0.670	0.768	0.568	0.675	0.559	0.673
Within R^2	0.508	0.354	0.445	0.318	0.433	0.315
Obs.	4,220	4,220	1,247	1,247	1,247	1,247

Note: The dependent variable is conflict incidence. All specifications include *peace years & war history*, *family*, *natural resources* controls (see notes to Table 6). The sample used in the specification is reported in the column headings. Standard errors are clustered at the country-year level. *** p<0.01, ** p<0.05, * p<0.1. Sources: for conflict incidence, peace years and war history UCDP-GED; for the continuous measure of mismatch, authors' data collection on cabinet members of ethnic groups and estimates of military power from the machine learning procedure; for the other controls, Ethnic Power Relations (EPR) Dataset Family.

These results on the particular relevance of a mismatch when $m > p$ is consistent with our general intuition on the conflict initiation incentives: even though conflict initiation data does not exist—and would be anyway unreliable—, $m > p$ cases of mismatch are those where it is clearly the militarily strong but politically weaker group that has the maximum incentive to challenge the status quo. Cases of the other type of mismatch (where $p > m$), may end up in conflict because of the power shift logic (Powell, 2012,2013), but it is a more indirect type of incentive, and hence, intuitively, less likely to show up in the data.

The second question we address is whether the effect of the mismatch is linear. Our simple theoretical framework implies that, for any realization of costs around the mean,

[41] 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.

only sufficiently high mismatches determine a rational incentive to attack. To investigate this, we leverage the continuous measure of mismatch and begin by dividing the sample using the median of the mismatch distribution as the cutoff value. We then estimate the empirical model separately for groups whose mismatch is below the median value and those whose mismatch exceeds it. The results are summarized in Table 10. Columns (1) and (2) present the results for the sample below the median, while columns (3) and (4) show the results for the sample above the median. Odd-numbered columns include specifications that contain the baseline set of controls (peace years, family, and natural resources controls) and country$\times$year fixed effects. Specifications in even-numbered columns are augmented with ethnic-group fixed effects.

The results reveal that the positive correlation between power mismatch and conflict is evident only in the sample above the median. This indicates a non-linear relationship between mismatch and conflict: marginal increases in the mismatch only lead to a higher probability of conflict for groups characterized by a relatively high asymmetry between the two dimensions of power. Consistent with the findings from the previous table, conflicts seem to be linked to mismatches primarily when military power surpasses political power. Specifically, in the sample above the median, a one-standard deviation increase in *Mismatch cont.* $m > p$ (0.14) is associated with a 9-percentage-point rise in the likelihood of conflict participation, which doubles the probability compared to the sample average of 0.094.

In columns (5)-(6) of Table 10, we corroborate this finding by incorporating a quadratic term of the mismatch measures into the baseline specification. The coefficients on these two variables indicate a convex relationship between mismatch and conflict: the coefficient on the linear term is negative, while the coefficient on the square of the mismatch is positive and statistically significant. The point estimates from column (6) suggest that the relationship between military mismatch and conflict is positive for values of the military mismatch higher than 0.09 (the median of *Mismatch cont.* $m > p$ in this sample is 0.16). This reinforces the idea that conflict probability increases significantly when military power exceeds political power by a substantial margin.

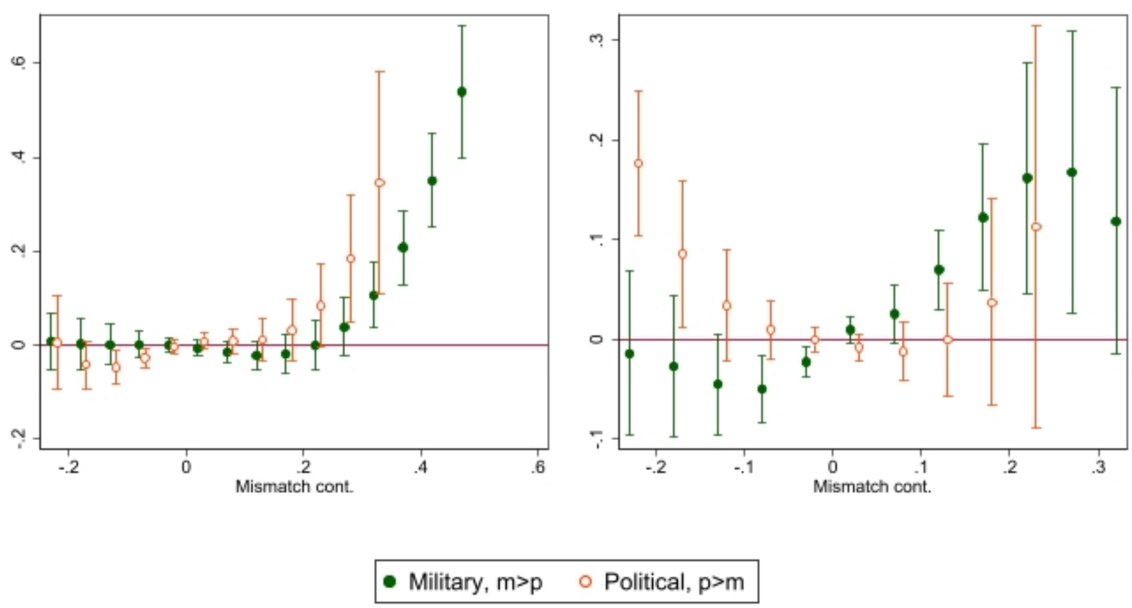
We delve deeper into the non-linearity of the effect using non-parametric regressions and visualize the results in Figure 7. Panel (a) presents the outcome of a simple non-parametric regression of conflict incidence on the mismatch variable after we condition on country$\times$year fixed effects, while in Panel (b) we additionally include ethnic group fixed effects. The plotted graphs in Figure 7 strongly support the notion that the relationship between power mismatch and conflict is indeed non-linear. They illustrate that an increase in mismatch raises the likelihood of conflict only for values of mismatch above a certain threshold. Moreover, the findings are consistent with our previous results, where we found that, after accounting for unobserved ethnic group characteristics, only military mismatch showed a positive correlation with conflict. As mentioned before, the non-linearity of the effect can be readily explained by referring back to the intuition of the simple theoretical framework. When costs are around the mean, only sufficiently high mismatches lead to a rational incentive to initiate conflict. In other words, when the mismatch is small, the expected costs may outweigh the benefits of conflict.

Table 10: Non-linearity of the effect

Dep. Var.: Conflict Incidence	Below Median (1)	(2)	Above Median (3)	(4)	Entire Sample (5)	(6)
Mismatch cont., $m > p$	-0.165	-0.266	0.673***	0.659***	-0.512***	-0.0627
	(0.260)	(0.255)	(0.175)	(0.208)	(0.153)	(0.199)
Mismatch cont., $p > m$	-0.00403	0.405	0.616***	0.201	-0.544***	-0.244
	(0.290)	(0.276)	(0.150)	(0.206)	(0.156)	(0.256)
Mismatch cont. squared, $m > p$					1.187***	0.714*
					(0.325)	(0.393)
Mismatch cont. squared, $p > m$					1.477***	0.337
					(0.354)	(0.541)
Fixed Effects						
Country×year	!	!	!	!	!	!
Ethnic group		!		!		!
Mean dep. var.	0.0689	0.0689	0.0942	0.0942	0.0818	0.0818
R^2	0.703	0.825	0.554	0.663	0.576	0.675
Within R^2	0.573	0.357	0.386	0.348	0.455	0.319
Obs.	610	610	637	637	1,247	1,247

Note: The dependent variable is conflict incidence. Odd-numbered columns reports specification with country ! year fixed effects, and even-numbered columns the specification with country ! year and ethnic group fixed effects. All specifications include peace years & war history, family controls, and natural resources controls (see notes to Table 6). Standard errors are clustered at the country-year level. *** p<0.01, ** p<0.05, * p<0.1. Sources: for conflict incidence, peace years and war history UCDP-GED; for the continuous measure of mismatch, authors' data collection on cabinet members of ethnic groups and estimates of military power from the machine learning procedure; for the other controls, Ethnic Power Relations (EPR) Dataset Family.

Figure 7: Non-Parametric regressions



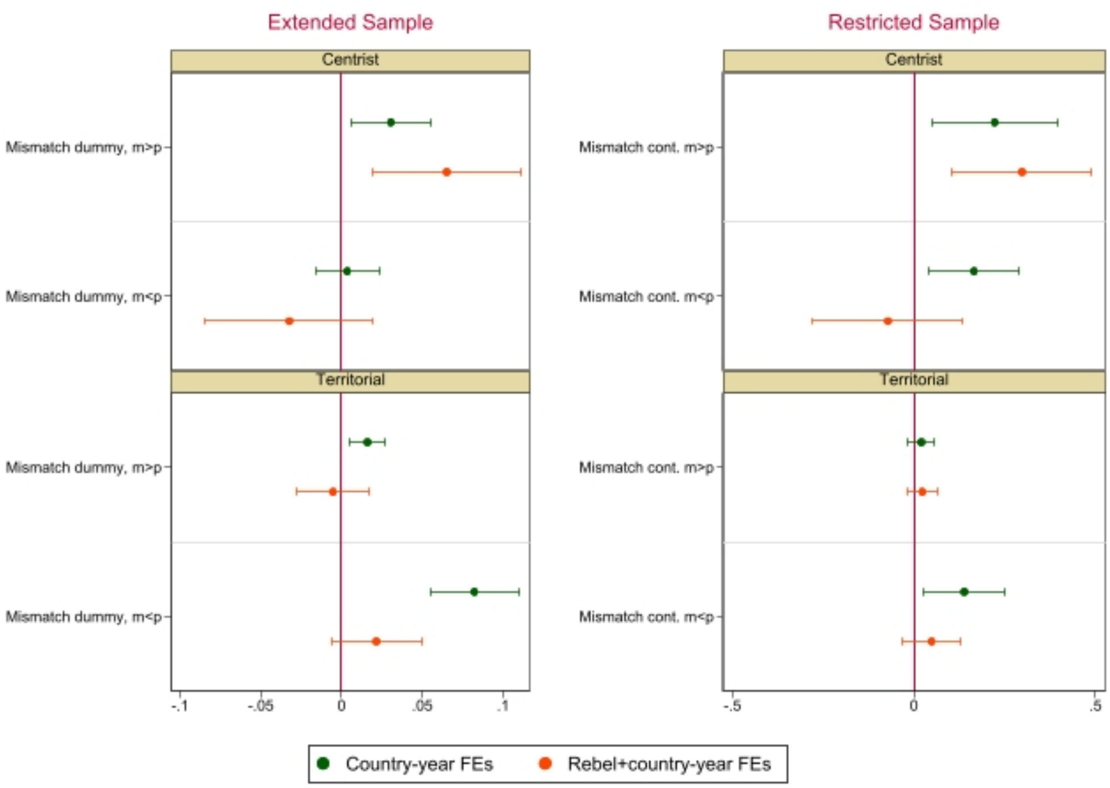
Note:: The left panel reports results of non parametric regression of conflict incidence on the signed continuous measures of mismatch after country-year fixed effects have been partialled-out while the panel on the right reports the result of non parametric regression after partialling out both country-year and ethnic-group fixed effects. Sources: for conflict incidence, peace years and war history UCDP-GED; for the continuous measure of mismatch, authors' data collection on cabinet members of ethnic groups and estimates of military power from the machine learning procedure. Sources: for conflict incidence, peace years and war history UCDP-GED; for the continuous measure of mismatch, authors' data collection on cabinet members of ethnic groups and estimates of military power from the machine learning procedure; for the other controls, Ethnic Power Relations (EPR) Dataset Family.

We now investigate whether the effect of power mismatch varies with different conflict characteristics, particularly conflict size and grievance type. To begin, we categorize conflicts based on the type of incompatibility underlying them. The GED dataset contains a variable that categorizes civil conflicts into 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). Using this information, we construct two new measures of conflict incidence, which we then use as dependent variables in our analysis. Next, we capitalize on the comprehensiveness of the GED dataset, which encompasses not only major civil wars but also smaller conflicts. We inquire whether power mismatch has different implications for major and minor conflicts. For this purpose, we define a conflict as *big* if the yearly average number of casualties exceeds 25.[42] Accordingly, we establish two separate dependent variables for conflict incidence: one for *big* conflicts and the other for *small* conflicts.[43]

[42]For example, if a conflict lasts for 3 years with reported casualties of 10, 20, and 100, respectively, the average number of casualties is 43.3, classifying the conflict as big. We adhere to the convention of UCDP/PRIO and use 25 as the cutoff value.

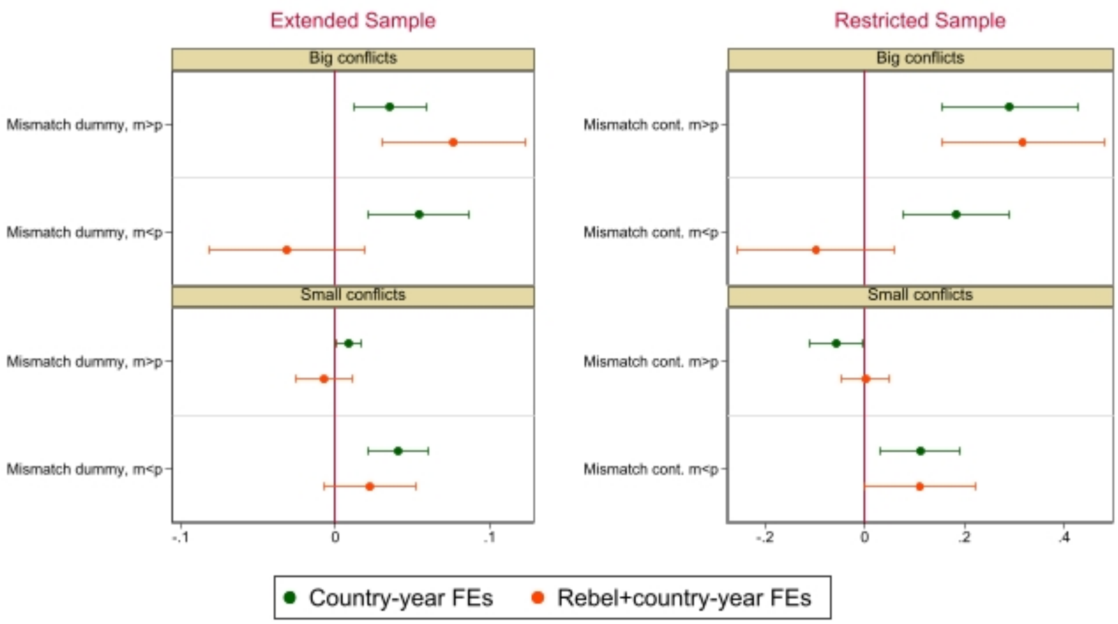
[43]It's important to note that the samples used in the regressions for big and small conflicts differ since groups experiencing big conflicts are not considered in the analysis for small conflicts and vice versa.

Figure 8: Territorial vs Centrist conflicts



Notes: the figure reports coefficients and standard errors of the mismatch variable. The dependent variable is conflict incidence either for territorial conflicts or for centrist conflicts, and it is specified in the graphs headings. The left panel shows the results of the regressions using the extended sample of countries and the indicator of mismatch, the right panel shows the results using the restricted sample of countries and the continuous mismatch variable. All specifications include peace years & war history, family controls, and natural resources controls (see notes to Table 6). Standard errors are clustered at the country-year level. Sources: for conflict incidence, peace years and war history UCDP-GED; for the continuous measure of mismatch, authors' data collection on cabinet members of ethnic groups and estimates of military power from the machine learning procedure; for the other variables, Ethnic Power Relations (EPR) Dataset Family.

Figure 9: Big vs Small conflicts



Notes: the figure reports coefficients and standard errors of the mismatch variable. The dependent variable is conflict incidence either for big conflicts (average yearly fatalities ¿25) or for small conflicts (average yearly fatalities ¡25), and it is specified in the graphs headings. The left panel shows the results of the regressions using the extended sample of countries and the indicator of mismatch, the right panel shows the results using the restricted sample of countries and the continuous mismatch variable. All specifications include peace years & war history, family controls, and natural resources controls (see notes to Table 6). Standard errors are clustered at the country-year level.Sources: for conflict incidence, peace years and war history UCDP-GED; for the continuous measure of mismatch, authors' data collection on cabinet members of ethnic groups and estimates of military power from the machine learning procedure; for the other variables, Ethnic Power Relations (EPR) Dataset Family.

The findings are summarized in Figures 8 and 9. These graphs display the coefficients obtained from regressing conflict incidence on both military and political mismatches for both the extended sample of countries (left panel) and the restricted sample (right panel). As usual, all the specifications include the standard group-level controls and country-year fixed effects, and we present results with and without ethnic-group fixed effects.

In Figure 8, we observe that power mismatch, especially when military power surpasses political power ($m > p$), has a more significant impact on conflict incidence when considering centrist conflicts. This pattern holds true regardless of the sample, the measure of mismatch used, or the variation employed to identify the coefficients. This finding aligns with the nature of the political power dimension captured by the mismatch variable. Whether using the EPR index or the continuous mismatch variable, political power is understood to be the group's participation in decision-making at the central level. These results are consistent with the theory proposed by Esteban et al. (2022), which emphasizes how dimensions such as cultural and religious group identity can heighten group preferences for autonomy and potentially lead to territorial conflicts even in the absence of a substantial power mismatch. Turning to Figure 9, the relationship between mismatch and the size of conflicts is less definitive. The coefficients of the mismatch variables for small conflicts are generally small and imprecisely estimated. On the other hand, power mismatch proves to be more significant for big conflicts. This suggests that if the mismatch is high and cannot be resolved through negotiation, addressing the grievance may entail large and likely prolonged conflicts.

The results of this section indicate that power mismatches are linked to an increased risk of conflict, particularly when the military power of an ethnic group excluded from power is substantial. Moreover, the impact of a mismatch increase is most notable for groups that initially experience relatively high levels of mismatch. Furthermore, the analysis shows that groups with a significant power mismatch tend to participate in larger conflicts and the underlying grievance for these conflicts often revolves around the division of central political power.

Next, we explore the correlation between power mismatch and specific country characteristics. Additionally, we examine whether the effect of an increase in mismatch varies across different countries, contingent upon crucial characteristics. This analysis aims to shed light on the complex interplay between power imbalances and various country attributes, contributing to a deeper understanding of the factors that shape conflict participation. In our investigation, we focus on two sets of characteristics.

The first set centers around institutional and cultural factors. Specifically, we consider the quality of democracy, the artificiality of the country's borders, ethno-linguistic fractionalization, ethnic and religious fragmentation, polarization, and genetic diversity. On the one hand, these factors are likely correlated with long-term determinants of power mismatch, as well as with the ease with which the mismatch can be resolved through negotiations. On the other hand, these same factors can interact with power imbalances, shaping perceptions of power disparities and exacerbating tensions among diverse groups within a country, thereby amplifying or moderating the effect of power mismatch in various countries. In the second set, we analyze economic characteristics, which encompass proxies for wealth, along with indicators for inequality and wealth distribution,

Table 11: Mismatch and country characteristics

	No Mismatched Groups		At Least One Mismatched Group		Difference
	N	Mean	N	Mean	
Institutional and Cultural Factors					
Polity IV Index	542	-0.683	432	0.0303	-0.713***
Free and Fair Elecions Index (V-DEM)	544	0.293	470	0.306	-0.013
Executive Accountability Index (V-DEM)	544	0.577	470	0.550	0.026*
Political Corruption Index (V-DEM)	544	0.697	457	0.67	0.023**
Border Artificiality (Alesina et al., 2010)	487	0.025	421	0.0282	-0.003**
Ethno-linguistic Fractionalization (Desmet et al., 2009)	544	0.713	470	0.703	0.009
Polarization Index (Reynal-Querol 2002)	544	0.130	470	0.127	0.003
Economic Factors					
Nightlight per capita 1992	544	0.026	470	0.016	0.009***
ln(GDP per capita, 2000)	544	7.539	470	7.536	0.003
Gini Index	259	48.29	312	48.27	0.023
Share Income Top 5%	259	51.56	312	53.38	-1.817**
Natural Resource Rents (% of GDP)	543	15.09	460	14.87	0.215
Oil Gini (Morelli Rohner, 2015)	397	0.188	329	0.250	-0.062**

Note: the table reports the results of a t-test of differences of means for the sample of country with no mismatched groups and the sample with at least one mismatch group. Observations are at country-year level, total number of observations 1,014.
*** p<0.01, ** p<0.05, * p<0.1. Sources: for Polity IV score, Polity IV Project (Marshall et al., 2014); for the indexes Free and Fair Eletion, Executive Accountability, Political Corruption, Varieties of Democracy (V-Dem) dataset (Coppedge et al., 2023); for Border Artificiality, Ethno-linguistic Fractionalization, Polarization indexes as well as for Gini index and Share of income in the top 5%, Alesina et al. (2016); for Natural Resource Rents, The World Bank; for Oil Gini, Morelli and Rohner (2015); for Nightlight per capita in 1992 and GDP per capita in 2000, GROWup Dataset (Girardin et al., 2015).

and the dependency on natural resources. These economic factors directly influence the material conditions of different ethnic groups and may interact with power mismatch to generate conflict by influencing the perceived disparities in economic opportunities and resource distribution among different ethnic groups.

To investigate the correlation between the presence of mismatched groups and specific country characteristics, we utilize the extended sample, which provides greater variation for analysis. Aggregating the data at the country-year level, we conduct t-tests to compare the averages of each characteristic between countries that have no mismatched groups in a year and countries that have at least one mismatched group. The results of these t-tests are presented in Table 11.[44] Upon examining these descriptive tests, we find that certain characteristics display the expected correlations with the presence of mismatched groups, while others show unexpected signs. Notably, the presence of mismatched groups appears to be less likely in highly autocratic countries. However, this correlation reverses when we disaggregate the democracy index. Specifically, countries with no mismatched groups tend to have higher indices for executive accountability and a lower incidence

[44] We sourced the Polity IV score from Polity IV Project (Marshall et al., 2014); the indexes Free and Fair Eletion, Executive Accountability, Political Corruption from the Varieties of Democracy (V-Dem) dataset (Coppedge et al., 2023), Border Artificiality, Ethno-linguistic Fractionalization, Polarization indexes as well as for Gini index and Share of income in the top 5% from Alesina et al. (2016); Natural Resource Rents from The World Bank, Oil Gini from Morelli and Rohner (2015), and Nightlight per capita in 1992 and GDP per capita in 2000 from GROWup Dataset (Girardin et al., 2015).

of political corruption.[45] Regarding the ethnic heterogeneity dimension, we observe that mismatches seem slightly more prevalent in countries with artificial borders, while measures of ethnic fractionalization and polarization show similar patterns in both groups. Turning to economic characteristics, we find that countries with mismatched groups tend to be poorer, as indicated by proxies such as pre-sample nightlight luminosity, and characterized by higher inequality in the distribution of resources. This inequality is evident in terms of income, as measured by the share of income accruing to the top 5% of the population, as well as in the distribution of natural resources, represented by the Oil Gini index.

While the results of Table 11 are descriptive in nature, they serve a valuable purpose by providing insights into the factors that are absorbed by the country×year fixed effect in our main specifications. Understanding these factors helps guide us in selecting crucial characteristics that may interact with power mismatch to produce heterogeneous effects. To conclude the section, we examine the presence of such heterogeneous effects. Building on previous analyses, we split the sample based on the median of each characteristic and independently run our main model (eq.5 with rebel-group fixed effect) on the two subsamples. The results of these analyses are summarized in Figure 10. The bar graphs represent the size of the coefficient on the *Mismatch Dummy* variable, while the segments denote the 90% confidence intervals. The results for the sample below the median of the characteristics are depicted in orange, while those for the sample above the median are shown in blue. Inspecting the results, we notice that from a statistical perspective, the coefficients of the mismatch variable are consistently not different between the two samples, regardless of the characteristics. However, when we examine the magnitude of the coefficients, a clear pattern emerges: if a group becomes mismatched, the likelihood of participating in a conflict increases more significantly in countries characterized by poverty, high levels of inequality, heavy dependence on natural resources, and political corruption. Interestingly, we also observe a somewhat unexpected finding: higher measures of the quality of democratic institutions are associated with a stronger impact of power mismatches on conflict participation. It is worth noting that even in the sample above the median, countries are not full democracies, as the value of the indexes is still quite low.[46]

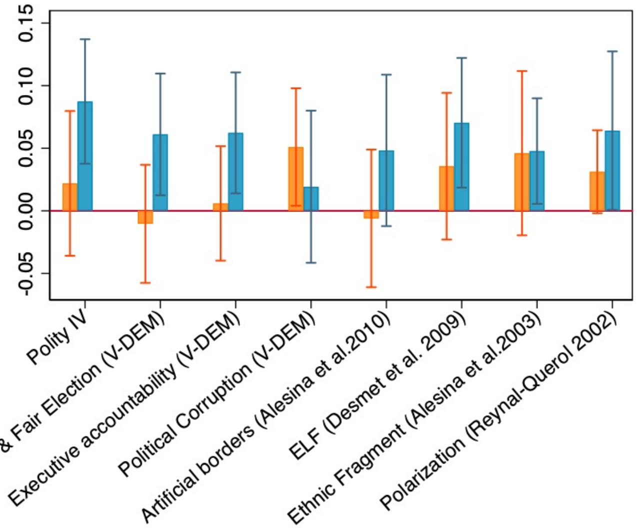
5.3 Interpreting the results

As previously stated, the primary objective of this paper is to provide a descriptive analysis. Nevertheless, in this section we interpret the results, acknowledging factors that hinder a straightforward causal interpretation of the findings. Although we have taken great care to control for numerous determinants of conflicts as proposed by the literature, country-level shocks are accounted for through the country×year fixed effect, and unobserved group characteristics are absorbed by ethnic-group fixed effects, establishing definitive causality between mismatch and conflict remains challenging. A key concern

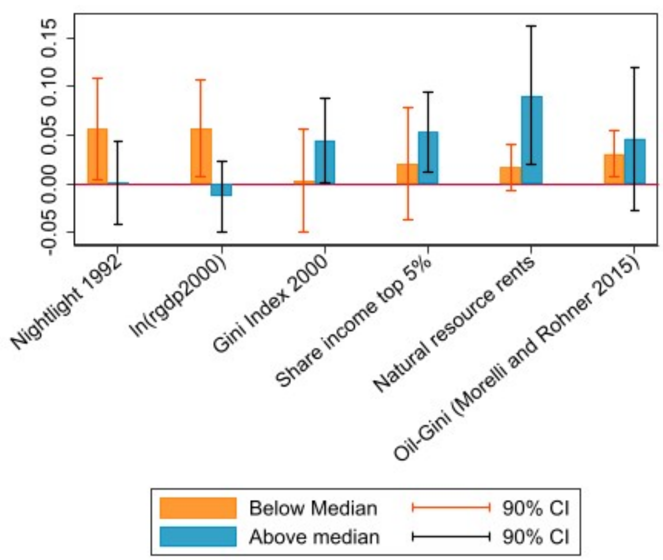
[45] Indexes of free and fair elections, executive accountability and political corruption are taken from the V-DEM dataset (Coppedge et al., 2023). Higher values of the indexes mean better institutional quality.

[46] The 90th percentile of the polity2 score is 7 (Range of the index [-10,10]), of the accountability index is 0.8 (range of the index [0,1]), and of the free and fair election is 0.64 (range of the index [0,1])

Figure 10: Heterogeneity of the Effect



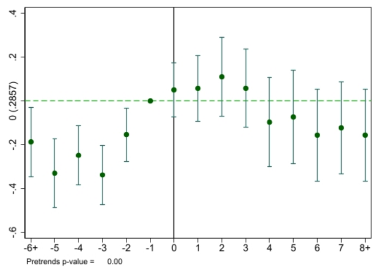
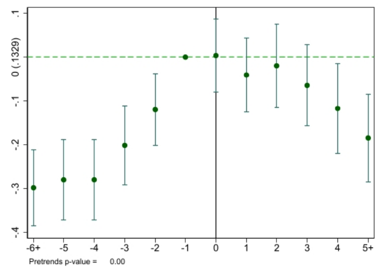
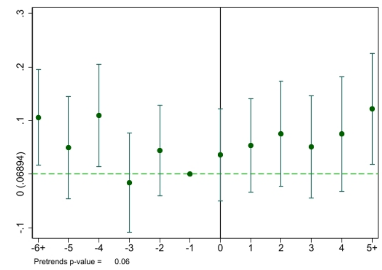
(a) Institutional and Cultural Factors



(b) Economic Factors

Notes: the figures report the coefficients and 90% confidence intervals on the mismatch dummy estimated using model 5. Each couple of bins is estimated on subsamples of the extended sample of countries above (blue) and below (orange) the median of the country characteristic reported on the x-axis. Panel (a) includes institutional and cultural characteristics, while panel (b) focuses on economic characteristics. Sources: for Polity IV score, Polity IV Project (Marshall et al., 2014); for the indexes Free and Fair Eletion, Executive Accountability, Political Corruption, Varieties of Democracy (V-Dem) dataset (Coppedge et al., 2023); for Border Artificiality, Ethno-linguistic Fractionalization, Polarization indexes as well as for Gini index and Share of income in the top 5%, Alesina et al. (2016); for Natural Resource Rents, The World Bank; for Oil Gini, Morelli and Rohner (2015); for Nightlight per capita in 1992 and GDP per capita in 2000, GROWup Dataset (Girardin et al., 2015).

Table 12: Event Study

Mismatch	Mismatch $m > p$	Mismatch $m < p$
		

Notes: The table of graphs reports the results of the event study in eq. 6. The dependent variable is reported in the column headings. All specifications include peace years & war history, family controls, and natural resources controls (see notes to Table 6). Standard errors are clustered at the country-year level. Sources: conflict onset, peace years and war history UCDP-GED; for the continuous measure of mismatch, authors' data collection on cabinet members of ethnic groups and estimates of military power from the machine learning procedure; for the other variables, Ethnic Power Relations (EPR) Dataset Family.

is the potential for reverse causality: groups that are excluded from power enter into conflict and, precisely because they are experiencing conflict, have higher military power than groups that are not. While it is not possible to completely eliminate this possibility, we undertake an examination of the temporal dynamics to explore whether mismatch or conflict comes first. We do this by performing an event study, where the event is defined as the first year of conflict participation. Specifically, we exploit the continuous measure of mismatch and estimate the following model:

$$\mathit{Mismatch\,cont.}_{ect} = !_{ct} + \$_e + \underset{s=\#5}{\overset{\#2}{!}} \text{"}_s Det^s + \underset{s=0}{\overset{5}{!}} \text{"}_s Det^s + X_{ect}\# + \%_{ect}, \tag{6}$$

where Dit^s are event-study dummies, which are equal to 1 if ethnic group e is s periods away from the first instance of conflict at time t, and 0 otherwise. $\$_e$ are ethnic-group fixed effects, $!_c t$ are country$\times$year fixed effects and X_{ect} is the matrix containing the usual set of controls. The coefficients "_s illustrate how power mismatch evolves over time within groups that participate in a conflict relative to groups that do not participate in conflicts over a 10 year window around the initial instance of conflict, $s = 0$. The table of graphs 12 collects the results for the continuous measure of mismatch, $M = |m - p|$, and for the split measures, $m > p$ and $p > m$. Notably, mismatch tends to increase three years before the conflict onset—reaching its peak usually in the year of the conflict—and then decreases. This pattern suggests that our mismatch variables precede the occurrence of conflict.

A second potential concern arises regarding whether groups with low political power might deliberately increase their military capabilities in anticipation of future conflicts with the government in order to gain political power. This forward-looking behavior could lead to the accumulation of military power in the years before the conflict breaks out. While it is challenging to fully eliminate this possibility, it is unlikely that our mismatch variable captures such forward-looking behavior. The predictors used in the machine learning procedure are based on structural characteristics of the ethnic groups, designed to capture their "potential" military power, and are not expected to respond to the forward looking decisions of actors.

In any case, to further allay this concern and ensure the robustness of our baseline results, we investigate different specifications that do not rely on yearly time variation in mismatch for identification. The results of this exercise are reported in Table 13. In column (1), we estimate the baseline model, replacing the *Mismatch Dummy* with its initial value at the beginning of the sample. Similarly, in column (2), we collapse our data in a cross section, with the main explanatory variable being the Mismatch Dummy at the beginning of the sample period, and the dependent variable being a dummy variable that takes a value of 1 if the ethnic group experienced conflict during the sample period. We further repeat this exercise by stacking different time windows in the remaining columns. The dependent variable is a dummy that takes a value of 1 if the groups experience conflict over the time window, and the main explanatory variable is the mismatch dummy computed at the beginning of the time window, with all controls computed at the beginning of the time window as well. Columns (3) and (4) stack 10-year windows, while columns (5) and (6) use 5-year windows. The advantage of stacking multiple time windows is that it provides multiple observations for each ethnic group, allowing us to include ethnic group fixed effects in the specification. The results are consistently robust, indicating that initial mismatch is always positively and significantly correlated with future conflict, confirming that it is the groups that are initially mismatched that are more likely to enter into a conflict.

A final concern pertains to the potential confounding effect of political power on the mismatch variable. Recent research (e.g., Deiwiks et al., 2012; Wimmer et al., 2009) highlights the significance of political exclusion as a catalyst for conflict. In our context, this could pose a challenge if the mismatch variable merely captures exclusion from political power. Specifically, our findings suggest that conflict participation is more likely when a group has low political power but high military power. If a group initially possesses average levels of both military and political power, but for some reason experiences a loss of political power, this would lead to an increase in the mismatch variable. Consequently, if conflict emerges after this "downgrade" in political power, it might be attributed to the increase in the mismatch, when in reality, the trigger for the conflict was the exclusion from political power. Even though we cannot completely rule out this possibility, a closer examination of episodes of power downgrading in our data offers valuable insights. In the extended sample, we identified 28 groups (30 cases) that experienced such a reduction in their power rank measure. For each of these events, we assessed whether the group participated in a conflict against the government in the year of the power downgrading or in the subsequent five years.[47] Descriptive statistics for these cases are presented in Table 14. Remarkably, only 10 out of the 30 groups that underwent a reduction in political power were involved in a conflict within the following five years. Notably, the characteristics related to political power before the downgrade and the magnitude of the downgrade were strikingly similar between the groups that experienced conflict and those that did

[47]We drop two cases where the downgrade happens at the end of the sample period.

Table 13: Long Differences

Dep. Var.:	Conflict Incidence	At least one conflict incidence in the last:				
		20 years	10 years	10 years	5 years	5 years
	(1)	(2)	(3)	(4)	(5)	(6)
Initial Mismatch Dummy	0.0295**	0.194*	0.120*	0.201**	0.108**	0.158**
	(0.0137)	(0.106)	(0.0696)	(0.0977)	(0.0485)	(0.0631)
Fixed Effects						
Country×Year	!					
Country		!	!		!	
Year			!	!	!	!
Ethnic Group				!		!
Mean dep. var.	0.114	0.288	0.213	0.213	0.182	0.182
R^2	0.667	0.529	0.457	0.825	0.466	0.737
Within R^2	0.503	0.202	0.235	0.312	0.279	0.193
Observations	4,220	236	445	445	848	848

Note: The dependent variable is reported in the column headings. The results of the table are obtained using the extended sample and the indicator of mismatch. The main explanatory variable is the mismatch dummy at the beginning of the relevant period (which is also reported in the column headings). All specifications include peace years, family controls, and natural resources controls (see notes to Table 6). Standard errors are clustered at the country-year level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Sources: for conflict incidence, peace years and war history UCDP-GED; for the measure of mismatch, EPR core dataset and authors' estimates of military power from the machine learning procedure; for the other controls, Ethnic Power Relations (EPR) Dataset Family.

not. On average, these groups transitioned from a political rank of 4 (i.e., junior partner in the government) to a political rank of 2 (powerless). The most significant distinction was observed in their military power: groups that experienced conflict following a political downgrade displayed considerably higher estimated military power compared to those that did not. To gauge the magnitude of this difference, downgraded groups that did not experience conflict had a median military power of 0.24, a value similar to the median of the entire sample. In comparison, the median military power for downgraded groups that did experience conflict was 0.55, surpassing the value of the 90th percentile of the military power distribution in the entire sample (0.50). This simple exercise suggests that being excluded from power is not enough to predict conflict, and that our mismatch variable is not just picking up the effect of power downgrading, but an actual asymmetry between political and military power.

6 Policy Discussion

Our paper suggests that power sharing policies should be designed and conceived in a manner completely different from the past, which is a history mostly filled with failures. The history of failures in power sharing is due to a wrong idea of what power sharing

Table 14: Political power downgrading

	Frequency	Rank before Downgrade		Δ Rank		Military Power	
		mean	median	mean	median	mean	median
Conflict	10	3.8	4	-2	-2	0.5	0.55
No Conflict	20	3.95	4	-2	-2	0.3	0.24

Notes: the table reports summary statistics for 30 groups in the extended sample that experienced episodes of power rank downgrading. Groups are divided (by rows) according to whether they participate in a conflict within 5 years of from downgrading episode. Sources: EPR-core dataset and result of the Machine learning procedure for military power.

should mean and imply: the existing simplistic concepts of power sharing range from grand coalitions to cooperation among elites that were previously at war, and are always unrealistic (see Spears (2000)). In this section we propose as anecdotal evidence of the relevance of our framework a description of why some power sharing designs have worked, and indeed how these few successes go in the direction of power mismatch reduction.

As discussed, for example, in Hartzell and Hoddie (2003) and Reilly (2012), the two opposing theories on how power sharing should be advocated to resolve or avoid conflicts are consociationalism and centripetalism: the former relates to grand coalitions, cooperation of elites, or "fixed" proportional access to political power by all relevant groups, as in Lebanon; the latter refers to attempts to create multi-ethnic parties competing for power, as in Kenya, Indonesia and Nigeria.

Consociationalism has the drawback of focusing exclusively on fixing the relative political power of groups, ignoring the need to consider the present and future adjustments that in our mismatch framework clearly appear as necessary when a group's relative military power and/or external support change. Fixing the relative political power of an ethnic group forever risks leading stronger groups to try to control the military and/or to sabotage political institutions.

At the opposite extreme, centripetalism experiments aimed to reduce the salience of ethnicity, but with mixed success. When a country's citizens shift (gradually) from ethnic voting to party voting, the relative political power that matters also likely shifts from the relative power of ethnic groups to the relative power of parties. Regardless of whether the transition is half way or complete, our mismatch framework still applies. The implication we draw is that even to design such a transition one should consider the simultaneous access to military power of different groups, whether ethnic or political. Also mismatched parties, for example, supported militarily by a foreign power with aligned political ideologies, can lead to conflict.

Both the of above extreme theories focus on the pros and cons of different types of distributions of political power "alone", and this is the most important mistake suggested by our framework, even more important than the problems of rigidity of quotas in the former case and multiple identities in the latter case. The whole debate must shift, we argue, in the direction of considering simultaneously all the relevant dimensions of power, not just political power. The most successful case of power sharing agreement, the Good Friday

Agreement of 1998, which was proposed to end the conflict in Northern Ireland, is indeed a case in which the power mismatch was addressed: deposition of weapons and access to political power and public sector jobs were part of the multidimensional deal – see, for example, O'Leary (2001). Guaranteeing job quotas in the public sector, police, and even the military, generated commitment to a balance of power that gave equal importance to incentives, political representation and access to resources, with some credibility.

Moving from the debate on power sharing design to the debate on political institutions, an important body of literature exists on the usefulness of proportional representation electoral systems, as a peace-inducing electoral mechanism – see e.g., Horowitz (1990, 2000, 2003, 2005). In a winner-takes-all system like plurality rule, democratic elections can still lead to high mismatches of power, as illustrated by the history of Angola: civil war against the MPLA government in Luanda provinces was justified by the UNITA group's exclusion from power at the time of independence in 1975. However, after signing a peace agreement in 1991 and losing the winner-takes-all type of elections in 1992, UNITA returned to war until it was finally induced to sign a power-sharing agreement in 1994. This also did not last because there remained no credible real access to resources (see Morelli and Rohner, 2015, on the role of resources as an additional source of difficulty in negotiating an agreement). This shows that elections are not a panacea, especially in winner-takes-all system. The frequent presence of mismatched groups in democratic countries that are rich in natural resources and where conflict is centrist – all highlighted in our heterogeneity analysis – are perfectly represented by the Angola example.

Intuitively, proportional representation allows for flexible rebalancing of the relative political power of ethnic groups, and this is the main reason to advocate proportional representation style electoral reforms. However, the intuitive appeal of proportional representation systems has two limitations:

(1) political power proportional to one source of strength, namely relative group sizes, works only when voters vote along ethnic lines. Interestingly, when the transition to party voting is underway, as advocated in the centripetalism literature, the effect of a proportional representation electoral system on the reduction of relative power differences across ethnic groups is reduced, and hence also the connection with power mismatch reduction is tenuous.[48] Moreover, the greater the transition to party politics, the more relevant the income inequality dimension may become, which we also saw was relevant in our heterogeneity analysis.

(2) Most importantly, even in a polity with pure ethnic voting and pure proportional representation, a proportional electoral system might guarantee that a group with 30 percent of ethnic group voters obtains 30 percent of political power, but if such a group has a probability of victory against the majority group that is much higher than 30 percent – for one of the many reasons behind our estimation of military power that often differs significantly from relative group size – then the mismatch is not eliminated.

Given that democratic elections do not guarantee the elimination of mismatches, 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.

[48]See Huber (2012) for a deep discussion of ethnic vs party voting.

Similarly, in Colombia, the demilitarization of the FARC had to go hand-in-hand with the concession of a political role. Also in Spain the reduction of the conflict and terror incentives of independentist movements had to go through a political representation channel.

That democracy alone is not enough to eliminate mismatches can also be observed by looking at the correlation of power mismatches with country characteristics: as we have seen, more democratic countries tend to have a larger share of high-mismatch groups, especially when political corruption is high. Contrary to popular belief, it seems that groups at high risk of conflict reside in countries that are traditionally believed not to be at risk of conflict – where this incorrect belief derives from the results of democratic peace found in literature on interstate wars. One potential reason for this is the above-mentioned transition from ethnic politics to party politics, which happens slowly during the democratic transition. Another reason could be 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. But perhaps the most important and widespread source of problems is insufficient realism: relative political power and relative strength of groups need to be balanced if the objective is peace, even if sometimes this implies giving up a desire for equality on each dimension of power.

Our results therefore highlight an important policy dilemma for future research: peace objectives and political or economic equality objectives are often made incompatible by the "short blanket" consequences of understanding the incentives created by power mismatches: in any country where an ethnic group exists with military power greater than proportional to its size, elimination of the mismatch helps reduce war incentives but exacerbates political and/or economic inequality.

7 Conclusions

This paper makes two contributions to the literature and future research: it provides new measures for the study of ethnic conflict, filling an important gap, especially on the military power dimension, and provides the first empirical analysis of the mismatch theory of power wars. Regarding 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 its corresponding government controlling group, that varies 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.

In terms of 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 this new measure of ethnic-group military power, we show the existence of a relationship between the likelihood of conflict 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 between 30 and 50% more likely to take part in a conflict against their government. Moreover, these conflicts tend to be more deadly 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 consistent with the theory of power wars (Herrera et al., 2022). Hence, we believe that our evidence on the key role of power mismatch should encourage further research, both in the precise identification of conflicts and the forecasting of future conflicts. In addition, our evidence has 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 would also be interesting to dig deeper into the origin of power mismatch and into the causes of its persistence. 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

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

Angrist, J. D. and Pischke, J.-S. (2009), *Mostly harmless econometrics: An empiricist's companion*, Princeton university press.

Angrist, J. D. and Pischke, J.-S. (2014), *Mastering'metrics: The path from cause to effect*, Princeton university press.

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

Bruk, S. I. and Apenchenko, V. (1964), 'Atlas narodov mira (atlas of the world's peoples)', *Moscow: Glavnoye Upravleniye Geodezii i Kartografii* .

Carroll, R. J. and Kenkel, B. (2019), 'Prediction, proxies, and power', *American Journal of Political Science* **63**(3), 577–593.

Cederman, L.-E. and Girardin, L. (2007), 'Beyond fractionalization: Mapping ethnicity onto nationalist insurgencies', *American Political science review* **101**(1), 173–185.

Cederman, L.-E., Gleditsch, K. S. and Buhaug, H. (2013), *Inequality, grievances, and civil war*, Cambridge University Press.

Cederman, L.-E., Wimmer, A. and Min, B. (2010), 'Why do ethnic groups rebel? new data and analysis', *World politics* **62**(1), 87–119.

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

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., Bernhard, M., Cornell, A., Fish, M. S., Gastaldi, L., Gjerløw, H., Glynn, A., Good God, A., Grahn, S., Hicken, A., Kinzelbach, K., Krusell, J., Marquardt, K. L., McMann, K., Mechkova, V., Medzihorsky, J., Paxton, P., Pemstein, D., Pernes, J., Ryden, O., von Römer, J., Seim, B., Sigman, R., Skaaning, S.-E., Staton, J., Sundström, A., Tzelgov, E., Wang, Y.-t., Wig, T., Wilson, S. and Ziblatt, D. (2023), 'V-dem country-year dataset v13'.

Deiwiks, C., Cederman, L.-E. and Gleditsch, K. S. (2012), 'Inequality and conflict in federations', *Journal of Peace Research* **49**(2), 289–304.

Eck, K. (2012), 'In data we trust? a comparison of ucdp ged and acled conflict events datasets', *Cooperation and Conflict* **47**(1), 124–141.

Esteban, J., Flamand, S., Morelli, M. and Rohner, D. (2022), 'Stay, split or strike: Theory and evidence on secessionist vs centrist conflict'.

Esteban, J., Mayoral, L. and Ray, D. (2012), 'Ethnicity and conflict: An empirical study', *American Economic Review* **102**(4), 1310–42.

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

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

Girardin, L., Hunziker, P., Cederman, L.-E., Bormann, N.-C. and Vogt, M. (2015), 'Growup-geographical research on war, unified platform', *ETH Zurich. http://growup. ethz. ch* p. 33.

Gleditsch, N. P., Wallensteen, P., Eriksson, M., Sollenberg, M. and Strand, H. (2002), 'Armed conflict 1946-2001: A new dataset', *Journal of peace research* **39**(5), 615–637.

Hartzell, C. and Hoddie, M. (2003), 'Institutionalizing peace: power sharing and post-civil war conflict management', *American Journal of Political Science* **47**(2), 318–332.

Herrera, H., Morelli, M. and Nunnari, S. (2022), 'A theory of power wars', *Quarterly Journal of Political Science* **17**(1).

Horowitz, D. L. (1990), 'Ethnic conflict management for policymakers', *Conflict and peacemaking in multiethnic societies* **115**.

Horowitz, D. L. (2000), *Ethnic groups in conflict, updated edition with a new preface*, Univ of California Press.

Horowitz, D. L. (2003), 'Electoral systems and their goals', *Journal of Democracy* **14**(4), 115–127.

Horowitz, D. L. (2005), *Power Sharing: New Challenges For Divided Societies*, Pluto Press.

Huber, J. D. (2012), 'Measuring ethnic voting: Do proportional electoral laws politicize ethnicity?', *American journal of political science* **56**(4), 986–1001.

Marshall, M. G., Gurr, T. R. and Jaggers, K. (2014), 'Polity iv project: Political regime characteristics and transitions, 1800–2013', *Center for Systemic Peace* **5**.

Michalopoulos, S. and Papaioannou, E. (2016), 'The long-run effects of the scramble for africa', *American Economic Review* **106**(7), 1802–48.

Morelli, M. and Rohner, D. (2015), 'Resource concentration and civil wars', *Journal of Development Economics* **117**, 32–47.

Moscona, J., Nunn, N. and Robinson, J. A. (2020), 'Segmentary lineage organization and conflict in sub-saharan africa', *Econometrica* **88**(5), 1999–2036.

O'Leary, B. (2001), 'The character of the 1998 agreement: results and prospects', *Aspects of the Belfast Agreement* pp. 49–83.

Posner, D. N. (2004), 'The political salience of cultural difference: Why chewas and tumbukas are allies in zambia and adversaries in malawi', *American Political Science Review* **98**(4), 529–545.

Rainer, I. and Trebbi, F. (2014), New tools for the analysis of political power in africa, *in* 'African Successes, Volume I: Government and Institutions', University of Chicago Press, pp. 145–212.

Reilly, B. (2012), 'Institutional designs for diverse democracies: Consociationalism, centripetalism and communalism compared', *European Political Science* **11**, 259–270.

Sarkees, M. R. and Wayman, F. (2010), *Resort to war: 1816-2007*, Cq Press.

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

Sundberg, R. and Melander, E. (2013), 'Introducing the ucdp georeferenced event dataset', *Journal of Peace Research* **50**(4), 523–532.

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

Tollefsen, A. F., Strand, H. and Buhaug, H. (2012), 'Prio-grid: A unified spatial data structure', *Journal of Peace Research* **49**(2), 363–374.

Van der Laan, M. J., Polley, E. C. and Hubbard, A. E. (2007), 'Super learner', *Statistical applications in genetics and molecular biology* **6**(1).

Vogt, M., Bormann, N.-C., Rüegger, S., Cederman, L.-E., Hunziker, P. and Girardin, L. (2015), 'Integrating data on ethnicity, geography, and conflict: The ethnic power relations data set family', *Journal of Conflict Resolution* **59**(7), 1327–1342.

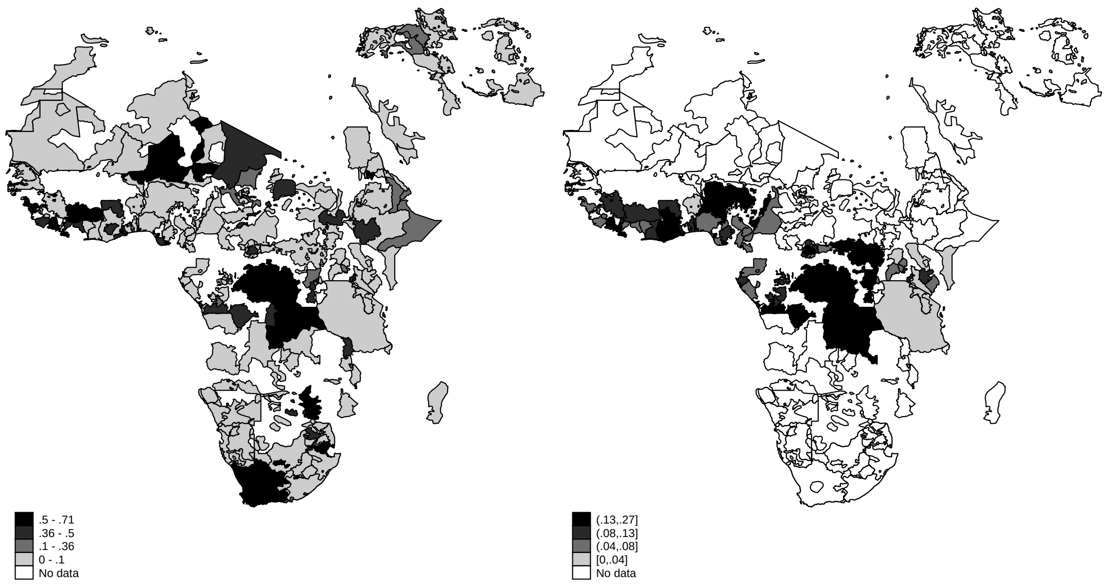
Wimmer, A., Cederman, L.-E. and Min, B. (2009), 'Ethnic politics and armed conflict: A configurational analysis of a new global data set', *American sociological review* **74**(2), 316–337.

Wucherpfennig, J., Weidmann, N. B., Girardin, L., Cederman, L.-E. and Wimmer, A. (2011), 'Politically relevant ethnic groups across space and time: Introducing the geoepr dataset', *Conflict Management and Peace Science* **28**(5), 423–437.

Appendix

A Additional Figures

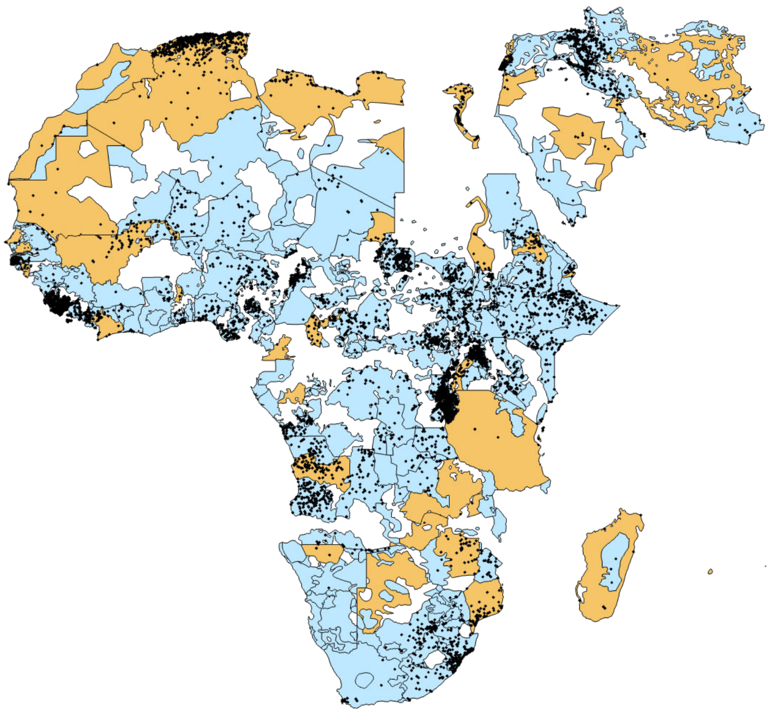
Figure A.1: Mismatch Measures Ð Within-Group Standard Deviation



(a) Mismatch Dummy (b) Continous measure of Mismatch

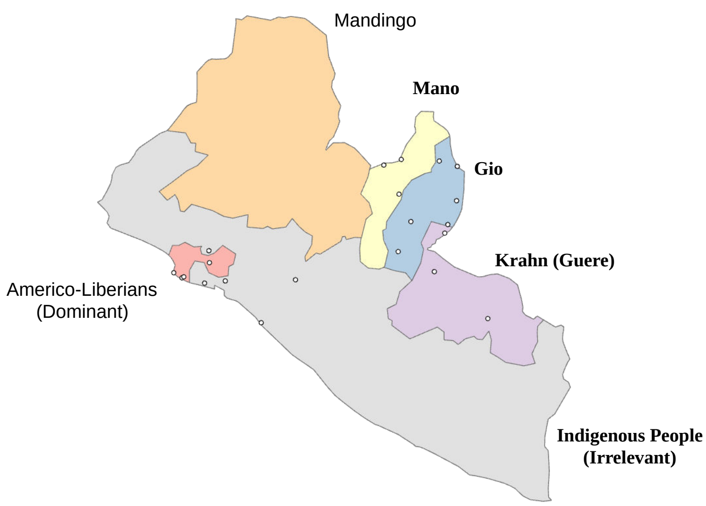
Notes: The Þgure plots the geographical distribution of the within-group standard deviation of the indicator of mismatch (panel a) and the continuous measure of mismatch (panel b). Sources: Polygons for the homeland of the ethnic group are sourced from the GEO-EPR dataset; data on the Mismatch Dummy are from authorsÕ computation based on the EPR-Core Dataset; data on the continuous measure of mismatch are based on authorsÕ data collection.

Figure A.2: Geo-matching between GED and GeoEPR in Africa and Middle East



Notes. The polygons are the homeland of each politically relevant ethnic group in Africa and the Middle East as coded in GeoEPR. Specifically, the golden-colored polygons are the homeland of the government ethnic groups, and the blue-colored ones are the ethnic homeland of the rebel groups. The irrelevant ethnic groups are not geo-coded in GeoEPR, so they are left blank on the map. The black dots are the conflict events recorded in GED.

Figure A.3: Rebel group coding: NPFL 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 NPFL.

B Sample

This section provides the list of countries in our extended and restricted sample, the distribution of the observations in the different EPR power rank categories and a description of the detailed manual checks on multi-ethnic governments.

Country list For each country Table B.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 B.1: Country list

Country	Avg # of groups	Restricted sample	Country	Avg # of groups	Restricted sample
Algeria	1		Lebanon	8	
Angola	4		Liberia	4	!
Bahrain	1		Libya	3	
Benin	3	!	Madagascar	1	
Botswana	9		Malawi	2	
Cameroon	5	!	Mali	2	
Central African Repub.	3.8		Mauritania	2	
Chad	5		Mauritius	5.2	
Comoros	2		Morocco	2	
Congo	5	!	Mozambique	2	
Congo, DRC	11.6	!	Namibia	11	
Cote d'Ivoire	4	!	Niger	3.9	
Djibouti	1		Nigeria	5	!
Egypt	1.6		Saudi Arabia	3	
Equatorial Guinea	4		Senegal	4	
Eritrea	3		Sierra Leone	3.7	!
Ethiopia	8.2		South Africa	12.5	
Gabon	3.2	!	Sudan	14.8	
Ghana	4	!	Syria	4	
Guinea	2		Tanzania	3.18	!
Guinea-Bissau	3.9		The Gambia	4	
Iran	10		Togo	1	
Iraq	4		Uganda	8	!
Jordan	2		Zambia	4	
Kenya	6.8	!	Zimbabwe	2	
Kuwait	2				

Power rank in EPR The Power rank index is provided by the EPR Core dataset and is defined as:

- The group rules alone:
  - Monopoly: Elite members hold monopoly power in the executive to the exclusion of members of all other ethnic groups
  - Dominance: Elite members of the group hold dominant power in the executive but there is some limited inclusion of ”token” members of other groups who however do not have real influence on decision making.
- The group shares power:
  - Senior Partner: Representatives of the group participate as senior partners in a formal or informal power-sharing arrangement. By power sharing, we mean any arrangement that divides executive power among leaders who claim to represent particular ethnic groups and who have real influence on political decision making.
  - Junior Partner: Representatives participate as junior partners in government
- The group is excluded from power:
  - Self-exclusion: The special category of self-exclusion applies to groups that have excluded themselves from central state power, in the sense that they control a particular territory of the state which they have declared independent from the central government
  - Powerless: Elite representatives hold no political power (or do not have influence on decision making) at the national level of executive power, without being explicitly discriminated against.
  - Discrimination: Group members are subjected to active, intentional, and targeted discrimination by the state, with the intent of excluding them from political power. Such active discrimination can be either formal or informal, but always refers to the domain of public politics (excluding discrimination in the socio-economic sphere).

Table B.2 shows the political power distribution of in the extended and restricted samples.

Table B.2: Power rank in EPR

	Government groups			Other groups			
Status	Monopoly	Dominant	Senior Partner	Junior Partner	Self-Excluded	Powerless	Discriminated
Rank	7	6	5	4	3	2	1
Share of obs. extended	4.79%	15.73%	79.48%	47.04%	0.55%	34.07%	18.31%
Share of obs. restricted	0	8.58%	91.52%	56.4%	0	30.69%	12.84%

Detailed manual checks on government's ethnicity

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 by the military against the Peul. Malinke and Peul have historically clashed. It appears right to depart from EPR classification and to assume that the Malinke were dominating the government in 2009. Further support for this is that the Malinke were the dominant group after 2009.

Liberia 1990-6: the First Liberian Civil War EPR classifies the country as in a state of 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 were 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. In 1991 the NPFL set up 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 launched an assault on Monrovia but was pushed back.

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

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

Liberia 2004-5 EPR states that all of the country's relevant ethnic groups were 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".

Sierra-Leone 1993-6 EPR states that the country is in state collapse. There was a military coup by 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 was a civil war from 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 a state of collapse. A new military coup against Kabbah (from the Mende group) and The Armed Forces Revolutionary Council (AFRC), allied with the RUF, was 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) was reinstalled by ECOMOG forces in March 1998 but the civil war continued, 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 was more moderate than the previous one: he had 4 vice-presidents each one representing a different faction. Kabila was supported by two related ethnic groups which were both coded as Senior Partners: Lunda-Yeke and Luba Shaba which have are related and formed Kabila's main power base. Here we consider them as allies.

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

Kenya 2008-11 Ethnicity was dominant in Kenyan politics: political parties were organized 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 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 were however marked by a fight against the Somali ethnic group suspected of supporting 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 held a larger share in government posts which was compensated by the Tutsi's hold over the army. This equal sharing, after a historical hostility between the two ethnic groups, was 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 were 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 was however "at least one interest group claiming to represent the interest of an ethnic group" until 2001. There was 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 was 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 to be different ethnic group. The three ethnicities corresponding to the three islands were Senior Partners in 2002. Each island had a president and they were vice-presidents in the Comorian Union government. The office of the president of the Union rotated among the three islands in 4-year terms.

C Military measure

C.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 incorporate 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 incorporate 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 incorporate 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 to be 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.

C.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 is increased;
- S_4: Random Forests' column sampling is decreased;
- S_5: Random Forests' column sampling is increased; Boosted Decision Tree's learning rate is increased; Boosted Decision Tree's total number of allowed trees is decreased;
- S_6: in both Random Forests and Boosted Decision Tree the row sampling rate is increased.

Table C.3: Robustness check: algorithm's parameters.

Models	S_1	S_2	S_3	S_4	S_5	S_6
$\mathrm{corr}(m_{S_0}, m_{S_i})$	99.09	99.10	99.18	99.87	98.86	99.52
$\Delta_{\mathrm{PRL}}^{S_o,S_i}$	2.97	2.62	2.94	0.01	0.15	0.07

C.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 the combination of 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 too scarce to use 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 $\mathcal{Z} \in \mathsf{M}^{N\$L}$ storing the predictions. The matrix $\mathcal{Z}$ and the original response vector $Y \in \mathsf{R}^N$ 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 $\Psi(\mathcal{Z}) = \mathsf{E}(Y|Z)$ 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 $\Psi = \arg\min_{\psi"\mathsf{P}} \sum_i \left(Y_i - \&(Z_i) \right)^2$ 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 $"_l \geq 0$ and $\|"\|_1 = 1$, that is a weighted average of the predictions, the weights (to each model) being determined to minimise the cross-validated risk.

[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 P is the space of multi-linear functions in X.

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 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 the base learner 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).

C.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 a 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 up 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.

C.3.2 Gradient Boosting Machine

A gradient boosting machine, like a 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, GBMs belong 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, the slower will the boosted model evolve and the lower will be the learning rate. The intuition behind a slow learning rate is that it allows slowly improvement of 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 slow 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 for with a high number of trees. The maximal depth of the trees allowed is spanned from 1 to 5.

C.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 to 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 $\| " \|_1 \leq t$ 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 attractive 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 $\| " \|_2 \leq t$ 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.