GROUPS IN CONFLICT:

Private and Public Prizes*

Laura Mayoral ⓡ Debraj Ray†

July 2019

This paper studies costly conflict over private and public goods. Oil is an example of the former, political and civil rights an example of the latter. Groups involved in conflict are likely to be small when the prize is private, and large when the prize is public. We examine these implications empirically by constructing a global dataset at the ethnic group level and studying conflict along ethnic lines. Our theoretical predictions find significant confirmation in an empirical setting.

1. Introduction

We study social conflict under multiple potential threats to peace. There are several potential groups, demarcated by one or more characteristics —- economic, ethnic, occupational or geographic. From these, a group might emerge to challenge the existing state of affairs. We address two issues:

1. Whether large groups or small groups are more likely to be involved in conflict against the State;

2. Whether our predictions regarding group size and conflict are supported by the data.

Which groups are likely to be involved in conflict? This is, of course, a question that cannot be answered in full generality, as the questions of identity and cohesion of various potential groupings are deep issues that can only be resolved through specific econometric and ethnographic research. But there is one aspect of a group that commands special attention, and that can be examined both theoretically and empirically: group *size*. Are large groups or small groups more likely to initiate conflict, or resist what are perceived to be the unfair incursions of the State?

The literature offers both answers. We are all aware of the "tyranny of the majority" (see, e.g. Tocqueville 1835), in which a larger group can impose its will on society even on issues that a relative minority might feel very strongly about. The tyranny expresses itself most clearly in a voting context, for after all, voting is an expression of ordinal preferences, and not the intensity of those preferences. But it is certainly not limited to voting. The suppression of minorities via extra-democratic channels, including coercive and violent means, is extremely common.

*We are grateful to Oeindrila Dube, William Easterly, Joan Esteban and Sahar Parsa for helpful comments on an earlier draft. Ray thanks the National Science Foundation for research support under grant numbers SES-1261560 and SES-1851758. Mayoral gratefully acknowledges financial support from the Generalitat de Catalunya, and the Ministry of Economy and Competitiveness Grant number ECO2015-66883-P.

†Mayoral: Instituto de Análisis Económico (CSIC), Barcelona GSE and University of Gothenburg; Ray: New York University and University of Warwick

But there is a contrasting view which argues that small groups may be more involved than large groups in lobbying or conflict (see Pareto 1927 and Olson 1965). For more on these matters, see, e.g., Chamberlin (1974), McGuire (1974), Marwell and Oliver (1993), Oliver and Marwell (1988), Sandler (1992), Taylor (1987) and Esteban and Ray (2001a). This literature studies the intensity of conflict displayed across small and large groups, *assuming that there is conflict to begin with*. In contrast, we ask whether a small or large group is willing to enter a conflict, or to resist a perceived act of aggression. This is a more subtle issue — after all, it is generally the case that large groups continue to have better chances of winning the conflict. But the group with the better chances is not necessarily the one to get involved. Rather, the entry into conflict depends on the expected payoff to a group, *relative to its received allocation otherwise*.

For concreteness, suppose that an ethnic group in a country has oil reserves located in its homeland. This is our leading case of a *private prize*. Suppose that the revenue from oil is distributed equally across the entire country. Or suppose that the homeland itself is settled by other ethnicities in the country. Then the "peacetime revenue share" of our ethnic group will correspond to the allocation of those revenues across the population and therefore, by implication, over ethnicities.

Likewise, suppose that the prize is *public*. It could represent political power and the payoffs that go with it. Or it could represent funding for secular versus religious infrastructure; e.g., public schools versus churches and temples. It could even represent private gains that are relatively undiluted by the number of recipients; e.g., changes in the tariff structure, public-sector job allocation or the awarding of government licenses to a favored ethnicity. In each of these last cases the good in question is excludable for the population as a whole, but as far as a particular ethnicity is concerned, it may have the dominant features of a public good. Then again, just as in the private case, we have a peacetime allocation of the public good.

In either case, if the group is involved in a war of secession to retain the private good, or in a war for control of the State apparatus, its chances of winning will be some function $p(m)$, where m is the population share of the group. The question is not whether $p(m)$ is small or large, but how large it is *relative* to m (and net of any costs of conflict). That will determine the decision to get involved in a conflict — either to initiate, or depending on the context at hand, to resist.

Our main theoretical result can be described as follows: under the assumption that the peacetime allocation is non-discriminatory — that is, contestable resources are equally allocated — conflict is more likely to be associated with small groups when the prize in question is private, but more likely to be associated with large groups when the prize is public. See Propositions 1 and 2. (We extend the analysis to discriminatory allocations in Section 3.3.) This is the central prediction — one that restricts the *interactions* across appropriate variables — that we take to the data.

Our empirical study focuses on groups that are defined along ethnic lines. Ethnic conflict is a natural choice for the study, as groups demarcated by ethnicity account for between 50–75% of internal conflicts since 1945 (Fearon and Laitin, 2003; Doyle and Sambanis, 2006). To conduct the analysis, we construct a panel dataset at the ethnic group level with global coverage. The dataset contains information for 145 countries and 1475 ethnic groups spanning the years 1960 to 2006.

The data is replete with examples of both public- and private-goods conflict; often mixtures of the two. The typical ethnic conflict could involve a struggle for political power or control (as in Burundi, Bosnia, Liberia, or Zimbabwe), but it can involve secessionist struggles by groups seeking to control their own land or resources (Chechnya, Kashmir, Tamils in Sri Lanka, the Casamance in Senegal, and many other examples). Land and oil are often central among these resources (e.g., the Ijaw conflict in Nigeria, the Darfur conflict, or the Second Civil War in the Sudan). Our empirical strategy, which we discuss in more detail later, is to allow for possible mixtures of public and private conflicts and then to tease out these private and public components of the conflict.

To obtain a proxy for private payoffs, we consider rents that are easily appropriable. Because appropriability is closely connected to the presence of resources, we approximate the degree of privateness in the prize by asking if the homeland of the ethnic group is rich in natural resources. In our baseline specification we use oil abundance in the homeland as a proxy for privateness, but we also consider alternative measures based on mineral and land abundance, again at the ethnic group level.

In similar vein, we use specific measures of public payoffs. Our proxies rest on the idea that the public payoff from conflict will be large when groups don't have access to their (preferred) public goods. Because of the non-excludable character of these goods, our proxies aim to capture the *lack* of public goods in society (rather than its abundance, as in the private payoffs case). In our baseline specification we capture this by using a pre-sample measure of the *lack of political and civil rights*, constructed by Freedom House, which is a country-level index measuring the rights effectively enjoyed by the citizens. We also employ alternative proxies based on group exclusion from power, lack of democracy (as in Esteban et al. 2012), or lack of basic public services.

Our results appear to firmly support the predictions of the theory: smaller ethnic groups are more likely to be involved in conflict when oil, minerals and/or land are abundant for the group. At the same time, using the specific measures of publicness just described, larger ethnic groups are more likely to participate in conflict when the valuation of the public payoff is high. Table 1 contains these core results across different specifications. It is particularly noteworthy that while we derive these results off interaction effects (between size and prize), the findings fully survive a more flexible specification that goes beyond interaction in an ambient linear model.

These core results are consistently backed up by a large number of robustness checks, and further supported by some ancillary results. Robustness checks include the consideration of alternative conflict variables, estimation strategies and ways of proxying for the prizes at stake, both private and public. There are a number of ancillary results. For instance, once the private prize and its interaction with group size have been accounted for, the coefficient on group size turns positive and significant (it is insignificant if entered on its own). In parallel vein, if per-capita payoffs are held constant, large groups are always more prone to fighting, irrespective of the type of payoff at stake. (Empirically, the role of group size changes dramatically depending on whether one controls for total or per-capita private payoffs, in line with the theoretical predictions.) We also find results in line with the extended analysis in Section 3.3: conflict over private payoffs is more likely as the number of potential threats to conflict increases.

Of course, it is well known in the empirical literature that the presence of natural resources — particularly oil — is correlated with conflict; see, for example, Le Billon (2001), Fearon (2005), Lujala (2010) and Dube and Vargas (2013). Morelli and Rohner (2015) show, additionally, that the *concentration* of those natural resources in ethnic homelands is related to conflict. As in the Morelli-Rohner paper, our empirical study is set at the ethnic group level. But the question we ask is different: our focus is on the *interaction* between group size and the homeland resource variable. In addition, as already described, we are equally interested in the public payoff variable and its interaction with group size. To our knowledge, neither interaction has been explored empirically in the literature. Together, they reconcile the Tyranny of the Majority with the Pareto-Olson thesis.

Section 2 introduces a baseline model of conflict. Section 3.1 analyzes the relation between group size and conflict when conflict is over a private prize. Sections 3.2 does the same for public prizes. Our main empirical results are presented in Section 5. Section **??** considers alternative explanations that could rationalise our empirical findings and provides evidence against them. Section 6 contains additional variations that examine the robustness of the results. See Appendix A for detailed definitions of all the variables considered in the empirical analysis as well as a table of summary statistics. Appendix B contains additional empirical results. Section 7 concludes.

2. A Model of Conflict

2.1. **Allocations.** Denote by v the total "appropriable resources" of society. This value is transferable to different degrees across a population of individuals, normalized to have unit mass. Let X be the resulting set of efficient payoff allocations $\mathbf{x} = \{x(i)\}$ that can be generated by v. A special, salient allocation of appropriable resources is the *non-discriminatory* allocation under which everyone receives an equal payoff. We assume that this allocation is feasible.

The value v could represent material resources such as oil from a particular geographical location within the society, or less material payoffs such as the acquisition of political or cultural power. Assume that the State (or society as a whole) seeks to allocate v widely over the entire community, or perhaps according to the wishes of a dominant group; say, a majority. But there is a subgroup, demarcated by ethnicity, geography, religion or occupation, which seeks to retain — or seize — the proceeds of v for itself. For instance, one might think of v as the value of oil reserves located within the homeland of an ethnic group. The State wants to distribute those revenues over the entire country, while the ethnic group might feel that this is "their oil."

2.2. **Conflict.** The group can accede to the peaceful allocation, or its members can engage in costly conflict. In the case of conflict, we suppose that society is partitioned into two subsets, one of size m (pertaining to the group in question) and the remainder — represented by the "State" — of size $\overline{m}$ ($m + \overline{m} = 1$), and that they engage in a bilateral conflict. In short, our group does battle against society as a whole, with the complementary group interpreted as the incumbent State.

We leave open the interpretation of whether our group "initiates" conflict or "defends itself" against what it perceives to be the incursions made by the State. That will depend on the situation at hand. For instance, if there is settlement on the group's territory, conflict may be interpretable as defense against State aggression. If the group is fighting to overthrow the State and seize power, then the group may be viewed as the aggressor. We sidestep these interpretations altogether and simply refer to the two groups as "Rebel" and "State."

Conflict involves — on each side — the expending of effort or resources. The utility cost to an individual from a contribution of r is given by

$$c(r) = (1/\alpha)r^{\alpha}$$

for some $\alpha > 1$.[3] We will presume that the winning party — Rebel or State — obtains full control over the appropriable resources. Therefore, it is assumed that a leader on each side extracts these resources from everyone to maximize the per-capita payoff of her coalition.[4] In particular, we proceed as if the State entirely abandons the interests of the Rebel group, though variants that allow the State to bring Rebels (partially or fully) back into the fold yield the same results. Because the cost of effort provision is strictly convex, the leader will ask for equal effort from each individual, and will make transfers if needed to compensate them.

To map efforts into win probabilities, we use contest success functions (Skaperdas 1996), so the probability that the Rebel will win is given by

$$p = \frac{mr}{R},$$

where r is contribution per person in the Rebel, and $R = mr + \overline{mr}$ is the sum of contributions made by both the groups. (Throughout, we use bars on the corresponding variables for the State.) As in the case of the cost function, this specification too can be substantially generalized.

Letting π stand for the per-capita payoff conditional on winning, and normalizing loss payoffs to zero, the Rebel seeks to maximize its expected payoff

$$\pi\frac{mr}{R} - c(r),$$

A similar problem is faced by the State, with payoff $\overline{\pi}$ conditional on winning and 0 conditional on losing. A conflict equilibrium is a Nash equilibrium of this game. Such equilibria are fully described by the first-order conditions

$$\pi m\overline{m} = R^2\frac{r^{\alpha-1}}{\overline{r}} \tag{1}$$

for the Rebel, and by

$$\overline{\pi} m\overline{m} = R^2\frac{\overline{r}^{\alpha-1}}{r} \tag{2}$$

[3]Nothing of substance hangs on the specific choice of cost function. Strict convexity of cost is important, however.

[4]To be sure, this neglects the free-rider problem or the question of intra-group *cohesion*, which is another aspect of small versus large groups worth studying, though we don't do so here. It is easy to write down variants of our model in which individuals unilaterally make resource contributions, provided that they at least partially internalize the payoffs of their fellow group members (see Esteban and Ray 2011).

for the State. Conditions (1) and (2) yield a simple expression for the provision of individual resources by the group, relative to its rival:

(3) $$\frac{r}{\overline{r}} = \left(\frac{\pi}{\overline{\pi}}\right)^{1/\alpha} \equiv \gamma.$$

We can use these conditions to describe the conflict payoff of each group. For the Rebel, rewrite (1) to observe that

$$r^{\alpha} = \pi p \overline{p},$$

so that the expected payoff from conflict is given by

(4) $$\pi p - c(r) = \pi p - (1/\alpha)\pi p \overline{p} = \pi[kp + (1-k)p^2],$$

where $k \equiv (\alpha - 1)/\alpha$, which lies in $(0, 1)$. Finally, note that

(5) $$p = \frac{mr}{mr + (1-m)\overline{r}} = \frac{m\gamma}{m\gamma + (1-m)},$$

where γ is defined in (3). Together, (3), (4) and (5) describe a full solution to the Rebel's payoff in conflict equilibrium. A parallel expression holds for the State.

Conflict is a threat to peace, and we seek conditions under which that threat might manifest itself. That will depend to some degree on what the peaceful allocation is. In the main analysis we take that peaceful allocation to be non-discriminatory, but we remark on extensions. Say that a peaceful allocation $\mathbf{x} \in X$ is *blocked* if the expected payoff to the Rebel under conflict exceeds its average payoff under the allocation:

(6) $$\pi[kp + (1-k)p^2] > \frac{1}{m}\int_{\text{Rebel}} x(i).$$

We wish to understand whether small or large Rebels are more likely to be involved in conflict. To address this question, we must first link the appropriable surplus proxied by v to the victory payoffs π and $\overline{\pi}$ for each group. We do so by conducting the exercise in more detail for two leading cases: one in which the prize is a divisible, private good, and the other in which the prize must be used to provide public goods. We take the resulting predictions to the data.

3. Group Size and Conflict

3.1. **Private Goods.** Little by way of additional interpretation is needed when the entire prize v is a private good; say, oil located on the homeland of the (potential) Rebel. Now X is just the set of all distributions of v among the population: $X = \{\mathbf{x} | \int x(i)di = v\}$.

We assume that the winning group seizes the resources v entirely and excludes losers from the division of the spoils. Therefore, with a Rebel of size m,

$$\pi = v/m \text{ and } \overline{\pi} = v/(1-m).$$

Using this information in (3), we see that

$$\gamma = \left(\frac{1-m}{m}\right)^{1/\alpha},$$

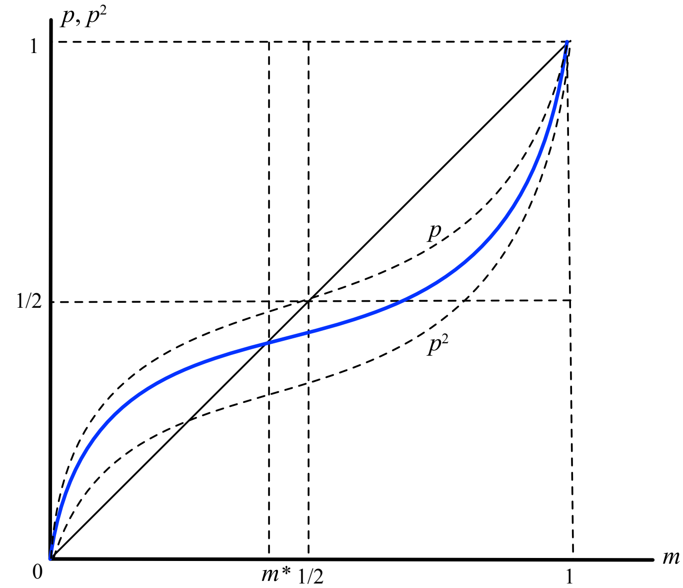


Figure 1. Threshold for Conflict with Private Prize and Non-Discriminatory Allocation.

so that by (5),

$$p = \frac{m^k}{m^k + (1-m)^k}, \tag{7}$$

where $k = (\alpha - 1)/\alpha$.

Notice from (7) that smaller Rebels are disadvantaged in conflict in the sense that they have a lower probability of winning; after all p is increasing in m and $p(1/2) = 1/2$. Nevertheless,

Proposition 1. *Assume that the prize is private. Then there exists $m^* \in (0, 1/2)$ such that a Rebel with $m < m^*$ will block the non-discriminatory allocation and engage in conflict. Thus society is conflict-prone in the presence of smaller Rebels.*

The proof that follows may be worth reading as part of the text, as it provides some intuition, tells us how m^* is calculated, and suggests how the results extend to the case of discriminatory peaceful allocations.

Proof. The non-discriminatory allocation gives v to every player. Using (4), conflict payoff is given by $\pi[kp + (1-k)p^2] = v[kp + (1-k)p^2]/m$. So a Rebel of size m will block if

$$kp(m) + (1-k)p(m)^2 > m, \tag{8}$$

where $p(m)$ is given by (7).

The function p has a "reverse-logistic" shape. It starts above the 45^0 line and at the point $n = 1/2$ crosses it and dips below. The derivatives at the two ends are infinite.[5] See Figure 1, which plots p, p^2 and the convex combination $kp + (1-k)p^2$. With this shape in mind, observe that the left-hand side of (8) starts out higher than the right-hand side for small values of m, but ends up lower. Note that

$$kp(m) + (1-k)p(m)^2 < m,$$

for *any* $m \geq 1/2$.[6] This observation, in conjunction with Figure 1, shows that there is a unique intersection (crossing from above to below) in the interior of $(0, 1/2)$.[7] The proof of the proposition is now complete. ■

Notice that what matters is *not* the level of win probabilities or whether it increases or falls with group size. In fact, it always increases with size. While small Rebels fight more intensely (the per-capita stakes are higher), this does not overturn the fact they have a lower probability of winning than big groups do. Thus small groups engage in conflict not because they have a high chance of winning. (They don't.) Rather, they do so because they have a high chance of winning *relative* to their share from the non-discriminatory allocation. That fact is reflected in the reverse-logistic shape of the win probability, derived in the proof of Proposition 1.

We reiterate that we do not interpret this result as a small Rebel deliberately initiating conflict in some "unprovoked fashion." Indeed, in the empirical implementation below, the prize will refer to resources located in the homeland of some ethnic group. The "non-discriminatory allocation," in which a State attempts to control these resources in order to redistribute its revenues to the country at large, can be viewed by the group in question as an unwarranted infringement of its rights (to the resource). In that case, the correct interpretation is not one of conflict initiation, but rather one of resistance.[8]

3.2. **Public Goods.** Suppose, now, that v is a budget for the production of public goods. As described in the Introduction, it could represent political power, the restoration of rights, or funding for secular versus religious infrastructure. It could even represent private gains that stem from public policies, such as protectionism or the reservation of public-sector jobs.

Suppose there are n *disjoint* groups, each with their favored public good on which the budget can be spent. (We will return to the disjointness assumption in Section 3.3.) A budget allocation is just $\mathbf{v} = (v_1, \ldots, v_n)$, representing resources going to each group and summing to v. Assume that an individual gets payoff 1 from each unit of the budget spent on a group where she has membership; otherwise, she gets zero. (This payoff structure is only for expositional ease and can be easily generalized.) Then, given that each person belongs to just one group, the payoff to

[5] To check these claims, note that $\frac{m^k}{m^k + (1-m)^k} \geq n$ if and only if $m \leq 1/2$ (simply cross-multiply and verify this), and that $p'(m) = \frac{km^{k-1}(1-m)^{k-1}}{[m^k + (1-m)^k]^2}$, which is infinite both at $n = 0$ and $n = 1$.

[6] Suppose this is false for some $1 > m \geq 1/2$. By the properties of p already established, we know that $m \geq 1/2$ implies $m \geq p(m)$, so that $km + (1-k)m^2 \geq m$, but this can never happen when $m < 1$, a contradiction.

[7] More formally, the derivative of $kp(m) + (1-k)p(m)^2$ is strictly smaller than 1 at any intersection, so that there can be only one intersection; we omit the details.

[8] We should be also careful not to take Proposition 1 as literally applying to *all* group sizes, however small. Obviously, the model ignores the fact that *some* minimum threshold size is needed to even pose a serious threat.

a person from a budget allocation $\mathbf{v}$ is v_j, where j is her group membership. So X is now the set of payoff allocations $\mathbf{x}$ that can arise from all budget allocations. These are all step functions across groups. The non-discriminatory allocation is given by dividing the budget equally across all groups, so that each individual obtains a payoff of v/n in peacetime.

A Rebel who wins a conflict gets to implement its own good, so that $\pi = v$. Normalize the value to the State to be zero. Assume that if the State wins, it excludes the Rebel and implements the non-discriminatory allocation for everyone else, with payoff $v/(n-1)$. The crucial point is that in the public prize case, group population size is eliminated as a determinant of per-capita payoff. An amount v_j spent on the favorite public good of a group j yields each member of that group v_j no matter what the group size is.

Proposition 2. *Assume that the prize is public and all relevant allocations are non-discriminatory. Then there exists $\hat{m} \in (0,1)$ such that a Rebel with $m > \hat{m}$ will block the resulting payoff allocation and engage in conflict. Society is conflict-prone in the presence of larger Rebels.*

Proof. Consider any conflict involving a Rebel of size m and the State of size $\overline{m} = 1 - m$. Then (5) tells us that

$$p(m) = \frac{m\gamma}{m\gamma + (1-m)}, \tag{9}$$

where $\gamma = [\pi/\overline{\pi}]^{1/\alpha} = (n-1)^{1/\alpha}$ is independent of m. Using (6) for the nondiscriminatory allocation with payoff v/n per-capita, we see that the Rebel will wish to engage in conflict if

$$kp(m) + (1-k)p(m)^2 > 1/n. \tag{10}$$

Given (9), the left-hand side of this inequality is monotonically increasing in m. For m close to zero, the inequality must fail because $p(m) \to 0$, and for m close to 1 the inequality must hold because $p(m) \to 1$. Define $\hat{m}$ by equality in the relationship above to complete the argument. ∎

Numerical calculations are easy to perform. Combining (9) and (10) and remembering that $\gamma = (n-1)^{1/\alpha}$, the blocking condition for conflict reduces to

$$k\frac{m(n-1)^{1/\alpha}}{m(n-1)^{1/\alpha} + (1-m)} + (1-k)\left[\frac{m(n-1)^{1/\alpha}}{m(n-1)^{1/\alpha} + (1-m)}\right]^2 > \frac{1}{n},$$

and some straightforward but tedious computation eventually reveals that

$$\hat{m} = \left[1 + (n-1)^{1/\alpha}\left\{\frac{(1+\alpha) - \sqrt{(\alpha-1)^2 + \frac{4\alpha}{n}}}{\sqrt{(\alpha-1)^2 + \frac{4\alpha}{n}} - (\alpha-1)}\right\}\right]^{-1}. \tag{11}$$

For instance, when there are just two groups and the cost function is quadratic, then the Rebel needs to exceeds 61.8% of the population. When there are three groups and $\alpha = 1.2$, then the Rebel needs to exceed 39.7% of the population. We can use (11) to perform these calculations for any number of groups and any curvature of the cost function, but the point should be clear: it is large groups (typically but not always larger than the average) that pose a threat when the potential conflict is over public goods.

3.3. **Non-Discriminatory Peaceful Allocations and Multiple Threats.** Our analysis so far presumes that peacetime allocations are non-discriminatory. Of course, Proposition 1 applies even more strongly if society has a reason to favor larger groups to begin with, as it will in a democratic (or voting) scenario. But if the initial allocation is chosen to appease the small groups, then it is the larger groups who will have to pay for that appeasement, and matters are more complex.[9]

Suppose that there is a variety of potential markers (religion, caste, occupation, ethnicity, geography, and so on) that might delineate a potential Rebel coalition. To formalize the idea of multiple threats, say that a finite collection $\mathcal{C}$ of groups (or potential Rebels) is *balanced* if there is a set of weights in $[0, 1]$, $\{\lambda(G)\}_{G\in\mathcal{C}}$, such that

$$\sum_{G\in\mathcal{C}_i} \lambda(G) = 1 \text{ for every } i \text{ in society,} \tag{12}$$

where $\mathcal{C}_i$ is the subcollection of all groups for which i is a member.

Essentially, balancedness implies that it is hard to "buy off" small groups of individuals who are central to all potential conflicts. It assures us that there is no such "central group." For instance, suppose that $\mathcal{C}$ is fully described by any collection of potential Rebels that contain the special set of individuals $[0, 1/2]$. Then that collection is not balanced: we relegate the details to a footnote.[10] It contains some distinguished group (in this example, $[0, 1/2]$) which is "over-represented" in the collection. In contrast, a balanced collection contains no "over-represented" group. For instance, any partition of $[0, 1]$ is a balanced collection (simply use $\lambda(G) = 1$ for all G and verify that the balancing condition is satisfied).[11] We can now state:

Proposition 3. *Assume that the prize is private. Suppose that the collection of all potential Rebels includes a balanced collection $\mathcal{C}$, with each member of cardinality $m < m^*$, where m^* is given by Proposition 1. Then every peaceful allocation, non-discriminatory or otherwise, is blocked by some member(s) of this collection.*

Proof. See Appendix A.

Because (as already noted) every partition is balanced, the following corollary applies:

Corollary 1. *Suppose that society can be partitioned into potential Rebels of size $m < m^*$. Then there is no allocation for society that is immune to conflict.*

---

[9] Discriminatory peacetime allocations are of separate interest because of the Coase Theorem. Because conflict is costly, for each conflictual outcome there is a "peaceful" outcome that Pareto-dominates it, provided that appropriate Coaseian transfers are available. But is there *one* outcome that can *simultaneously* withstand all threats? It is true that conflict is inefficient, but if the variety of potential threats is large relative to the degree of inefficiency, *every* peacetime allocation, discriminatory or not, may be blocked by *some* coalition. This is akin to the problem of an empty core in characteristic function games (Bondareva 1963, Shapley 1967, and Scarf 1967).

[10] For suppose we could find "balancing weights" $\{\lambda(G)\}$; then, in particular, (12) must hold for any $i \in [0, 1/2]$, but since i is contained in every $G \in \mathcal{C}$, this implies that the *entire* set of weights add to 1: $\sum_{G\in\mathcal{C}} \lambda(G) = 1$. Now pick any G' with $\lambda(G') > 0$. Because G' is a strict subset of $[0, 1]$, there is some individual $j \notin G'$. Given (14), it must be the case that $\sum_{G\in\mathcal{C}_j} \lambda(G) < 1$, which contradicts balancedness.

[11] Or, if $[0, 1]$ is the union of K equally-sized intervals of the form $[i/K, (i + 2)/K]$, for $i = 0, \ldots, K - 1$, then the collection $\{[0, 2/K), [1/K, 3/K), [2/K, 4/K), \ldots, [(K-2)/K, 1), [(K-1)/K, 1/K)\}$ has "overlaps" but is also balanced.

Notice that we do not place any assumptions on the peacetime allocations. They could be *any* allocation of the private good, perhaps discriminating across individuals in the same coalition. And yet, if there is a sufficiently varied multiplicity of small groups all challenging the private prize, society is necessarily unable to find a peaceful allocation that buys off all potential Rebels.[12] Of course, it is possible that for some particularly unequal allocations, a large group may also want to instigate a conflict. The point is that in such a case, *some* small group also will — under the conditions of Proposition 3.

There is a parallel analysis to private goods in the spirit of an "empty core" that we can easily conduct for public goods. In fact some of that analysis is simpler. That is because for any arbitrary peacetime allocation of the budget across different public goods, it always remains the case that both the per-capita payoff gap between victory and defeat, as well as the defeat or victory payoffs on either side, are independent of the size of the Rebel m. For this reason, Proposition 2 survives with no essential change whether or not the initial allocation is non-discriminatory. Of course, the threshold $\hat{m}$ will change with the relevant parameters, including the size of the peacetime offer, but the qualitative result survives with no alterations.

Once we allow for arbitrary allocations, there is also no need to assume that the groups are pairwise disjoint. That assumption was only used to assure ourselves that a non-discriminatory allocation always exists. With group intersections, a non-discriminatory allocation may not exist in the first place,[13] but the brief discussion here assures us that it does not matter.

But we also take note of an importance difference between the two cases. With public goods, we need to be especially careful about the transferability of payoffs and exactly what it entails. We have restricted ourselves to the case in which budgets are transferable across groups, not in units of money, but by changing the allocation of public goods. One might allow for a broader class of transfers in which compensatory side-payments of money are made from one group to another in exchange for an uneven distribution of public goods.[14] The analysis of this case is somewhat different but yields similar results. (The details are available on request from the authors.)

4. GROUP SIZE AND CONFLICT: EMPIRICS

We now turn to the empirical relationship between group size, the public or private nature of the payoffs, and conflict. Our theory implies that the impact of group size on conflict is negative if

[12]It should be noted that the balancedness condition on potential Rebels, while sufficient, is not necessary for the conflict result. For instance, for groups that are smaller than the threshold m^*, extra per-capita surplus is available in the event of conflict. For instance, suppose that the cost function is quadratic (so that $\alpha = 2$). It is then easy to verify that $m^* = 1/4$. However, groups of size 10% make a strict gain from blocking a non-discriminatory allocation. It is possible to check that if there are six such pairwise disjoint groups, conflict is inevitable regardless of the baseline allocation: no such allocation can be stable.

[13]For instance, if there are two groups that intersect but neither group is a subset of other, a non-discriminatory allocation will not exist, as members of the intersection will benefit to a greater degree from any allocation.

[14]This possible alternative approach with transferability should be used with caution. Public goods are not like oil revenues. Think of ethnic or religious representation, or the sharing of political power. The relative price across objects such as these may be very hard to define. So it may be impossible to conceive of "classical" financial transfers as compensation for the loss of power or culture; see, e.g., Kirshner (2000). What price would those who are thus negated accept as compensation?

the prize is private, and positive if the prize is public. There are several considerations that arise when using the data to address the theory. These include, but are not limited to, a suitable definition of "groups," as well as a classification of conflicts into their "private" and "public" payoff components. We also need to be careful about transplanting the notion of "conflict initiation" to the data. Below is a discussion of these and related issues.

4.1. **Empirical Overview.** The first empirical question is how to choose the social cleavages that define potential Rebel groups. We settle for ethnicity, and study ethnic conflicts. Given that such conflicts account for between 50–75% of internal conflicts since 1945 (Fearon and Laitin 2003, Doyle and Sambanis 2006), this appears to be a natural and relatively tractable choice. Esteban and Ray (2008) provide theoretical reasons for ethnic markers to be salient in conflict.

The second question has to do with the definition of a private or public prize in conflict. As a proxy for a private prize, we consider resources that are located in the homeland of each ethnic group. In our baseline specification the resource is oil, but we also consider other minerals as well as the land area of the homeland itself. The underlying presumption is that the State seeks to extract those resources (or output from the homeland) and distribute them more evenly across the country, and that the ethnic group in question can either accept the State policy, or reject it. Note that ethnic homelands are *not* countries; a variant would be to consider resources at the national level. We do so in Table B6, obtaining similar results.

To implement the idea of a public prize, we consider several proxies for the *lack of access* to public goods. The public prize is then represented by the potential expansion of such public goods if the battle against the State is won.[15] In our baseline specification we consider the extent of rights and freedoms available to the citizens and use a pre-sample index capturing the *lack of political and civil rights* (Freedom House); see details below.

This baseline specification for public prizes is augmented by alternative proxies. As in our baseline, each such proxy seeks to measure the gains from seizure of political power at the Center. The first is the pre-sample average of an autocracy index from Polity IV. The idea is that the more autocratic the State, the less is the sharing of power, and consequently, the higher is the valuation — to an outsider group — of controlling the State. Or perhaps the disaffected who seize power simply want to get rid of the shackles of authoritarian rule and install a democratic government. Either interpretation will do. A second proxy — with the same interpretation — is a group-level variable that directly measures whether a group is excluded from executive power at the national level. (Only pre-sample information is used.) A third measure aims to capture the low provision of basic public services, as measured by infant mortality rates pre-sample.

The conflict data we employ is a subset of the UCDP/PRIO Armed Conflict Dataset. It records conflicts between ethnic groups and the State. The conflict measure comes in two flavors: *incidence* or *onset*. Briefly, a case of incidence records all conflict in a given time period, whether it is new or ongoing, while a case of onset records just the former. Below, we take our baseline model to be one of incidence, though we explore variations that use onset and alternative conflict variables (see Table B5). We now turn to a more detailed description of the data.

[15] Notice the contrast with the case of private payoffs, where groups fight for the appropriation of *existing* resources.

4.2. **Details.** We construct a panel dataset at the ethnic group level with global coverage.[16] It contains information for 145 countries and 1475 ethnic groups over 1960 to 2006.[17]

4.2.1. *Ethnic Groups.* We use the sample of ethnic groups from the dataset "Geo-Referencing of Ethnic Groups" (GREG); see Weidman, Rod and Cederman (2010). The GREG dataset provides detailed geographical location of ethnic groups for the whole world. This last feature enables us to merge with it other geo-referenced datasets needed for the computation of some of our key group-level variables. The GREG is based on the *Atlas Narodov Mira* or ANV (Bruk and Apenchenko, 1964), which was created by Soviet ethnographers in the early 1960 with the aim of charting ethnic groups worldwide. It provides information on the homelands of 929 groups and it employs a consistent classification of ethnicity with a uniform group list that is valid across state borders.[18] Most homelands are coded as pertaining to one group only, but in some instances up to three ethnic groups share the same territory.

The GREG extension of ANV permits us to create units that are group-country pairs: that is, we assign ethnic groups to countries depending on the land area occupied by them in each country.[19] When all is said and done, GREG contains a larger number of groups than alternative sources (such as the Geo-Ethnic Power Relations dataset) as it contains many small-language groups. There are 1475 distinct group-country pairs in the dataset, to be referred to from now on simply as "group." Our central variable, SIZE, is the size of the (country-specific) group relative to that of the country population.[20]

The fact that GREG's settlement patterns — and our consequent classification of groups — are a snapshot from the late 1950s and early 1960s has advantages and disadvantages. On the negative side, settlement patterns may be outdated for some parts of the world. Also, as ethnic maps were chartered by Soviet ethnographers during the Cold War, the level of accuracy and resolution varies considerably for different regions in the world. On the positive side, it alleviates concerns that ethnic group locations are endogenous to the conflicts we aim to explain.

4.2.2. *Conflict.* Data on group-level conflict has been taken from Cederman, Buhaug and Rod (2009), CBR henceforth.[21] Our baseline measure is group-level conflict *incidence*, which is equal to 1 in a given year if that group is involved in an armed conflict against the state, resulting

[16]This dataset is similar to that employed by Morelli and Rohner (2015) who consider similar sources for ethnic group location and oil fields.

[17]We focus on the post-1960 period as our data on ethnic group location and population are drawn from the start of the 1960s. In most of our regressions containing public prize proxies, the sample period is further restricted to post 1975 observations due to data availability restrictions, see Section 5 for details.

[18]The ANV actually contains information for 1248 groups, but 319 of them do not have any territorial basis.

[19]The definition of ethnic group is not clearly stated anywhere in the ANV so it is only possible to infer the coding criteria by comparison with existing data sources on ethnic groups. Fearon (2003) argues that the main criterion in the ANV for distinguishing between two groups is the historic origin of language.

[20]Population figures correspond to the early 60's, see Cederman, Buhaud and Rod (2009) for details.

[21]CBR use the UCDP/PRIO Armed Conflict Dataset (Gleditsch et al. 2002) and check this list against previous sources that identify ethnic civil wars (such as Fearon and Laitin 2003, Licklider 1995 and Sambanis 2001). Ethnic conflicts are coded based on whether mobilization was shaped by ethnic affiliation. Once a list of plausible conflicts was established, CBR code the various groups involved in each case.

in more than 25 battle-related deaths in that year. To explore the robustness of our results other conflict measures are employed. Group-level conflict *onset* is equal to 1 in a given year if an armed conflict against the state resulting in more than 25 battle-related deaths *begins* in that year.[22] For ongoing conflicts, onset is coded as missing. Finally, we collapse the time dimension of the data and compute for each group the share of years it has been involved in conflict against the State as well as the share of onset years.

4.2.3. *Prizes.* A key prediction of our theory is that the size of the group in conflict depends on whether the payoff is private or public. In order to test this hypothesis, proxies for the nature of the prize (or prizes) at stake are needed. To construct such proxies, we closely follow the approach in Esteban et al. (2012).

Private Prize. We ask if the ethnic homeland is rich in natural resources. In our baseline specification we use oil in the homeland as a proxy for "private prize". We also consider mineral availability and land abundance (see Table 3) as well as an index of group "resource wealth" (see Table 1), computed by applying factor analysis on the above mentioned proxies of private payoffs. The baseline measure, OIL, is computed as follows. First, geo-referenced information on the location of oil fields and associated discovery dates is obtained from Petrodata (Lujala, Rod and Thieme, 2007). Next, we combine the information on group and oil location from GREG and Petrodata, respectively, to construct maps of oil fields at the ethnic group level. Finally, OIL is computed as the log of the ethnic homeland area covered by oil (in thousands of square kilometres) times the international price of oil. Our results are robust to alternative ways of measuring oil abundance (see Table B6).

Notice that private prizes are firmly tied to ethnic homelands. While one ethnic group cannot directly reach out to seize resources located in another group's homeland, the State as a whole can, of course, attempt to redistribute the revenues from those resources over the country as a whole, or settle relatively abundant lands with other ethnicities. If violent conflict occurs in that process, our data will pick it up.

Public Prize. To obtain proxies for the public payoff, we ask if groups lacks access to several types of public goods. Our baseline specification is LACK RIGHTS, which is a composite index from Freedom House that measures lack of political and civil rights. It is designed to capture the real-world rights and freedoms enjoyed by individuals. In formulating the index, both the legal guarantees of rights as well as actual practices are taken into account. The index is measured on a scale from 0 to 7, where 7 indicates the lowest rights provision. We normalize this index to be between 0 and 1. LACK RIGHTS is the average of two individual indices, the lack of political

[22]In practice, conflict onset as defined by the PRIO threshold is far from a sharp concept. Before the threshold is crossed, we might have several manifestations of serious conflict (a breakdown in negotiations, an insurgency, a crackdown). Thus, a year of onset is arguably no different from a year of incidence, though to be sure, the factors that contribute to the outbreak of a conflict do not coincide with the ones that continue to feed it (Schneider and Wiesehomeier 2006). This is why we control for lagged conflict in our incidence regressions.

rights and the lack of civil rights; see Freedom House for additional details.[23] For robustness we will also consider these indices individually (see Table 4).

We deliberately take this measure off the shelf so as to avoid any implication that the components or weights are chosen to suit our purpose. We are also aware that there are concerns of endogeneity: for instance, conflict can lead to changes in the level of rights. Therefore, we *only* consider pre-sample values of the index (and in addition we control for past conflict in all our regressions). Specifically, our main "publicness" measure, LACK RIGHTS, is computed by averaging the values of the Freedom House index from the first year where it is available (which is 1972) to 1975, and is then employed in regressions using post-1975 data only.The resulting measure is "assigned" to all the ethnic groups in the country, so that LACK RIGHTS is a time-invariant country-level index, a price to be paid for achieving some acceptable degree of exogeneity. We check the robustness of our results by considering alternative proxies for the publicness index as well as alternative ways of operationalizing the above-described measures; see Table 4. For instance, we use a group-level proxy for publicness, which is group exclusion, and is defined as the average over the period 1960–1975 of a dummy variable that captures whether a group is excluded from national power (Cederman et al, 2009). Our results are robust to using these alternative definitions.

4.2.4. *Additional Controls.* We also consider a number of control variables, both at the group and at the country level. Group-level controls are obtained from Cederman et al. (2009) or directly computed from the GREG dataset. MOUNT is an index that captures the group's share of mountainous terrain. GROUPAREA is homeland area (in thousands of square km). DISTCAP measures the group's distance to the country capital. GIP is 1 if the ethnic group is in power in the country, lagged one year. SOILCONST measures the limitations of homeland soil for agriculture. PARTITIONED is 1 if the group's homeland is located in two or more countries. PEACEYRS is the number of years since the last group-level onset and LAG is lagged conflict incidence. At the country level we control for the log of GDP per capita, lagged one year (GDP), the polity index (POLITY), lagged one year, and the log of total population (POP), also lagged one year. GDP and population are taken from the Penn World Tables while POLITY is taken from Polity IV.

4.3. **Estimation.** We examine the effect of group size on conflict by running variants of the following specification:

$$\text{CONFLICT}_{c,g,t} = \beta_1 \text{SIZE}_{c,g} + \beta_2 \text{SIZE}_{c,g} \times \text{PRIV}_{c,g,t} + \beta_3 \text{PRIV}_{c,g,t} + \beta_4 \text{SIZE}_{c,g} \times \text{PUB}_c + X'_{c,g,t}\alpha + Y'_{c,t}\delta + Z'_c\gamma + W'_t\eta + \epsilon_{c,g,t}, \tag{13}$$

[23]The political rights index is elaborated taking into account i) the freedom in the electoral process, ii) political pluralism and participation, and iii) the functioning of the government. The civil rights index evaluates i) the freedom of expression and belief, ii) associational and organizational rights, iii) the rule of law, and iv) personal autonomy and individual rights. See `https://freedomhouse.org/report/methodology-freedom-world-2018` for additional details.

for countries $c = 1, \ldots, C$, groups $g = 1, \ldots, G_c$, and dates $t = 1, \ldots, T$. Our main outcome variable CONFLICT is "conflict incidence", as described above.[24] The variables PRIV and PUB are our measures of privateness and publicness, respectively, and their interactions with size are of particular interest. Our theory predicts that β_2, the coefficient associated with SIZE $\times$ PRIV, is negative, implying that smaller groups are more likely to be involved in conflict as the private prize in the homeland becomes more abundant.

As for the public prize, the measures use lack of political and civil rights, autocracy, group exclusion, infant mortality rates, as well as an index of publicness that summarizes all these variables, and predicts that β_4, the coefficient associated with the interaction of group size and PUB, is expected to be positive, so that the impact of group size on conflict increases as the public prize gets larger.

Unless otherwise stated, we always employ group- and country-level controls ($X_{c,g,t}$ and $Y_{c,t}$ respectively), a vector Z_c of country fixed effects and year dummies W_t. Identification for the interaction term SIZE$\times$OIL is achieved both because we have variation in ethnic groups within countries — so that size varies — and intertemporal variation in oil prices or in known reserves. However, the imposition of country fixed effects means that the only source of variation for the interaction term SIZE$\times$AUTOC is changes in ethnic groups within countries, because AUTOC is a *country-level*, time-invariant indicator. This is the main reason why we do not use group fixed effects, though in one version (see Section 6.2) we explore this case, as group-level variation in SIZE$\times$OIL is still possible through the OIL component.

We estimate equation (13) by OLS. The reason for fitting a linear probability model (rather than a non-linear specification, such as probit or logit) is that our key variables are interactions and interpreting them in nonlinear models isn't straightforward, as Ai and Norton (2003) point out.[25] For completeness, we study nonlinear variants in Section 6.8.

Finally, robust and clustered standard errors have been computed in all cases. We follow Abadie et al. (2017) and cluster errors according to the level of clustering of the assigned treatment. That is, whenever the "treatment" of interest is assigned at the group (country) level, we cluster errors at the group (country) level as well. This implies that in regressions where only private payoffs are considered (which are assigned at the ethnic group level), we cluster errors at the group level. When public payoffs are also in the regression, standard errors are clustered at the country level, as public payoffs are typically assigned at the country level. Our results are very robust to other clustering strategies as, for instance, two-way clustering (at the country and ethnic group level, where the latter considers all the territories occupied by the same group, even if they belong to different countries). See Table B7 in Appendix B.

[24] In the robustness check section, we consider alternavie we also consider an alternative specification where the unit of analysis is the group — not the group-year as in our main analysis — and the dependent variable is the share of conflict years, see Table B5.

[25] In linear models, the coefficient of the interaction term has a direct interpretation, as it is just the value of the cross derivative of the dependent variable with respect to the variables in the interaction. However, this logic does not extend to nonlinear models: the cross derivative in this case is a more complicated object. As shown by Ai and Norton (2003), its value depends on all the covariates of the model and the sign does not necessarily coincide with the sign of the coefficient of the interaction, see Appendix B.1 for a more detailed discussion.

	Dependent Variable: **Conflict Incidence**							
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
SIZE	-0.015	0.032	0.066***	0.066**	-0.073***	-0.002	0.011	0.084**
	(0.307)	(0.101)	(0.001)	(0.013)	(0.007)	(0.915)	(0.656)	(0.019)
OIL	0.448**	0.684***	0.771***		0.631*	0.887*	0.828*	
	(0.040)	(0.009)	(0.007)		(0.095)	(0.062)	(0.069)	
SIZE× OIL		-13.628***	-14.433***			-14.455**	-12.836**	
		(0.000)	(0.000)			(0.036)	(0.026)	
SIZE× PRIVATE INDEX				-0.049***				-0.046**
				(0.001)				(0.016)
SIZE× LACK RIGHTS					0.097**	0.068*	0.083**	
					(0.028)	(0.062)	(0.035)	
SIZE× PUBLIC INDEX								0.023*
								(0.080)
PRIVATE INDEX				0.002**				0.002
				(0.017)				(0.122)
PUBLIC INDEX								0.110
								(0.386)
POLITY			-0.002**	-0.002**			-0.001	-0.001
			(0.016)	(0.020)			(0.511)	(0.437)
GIP			-0.004**	-0.004**			-0.003	0.001
			(0.044)	(0.036)			(0.246)	(0.867)
GDP			0.001	0.001			0.003	0.003
			(0.156)	(0.193)			(0.324)	(0.356)
POP			0.001	0.001			-0.001	-0.001
			(0.589)	(0.744)			(0.878)	(0.919)
GROUPAREA			0.000	0.000			0.000	0.000
			(0.271)	(0.333)			(0.289)	(0.340)
SOILCONST			-0.001	-0.001			-0.001	-0.001
			(0.299)	(0.256)			(0.399)	(0.387)
DISTCAP			0.002***	0.002***			0.003*	0.003*
			(0.000)	(0.000)			(0.085)	(0.088)
MOUNT			0.002*	0.002*			0.002	0.002
			(0.069)	(0.093)			(0.107)	(0.168)
PARTITIONED			-0.001	-0.001			-0.001	-0.001
			(0.296)	(0.281)			(0.444)	(0.477)
LAG	0.895***	0.895***	0.894***	0.894***	0.898***	0.898***	0.900***	0.900***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
c	-0.002	-0.005***	-0.035	0.008	-0.008	-0.010	-0.014	-0.142
	(0.207)	(0.006)	(0.288)	(0.876)	(0.269)	(0.172)	(0.927)	(0.493)
R^2	0.844	0.844	0.847	0.847	0.850	0.850	0.855	0.853
Obs	64839	64839	55289	54486	41314	41314	38341	35755

Table 1. Group Size and Conflict: Baseline. This table regresses conflict incidence on group size and indices of private and public prizes, along with interactions between subsets of these variables as suggested by the theory. All regressions contain year dummies and country fixed effects. The time period considered is 1960-2006 (1975-2006) in regressions 1–3 (4–8). p-values are reported in parentheses. $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

5. Main Results

Tables 1 to 5 contain our main results. Table 1 examines the baseline specification and some core variants. Table 2 studies a far more flexible specification for interaction across group size and private prizes. Tables 3 and 4 consider other ways of proxying these prizes, and examine additional implications of our theory. Table 5 examines an additional implication of our theory when there are multiple threats to peace. Later, Section 6 examines the robustness of the results presented in these tables.

5.1. **Baseline.** In Table 1, the dependent variable is conflict incidence. Each column reports on a different linear probability specification, all with country and year fixed effects, and lagged conflict. Column 1 regresses INCIDENCE on only two variables: group size (SIZE) and group-level oil abundance (OIL).[26] The abundance of oil in the ethnic homeland is positively associated with conflict incidence involving that ethnicity. As already observed, this is a well-established correlation. The coefficient of SIZE is small and not significant. This is precisely what the theory would lead us to expect, as it predicts that the *unconditional* effect of group size on conflict is ambiguous. Column 2 introduces the interaction of SIZE and OIL. The coefficient of the interaction term is negative and significant, as predicted by the theory. Column 3 adds on controls to the regression in Column 2. The results are the same. Column 4 replaces OIL by an alternative proxy of privateness, PRIVATE INDEX, computed using factor analysis on the three main indicators of resource abundance employed in this paper: oil, land and mineral abundance at the group level; See Table 3 for additional results using each of these variables individually. Similar results are obtained.

Column 5 turns the spotlight on public prizes. It contains three regressors: OIL, SIZE and the interaction of SIZE with LACK RIGHTS.[27] Recall that to allay concerns of reverse causality, LACK RIGHTS is a pre-sample time-invariant index computed by averaging its values from 1972 (the first year it exists) to 1975. Then, the resulting index is employed in regressions including post-1975 observations only. (Note that LACK RIGHTS is not an independent regressor, as it is a time-invariant country-level variable and so is subsumed in the country fixed effects.) The interaction of SIZE and LACK RIGHTS has the predicted positive sign and is highly significant. Column 6 introduces *both* the interaction of SIZE and OIL, as well as SIZE and LACK RIGHTS while Column 7 is a variant of 6 with all the additional controls reintroduced. Both interactions return the predicted signs and are significant. Finally, Column 8 is similar to Column 5 but introduces an alternative proxy of publicness, PUBLIC INDEX, obtained by applying factor analysis on five indicators of (pre-sample) lack of public goods: lack of political rights, lack of civil rights, the level of autocracy, group exclusion from central power and infant mortality rates; See Section 5.4 below and Table 4 for further results using each of these variables individually. Our conclusions remain unchanged.

We provide a sense of the magnitudes of the coefficients using the estimates in Column 7, our baseline specification henceforth. These depend on the values of LACK RIGHTS and OIL, so we give a couple of examples. For LACK RIGHTS $= 0$ and a high value of oil (at the 95th percentile) an increase of one standard deviation in SIZE decreases the unconditional probability of conflict

[26]For convenience, the coefficients of SIZE and its interactions have been multiplied by 10 in all tables.

[27]As mentioned before, following Abadie et al. (2017), standard errors in regressions 6–10 have been clustered at the country level. This is so because in the case the "treatment" (absence of political and civil rights) is assigned at the country level. For the sake of robustness, we have re-estimated Table 1 using two-way clustering, with errors clustered at the group (as opposed to country-group) *and* country level. Our conclusions remain identical; see Table B7 in Appendix B.

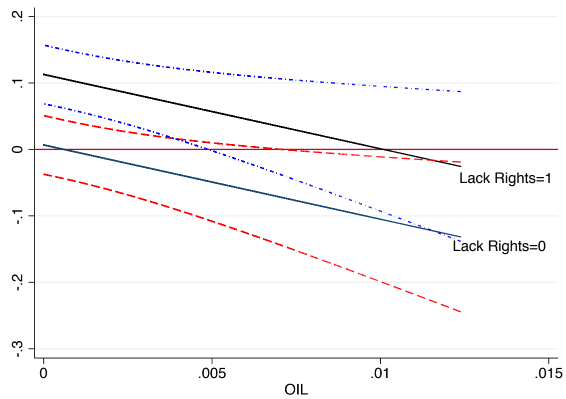


Figure 2. Marginal Effect of Group Size on Conflict INCIDENCE as a Function of OIL
This graph depicts the marginal effect of group size on conflict incidence as a function of OIL for two different values of LACK RIGHTS: LACK RIGHTS=0 (bottom solid line) and LACK RIGHTS$=1$ (top solid line). Confidence bands at the 95% level are also depicted. Estimates from Table 1 (Column 7) have been employed to compute the estimates.

incidence by 4.9%. Similarly, if OIL $= 0$ and LACK RIGHTS is high ($= 1$), an increase of one standard deviation in SIZE increases the probability of conflict by 7.3%.

This baseline table — as well as a sizable majority of the variations to come — contains an additional result. Observe that once SIZE$\times$ PRIV is in place (columns 2 to 4), the coefficient of SIZE *alone* now captures the effect of group size on conflict in the absence of the private prize that has been "removed" by the interaction term. In principle, this residual effect is ambiguous, because it will reflect the joint impact of any public prize and other, un-removed private prizes. But *if* it is positive when a private good is removed, it is highly indicative of a positive interaction between group size and conflict in the presence of public goods. And indeed, the coefficient of SIZE in those columns is positive, generally significant and its magnitude is much larger than in columns that do not contain that interaction. For instance, the coefficient of SIZE in Columns 3–4 is almost 4 times larger than that in Column 1. Comparing Columns 1 and 5, one can see that a parallel argument applies to regressions that only consider the interaction between of SIZE and LACK RIGHTS. Column 5 shows that when the latter interaction is included, the coefficient of SIZE, which now captures the effect of size in groups with a value of LACK RIGHTS equal to zero, becomes negative and significant.

In closing our baseline analysis, we note the obvious: while the data are replete with conflicts over private and public payoffs, the two are sometimes closely intertwined. For instance, even a conflict as seemingly primordial as Rwanda was permeated with economic looting, such as land grabs under the cover of ethnic violence. The Second Civil War in the Sudan is about different cultural and religious identities, but it is also — to some degree — about oil; so is the Chechnyan War. The Zimbabwean conflict is about identity and political power, but it is also about land, and so on. In the light of these expected complications, it is of interest that the two interaction predictions made by the theory hold up separately and robustly. This robustness will continue to be on display as we move to the variations.

	Dependent Variable: **Conflict Incidence**				
	[1]	[2]	[3]	[4]	[5]
SIZE	-0.019	0.049**	0.080***	0.011	0.022
	(0.221)	(0.017)	(0.000)	(0.659)	(0.412)
SIZE× LACK RIGHTS				0.071*	0.083*
				(0.079)	(0.050)
SIZE× OIL_{0-25}		-0.014	0.047	-0.022	0.033
		(0.831)	(0.539)	(0.822)	(0.761)
SIZE× OIL_{25-50}		-0.063	-0.058	0.134	0.182
		(0.151)	(0.398)	(0.585)	(0.551)
SIZE× OIL_{50-75}		-0.152***	-0.154***	-0.183***	-0.166**
		(0.000)	(0.000)	(0.010)	(0.027)
SIZE× $\text{OIL}_{>75}$		-0.130***	-0.135***	-0.136***	-0.118***
		(0.000)	(0.000)	(0.009)	(0.005)
OIL_{0-25}	-0.002	-0.002	-0.004***	-0.002	-0.003
	(0.176)	(0.153)	(0.007)	(0.488)	(0.393)
OIL_{25-50}	-0.000	-0.000	-0.001	-0.001	-0.002
	(0.798)	(0.803)	(0.701)	(0.769)	(0.600)
OIL_{50-75}	0.003*	0.004**	0.004**	0.005*	0.005*
	(0.082)	(0.017)	(0.023)	(0.061)	(0.064)
$\text{OIL}_{>75}$	0.004**	0.006***	0.007***	0.007**	0.007**
	(0.041)	(0.010)	(0.007)	(0.026)	(0.029)
POLITY			-0.002**		-0.001
			(0.016)		(0.512)
GIP			-0.003**		-0.003
			(0.049)		(0.253)
GDP			0.001		0.003
			(0.195)		(0.320)
POP			0.002		-0.001
			(0.462)		(0.884)
GROUPAREA			0.000		0.000
			(0.530)		(0.338)
SOILCONST			-0.001		-0.001
			(0.235)		(0.358)
DISTCAP			0.002***		0.002*
			(0.000)		(0.085)
MOUNT			0.003*		0.002
			(0.064)		(0.100)
PARTITIONED			-0.001		-0.001
			(0.337)		(0.431)
LAG	0.895***	0.894***	0.894***	0.898***	0.900***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
c	-0.002	-0.006***	-0.033	-0.012***	-0.019
	(0.289)	(0.001)	(0.325)	(0.004)	(0.903)
R^2	0.844	0.844	0.848	0.850	0.855
Obs	64839	64839	55289	41314	38341

Table 2. Group Size and Conflict: a more flexible specification for oil abundance. This table regresses conflict incidence on group size and indices of private and public prizes, along with interactions between subsets of these variables as suggested by the theory. All regressions contain year dummies and country fixed effects. The dummy variables OIL_{i-j} are equal to 1 if the value of oil reserves is between the ith and the jth percentile of the distribution of OIL, conditional on having oil in the homeland. The time period considered is 1960-2006 (1975-2006) in regressions 1–3 (4 and 5). p-values are reported in parentheses. $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

5.2. **A Flexible Variant for Group Size and Private Prizes.** The specification in Table 1 imposes the restriction that the marginal effect of SIZE on conflict is a linear function of oil. We've seen that the presence of oil attenuates the effect of size on conflict. But the theory makes a sharper prediction: the marginal effect of SIZE actually turns *negative* as the prize becomes increasingly private. (The opposite is true when the prize is public.) Figure 2 plots the marginal effect of SIZE on INCIDENCE computed using the estimates from Column 7 in Table 1. The marginal effect is a function of both OIL and LACK RIGHTS, and the plot displays this marginal effect as a function of OIL (in the X axis), for the minimum and maximum values of LACK RIGHTS (i.e., LACK RIGHTS=$\{0, 1\}$). The dashed lines represent 90% confidence bands. In line with the theory, the figure shows that the marginal effect of size can be negative or positive, depending on the values of the public and private payoffs. For a small value of LACK RIGHTS and moderate or large values of OIL, the effect of an increase in group size has a negative and significant effect on conflict incidence. The opposite is true when LACK RIGHTS is high and OIL is small: in this case the marginal effect of SIZE is positive and significant. However, it is not significantly different from zero when either both prizes are small or when both are large. Section B.2 in the Online Appendix makes these points in an even simpler way by taking binary cuts for private and public prizes and comparing the four cells that result.

One might respond that this observation only stems from the assumed linearity in the interaction of SIZE and OIL. Therefore, we re-do Table 1 using a more flexible specification, one in which linearity is not imposed. In Table 2, we employ four dummies that correspond to the quartiles of the distribution of OIL for the groups that have oil in their homeland (thus, the omitted category corresponds to groups that do not have oil). In this case, the marginal effect of SIZE on conflict is given by the sum of the coefficient of SIZE and that of the variables SIZE$\times$ OIL$_j$, where OIL$_j$ is equal to 1 if the group's oil is in quartile j. The marginal effect of SIZE on conflict is positive and significant in the absence of oil, but the effect decreases as the amount of oil in the homeland becomes more abundant and it eventually becomes negative for groups with abundant oil reserves. In particular, we can reject at the 1% level the assertion that the sum of the coefficient of SIZE and that of SIZE$\times$ OIL$_{50-75}$ (or SIZE$\times$ OIL$_{\geq 75}$) is greater than or equal to zero.[28] So this effect is not driven by merely extrapolating a linear specification.

Table 2 is subjected to a number of robustness checks in Table B3 of the Online Appendix. Among these, we drop various regions of the world and show that the same effect survives in our flexible specification.

This negative relationship between size and conflict when the prize is private — morphing into a positive relationship as the prize turns public — is central to our theory. We find it hard to think of any alternative explanation that would generate the same joint pattern. One could, of course, posit something *ad hoc*: that oil is somehow special for smaller groups, or that small groups have a better conflict technology, or that they seek secession and so fight harder. These *ad hoc* alternatives must all contend with one simple observation that we explore in columns 8-10 of Table 3: that controlling for the *per-capita value* of a private prize, group size is positively related to

[28]Similar results are also true if the estimates in Column 7 (with a moderate value of LACK RIGHTS, as suggested by the theory) are considered.

conflict (more detail below). None of the above *ad hoc* explanations can simultaneously explain why the relationship turns positive when per-capita controls for private wealth are imposed.

The one reasonable argument that does generate a negative relationship between group size and conflict is the free-rider theory. Small groups are better capable of cohesion in a fight. That said, such a theory works well for one side of our observations but not the other. They would have no prediction for group size and *public* prizes, where the observed relationship is positive.

	Dependent Variable: **Conflict Incidence**									
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
SIZE	0.013	0.053*	0.004	0.033	0.168***	0.110	0.020	0.038**	0.034**	0.034**
	(0.586)	(0.091)	(0.893)	(0.576)	(0.005)	(0.118)	(0.228)	(0.017)	(0.028)	(0.028)
MINES	-0.000	0.000	0.000							
	(0.672)	(0.995)	(0.938)							
SIZE× MINES		-0.015**	-0.013*							
		(0.036)	(0.099)							
SIZE× LACK RIGHTS			0.086*			0.067				
			(0.086)			(0.134)				
AREA(SHARE)				-0.002	0.016	0.014				
				(0.823)	(0.111)	(0.186)				
SIZE×AREA(SHARE)					-0.370***	-0.331**				
					(0.001)	(0.034)				
OIL PC								-0.001		0.012
								(0.971)		(0.611)
LAND PC									-0.001**	-0.001**
									(0.036)	(0.035)
OIL	0.551*	0.521*	0.482	0.551**	0.407*	0.515	0.570**			
	(0.073)	(0.087)	(0.251)	(0.022)	(0.092)	(0.165)	(0.030)			
POLITY	-0.001	-0.001	-0.001	-0.002**	-0.002**	-0.001		-0.002**	-0.002**	-0.002**
	(0.169)	(0.168)	(0.494)	(0.017)	(0.019)	(0.532)		(0.017)	(0.018)	(0.018)
GIP	-0.002	-0.002	-0.002	-0.004**	-0.005***	-0.004	-0.003*	-0.003*	-0.003*	-0.003*
	(0.344)	(0.347)	(0.380)	(0.050)	(0.006)	(0.101)	(0.065)	(0.064)	(0.054)	(0.054)
AREA	-0.000	0.000	0.000				-0.000	0.000		
	(0.796)	(0.339)	(0.340)				(0.680)	(0.318)		
GDP	0.004*	0.004*	0.005	0.001	0.001	0.003	0.001	0.001	0.001	0.001
	(0.095)	(0.095)	(0.199)	(0.162)	(0.160)	(0.346)	(0.134)	(0.133)	(0.133)	(0.133)
POP	-0.004	-0.004	-0.003	0.001	0.001	-0.001	0.001	0.001	0.001	0.001
	(0.545)	(0.549)	(0.786)	(0.538)	(0.566)	(0.864)	(0.495)	(0.609)	(0.774)	(0.771)
SOILCONST	-0.001*	-0.001**	-0.001	-0.000	-0.001	-0.001	-0.000	-0.000	0.000	0.000
	(0.062)	(0.027)	(0.264)	(0.402)	(0.224)	(0.333)	(0.685)	(0.625)	(0.719)	(0.733)
DISTCAP	0.002***	0.002***	0.002	0.002***	0.002***	0.003*	0.002***	0.002***	0.002***	0.002***
	(0.002)	(0.003)	(0.123)	(0.000)	(0.000)	(0.074)	(0.000)	(0.000)	(0.000)	(0.000)
MOUNT	0.002	0.003	0.002	0.002*	0.002*	0.002	0.002	0.002	0.002	0.002
	(0.162)	(0.140)	(0.191)	(0.084)	(0.075)	(0.124)	(0.131)	(0.137)	(0.193)	(0.193)
PARTITIONED	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.653)	(0.637)	(0.693)	(0.303)	(0.281)	(0.431)	(0.296)	(0.342)	(0.356)	(0.357)
LAG	0.887***	0.887***	0.890***	0.894***	0.894***	0.900***	0.894***	0.895***	0.895***	0.895***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
c	0.026	0.026	-0.008	-0.007	-0.003	0.004	-0.043	-0.036	-0.035	-0.035
	(0.830)	(0.830)	(0.958)	(0.884)	(0.955)	(0.979)	(0.295)	(0.271)	(0.280)	(0.279)
R^2	0.838	0.838	0.843	0.847	0.847	0.855	0.846	0.847	0.847	0.847
Obs	33325	33325	32026	54486	54486	37659	57559	55289	55289	55289

Table 3. Alternative Private Prize Specifications. This table regresses conflict incidence on group size and indices of private and public prizes, along with interactions between subsets of these variables. Columns 1–3 use mineral availability and Columns 4–6 use land-based measures, as described in the text. Columns 7–10 show how the magnitude and the significance of SIZE change if one controls for *total* or *per capita* private payoffs. All columns contain country fixed effects and year dummies, and have been estimated by OLS. p-values are reported in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

5.3. **Alternative Proxies for Privateness.** Table 3 considers different proxies of privateness (columns 1–6) as well as an additional implication of our theory (columns 7–10).

Columns 1 to 3 consider mineral availability in the ethnic homeland as a proxy for "privateness." We use geo-referenced data on the location of mining activities around the world since 1980.[29] For each year and mine, we know whether that mine is active or not, and the specific minerals produced by it. As in Berman et al (2015), we focus on 13 minerals for which we have world price data,[30] which we take from the World Bank's commodity price database. The MINES index is constructed as follows: for each group, year and mineral, we create a dummy variable that is one if the group has at least one active mine of that mineral. To introduce information on mineral prices, we multiply each of the mineral dummies by (the log of) its international price, normalized by (the log of the) price in 1980 (the year when the mineral data starts). The variable MINES is constructed as the sum of the resulting quantities for each group and year. Column 1 introduces MINES as an additional regressor and, as in previous cases, the coefficient of SIZE is not significant. Column 2 adds to this specification the interaction of the mines index and group size which has a negative and significant coefficient, as predicted.[31] Column 3 introduces the interaction of SIZE and LACK RIGHTS, and shows that results are robust.

In Column 4, the variable AREA(SHARE) measures the share of the ethnic homeland area as a fraction of the total area of the country. This available land can be seen as a private payoff, the valuation of which is effectively higher if land is relatively scarce in the rest of the country. Column 4 shows that the coefficient of SIZE is small and insignificant in a regression that contains both AREA(SHARE), OIL and the controls but doesn't include any of the interactions. Column 5 introduces the interaction of SIZE and AREA(SHARE), which is negative and significant, suggesting that small groups are more likely to be involved in conflict as the value of AREA(SHARE) increases. Column 6 shows that a similar result is found when the interaction of group size and our proxy of publicness is included (although the estimation of the latter interaction is noisier now, and has a p-value of 0.13).

Interestingly, a similar pattern as in Table 1 is found: when only SIZE is in the regression (Columns 1 and 4), it is insignificant. But it becomes positive and significant when the private prize interaction is introduced (columns 2 and 5), implying that larger groups are more likely to fight when the private prize is small. When both interactions are introduced (columns 3 and 6), SIZE becomes insignificant again, suggesting that our measures for public and private payoffs are indeed good proxies.

Columns 7 to 10 explore our main result in an alternative specification. The theory rests on the idea that large groups are more likely to fight because they have a relatively higher probability of winning, but less likely to fight because their per capita payoff is relatively low. So if one controls for *per capita* payoffs (rather than for total payoffs, as in the specifications above), large

[29]The source is the *Raw Material Data* (IntierraRMG, 2015). Since data on mining activity starts in 1980, our sample in these regressions focuses on the period 1980–2006.

[30]These are Bauxite, Coal, Copper, Diamond, Gold, Iron, Lead, Nickel, Platinum, Phospate, Silver, Tin and Zinc.

[31]Similar results hold in only information on mine availability is used to compute the mines proxy.

groups should *unambiguously* be more prone to conflict. To explore this prediction, Columns 8–10 control for *per capita* private payoffs.[32] Column 7 shows that group size is not significant in a regression that includes all the controls and OIL, (our baseline measure of total private payoff). Column 8 replaces OIL by OIL PC, computed by dividing OIL by group population. In this case, the coefficient of SIZE is significant and more than 70% larger than in Column 7, suggesting that larger groups are more prone to conflict once per capita payoffs are held constant. Columns 9 and 10 show that a similar result also holds when land per capita (Column 9) and land and oil per capita (Column 10) are introduced in the regression.[33]

5.4. **Alternative Proxies for the Public Prize.** Table 4 presents results using alternative proxies for the public prize. More specifically, it considers the 5 indices employed in the computation of PUBLIC INDEX, (see Column 10 in Table 1). Each of these indices aims to proxy the lack of access to public goods, where the notion of "public good" is approached in different ways.

Columns 1–4 consider separately the two indices employed to compute LACK RIGHTS, our baseline measure of publicness. The "subindices" LACK CIVIL RIGHTS and LACK POLITICAL RIGHTS measure the lack of civil rights and the lack of political rights, respectively, (see Appendix A for definitions of these variables). Identical results as in Table 1 are obtained.

Columns 5–8 consider proxies that aim to capture lack of access to political power. The first proxy (Columns 5 and 6) is similar to that employed by Esteban et al. (2012a, 2012b) and rests on the idea that there are large gains to seizing power when groups are excluded from it. Specifically, the more "autocratic" a country is, the less is the sharing of power and the larger the number of citizens/groups that are excluded from power, and consequently, the higher is the value of controlling the State. This may be because such groups are interested in seizing autocratic power themselves, or it may be because those groups want to install a democracy. To proxy lack of access to power we use AUTOC, a composite measure of autocracy from Polity IV.[34] To avoid reverse causality concerns, we *only* consider pre-sample values of the autocracy index (and in addition we control for past conflict in all our regressions). Specifically, AUTOC is computed by averaging the values of the autocracy index from 1960 to 1975, which is then employed in regressions using post 1975 data. The resulting measure is "assigned" to all the ethnic groups in the country, so that AUTOC is a time-invariant country-level index.

Whereas LACK RIGHTS and AUTOC are time-invariant and defined at the country-level, OIL is group-specific. To account for this asymmetry, we also consider a group-level measure of publicness, based on whether the group is excluded from State power. We construct an index for

[32]Notice that there is no need to perform a similar computation with public payoffs, as they are not diluted by group size.

[33]We cannot perform a similar calculation using MINES as we don't have a measure of total mineral production.

[34]This index is measured on a scale from 0 to 10, where 10 indicates the highest degree of autocracy, see Polity IV for details about its construction. We normalize this index to be between 0 and 1. The Polity IV manual summarizes the index thus: "[We] define [autocracy] operationally in terms of the presence of a distinctive set of political characteristics. In mature form, autocracies sharply restrict or suppress competitive political participation. Their chief executives are chosen in a regularized process of selection within the political elite, and once in office they exercise power with few institutional constraints . . . Our operational indicator of autocracy is derived from codings of the competitiveness of political participation, the regulation of participation, the openness and competitiveness of executive recruitment, and constraints on the chief executive."

	Dependent Variable: **Conflict Incidence**									
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
SIZE	-0.044	0.011	-0.041	0.014	-0.037	0.014	-0.020	0.031	-0.047*	0.012
	(0.173)	(0.665)	(0.160)	(0.543)	(0.186)	(0.587)	(0.493)	(0.305)	(0.098)	(0.656)
SIZE× CIVIL RIGHTS	0.112**	0.089**								
	(0.028)	(0.048)								
SIZE× POL. RIGHTS			0.095**	0.072**						
			(0.013)	(0.029)						
SIZE× AUTOC					0.116***	0.097**				
					(0.007)	(0.013)				
SIZE× EXCLUDED							0.112***	0.098**		
							(0.008)	(0.015)		
SIZE× INFANT MORTALITY									0.006**	0.004
									(0.032)	(0.137)
SIZE× OIL		-12.928**		-12.881**		-12.554**		-13.350***		-11.696**
		(0.024)		(0.026)		(0.025)		(0.001)		(0.046)
EXCLUDED							0.002	0.003		
							(0.305)	(0.233)		
POLITY	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	0.001	0.001	-0.001	-0.001
	(0.516)	(0.511)	(0.516)	(0.511)	(0.514)	(0.509)	(0.396)	(0.405)	(0.456)	(0.452)
OIL	0.642	0.826*	0.651	0.832*	0.673*	0.841*	0.519*	0.710**	0.647	0.796*
	(0.112)	(0.070)	(0.108)	(0.069)	(0.083)	(0.052)	(0.053)	(0.018)	(0.103)	(0.074)
GIP	-0.003	-0.003	-0.002	-0.002	-0.003	-0.003			-0.002	-0.002
	(0.249)	(0.242)	(0.259)	(0.252)	(0.250)	(0.243)			(0.357)	(0.361)
GDP	0.003	0.003	0.003	0.003	0.002	0.002	0.003**	0.003**	0.003	0.003
	(0.324)	(0.324)	(0.324)	(0.324)	(0.339)	(0.340)	(0.025)	(0.025)	(0.352)	(0.352)
POP	-0.001	-0.001	-0.001	-0.001	-0.002	-0.002	-0.001	-0.001	-0.001	-0.001
	(0.877)	(0.878)	(0.878)	(0.878)	(0.821)	(0.823)	(0.726)	(0.711)	(0.877)	(0.879)
GROUPAREA	-0.000	0.000	-0.000	0.000	-0.000	0.000	-0.000	0.000	-0.000	0.000
	(0.867)	(0.305)	(0.914)	(0.274)	(0.520)	(0.684)	(0.772)	(0.297)	(0.698)	(0.582)
SOILCONST	-0.000	-0.001	-0.000	-0.001	-0.001	-0.001	-0.001	-0.001*	-0.001	-0.001
	(0.512)	(0.401)	(0.510)	(0.402)	(0.186)	(0.121)	(0.117)	(0.067)	(0.209)	(0.153)
DISTCAP	0.002*	0.003*	0.002*	0.003*	0.002	0.002	0.002***	0.002***	0.002	0.002
	(0.090)	(0.086)	(0.088)	(0.084)	(0.128)	(0.121)	(0.001)	(0.001)	(0.118)	(0.112)
MOUNT	0.002	0.002	0.002	0.002	0.003*	0.003*	0.003*	0.003*	0.003*	0.003*
	(0.112)	(0.103)	(0.123)	(0.112)	(0.063)	(0.057)	(0.071)	(0.061)	(0.077)	(0.071)
PARTITIONED	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.440)	(0.444)	(0.440)	(0.444)	(0.409)	(0.411)	(0.529)	(0.531)	(0.446)	(0.450)
LAG	0.900***	0.900***	0.900***	0.900***	0.899***	0.899***	0.898***	0.898***	0.898***	0.898***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
c	-0.011	-0.015	-0.009	-0.013	0.009	0.004	-0.015	-0.013	-0.002	-0.007
	(0.940)	(0.921)	(0.954)	(0.933)	(0.953)	(0.978)	(0.767)	(0.803)	(0.985)	(0.955)
R^2	0.855	0.855	0.855	0.855	0.852	0.852	0.855	0.855	0.850	0.850
Obs	38341	38341	38341	38341	39394	39394	45330	45330	36839	36839

Table 4. Alternative Public Prize Specifications. This table regresses conflict incidence on group size and indices of private and public prizes. Alternative proxies are considered for the public prize. All regressions contain year dummies and country fixed effects, and have been estimated by OLS. p-values are reported in parentheses. For convenience, the coefficients of SIZE×INFANT MORTALITY are multiplied by 10. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

exclusion, EXCLUDED, in a similar fashion as AUTOC, i.e., by averaging the values of a yearly dummy for exclusion over the period 1945-1970. The interactions of SIZE and the different prize proxies in Columns 5–8 keep their expected signs and are significant in all these variations.

Columns 9 and 10 employ INFANT MORTALITY as a proxy for publicness. The idea is that, ceteris paribus, this variable can be interpreted as a measure of low provision of health services (see Appendix A for the exact definition and data sources). As before, we only use pre-sample values to compute this measure, to alleviate reserve causality concerns. Column 9 shows that the interaction of group size and infant mortality rates is positive and significant; when the interaction between SIZE and OIL is also included (Column 10), the sign is still positive but the estimation is a bit noisier (p-value is 0.13).

It can be argued that there is no perfect measure of a public prize, and we agree. Yet it is telling that our results exhibit robustness to a number of proxies, each of which attempt to capture in some way the absence of rights, freedoms, or publicly-provided services.

	Dependent Variable: **Conflict Incidence**					
	[1]	[2]	[3]	[4]	[5]	[6]
SIZE	0.041**	0.071***	0.016	0.022	0.072***	-0.005
	(0.030)	(0.001)	(0.553)	(0.272)	(0.003)	(0.887)
FRAC×OIL	3.267***	2.979***	2.879*			
	(0.000)	(0.001)	(0.056)			
POL×OIL				1.645	1.131	3.872
				(0.735)	(0.814)	(0.494)
SIZE× OIL		-10.137**	-8.187		-16.410***	-14.312**
		(0.014)	(0.230)		(0.000)	(0.034)
SIZE× LACK RIGHTS			0.081*			0.125**
			(0.061)			(0.016)
OIL	-1.304***	-0.991*	-0.925	0.609*	0.879**	0.767
	(0.009)	(0.074)	(0.427)	(0.081)	(0.018)	(0.211)
POLITY	-0.002**	-0.002**	-0.001	-0.002***	-0.002***	-0.001
	(0.017)	(0.017)	(0.506)	(0.009)	(0.009)	(0.268)
GIP	-0.004**	-0.004**	-0.003	-0.003	-0.003*	-0.002
	(0.032)	(0.029)	(0.205)	(0.110)	(0.099)	(0.518)
GDP	0.001	0.001	0.003	-0.000	-0.000	0.001
	(0.142)	(0.148)	(0.343)	(0.794)	(0.781)	(0.863)
POP	0.001	0.001	-0.001	0.005*	0.005	0.009
	(0.772)	(0.800)	(0.881)	(0.099)	(0.116)	(0.452)
GROUPAREA	0.000	0.000	0.000	-0.000	0.000	0.000
	(0.516)	(0.114)	(0.192)	(0.598)	(0.369)	(0.240)
SOILCONST	-0.000	-0.000	-0.000	-0.001	-0.001	-0.001
	(0.482)	(0.382)	(0.511)	(0.478)	(0.379)	(0.469)
DISTCAP	0.002***	0.002***	0.003*	0.002***	0.002***	0.003*
	(0.000)	(0.000)	(0.083)	(0.000)	(0.000)	(0.088)
MOUNT	0.002	0.002	0.002	0.003*	0.003**	0.003**
	(0.130)	(0.113)	(0.177)	(0.052)	(0.042)	(0.045)
PARTITIONED	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.350)	(0.344)	(0.482)	(0.168)	(0.169)	(0.283)
LAG	0.894***	0.894***	0.900***	0.896***	0.896***	0.903***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
c	-0.020	-0.016	-0.017	-0.058	-0.053	-0.173
	(0.640)	(0.712)	(0.890)	(0.363)	(0.408)	(0.419)
R^2	0.847	0.847	0.855	0.851	0.851	0.857
Obs	54762	54762	37957	45362	45362	29387

Table 5. Private Payoffs, Fractionalization, Polarization and Conflict.
Notes. This table regresses conflict incidence on group size, indices of private and public prizes, along with interactions of a Fractionalization index (FRAC) and OIL (Columns 1–3) as well as a Polarization index (POL) and OIL (Columns 4–6). All regressions contain year dummies and country fixed effects. FRAC and POL are country-level indices of ethnic fractionalization and polarization, respectively. p-values are reported in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

5.5. **Multiple Threats and Conflict.** Because fighting is costly, individual threats can always be appeased provided that appropriate Coaseian transfers are available. However, Section 3.3 shows that in conflicts over private payoffs the existence of a variety of potential threats might make conflict unavoidable. This is so because it might be impossible to find a set of transfers that *simultanously* appeases all threats. The empirical implication of this result is that, in the presence of private payoffs, we should observe a positive relationship between conflict and the number of groups that can threaten peace, where these groups can be defined along several dimensions (ethnicity, class, geography, etc). Without pretending to explore this issue in depth here — as we lack information on group affiliation on dimensions other than ethnicity — our empirical setting allows us to partially examine this prediction. Because attention is restricted to groups defined along ethnic lines, we proxy the multiplicity of threats by an ethnic fractionalization index. Then, we explore whether groups in more ethnically fragmented countries are more likely to be involved in conflicts over private payoffs.[35] Results are presented in Table 5. Column 1 introduces the interaction of an index of fractionalization, FRAC, and OIL in a regression of conflict incidence on all the controls. The interaction is positive and significant, as predicted by the theory.[36] This conclusion is robust to introducing in the regression the interactions of SIZE and OIL (Column 2) and SIZE and LACK RIGHTS (Columns 3).[37] For comparison, Columns 4–6 reestimate Columns 1–3 replacing FRAC×OIL by the interaction of a polarization index, POL, and OIL. Our theory doesn't have any prediction for this interaction, which turns out to be insignificant in our regressions.

6. VARIATIONS

The evidence so far shows a robust link between the probability of conflict, group size and the nature of the payoffs. We've employed country and year fixed effects throughout and also controlled for several group-level characteristics. In this section, we consider several variations on the baseline exercise. First, we consider alternative explanations that could rationalize our empirical findings and provide evidence against them (Section 6.1). Second, we consider specifications with group fixed effects (Section 6.2). These reduce the likelihood of omitted variable bias as they control for all group-level characteristics that are time invariant. Third, we address the possibility that groups could form alliances in conflict (Section 6.3). We then go on to consider alternative measures of conflict (Section 6.4), other ways of measuring oil wealth (Section 6.5), and the possibility that our results are due to omitted variables more generally, finding little support for this (Section 6.6). We end with some statistical variations: two-way clustering of standard errors instead of country- or group-level clustering (Section 6.7), the use of a nonlinear model; specifically logit (Section 6.8), and robustness to dropping different regions of the world (Section 6.9). Some tables are provided in the Online Appendix (Appendix B).

[35]The fractionalization index measures the probability that two individuals chosen at random in country j belong to two different ethnic groups. The polarization index is a measure of social "antagonism," and its key ingredients are intergroup distances (how alien groups are from each other) and group size (an indicator of the level of the group identification). Both indices are computed using data from Fearon (2003), and linguistic distances are employed to capture intergroup distances, see Esteban et al. (2012) for details on their computation.

[36]This result is closely related to Esteban et al. (2012), who show that more fragmented countries tend to fight more on account of oil. The results in Table 5 show that a similar result also holds when considered at the ethnic group level.

[37]Similar results are found if other ways of defining the private payoff are considered, such as oil at the country rather than at the group level or the index of private payoffs introduced in Table 1.

	Dependent Variable: **Conflict Incidence**							
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
SIZE	0.065***	0.014	0.065***	0.014	0.142***	0.033	0.077***	0.025
	(0.006)	(0.576)	(0.005)	(0.578)	(0.002)	(0.602)	(0.001)	(0.339)
SIZE×OIL	-15.007**	-14.491**	-16.306*	-17.488	-26.752***	-18.602*	-17.060***	-16.350**
	(0.016)	(0.027)	(0.063)	(0.112)	(0.000)	(0.063)	(0.000)	(0.016)
SIZE×LACK RIGHTS		0.081**		0.077**		0.199*		0.094**
		(0.033)		(0.045)		(0.094)		(0.027)
OIL	0.794*	0.900*	0.768*	0.829*	0.908**	1.096**	0.925***	0.980**
	(0.061)	(0.050)	(0.071)	(0.077)	(0.011)	(0.027)	(0.004)	(0.019)
OIL CONCENTRATION	-0.014	0.013						
	(0.158)	(0.304)						
OIL (SHARE)			0.001	0.003				
			(0.820)	(0.649)				
GIP	-0.003	-0.002	-0.003	-0.002			-0.004*	-0.003
	(0.141)	(0.304)	(0.127)	(0.306)			(0.068)	(0.208)
GDP	0.000	0.003	0.001	0.003	0.002*	0.003	0.002*	0.004
	(0.790)	(0.252)	(0.393)	(0.289)	(0.059)	(0.251)	(0.066)	(0.173)
POP	0.001	-0.000	0.001	-0.000	-0.002	0.000	-0.000	-0.005
	(0.822)	(0.985)	(0.797)	(0.968)	(0.632)	(0.991)	(0.879)	(0.623)
GROUPAREA	0.000	0.000	0.000	0.000	0.000	-0.000	0.000	0.000
	(0.288)	(0.247)	(0.268)	(0.256)	(0.689)	(0.555)	(0.296)	(0.270)
SOILCONST	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.001
	(0.645)	(0.588)	(0.636)	(0.590)	(0.366)	(0.850)	(0.518)	(0.416)
DISTCAP	0.002*	0.002*	0.002*	0.002*	0.002***	0.003*	0.002***	0.002*
	(0.098)	(0.084)	(0.099)	(0.085)	(0.000)	(0.063)	(0.000)	(0.092)
MOUNT	0.002**	0.002	0.002**	0.002	0.002	0.001	0.002	0.002
	(0.034)	(0.125)	(0.033)	(0.116)	(0.292)	(0.373)	(0.129)	(0.160)
PARTITIONED	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.282)	(0.546)	(0.280)	(0.542)	(0.303)	(0.328)	(0.219)	(0.650)
LAG	0.894***	0.900***	0.893***	0.900***	0.896***	0.900***	0.889***	0.895***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
c	-0.010	-0.036	-0.033	-0.031	-0.011	-0.046	-0.013	0.024
	(0.917)	(0.734)	(0.699)	(0.777)	(0.836)	(0.790)	(0.794)	(0.854)
R^2	0.846	0.854	0.846	0.854	0.860	0.862	0.839	0.847
Obs	57559	39992	57559	39992	41158	34198	49892	34647

Table 6. Variations: Oil Concentration, Excluded Groups and Small Ruling Elites. This table regresses conflict incidence on group size and indices of private and public prizes, along with interactions between subsets of these variables as suggested by the theory. OIL CONCENTRATION is computed as the Herfindahl index of oil reserve distribution across groups. OIL SHARE is the share of oil in the homeland of the group. To compute Columns 5 and 6, excluded groups have been dropped from the sample. Excluded groups are those with a value of (pre-sample) EXCLUDED larger than 0.5. To compute Column 7, country/years in which the size of the ruling elite in autocraticies is small (as compared to the ruling elite in non-autocracies) have been dropped from the sample (see the main text for details). All regressions contain year dummies and country fixed effects. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

6.1. **Ruling out Alternative Explanations.** An alternative interpretation points to differences in conflict technology rather than to differences in expectations of individual payoffs. According to this view, large groups (with or without oil) would have an easier access to the funds needed to engage in conflict against the State. However, small groups would find it particularly useful to have oil in their homeland to purchase weapons, hire mercenaries, etc., which otherwise would be beyond their means. As in the case of our theory, this explanation would generate a heterogeneous impact of group size on conflict: small groups would tend to fight less than large groups, unless they have oil. However, this alternative explanation would *fail to generate the negative relationship between size and conflict* shown in Table 1 and Figure 2. While the effect of size on conflict would be attenuated by oil, the net effect of group size must always remain positive.

Another interpretation of our findings would focus on the potential confounding role of oil concentration. When a small ethnic group has oil in its homeland, then it is likely, ceteris paribus, that oil reserves are contained within multiple ethnic homelands, rather than within a single large homeland. That in itself could lead to violence. Now, to the extent that the distribution of oil reserves in the country is stable over time, the country fixed effects included in all our specifications will partly eliminate this effect. Nevertheless, as our regressions consider a large time span, new oil discoveries could change distributional patterns within the country. Columns 1 and 2 in Table 6 add to our baseline specifications (Columns 3 and 7 in Table 1) a (time-varying) country-level oil concentration index. We do not find a significant effect of concentration and this variable does not change the interaction effect of group size and oil: the coefficients and their significance are very similar as in the baseline table.

Morelli and Rohner (2015) study the relationship between conflict and the *concentration* of natural resources in ethnic homelands. They show that the larger the group's share of oil, the larger the probability of conflict onset. This is a completely different prediction from ours; it is orthogonal to what we do. That said, and because bigger groups are more likely to have a larger share of national oil, we check that oil share is not a confounder in our regressions. Columns 3 and 4 add to our baseline specifications the share of oil as computed by Morelli and Rohner (i.e., the surface of an ethnic group's territory covered with oil and gas as a percentage of total country surface covered with oil and gas). Oil share is far from being significant in these specifications. Our conclusions survived unchanged, although the coefficient of SIZE$\times$ OIL is estimated more noisily in Column 4 (p-value is .11), when the public prize interaction is also in the regression.[38]

Our theoretical results stress the fact that if the initial allocation is non-discriminatory — or if it is biased *against* small groups — then the latter are more likely to be involved in conflict if the payoff is private. An alternative interpretation, however, would run as follows: it's reasonable to think that large groups are stronger and, as a result, more likely to be in power. Thus, they can get a disproportionately large share of the rents from the center, and therefore they are less likely to rebel. Although this argument is clearly related to ours, the underlying mechanism is somewhat different: small groups rebel because they are more likely to be excluded from power and, as a result, they do badly under the initial allocation. This alternative explanation is, however, at odds with our empirical results: in that case we should see small groups fighting more, not just on account of oil, but simply because they are treated worse. But there is no evidence of that at all: in *all* our specifications, group size per se is either insignificant or positive whenever significant.

Absent a direct measure of the initial allocation, we do control in all our regressions for whether a group is included or excluded from political power. In addition, we have also considered whether

[38] Morelli and Rohner (2015) consider onset rather than incidence. We've also checked that in onset regressions the significance of the share of oil vanishes once one controls for total group oil, but again, this is not our focus.

our results are robust when only *excluded* groups are considered.[39] Columns 5 and 6 in Table 6 restrict the sample to groups excluded from power and shows that our conclusions continue to hold when only those groups are in the sample.

Finally, it could be argued that the index LACK RIGHTS is typically high in authoritarian regimes, which often tend to have a ruling elite made up of minorities. So it would be possible that in those regimes conflict is initiated by majorities that want to take over power from these minorities. Indeed, the average size of the group(s) in power in autocracies is smaller than in less autocratic regimes. To rule out this possibility, we have dropped from the sample countries whose ruling elites are small (as compared to ruling elites in non-autocratic countries). More specifically, we have divided the sample into autocratic and non-autocratic countries (defined as those with autocracy index higher/lower than 5) and we have dropped from the sample autocratic countries where the size of the ruling elite is smaller than the median of the size of the group(s) in power in non-autocracies. Then, we've re-run our baseline specification with this reduced sample. Our results remain robust to this variation, see Columns 7 and 8 in Table 6.

6.2. **Group Fixed Effects.** Group fixed effects would contribute to the reduction of bias from omitted variables. The reason why we did not use them in the first place is that we need variation in group size in order to identify the effect of SIZE×LACK RIGHTS, given that LACK RIGHTS is already subsumed in the country-fixed effects. With group fixed effects, all time-invariant controls drop out from the regression, including two of our key variables (SIZE and SIZE×LACK RIGHTS). Nevertheless, it is still possible to test one of the two key hypotheses, that pertaining to SIZE× OIL. Columns 1– 3 in Table 7 do just that. Note that the three columns contain group fixed effects but are still different, because Column 1 excludes lagged conflict, while Columns 2 and 3 include this variable and are estimated by OLS and system GMM (Blundell and Bond 1998), respectively. In all cases, the interaction of SIZE and OIL remains negative and significant.

6.3. **Alliances in Conflict.** It may so happen that in some cases, *alliances* of groups could form. For instance, in the First Sudanese Civil War, also known as the Anyanya Rebellion, a conglomeration of the Acholi, Bari, Dinka, Lotuko, Madi, Nuer and the Zande from South Sudan came together, albeit in an alliance marked by substantial infighting. Other alliances are not hard to find: e.g., ethnic alliances exist in the Casamance conflict in Senegal or in the Liberian war that toppled the Taylor government.

As already described, the data we use code ethnic groups in conflict against the State. In the case of alliances, *each* ethnic group is so coded. As expected, the dataset has a number of such conflicts. Now, several of these conflicts are genuinely separate conflicts, and some are not. It is unclear how one might approach this problem comprehensively without running into severe issues of endogeneity in the definition of a "coalition."

Without pretending to satisfactorily solve this dilemma, one can run a rough variant of our exercise by mechanically combining all multiple instances of conflict. Table B4 is similar to Table 1 but uses an alternative definition of group size, $\text{SIZE}_{\text{COAL}}$. This variable is defined as follows: for

[39] More specifically, we drop from the sample groups with a value of EXCLUDED (a pre-sample average of an exclusion dummy over the years 1960–1975) larger than 0.5. Results are very robust to other ways of defining exclusion.

	Dependent Variable: **Conflict Incidence**		
	[1]	[2]	[3]
OIL	6.253	1.348	-0.233
	(0.191)	(0.289)	(0.523)
SIZE× OIL	-115.074*	-27.901*	-9.609***
	(0.052)	(0.086)	(0.003)
GDP	-0.006	0.001	0.004**
	(0.194)	(0.528)	(0.011)
POP	0.017	0.003	-0.000
	(0.141)	(0.304)	(0.646)
LAG		0.797***	0.919***
		(0.000)	(0.000)
c	-0.230	-0.056	
	(0.278)	(0.268)	
Estimation	OLS	OLS	System GMM
GFE	✓	✓	✓
R^2	0.011	0.640	–
Obs	57559	57559	57559

Table 7. Variations: Group Fixed Effects. This table regresses conflict incidence on group size and indices of private and public prizes. All regressions contain year dummies. GFE denotes group fixed effects. Columns 1 and 2 have been estimated by OLS and Column 3 by system GMM (Blundell and Bond, 1998). p-values are reported in parentheses. $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

peace years, $\text{SIZE}_{\text{COAL}}$ and SIZE coincide. For years where *some* group is in conflict, $\text{SIZE}_{\text{COAL}}$ adds up the size of *all* the groups in conflict in the same country and year. In this way we try to capture the possibility that there exists an alliance between the fighting groups. The variable OIL_{COAL} is defined in a similar way: in peace years, OIL and OIL_{COAL} are identical. In case of conflict, $\text{SIZE}_{\text{COAL}}$ adds up the oil in the homelands of all the groups in conflict in that country and year. Our conclusions are robust to this variation.

6.4. **Alternative Measures of Conflict.** Table B5 considers alternative measures of conflict: Columns 1–3 use conflict *onset*, while columns 4–7 collapse the time dimension of the data. The dependent variable is the share of years over 1975–2006 in which a group has been involved in conflict against the State (Columns 4–6), and the share of onset years (Column 7).

Qualitatively, the results are very similar to those described above. The interactions of group size and the publicness/privateness indicators have the predicted sign and are highly significant. There is only one exception: when the dependent variable is conflict *onset* and the two interactions are in the regression (Column 3), both interactions maintain their predictive signs but become insignificant. It is important to notice, however, that the number of onset observations in the post 1975 sample employed in that regression is only 130 (0.35% of all observations). In contrast, in the same period there are 1,680 conflict observations, around 4.5% of all observations. This implies that onsets in the post 1975 sample are rare events. As a result coefficients are estimated with considerable error. Also notice that both interactions are significant in the cross-section, where the dependent variable is the share of onset years (Column 7).

6.5. **Alternative Ways of Measuring Oil Wealth.** Table B6 considers alternative ways of measuring oil wealth. Columns 1–3 consider oil rents at the national level, rather than just the oil that lies in the homeland of the group. Column 1 shows that the same results on the interaction between group size and oil goes through for national oil rents. That said, the level effects suggest clearly that it is resources under the spatial control of a particular group that are highly linked to conflict involving *that group*. Group oil matters for conflict involving that group; controlling for that, national oil does not. Moreover, columns 1 and 2 show that, once the interaction between group oil and group size is introduced in the regression, it is significant, while the corresponding interaction between group size and national oil falls silent. This suggests that, although our conclusions are robust to considering national oil rents, group-level oil seems to be a better proxy for the private payoff. Columns 4 and 5 use an alternative measure of oil abundance, OIL (SURFACE), based on the ethnic homeland's area covered by oil. Identical conclusions are obtained.

6.6. **Assessing the Importance of the Omitted Variable Bias.** Despite our attempts to control for a large number of potential confounders, we still cannot completely rule out the possibility that unobserved variables are biasing our results. However, it is possible to assess the likelihood that our observed effect is solely due to selection bias. To that effect, we apply a technique recently developed by Oster (2017), which builds on the work by Altonji, Elder and Taber (2005) and Bellows and Miguel (2008). This method allows to determine the degree of selection on unobservables relative to observables (denoted by δ) that would be necessary to explain away the result. If the set of observed controls is representative of all possible controls, then a large value of δ suggests that it is implausible that omitted variable bias explains away the entire effect. Altonji et al. (2005) and Oster (2017) suggest the use of a cut-off of δ (δ^*) equal to 1. This value means that selection on unobservables would need to be at least as important as that on observables to produce a treatment effect of zero. Thus, if for example a value of $\delta = 2$ is obtained, it would suggest that the unobservables would need to be twice as important as the observables to produce a treatment effect of zero. One reason to favor the cut-off $\delta^* = 1$ is that researchers typically choose the controls they believe ex ante are the most important (Angrist and Pischke, 2010) and thus situations where selection on unobservables is larger than that of the observed controls are deemed unlikely.

The statistic employed to compute δ is designed to evaluate the stability of the variable(s) of interest to the introduction of controls. More specifically, it is a function of the coefficient of the variable of interest estimated in a full model (that contains all controls), the same coefficient obtained in a restricted model with no (or few) controls, the R^2s obtained in these regressions and R^2_{max}, the R^2 from a hypothetical regression of the outcome on treatment and both observed and unobserved controls. If the outcome can be fully explained, then $R^2_{\text{max}} = 1$. However, as acknowledged by Oster (2017), in most empirical settings it seems likely (due, for example, to measurement error) that the outcome cannot be fully explained even if the full control set is included.

In our case, the variables of interest are either SIZE×OIL or SIZE×LACK RIGHTS. The full model corresponds to our baseline specification (Column 7 in Table 1). Since our variables of interest are interactions, the restricted model is one where the only controls are the variables included in the corresponding interaction.[40] The values of δ are quite sensitive to the choice of the (unobserved) value of R^2_{max} . Then, we have computed the maximum value of R^2_{max} we can use in order to obtain values of δ larger than 1. This value turns out to be quite large (around 0.93 for both interactions). Since the variables employed in our regressions are clearly not perfectly measured, we believe that this is a reasonable value for the maximum R^2 that can achieved given the quality of the available data. Therefore, we argue that the latter value of R^2_{max} gives support to our claim that omitted variable bias is not likely to explain our results.

6.7. **Alternative Clustering Strategies.** As explained in Section 4.3, throughout the paper we follow Abadie et al (2017) and cluster errors at the group or at the country level, depending on whether the corresponding treatment is assigned at the group or at the country level. Our results are firmly robust to other clustering strategies; for instance, to two-way clustering. Table B7 reproduces Table 1 but this time standard errors are adjusted for clustering at the ethic homeland *and* at the country level. Notice that since ethnic homelands are often split by an international border, the latter dimensions are not nested. Our conclusions are robust to considering alternative clustering schemes.

6.8. **Nonlinear Models.** Since our baseline dependent variable is binary, we have re-estimated our baseline specifications using a logit specification. Table B1 in Appendix B.1 presents the results. All equations contain the full list of controls as well as country and year fixed effects, but differ on the interactions included in them: Column 1 includes the interaction of SIZE and OIL, Column 2 that of SIZE and LACK RIGHTS, while Column 3 considers both of them. The coefficients of the interactions of SIZE and the public and private payoffs maintain the expected signs and remain significant. In nonlinear specifications, however, one has to be cautious when interpreting the change in two interacted variables, as Ai and Norton (2003) pointed out. Appendix B.1 discusses this issue in more detail and shows that our conclusions still hold when nonlinear estimation is considered.

6.9. **Dropping Regions of the World.** Table B8 drops observations from particular regions of the world. Those regions are: former USSR countries (columns 1 and 2), Asia (columns 3 and 4), Middle East (columns 5 and 6), West-South Africa (columns 7 and 8), East and Central Africa (columns 9 and 10), and Latin America (columns 11 and 12). For each region, the first (second) Column replicates Column 4 (5) in Table 1, that is, we consider specifications with and without the interaction of group size and the public prize. Results are generally robust. The only exception is when East and Central Africa observations are dropped. In this case, the interaction of SIZE and OIL is significant in Column 9, but when the interaction of SIZE and LACK RIGHTS is introduced as well (Column 10), it ceases to be so (while the p-value of the public interaction is 0.14). To put this result in perspective, however, notice East and Central Africa is by far the most conflictual region in our sample: it represents 15% of all observations but 30% of all conflict

[40]That is, SIZE and OIL or SIZE for the private and public interactions, respectively.

observations. Thus, it is not surprising that the coefficients are estimated less precisely when a large number of conflict observations is dropped from the sample.

7. CONCLUSION

Group size matters in social conflict. But there is more than one view on just how it matters. In the introduction to his essay, "On Liberty," John Stuart Mill (1859) writes:

"Society . . . practices a social tyranny more formidable than many kinds of political oppression, since, though not usually upheld by such extreme penalties, it leaves fewer means of escape, penetrating much more deeply into the details of life, and enslaving the soul itself. Protection, therefore, against the tyranny of the magistrate is not enough; there needs protection also against the tyranny of the prevailing opinion and feeling, against the tendency of society to impose, by other means than civil penalties, its own ideas and practices as rules of conduct on those who dissent from them . . . "

Mill is referring to the tyranny of the majority, a notion that also appears in the writings of John Adams and in the Federalist Papers, in the 18th century, and then amplified and used more extensively by Alexis de Tocqueville (1835).

Arrayed against this distinguished company are Wilfredo Pareto and Mancur Olson, who emphasize the power of minorities to cohere around a cause. In the words of Pareto (1927, p. 379), who was remarking on protectionist tendencies in trade,

"[A] protectionist measure provides large benefits to a small number of people, and causes a very great number of consumers a slight loss. This circumstance makes it easier to put a protection measure into practice."

In this paper we've studied a model of social conflict, in which the conflict may be over a *public* or a *private* good. The main result, that we explore empirically through a variety of specifications, is that conflict is more likely in the presence of a private prize when the group is small, and it is more likely in the presence of a public prize when the group is large. By using a global panel dataset at the ethnic group level we find powerful and robust empirical support for these claims. This is our approach to reconciling Tocqueville-Mill with Pareto-Olson.

Our approach can be extended to other questions of interest. Specifically, as already hinted at in Section 3.3, one can develop a theory of conflict in which there are *multiple* potential threats to peace from different groups, thereby generating conflict "in equilibrium" even in the presence of inefficiencies. Because a multiplicity of groups are typically formed using ethnic markers, such a theory could also help us understand why ethnic conflict might be salient.[41]

[41]Other factors that bear on the salience of ethnic violence includes the greater visibility of ethnicity (Caselli and Coleman 2013), or the ability of an economically unequal ethnic group to to exploit the synergy of money and labor when engaging in conflict (Esteban and Ray 2008).

REFERENCES

Abadie, A., S. Athey, G. W. Imbens and J. Wooldridge (2017), "When Should You Adjust Standard Errors for Clustering?", Mimeo.

Ai, C. R. and E. C. Norton (2003), "Interaction terms in logit and probit models," *Economics Letters* **80**: 123–129.

Altonji, J. G., T. E. Elder, and C. R. Taber (2005), "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy* **80**:151–184.

Angrist, J. D., and J. Pischke (2010), "The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics." *Journal of Economic Perspectives*, **24**: 3-30.

BALIGA, S. AND T. SJOSTROM (2004), "Arms Races and Negotiations, *Review of Economic Studies* **71**, 351–369.

Bellows, J., and E. Miguel (2009), "War and Local Collective Action in Sierra Leone". *Journal of Public Economics* **93**:1144?1157.

Berman N., M. Couttenier, D. Rohner and M. Thoenig (2015), "This mine is mine! How minerals fuel conflict in Africa," Mimeo.

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

Bester, H. and K. Wärneryd (2006), "Conflict and the Social Contract," *Scandinavian Journal of Economics* **108**, 231–249.

Blundell R. and S. Bond (1998), "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics* **87**: 115–143.

Bondareva, O. N. (1963), "Some applications of linear programming methods to the theory of cooperative games", *Problemy Kybernetiki* **10**: 119–139.

Bruk, S., and V. S. Apenchenko (eds) (1964), *Atlas Narodov Mira [Atlas of the Peoples of the World]*. Moscow: Glavnoe Upravlenie Geodezii i Kartogra.

Cederman, L.-E., H. Buhaug and Rod, J. K. (2009), "Ethno-Nationalist Dyads: A GIS-Based Analysis," *Journal of Conflict Resolution* **53**: 496-525.

Chamberlin, J. (1974), "Provision of Collective Goods as a Function of Group Size," *American Political Science Review* **68**: 707–716.

Doyle, M. and N. Sambanis (2006), *Making War and Building Peace*. Princeton: Princeton University Press.

Dube, O. and J. Vargas (2013), "Commodity Price Shocks and Civil Conflict: Evidence from Colombia," *Review of Economic Studies* **80**, 1384–1421.

Esteban, J. and D. Ray (2008), "On the Salience of Ethnic Conflict," *American Economic Review* **98**: 2185–2202.

Esteban, J. and D. Ray (2001), "Collective Action and the Group Size Paradox," *American Political Science Review* **95**: 663–672.

Esteban, J. and D. Ray (2011), "Linking Conflict to Inequality and Polarization," *American Economic Review* **101**: 1345–74.

Esteban, J., Mayoral, L. and D. Ray (2012),"Ethnicity and Conflict: An Empirical Investigation," *American Economic Review* **102**: 1310–1342.

Fearon, J. (1995), "Rationalist Explanations for War, *International Organization* **49**, 379–414.

Fearon, J. (2003), "Ethnic and cultural diversity by country," *Journal of Economic Growth* **8**:195-222.

Fearon, J. (2005), "Primary Commodity Exports and Civil War," *Journal of Conflict Resolution* **49**, 483–507

Fearon, J. and D. Laitin (2003), "Ethnicity, Insurgency, and Civil War," *American Political Science Review* **97**: 75-90.

Fischer, G., F. Nachtergaele, S. Prieler, H. van Velthuizen, L. Verelst, and D. Wiberg (2008), *Global Agro-ecological Zones Assessment for Agriculture*. IIASA, Laxenburg, Austria and FAO, Rome, Italy.

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

Freedom House (2014), *Freedom of the World*, dataset, `https://freedomhouse.org/report/freedom-world-aggregate-and-subcategory-scores`.

Garfinkel, M. and S. Skaperdas (2000), "Conflict Without Misperceptions or Incomplete Information: How the Future Matters," *Journal of Conflict Resolution* .**44**, 793–807

Gleditsch, N. P., P. Wallensteen, M. Eriksson, M. Sollenberg, and H. Strand (2002), "Armed Conflict 1946-2001: A New Dataset," *Journal of Peace Research* **39**: 615–637.

IntierraRMG (2015), "SNL Metals and Mining", dataset, `http://www.snl.com/Sectors/metalsmining/Default.aspx`.

Jackson, M. and M. Morelli (2007), "Political Bias and War," *American Economic Review* **97**, 1353–1373.

Kirshner, J. (2000), "Rationalist Explanations for War?," *Security Studies* **10**: 143–150.

Le Billon, P. (2001), "The Political Ecology of War: Natural Resources and Armed Conflicts," *Political Geography* **20**, 561–584.

Licklider, R. (1995), "The Consequences of Negotiated Settlements in Civil Wars, 1945-1993," *American Political Science Review* **89**: 681-690.

Lujala, P. (2010), "The Spoils of Nature: Armed Civil Conflict and Rebel Access to Natural Resources," *Journal of Peace Research* **47**, 15–28.

Lujala, P., J. K.Rod and N. Thieme (2007), "Fighting over Oil: Introducing a New Dataset," *Conflict Management and Peace Science* **24**: 239-56.

McGuire, M.C. (1974), "Group Size, Group Homogeneity and the Aggregate Provision of a Pure Public Good under Cournot Behavior," *Public Choice* **18** (Summer): 107–126.

Marwell, G. and P. Oliver (1993), *The Critical Mass in Collective Action. A Micro-Social Theory*. Cambridge, UK: Cambridge University Press.

Mill, J. S. (1859), *On Liberty*, London: Parker and Son.

Morelli, M. and D. Rohner (2015), "Resource Concentration and Civil Wars," *Journal of Development Economics* **117**: 32–47.

Oliver, P. and G. Marwell (1988), "The Paradox of Group Size in Collective Action: A Theory of the Critical Mass. II," *American Sociological Review* **53**: 1–8.

Olson, M. (1965), *The Logic of Collective Action*. Cambridge, MA: Harvard University Press.

Oster, E. (2016), "Unobservable Selection and Coefficient Stability: Theory and Evidence". *Journal of Business and Economic Statistics*, forthcoming.

Pareto, V. [1906] (1927), *Manual of Political Economy*. New York: A.M. Kelley.

Polity IV (2014), "Political Regime Characteristics and Transitions", dataset, `http://www.systemicpeace.org/inscrdata.html`.

Powell, R. (2004), "The Inefficient Use of Power: Costly Conflict with Complete Information," *American Political Science Review* **98**, 231–241.

Powell, R. (2006), "War as a Commitment Problem," *International Organization* **60**, 169–203.

Religion and State Project ARDA, Round 2, Dataset: `http://www.thearda.com/ras/`.

Sambanis, N. (2001), "Do Ethnic and Non-Ethnic Civil Wars Have the Same Causes? A Theoretical and Empirical Enquiry," Journal of Conflict Resolution **45**: 259-282.

Sánchez-Pagés, S. (2009), "Conflict as Part of the Bargaining Process," *Economic Journal* **119**, 1189–1207.

Sandler, T. (1992), *Collective Action: Theory and Applications*. Ann Arbor, MI: University of Michigan Press.

Scarf, H. (1967), "The Core of an n-Person Game," *Econometrica* **35**, 50–69.

Schneider, G.and N. Wiesehomeier (2006), "Ethnic Polarization, Potential Conflict, and Civil Wars: Comment," unpublished manuscript, University of Konstanz.

Shapley, Ll. S. (1967), “On Balanced Sets and Cores,” *Naval Research Logistics Quarterly* **14**: 453–460.

Skaperdas, S. (1996) “Contest Success Functions,” *Economic Theory* **7**: 283–290.

Slantchev, B. (2003), “The Power to Hurt: Costly Conflict with Completely Informed States,” *American Political Science Review* **47**, 123–135.

Taylor, M. (1987) *The Possibility of Cooperation.* Cambridge, UK: Cambridge University Press.

Tocqueville, A. de (1835), *Democracy in America*, Chicago, IL: University of Chicaco Press (2000 edition).

Vreeland, J. R. (2008). “The Effect of Political Regime on Civil War,” *Journal of Conflict Resolution*, **52**: 401–425.

Weidmann, N. B., Rod, J.K. and Cederman, L. -E. (2010), “Representing Ethnic Groups in Space: A New Dataset,” *Journal of Peace Research* **47**: 491-499.

Wooldridge, J.M. (2003), “Cluster-Sample Methods in applied Econometrics,” *The American Economic Review* **93**, Papers and Proceedings of the One Hundred Fifteenth Annual Meeting of the American Economic Association, 133–138.

Wucherpfennig, J., Weidmann, N., Girardin, L., Cederman, L. and A. Wimmer (2011), “Politically relevant ethnic groups across space and time: Introducing the GeoEPR dataset,” *Conflict Management and Peace Science* **28**: 423–437.

Appendix A

This Appendix contains a proof of Proposition 3 (Section A.1), definitions of all the variables in the empirical analysis (Section A.2) as well as a table of summary statistics (Section A.3).

A.1. **Proof of Proposition 3.** Suppose that the conditions in the proposition are met, but that there is indeed an unblocked allocation $\mathbf{x}$. For every group $G \in \mathcal{C}$, we have

$$\int_{j \in G} x(j) \geq v[kp(m) + (1-k)p(m)^2] > vm. \tag{14}$$

Pick a collection of balancing weights $\{\lambda(G)\}_{G \in \mathcal{C}}$. Multiplying each side of (14) by $\lambda(G)$, and adding over all groups in $\mathcal{C}$, we see that

$$\sum_{G \in \mathcal{C}_j} \lambda(G) \int_{j \in G} x(j) > \sum_{G \in \mathcal{C}_j} vm\lambda(G).$$

Because $\{\lambda(G)\}_{G \in \mathcal{C}}$ are balanced weights, this implies $\int_j x(j) > v$, a contradiction. ■

A.2. **Variable Definitions.** Conflict *incidence*: group-level dummy variable equal to 1 for those years where an ethnic group is involved in armed conflict against the state resulting in more than 25 battle-related deaths. Source: CBR.

Conflict *onset*: group-level dummy variable that equals 1 in a given year if an armed conflict against the state resulting in more than 25 battle-related deaths involving that ethnic group newly starts. For ongoing wars, ONSET is coded as missing. Source: CBR.

Share of conflict years: group-level variable that captures the share of years a group has been in conflict against the State in the period 1960–2006. Source: CBR.

Share of onset years: group-level variable that captures the share of years a group has started in conflict against the State (onset years) in the period 1960–2006. Source: CBR.

SIZE: relative size of the group, source: CBR.

OIL: log of the homeland area covered by oil (in thousands of square kilometres) times the international price of oil. To avoid taking the log of zero, 1 has been added to all observations. Source: Oil fields: Petrodata (Lujala et al. 2007). Oil prices: the World Bank.

OIL(AREA): log of homeland area covered by oil (in ’000 km^2). To avoid taking the log of zero, 1 has been added to all observations. Source: Petrodata (Lujala et al. 2007).

OIL(SHARE): ratio of OIL(AREA) to homeland area. Source: Petrodata and GREG.

OIL CONCENTRATION: Herfindahl index of oil reserve distribution across groups. Source: Petrodata and GREG.

OIL (COUNTRY): log of the area of the country covered by oil (in thousands of square kilometres) times the international price of oil. To avoid taking the log of zero, 1 has been added to all observations. Source: information on oil fields comes from Petrodata (Lujala et al. 2007). Data on oil prices comes from the World Bank.

MINES: measures mineral availability in the ethnic homeland and is computed as follows. First, we consider 13 minerals (bauxite, coal, copper, diamond, gold, iron, lead, nickel, platinum, phosphate, silver, tin and zinc) for which international price data is readily available. For each mineral, year and ethnic group, we create a dummy variable that is one if the group has at least one active mine of that mineral. Then, each of these dummies is multiplied by its normalized international price. The latter is constructed as the log of its international price divided by the log of its price in 1980 (the year when the data starts). The variable MINES is computed as the sum of the resulting products. Data on mineral availability comes from the *Raw Material Data* dataset, (IntierraRMG, 2015) whereas data on mineral prices is provided by the World Bank.

AUTOC: country average of the Polity IV autocracy index for the years 1960 to 1975.

EXCLUDED: average over the period 1960-1975 of *excluded*, a dummy variable that is 1 if the ethnic group is in power in a given country and year (source: CBR).

CIVIL RIGHTS: Lack of civil liberties from Freedom house. We rescale the original index so it is measured between 0 and 1 (where 0 indicates highest level of civil liberties). For each country, we average its value from 1972 to 1975 and assign the resulting quantity to all post 1975 years.

POLITICAL RIGHTS: Lack of political liberties from Freedom House, rescaled to lie between 0 and 1 (where 0 indicates highest level of civil liberties). For each country, we average its value from 1972 to 1975 and assign the resulting quantity to all post 1975 years.

LACK RIGHTS: average of CIVIL RIGHTS and POLITICAL RIGHTS.

INFANT MORTALITY: Number of deaths of children under 5 per 1000 live births. For each country, we consider the average of this quantity over the period 1960 to 1975 and assign it to all subsequent years. Source: Unicef Global Databases.

GIP: dummy variable that is 1 if the ethnic group is in power in a given country and year (lagged one year), source: CBR.

GROUPAREA: area of the ethnic homeland (in thousands of square kilometres), source: GREG.

AREA(SHARE): area of the ethnic homeland relative to total area of the country, source: GREG.

LAND PC: log of the total area divided by group population, source: GREG.

SOILCONST: measures limitations that homeland soil presents to agriculture, constructed using the Harmonized World Soil Database from Fischer et al., (2008). Fisher et al. (2008) construct a global grid of 38 nutrient availability ranked from 1 (no or slight constraints) to 4 (very severe constraints), and also including categories 5 (mainly non-soil), 6 (permafrost area) and 7 (water bodies). SOILCONST is the average of the cell values pertaining to the group's homeland.

DISTCAP: group's distance to the country capital, source: CBR.

MOUNT: 0-1 index capturing the group's share of mountainous terrain, source: CBR.

PEACEYEARS: number of years since the last group-level onset and LAG is lagged conflict incidence, source:CBR.

PARTITIONED: dummy variable that is 1 if the ethnic homeland covers two or more countries, source: GREG.

GDP: log of (country-level) GDP per capita, lagged one year. Source: Penn World Tables.

POP: log of total country population (POP), lagged one year. Source: Penn World Tables.

POLITY: Polity 2 index, lagged one year. Source: Polity IV.

A.3. **Summary Statistics.** Table A1 provides summary statistics of the main variables employed in the empirical analysis.

	Obs	Mean	SD	Min	Max
INCIDENCE	64001	0.04	0.19	0.00	1.00
ONSET	61928	0.00	0.06	0.00	1.00
SHARE CONFLICT	1475	.030	.123	0	.982
SIZE	64001	0.10	0.23	0.00	1.00
OIL	64001	0.00	0.001	0.002	0.01
LACK RIGHTS	42950	0.64	0.28	0.00	1.00
PUBLIC INDEX	38049	0.00	0.96	-2.17	1.32
PRIVATE INDEX	61968	-0.00	0.70	-0.41	5.37
CIVIL RIGHTS	42950	0.62	0.27	0.00	1.00
POLITICAL RIGHTS	42950	0.66	0.31	0.00	1.00
INFANT MORTALITY	60669	100.88	50.60	12.54	211.01
AUTOC	45870	0.53	0.29	0.00	1.00
EXCLUDED	63544	0.86	0.34	0.00	1.00
MINES	65639	0.57	1.42	0.00	13.00
GIP	64001	0.14	0.35	0.00	1.00
AREA	64001	84.28	406.74	0.04	7354.72
AREA (SHARE)	61968	0.09	0.20	0.00	1.01
AREA PC	62103	-9.73	2.00	-15.90	1.55
OIL PC	62103	0.00	0.00	0.00	0.17
SOILCONST	64001	1.62	0.78	0.00	6.15
DISTCAP	64001	0.92	1.03	0.00	7.97
MOUNT	64001	0.37	0.36	0.00	1.00
PARTITIONED	64001	0.62	0.48	0.00	1.00
GDP	56945	7.75	1.16	5.08	11.16
POP	61893	17.08	1.81	11.73	20.98
POLITY	58120	-0.09	0.70	-1.00	1.00

Table A1. Summary Statistics

APPENDIX B. (ONLINE APPENDIX [NOT FOR PUBLICATION])

This Appendix presents additional discussion and tables probing the robustness of the results in the main text; see Sections 5.2 and 6 for more details. Section B.1 and Table B1 present results estimated by maximum likelihood in a logit specification; this complements Section 6.8 in the main text. Section B.2 divides the sample in four groups (countries with/without private/public prizes) and compares for each of the bins the average sizes of groups that have experienced conflict over the sample with those that haven't. Results are provided in Table B2. Table B3 considers variations of the flexible specification for group size, oil and conflict in Table 2. Table B4 allows for the possibility that groups form coalitions (see Section 6.3). Table B5 considers additional conflict measures. Table B6 considers alternative ways of measuring oil wealth. Table B7 reestimates Table 1 clustering errors at the ethnic homeland and country level (two-way clustering). Table B8 considers the robustness of our results to dropping regions of the world.

B.1. **Interactions in Nonlinear Models.** Table B1 re-estimates our baseline models using a logit specification. The coefficient of the interactions has the expected signs and are highly significant. Interpreting the coefficients associated with interactions is straightforward in linear models, as they are simply the appropriate cross-partial derivatives of the dependent variable with respect to the relevant variables in the interaction. However, this logic does not extend to nonlinear models, as shown by Ai and Norton (2003). In non-linear models, the cross-partial derivative does not admit a simple interpretation, and important differences arise with respect to the linear case. First, the "true sign" of the interaction does not need to equal the sign of the cross-partial derivative. Second, the significance of that interaction cannot be tested with a simple t-test on the coefficient of the interaction term (in the regression). Third, given the nonlinearity, the value of the interaction term depends on all the independent variables of the model. See Ai and Norton(2003) for a discussion.

To overcome these difficulties and in order to facilitate the interpretation of the interactions reported in Table B1, we have evaluated the cross-partial derivative at each of the points in our sample. Panel (a) in Figure B1 plots the derivative of the dependent variable with respect to SIZE and OIL, using the specification in Column 1, Table B1. This figure shows that the cross-derivative is negative for most observations in our sample, a result that mimics the one obtained for the linear case. Panel (b) in Figure B1 plots the z-statistics associated with the cross-partial derivative for each of the points in the sample, together with confidence bands (at the 90 per cent level). This figure shows that the effect is significant in most cases. Similar results are found when interpreting the interaction of SIZE and LACK RIGHTS. In this case, the cross-partial derivative is positive and significant for most of the observations.[42]

B.2. **Average Percentage Sizes of Groups in Conflict.** The informal analysis here helps visualize our data in the simplest possible way. We have classified countries in four groups, depending on whether they are rich/poor in oil reserves and on whether they have low/high levels of LACK

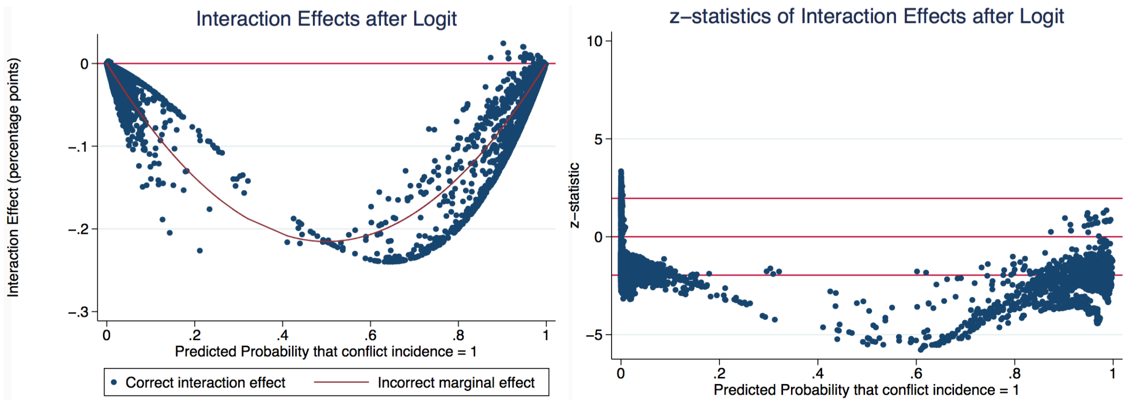
[42] For the sake of brevity, we don't report the corresponding graphs as they are very similar to those associated with SIZE and OIL, but they are available upon request.

	Dependent Variable: **Conflict Incidence**		
	[1]	[2]	[3]
SIZE	23.387***	-15.634	-0.179
	(0.000)	(0.184)	(0.988)
OIL	243.897***	152.334***	213.912***
	(0.000)	(0.001)	(0.000)
SIZE× OIL	-8383.401***		-7403.100***
	(0.000)		(0.000)
SIZE×AUTOC		40.962**	35.839**
		(0.018)	(0.048)
GDP	0.376**	0.338	0.350*
	(0.024)	(0.101)	(0.093)
POP	1.772**	0.828	0.850
	(0.039)	(0.447)	(0.436)
GROUPAREA	0.000	-0.000	-0.000
	(0.523)	(0.270)	(0.975)
GIP	-0.409	-0.094	-0.065
	(0.106)	(0.720)	(0.799)
SOILCONST	-0.193	-0.118	-0.212*
	(0.167)	(0.311)	(0.081)
DISTCAP	0.565***	0.486***	0.508***
	(0.000)	(0.002)	(0.001)
MOUNT	0.613***	0.587**	0.686***
	(0.009)	(0.019)	(0.007)
PARTITIONED	-0.150	-0.099	
	(0.323)	(0.540)	
LAG	7.270***	7.255***	7.229***
	(0.000)	(0.000)	(0.000)
c	-46.125***	-27.108	-27.751
	(0.007)	(0.211)	(0.201)
R^2			
Obs	27344	20960	20960

Table B1. Group Size and Conflict: Non-Linear Models. This table regresses conflict incidence on group size and indices of private and public prizes, along with interactions between subsets of these variables as suggested by the theory. Estimation has been carried out my maximum likelihood in a Logit model. All regressions contain year dummies and country fixed effects. Robust standard errors clustered at the group level have been computed. p-values are reported in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

RIGHTS.[43] For each category, Table B2 provides the average size of groups that have never experienced conflict and, in parentheses, similar information corresponding to groups that have

[43]To do that, we have first collapsed the time dimension of the data by considering for each group the first non-missing value of OIL and LACK RIGHTS. A country is classified as oil rich (poor) if the the oil reserves corresponding to the group with largest reserves in the country are above the overall median. Analogously, a country is classified as low (high) level of LACK RIGHTS if the first non-missing value in the sample is below (above) the median.



(a) Interpreting Interactions in Nonlinear Models: Cross-Partials

(b) Interpreting Interactions in Nonlinear Models: z-Statistics

Figure B1. These graphs depict the value of the cross-partial derivative of conflict incidence with respect to OIL and SIZE (Panel a) and the value of the cross-partial derivative of conflict incidence with respect to OIL and SIZE (Panel b), for each of the points in the sample. Estimates from Table 7 (Column 1) have been employed to compute the estimates.

experienced conflict in the period considered in our sample. Our theory makes predictions for the main diagonal and this simple analysis confirms the main implications of the theory: the average size of groups in conflict changes dramatically when conflict is over public prizes (where the average size of groups in conflict is .16) or over private prizes (where the same figure falls to .04).

	AVERAGE %-SIZE OF GROUPS	
	LACK RIGHTS low	LACK RIGHTS high
OIL high	6.7% (**4.0%**)	8.1% (**7.0%**)
OIL low	18.1% (**6.3%**)	8.0% (**16.0%**)

Table B2. Average Sizes of Groups in Peace and in Conflict. This table divides all countries in our sample in four categories depending on whether they have high/low values of oil reserves and on whether they have high/low values of LACK RIGHTS. Then, for each bin it reports the average percentage size of groups that have never experienced conflict in our sample as well as the average percentage size of groups that *have* experienced conflict (in parentheses). See Footnote 43 for details on how countries have been assigned to bins.

	Dependent Variable: **Conflict Incidence**						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
SIZE	0.016	0.012	0.018	0.033	0.025	0.050	0.262
	(0.553)	(0.660)	(0.495)	(0.252)	(0.293)	(0.218)	(0.392)
SIZE× LACK RIGHTS	0.093*	0.096*	0.079*	0.075*	0.042	0.077	1.103**
	(0.052)	(0.058)	(0.086)	(0.082)	(0.159)	(0.112)	(0.015)
SIZE× OIL_{0-25}	0.044	-0.152	0.028	0.033	0.098	0.094	0.298
	(0.704)	(0.199)	(0.805)	(0.742)	(0.361)	(0.460)	(0.810)
SIZE× OIL_{25-50}	0.181	0.365	0.127	-0.050	0.210	0.219	3.976
	(0.554)	(0.380)	(0.760)	(0.754)	(0.514)	(0.525)	(0.608)
SIZE× OIL_{50-75}	-0.170**	-0.115	-0.150**	-0.146*	-0.157**	-0.218**	-2.417**
	(0.044)	(0.125)	(0.040)	(0.061)	(0.034)	(0.015)	(0.044)
SIZE× $\text{OIL}_{\geq 75}$	-0.112***	-0.137***	-0.110**	-0.114**	-0.062*	-0.166***	-0.970*
	(0.010)	(0.006)	(0.013)	(0.013)	(0.051)	(0.006)	(0.057)
OIL_{0-25}	-0.003	0.002**	-0.003	-0.003	-0.004	-0.005	-0.028
	(0.377)	(0.016)	(0.386)	(0.419)	(0.237)	(0.235)	(0.464)
OIL_{25-50}	-0.002	0.001	-0.002	-0.002	-0.002	-0.002	-0.041
	(0.605)	(0.542)	(0.578)	(0.525)	(0.583)	(0.687)	(0.257)
OIL_{50-75}	0.005*	0.005*	0.005	0.005*	0.004	0.008**	0.072**
	(0.064)	(0.062)	(0.125)	(0.078)	(0.170)	(0.042)	(0.029)
$\text{OIL}_{\geq 75}$	0.007**	0.009**	0.006**	0.005*	0.005*	0.009**	0.052*
	(0.031)	(0.019)	(0.050)	(0.099)	(0.068)	(0.031)	(0.081)
GIP	-0.003	-0.001	-0.002	-0.002	-0.005**	-0.002	-0.049**
	(0.257)	(0.667)	(0.424)	(0.390)	(0.027)	(0.349)	(0.021)
POLITY	-0.001	-0.000	-0.001	-0.001	-0.001	-0.000	-3.262***
	(0.437)	(0.837)	(0.555)	(0.261)	(0.241)	(0.975)	(0.000)
GDP	0.003	0.004	0.002	0.001	0.004	0.002	-1.057***
	(0.298)	(0.376)	(0.373)	(0.793)	(0.189)	(0.410)	(0.000)
POP	0.001	-0.001	-0.001	-0.010	-0.005	0.004	-0.394
	(0.931)	(0.885)	(0.904)	(0.300)	(0.605)	(0.539)	(0.389)
GROUPAREA	0.000	0.000	0.000	0.000	0.000*	0.000	0.000
	(0.416)	(0.415)	(0.440)	(0.392)	(0.078)	(0.426)	(0.195)
SOILCONST	-0.001	0.000	-0.000	-0.000	-0.002***	-0.001	-0.003
	(0.351)	(0.801)	(0.582)	(0.488)	(0.002)	(0.290)	(0.750)
DISTCAP	0.002*	0.001	0.003*	0.002*	0.002	0.003*	0.019
	(0.085)	(0.292)	(0.078)	(0.086)	(0.176)	(0.090)	(0.190)
MOUNT	0.002*	0.001	0.002	0.002	0.004**	0.003*	0.017**
	(0.100)	(0.669)	(0.178)	(0.242)	(0.019)	(0.080)	(0.039)
PARTITIONED	-0.001	-0.001	-0.002	-0.001	0.001	-0.002	-0.010
	(0.434)	(0.382)	(0.176)	(0.464)	(0.160)	(0.327)	(0.340)
LAG	0.900***	0.893***	0.899***	0.911***	0.890***	0.898***	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	
c	-0.048	-0.014	-0.012	0.162	0.040	-0.095	12.826**
	(0.706)	(0.919)	(0.921)	(0.355)	(0.797)	(0.375)	(0.033)
R^2	0.855	0.856	0.854	0.870	0.849	0.841	0.419
Obs	37504	26829	36473	31513	31580	31092	1294

Table B3. Group Size and Conflict: Variations on a Flexible Specification for Private Prizes. This table regresses conflict incidence on group size and indices of private and public prizes, along with interactions between subsets of these variables as suggested by the theory. The dummy variables OIL_{i-j} are equal to 1 if the value of oil reserves is between the ith and the jth percentile of the distribution of OIL, conditional on having oil in the homeland. Regressions 1–6 drop regions of the world in the same order as Table B8 (more specifically, the regions dropped are former USSR countries, Asia, the Middle East, West-South Africa, East and Central Africa, and Latin America in columns 1–6, respectively. Column 7 collapses the time variation of the data and runs a similar regression in the cross-section, where the unit of analysis is the group and the dependent variable is the share of conflict years in the sample. All regressions contain country and year dummies, except for Column 7 which only contains country fixed effects, as the time dimension of the data has been removed. p-values are reported in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

	Dependent Variable: **Conflict Incidence**					
	[1]	[2]	[3]	[4]	[5]	[6]
$\text{SIZE}_{\text{COAL}}$	0.221***	0.405***	0.623***	0.765**	1.100***	1.193***
	(0.000)	(0.000)	(0.000)	(0.015)	(0.002)	(0.002)
OIL_{COAL}	0.002***	0.003***	0.004***	0.003**	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.022)	(0.003)	(0.004)
$\text{SIZE}_{\text{COAL}} \times \text{OIL}_{\text{COAL}}$		-0.054***	-0.047***		-0.069***	-0.048*
		(0.000)	(0.000)		(0.005)	(0.079)
$\text{SIZE}_{\text{COAL}} \times$LACK RIGHTS				-1.147***	-1.285***	-1.227***
				(0.006)	(0.003)	(0.003)
POLITY			-0.002**			-0.000
			(0.022)			(0.702)
GIP			-0.021***			-0.016***
			(0.000)			(0.002)
GDP			0.001			0.002
			(0.274)			(0.342)
POP			0.002			0.000
			(0.360)			(0.947)
GROUPAREA			-0.000***			-0.000**
			(0.002)			(0.039)
SOILCONST			-0.001			-0.001
			(0.202)			(0.605)
DISTCAP			0.004***			0.005**
			(0.000)			(0.027)
MOUNT			0.003			0.002
			(0.125)			(0.387)
PARTITIONED			-0.001			-0.000
			(0.570)			(0.942)
LAG	0.886***	0.885***	0.876***	0.877***	0.873***	0.870***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
c	-0.027***	-0.039***	-0.095***	-0.024***	-0.034***	-0.067
	(0.000)	(0.000)	(0.005)	(0.007)	(0.001)	(0.608)
R^2	0.845	0.845	0.850	0.853	0.853	0.859
Obs	64839	64839	55289	41314	41314	38341

Table B4. Groups and Conflict: Alliances. This table regresses conflict incidence on group size (allowing for the possibility of coalitions) and indices of private and public prizes, along with interactions between subsets of these variables as suggested by the theory. All regressions contain year dummies and country fixed effects. p-values are reported in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

	Dependent Variable: **Conflict Onset** [1–3]; **% of Years in Conflict** [4-6]; **% of Onset Years** [7]						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
SIZE	0.018	0.051***	0.019	0.226	0.613**	0.091	0.080
	(0.191)	(0.001)	(0.382)	(0.318)	(0.030)	(0.766)	(0.199)
OIL	0.698***	0.867***	0.685*	7.549**	9.340**	10.320**	3.169*
	(0.004)	(0.002)	(0.072)	(0.038)	(0.024)	(0.028)	(0.073)
SIZE×OIL		-11.006***	-5.179		-126.967**	-130.765**	-32.260*
		(0.000)	(0.294)		(0.028)	(0.040)	(0.051)
SIZE×LACK RIGHTS			0.039			1.003**	0.170*
			(0.244)			(0.014)	(0.054)
POLITY	-0.001	-0.001	-0.000				
	(0.172)	(0.172)	(0.915)				
GDP	0.001	0.001	0.002	-0.039	-0.027	-0.070	-0.027
	(0.288)	(0.307)	(0.391)	(0.358)	(0.517)	(0.336)	(0.250)
POP	0.002	0.002	0.007	-0.237	-0.242	-0.494	-0.185
	(0.174)	(0.203)	(0.165)	(0.401)	(0.393)	(0.294)	(0.234)
GROUPAREA	-0.000*	-0.000	-0.000	-0.000	0.000	0.000	0.000
	(0.072)	(0.682)	(0.283)	(0.558)	(0.599)	(0.270)	(0.621)
GIP	-0.002	-0.002	-0.001	-0.045**	-0.046**	-0.049**	-0.012*
	(0.125)	(0.112)	(0.460)	(0.027)	(0.023)	(0.020)	(0.074)
SOILCONST	-0.000	-0.000	-0.001	-0.006	-0.006	-0.004	-0.001
	(0.633)	(0.486)	(0.300)	(0.287)	(0.201)	(0.619)	(0.486)
DISTCAP	0.001***	0.001***	0.001	0.013	0.013	0.024	0.008
	(0.000)	(0.000)	(0.179)	(0.193)	(0.186)	(0.143)	(0.141)
MOUNT	0.002	0.002*	0.002	0.024**	0.025**	0.026**	0.007*
	(0.100)	(0.085)	(0.242)	(0.022)	(0.016)	(0.033)	(0.070)
PARTITIONED	-0.001	-0.001	-0.000	-0.009	-0.009	-0.012	-0.004*
	(0.134)	(0.130)	(0.973)	(0.293)	(0.282)	(0.268)	(0.098)
PEACEYRS	-0.001***	-0.001***	-0.002***				
	(0.000)	(0.000)	(0.000)				
c	0.020	0.022	-0.061	3.209	3.164	6.600	2.484
	(0.999)	(.)	(0.426)	(0.401)	(0.410)	(0.302)	(0.238)
R^2	0.033	0.034	0.048	0.373	0.375	0.377	0.240
Obs	53357	53357	36761	1427	1427	1297	1296

Table B5. Group Size and Conflict: Alternative Dependent Variables. This table regresses conflict onset (columns 1–3), the share of years in conflict (columns 4–6) and the share of onset years (Column 7) on group size and indices of private and public prizes, along with interactions between subsets of these variables as suggested by the theory. All regressions contain country fixed effects and Columns 1–3 include year dummies as well. p-values are reported in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

	Dependent Variable: **Conflict Incidence**				
	[1]	[2]	[3]	[4]	[5]
SIZE	0.053***	0.054***	-0.005	0.056***	-0.000
	(0.006)	(0.005)	(0.834)	(0.006)	(0.986)
OIL	0.741***	0.844***	0.856*		
	(0.008)	(0.003)	(0.057)		
OIL COUNTRY	-1.143	-1.238*	0.615		
	(0.105)	(0.077)	(0.823)		
OIL (SURFACE)				0.002***	0.003**
				(0.008)	(0.036)
SIZE× OIL COUNTRY	-10.171***	21.960	31.127		
	(0.006)	(0.182)	(0.177)		
SIZE× OIL		-34.888**	-41.824*		
		(0.038)	(0.086)		
SIZE× LACK RIGHTS			0.083**		0.084**
			(0.034)		(0.032)
SIZE×OIL (SURFACE)				-0.029***	-0.031**
				(0.001)	(0.030)
GDP	0.001	0.001	0.003	0.001	0.003
	(0.192)	(0.159)	(0.312)	(0.134)	(0.326)
POLITY	-0.002**	-0.001**	-0.001	-0.002**	-0.001
	(0.021)	(0.022)	(0.548)	(0.017)	(0.515)
POP	0.001	0.001	-0.001	0.001	-0.001
	(0.773)	(0.746)	(0.892)	(0.614)	(0.870)
GROUPAREA	0.000	0.000	0.000	0.000	0.000
	(0.570)	(0.470)	(0.363)	(0.572)	(0.634)
GIP	-0.003*	-0.004**	-0.003	-0.004**	-0.003
	(0.060)	(0.034)	(0.179)	(0.041)	(0.236)
SOILCONST	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.320)	(0.284)	(0.407)	(0.301)	(0.429)
DISTCAP	0.002***	0.002***	0.003*	0.002***	0.002*
	(0.000)	(0.000)	(0.072)	(0.000)	(0.098)
MOUNT	0.002*	0.002*	0.002	0.002*	0.002
	(0.067)	(0.068)	(0.122)	(0.076)	(0.104)
PARTITIONED	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.305)	(0.270)	(0.414)	(0.316)	(0.475)
LAG	0.895***	0.894***	0.900***	0.894***	0.900***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
c	-0.026	-0.026	-0.016	-0.036	-0.014
	(0.475)	(0.460)	(0.916)	(0.269)	(0.930)
R^2	0.847	0.847	0.855	0.847	0.855
Obs	55289	55289	38341	55289	38341

Table B6. Group Size and Conflict: Alternative Ways of Measuring Oil Wealth. This table regresses conflict incidence on group size and indices of private and public prizes, along with interactions between subsets of these variables as suggested by the theory. All regressions contain year dummies and country fixed effects. p-values are reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

	Dependent Variable: **Conflict Incidence**									
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
SIZE	-0.015	0.032	0.066***	0.080***	0.066**	-0.073***	-0.002	0.011	0.022	0.084**
	(0.372)	(0.192)	(0.006)	(0.002)	(0.019)	(0.007)	(0.916)	(0.658)	(0.413)	(0.019)
OIL	0.448	0.684*	0.771*			0.631*	0.887*	0.828*		
	(0.143)	(0.082)	(0.072)			(0.099)	(0.065)	(0.075)		
SIZE× OIL		-13.628**	-14.433**				-14.455**	-12.836**		
		(0.023)	(0.014)				(0.038)	(0.028)		
SIZE× OIL_{0-25}				0.047					0.033	
				(0.582)					(0.763)	
SIZE× OIL_{25-50}				-0.058					0.182	
				(0.392)					(0.549)	
SIZE× OIL_{50-75}				-0.154***					-0.166**	
				(0.004)					(0.026)	
SIZE× $\text{OIL}_{>75}$				-0.135***					-0.118***	
				(0.001)					(0.006)	
SIZE× PRIV. IND.					-0.049**					-0.046**
					(0.017)					(0.018)
SIZE× LACK RIGHTS						0.097**	0.068*	0.083**	0.083**	
						(0.029)	(0.065)	(0.035)	(0.050)	
SIZE× PUBLIC INDEX										0.023*
										(0.080)
$\text{OIL}_{>75}$				0.007**					0.007**	
				(0.014)					(0.024)	
OIL_{50-75}				0.004*					0.005*	
				(0.095)					(0.067)	
OIL_{25-50}				-0.001					-0.002	
				(0.834)					(0.600)	
OIL_{0-25}				-0.004					-0.003	
				(0.202)					(0.399)	
PRIVATE INDEX					0.002*					0.002
					(0.067)					(0.134)
PUBLIC INDEX										0.110
										(0.356)
POLITY			-0.002	-0.002	-0.002			-0.001	-0.001	-0.001
			(0.213)	(0.215)	(0.233)			(0.512)	(0.513)	(0.437)
GIP			-0.004*	-0.003	-0.004*			-0.003	-0.003	0.001
			(0.098)	(0.105)	(0.084)			(0.224)	(0.231)	(0.856)
GDP			0.001	0.001	0.001			0.003	0.003	0.003
			(0.433)	(0.476)	(0.469)			(0.321)	(0.317)	(0.354)
POP			0.001	0.002	0.001			-0.001	-0.001	-0.001
			(0.810)	(0.739)	(0.883)			(0.878)	(0.883)	(0.919)
GROUPAREA			0.000	0.000	0.000			0.000	0.000	0.000
			(0.331)	(0.529)	(0.354)			(0.290)	(0.324)	(0.321)
SOILCONST			-0.001	-0.001	-0.001			-0.001	-0.001	-0.001
			(0.455)	(0.407)	(0.396)			(0.409)	(0.368)	(0.399)
DISTCAP			0.002	0.002	0.002			0.003*	0.002*	0.003*
			(0.117)	(0.125)	(0.129)			(0.084)	(0.084)	(0.087)
MOUNT			0.002**	0.003**	0.002**			0.002	0.002	0.002
			(0.022)	(0.016)	(0.030)			(0.109)	(0.102)	(0.169)
PARTITIONED			-0.001	-0.001	-0.001			-0.001	-0.001	-0.001
			(0.306)	(0.348)	(0.296)			(0.445)	(0.431)	(0.478)
LAG	0.895***	0.895***	0.894***	0.894***	0.894***	0.898***	0.898***	0.900***	0.900***	0.900***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
c	-0.002	-0.005**	-0.035	-0.033	-0.028	-0.002	-0.010	-0.015	-0.020	-0.142
	(0.355)	(0.021)	(0.601)	(0.626)	(0.777)	(0.800)	(0.166)	(0.924)	(0.902)	(0.473)
R^2	0.844	0.844	0.847	0.848	0.847	0.850	0.850	0.855	0.855	0.853
Obs	64839	64839	55289	55289	54486	41314	41314	38341	38341	35755

Table B7. Group Size and Conflict: Baseline with errors clustered at the country level and at the country and group level (two way clustering). This table regresses conflict incidence on group size and indices of private and public prizes, along with interactions between subsets of these variables as suggested by the theory. All regressions contain year dummies and country fixed effects. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

	Dependent Variable: **Conflict Incidence**											
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
SIZE	0.067***	0.007	0.067***	0.006	0.060***	0.004	0.048**	0.018	0.054**	0.016	0.094***	0.033
	(0.001)	(0.795)	(0.002)	(0.801)	(0.003)	(0.855)	(0.024)	(0.486)	(0.014)	(0.483)	(0.001)	(0.329)
OIL	0.775**	0.838*	0.944***	1.102***	0.671**	0.707	0.592**	0.649	0.627**	0.633	1.054***	1.132*
	(0.011)	(0.070)	(0.006)	(0.010)	(0.022)	(0.143)	(0.032)	(0.136)	(0.023)	(0.209)	(0.004)	(0.058)
SIZE× OIL	-15.148***	-12.398**	-16.180***	-16.704***	-13.430***	-10.574*	-11.333***	-11.640**	-8.499**	-6.772	-21.314***	-18.699**
	(0.000)	(0.028)	(0.000)	(0.004)	(0.001)	(0.071)	(0.004)	(0.043)	(0.019)	(0.211)	(0.000)	(0.021)
SIZE× LACK RIGHTS		0.092**		0.096**		0.082*		0.074*		0.041		0.079*
		(0.049)		(0.045)		(0.050)		(0.059)		(0.140)		(0.074)
POLITY	-0.002**	-0.001	-0.001	-0.000	-0.002**	-0.001	-0.003***	-0.001	-0.001*	-0.001	-0.001	-0.000
	(0.010)	(0.435)	(0.133)	(0.845)	(0.025)	(0.557)	(0.000)	(0.261)	(0.079)	(0.240)	(0.364)	(0.974)
GIP	-0.004**	-0.003	-0.002	-0.001	-0.003	-0.002	-0.003	-0.002	-0.007***	-0.005**	-0.003*	-0.002
	(0.047)	(0.250)	(0.241)	(0.639)	(0.116)	(0.426)	(0.210)	(0.361)	(0.001)	(0.027)	(0.090)	(0.346)
GDP	0.002	0.003	0.000	0.004	0.001	0.002	0.000	0.001	0.002**	0.004	0.002	0.002
	(0.109)	(0.302)	(0.886)	(0.377)	(0.179)	(0.374)	(0.650)	(0.805)	(0.023)	(0.191)	(0.135)	(0.414)
POP	0.003	0.001	0.001	-0.001	0.001	-0.001	-0.003	-0.010	-0.003	-0.005	0.005**	0.004
	(0.275)	(0.935)	(0.669)	(0.883)	(0.577)	(0.906)	(0.225)	(0.297)	(0.183)	(0.599)	(0.023)	(0.541)
GROUPAREA	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.156)	(0.353)	(0.517)	(0.301)	(0.328)	(0.395)	(0.341)	(0.372)	(0.181)	(0.114)	(0.441)	(0.374)
SOILCONST	-0.000	-0.001	0.001	0.000	-0.000	-0.000	-0.000	-0.000	-0.002***	-0.002***	-0.001	-0.001
	(0.697)	(0.396)	(0.416)	(0.774)	(0.459)	(0.675)	(0.370)	(0.550)	(0.000)	(0.004)	(0.243)	(0.301)
DISTCAP	0.002***	0.003*	0.000	0.001	0.002***	0.003*	0.002***	0.002*	0.002***	0.002	0.002***	0.003*
	(0.000)	(0.085)	(0.435)	(0.291)	(0.000)	(0.077)	(0.000)	(0.086)	(0.002)	(0.173)	(0.000)	(0.094)
MOUNT	0.002	0.003	0.002	0.001	0.002	0.002	0.002	0.002	0.003**	0.004**	0.003*	0.003*
	(0.108)	(0.102)	(0.167)	(0.664)	(0.118)	(0.198)	(0.102)	(0.228)	(0.042)	(0.021)	(0.077)	(0.089)
PARTITIONED	-0.001	-0.001	-0.002	-0.001	-0.002	-0.002	-0.001	-0.001	0.001	0.001	-0.002	-0.001
	(0.291)	(0.446)	(0.186)	(0.380)	(0.105)	(0.194)	(0.429)	(0.475)	(0.371)	(0.176)	(0.196)	(0.378)
LAG	0.895***	0.900***	0.885***	0.893***	0.893***	0.899***	0.903***	0.911***	0.886***	0.891***	0.895***	0.899***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
c	-0.055	-0.043	-0.022	-0.013	-0.036	-0.010	0.038	0.164	0.028	0.043	-0.105**	-0.094
	(0.118)	(0.724)	(0.621)	(0.925)	(0.409)	(0.929)	(0.303)	(0.333)	(0.484)	(0.786)	(0.017)	(0.384)
Dropped Reg.	EX-USSR	EX-USSR	ASIA	ASIA	MID. EAST	MID. EAST	W-SAFR.	W-S AFR.	E-C AFR.	E-C AFR.	LATIN AM.	LATIN AM.
R^2	0.850	0.854	0.851	0.856	0.846	0.854	0.861	0.869	0.844	0.849	0.829	0.841
Obs	53181	37504	39004	26829	52637	36473	45984	31513	45840	31580	44673	31092

Table B8. Dropping Regions of the World. This table reproduces columns 3 and 8 from Table 1 dropping regions of the world. Regions dropped are: former USSR countries (columns 1 and 2), Asia (columns 3 and 4), the Middle East (columns 5 and 6), West-South Africa (columns 7 and 8), East and Central Africa (columns 9 and 10), and Latin America (columns 11 and 12). All regressions contain year dummies and country fixed effects. p-values are reported in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.