Resource Concentration and Civil Wars*

Massimo Morelli† and Dominic Rohner‡

June 12, 2015

Abstract

This paper highlights the importance of natural resource concentration and ethnic group regional concentration for ethnic conflict. The existence of multiple conflict terrains (and hence multiple threat points) is the source of bargaining failure, similar to the one determined by the presence of offensive advantages. The theory predicts war to be more likely when resource and group concentration are high, and the empirical analysis, both at the country level and at the ethnic group level, confirms the essential role of geographic concentration variables for civil war.

Keywords: Natural Resources, Civil War, Conflict, Secession, Bargaining Failure.

JEL Classication: C72, D74, O12, Q34.

Published in Journal of Development Economics:

https://doi.org/10.1016/j.jdeveco.2015.06.003

*An earlier version of this paper has circulated under the title "Natural Resource Distribution and Multiple Forms of Civil War". We wish to thank the editor Gerard Padro i Miquel and two anonymous referees for very helpful comments. We further thank Peter Van der Windt and David Schoenholzer for excellent research assistance and Dawn Brancati, Lars-Erik Cederman, Oeindrila Dube, Tanisha Fazal, Benedikt Goderis, Matt Jackson, Cathy Hafer, John Huber, Päivi Lujala, Egil Matsen, Guy Michaels, Kalle Moene, Costantino Pischedda, Robert Powell, Rick van der Ploeg, Petros Sekeris, Tommaso Sonno, Jonathan Temple, Mathias Thoenig, Ragnar Torvik, Nicolas van de Sijpe, Peter van der Windt, Fabrizio Zilibotti and seminar and workshop participants in Washington University St. Louis, Geneva, Zurich, Columbia, Bristol, Oxford, York, Purdue, London School of Economics, Gerzensee, Trondheim, UC Berkeley, Mannheim, Namur, Rotterdam, Seattle (APSA), Milano Bicocca, Oslo, Barcelona Summer Forum, and Montpellier for helpful discussions and comments. The usual disclaimer applies.

†Bocconi University, IGIER and NBER. Email: massimo.morelli@unibocconi.it.

‡University of Lausanne and CEPR. Email: dominic.rohner@unil.ch.

1 Introduction

This paper provides a theoretical explanation and novel empirical investigation of the importance of the geography of natural resources for civil war. We find that civil wars should be expected to take place more frequently when the homeland of a concentrated minority group is particularly resource rich. The paper then tests the predictions of the theory using both panel data at the country and ethnic group level.

Two things seem to matter in general for civil war incentives: balance of strength and balance of control on resources. When trying to resolve a conflict between two groups over control of resources, one difficulty is that the *relative strength* of the two groups may differ from the *relative wealth* of natural resources of the territories they occupy. Having surplus sharing reflect relative strength eliminates the incentives to "nationwide" wars, but "secessionist" wars could then materialize; on the other hand, making surplus sharing depend on the groups' relative endowments of natural resources avoids secession tensions but may cause incentives for the majority group to use their strength to gain more power.

Recognizing this tension between the two most important determinants of bargaining power, we have decided to focus attention on a connected observation about bargaining games: while in a standard bargaining game there is a unique "threat point" (for example a unique type of war that players could fall into if bargaining breaks down), in reality there are multiple threat points, which depend on the balance of strength and geographic distribution of natural resources. If an ethnic group is particularly influential for the government of a country but another group has an important presence (in terms of population size and rootage to the territory) in a region of the country that is particularly rich in terms of natural resources, the tensions between the two criteria of surplus sharing mentioned above are maximized, and are exacerbated by the fact that the two groups have access to different threats: the powerful group controlling government forces should typically be stronger in a nationwide ethnic conflict, but the minority group could sustain the secession threat with guerrilla war and focus its lower total strength on the defense of the area where it is locally stronger.

We model these tensions in the following way: in a country divided into two regions and populated by two major groups, we assume that the nationwide stronger group has a realistic offensive advantage when starting a nationwide conflict, while a minority group mostly concentrated in one region has a probability of winning in a secessionist civil war that exceeds the probability of winning in a nationwide civil war. When groups are unable to commit not to use their favorite type of fighting as deviations from peace, bargaining may break down for analogous reasons to the ones

put forward in the offensive advantage literature (see e.g. Chassang and Padro i Miquel, 2009).[1] The most conflict prone situations are shown to be those in which the mineral resources of value are mostly concentrated in the minority group region, and the risk is especially high in case of low state capacity, high regional concentration of the minority group in question, and large geographic distance of the minority region from the capital.

There are many cases where, when the presence of a concentrated ethnic group coincides with large natural resource abundance concentrated in its region, the concentrated minority group could be financially better off if it were independent and may under some conditions have incentives to start secessionist rebellion. This corresponds for example to the separatist movement in the now independent Timor-Leste, and the recent turmoil in the oil-abundant regions of Nigeria. Also the rebellion of the Aceh Freedom Movement in Indonesia starting in 1976 and the armed fight of the Sudan People's Liberation Army beginning in 1983 can to a large extent be explained by the abundance of natural resources in these separatist regions.[2] Other countries where secessionist movements have been linked to large local natural resources include Angola, Myanmar, Democratic Republic of Congo, Morocco and Papua New Guinea. In all these cases an uneven natural resource distribution has been amplified by ethnic divisions. In contrast, if natural resources are absent or if natural resources (and political power) are evenly dispersed in a country, there are typically fewer conflict incentives, even when there are ethnic divisions.[3] Similarly, when there are large amounts of natural resources available, but the society is ethnically homogeneous, war incentives are weak.[4]

In a country level empirical analysis, we study how the unevenness of geographical distribution of petrol fields across ethnic groups in a given country affects the likelihood of conflict. For this purpose we have put together a panel of 157 countries with sample period 1960-2008, and have constructed a new variable, *Oil Gini*, which captures how unevenly oil holdings are spread between different ethnic groups in a country.[5] To the best of our knowledge we are the first to have constructed such a measure of inter-ethnic inequality in abundance of petrol fields. In the regression analysis we include –in addition to our main, new variable of interest– the standard battery of control variables, as well as country fixed effects and annual time dummies. We find that our novel *Oil Gini* measure

[1] Ray (2010) also studies multiple threats to peace. However, he focuses on the important issue of multiple cleavages, while in our paper the cleavage is unique (for example an ethnic or religious cleavage), but the groups have different relative strengths in different types of conflict.

[2] For a discussion of these cases see Ross (2004b).

[3] This is for example the case of countries like Benin, which has only few natural resources, or of small oil-rich countries like Brunei or Qatar, where natural resources are evenly spread.

[4] Examples for this include Chile and Mongolia.

[5] As discussed in detail in section 3.1.1, we have used the GIS-coordinates of all ethnic groups in the "Geo-referencing of ethnic groups" (GREG) dataset (Weidmann, Rod and Cederman, 2010), and have merged them with the geo-referenced petroleum dataset (PETRODATA) from Lujala, Rod and Thieme (2007), which allowed us to construct a time-varying measure of how relatively petrol-rich the homelands of a given ethnic group are. Using this information, we have been able to apply the Gini formula to capture geographical oil unevenness.

has a statistically significant and quantitatively strong positive effect on the likelihood of civil war onsets, as predicted by our theory.

Next, we move to a more disaggregated level of analysis: we study the effect of natural resource unevenness on civil war with a panel dataset *at the ethnic group level*, covering 1120 ethnic groups and spanning over the period 1960-2006. This has the advantage that unobserved heterogeneity problems are reduced, and ethnic group level data allows us to better discriminate between our theory and competing explanations: our model predicts that conflicts are fuelled by non-governing ethnic minority groups living in oil rich regions, while alternative mechanisms (discussed below) predict conflict onsets in the presence of oil-rich ethnic groups that control the government. While our country level regressions only establish the link between inter-group oil unevenness and civil war onset, the ethnic group level regressions distinguish such different mechanisms, and show that it is indeed oil abundance in the regions mostly inhabited by powerless groups, rather than in the homelands of governing ethnic groups, that drives civil wars.

Our main, novel independent variable on the ethnic group level is the surface of an ethnic group's territory covered with petrol (i.e. oil and gas) as a percentage of the country's total surface covered with petrol. To the best of our knowledge we are the first to study civil conflict using an ethnic group panel with natural resource variables that vary for different ethnic groups and over time.[6] We find a statistically significant and quantitatively strong positive effect of the relative resource abundance of a non-governing ethnic group on the likelihood that this group is involved in a civil war onset. We also find that the interaction term of an ethnic group's relative resource abundance with its group concentration is positive, very sizeable and statistically significant, which is in line with our theoretical predictions. The results are robust for very demanding specifications that control for ethnic group fixed effects, annual time dummies, time-varying ethnic group level controls and all country-level control variables used in the country-level regressions. Indeed civil war is more likely when resource discoveries happen in regions that are mostly populated by groups that do not belong to the governing coalition in the country.

Related literature

Natural resources and ethnic divisions are known to be correlated with civil conflict in one way or another (see e.g., Le Billon 2001, Collier and Hoeffler 2004, Ross 2004, Montalvo and Reynal-Querol 2005, Fearon 2005, Lujala, Gleditsch and Gilmore 2005, Humphreys 2005, Lujala 2010, Dube and

[6] The only exception is the paper by Sorens (2011) which also uses ethnic group level natural resource data to explain territorial conflict. In contrast to our paper which runs fixed effects regressions on a global sample of all ethnic groups, Sorens' study only contains the selected sample of discriminated groups from the "Minorities at Risk" (MAR) project and runs pooled panel regressions which do not control for country/ethnic group fixed effects.

Vargas 2013, Berman et al. 2014), but the literature does not emphasize the particularly important role of resource concentration and ethnic concentration, independently and jointly. The existing theoretical studies about the effect of natural resources on conflict, by and large do not relate to geographic concentration: Caselli and Coleman (2013) focus on the decision of the dominant ethnic group to exploit or not the other groups in terms of the proceeds from extraction of natural resources, but do not take into account how the geographic distribution and the economic features of natural resources affect the risk of ethnic conflict of different kinds; Reuveny and Maxwell (2001) and Grossman and Mendoza (2003) use a dynamic framework to predict that present resource scarcity and future resource abundance cause appropriative competition; Hodler (2006) finds that natural resources lead to more conflicts in fractionalized countries; Rohner, Thoenig and Zilibotti (2013) predict natural resources to have a particularly detrimental effect if initial trust in a country is low; Fearon (2005) argues that natural resources can foster conflict by weakening state capacity; Besley and Persson (2011) and Bell and Wolford (2014) emphasize that weak institutions, low income and large natural resources lead to a greater risk of civil war; van der Ploeg and Rohner (2012) and Rohner (2014) study the two-way interaction between natural resource extraction and civil war, focusing on depletion speed and optimal investments of windfalls. To repeat, none of these papers consider geographic concentration of resources and how it overlaps with the geographic concentration of minority groups.

Horowitz (1985) did make the anecdotal observation that backward concentrated minorities with resource concentration may have the highest benefit-cost ratio from rebellion, and Walter (2006b) proposed a reputation building theory for why governments tend to repress rebellions in areas with high value of land. However, Walter's reputation theory predicts that the group secession attempts that tend to be blocked are in countries where other groups could otherwise make similar attempts for other resources in other areas. Thus, the explanatory variables with that type of theory would be the number of other relevant groups in the country and the general value of land, more than the value of land in that given group's homelands specifically. The current paper shows that the cases with natural resources concentrated in the homelands of ethnic minorities are exactly those where bargaining can break down, even without any reputation building factor.

The paper relates also to the literature connecting income and conflict incentives. An increase in income has two countervailing effects: it increases the opportunity cost of fighting, but it may also raise the value of the appropriable prize. As shown by Fearon (2007) and Chassang and Padro i Miquel (2009), when there is only a single source of income, these two countervailing effects are likely to cancel out, making the total effect of a higher income ambiguous. In contrast, in

countries that are abundant in natural resources, but have low labor productivity in other sectors, an increase in the value of natural resources will only affect the appropriable prize, while keeping the opportunity cost unchanged (Dube and Vargas, 2013; Esteban, Morelli and Rohner, 2015). In the current model we adopt this distinction between natural resource rents and incomes from other sectors.

Preemptive and preventive war incentives due to commitment problems are a well established phenomenon (see e.g. Taylor 1954, Levy 1987, Fearon 1995, Powell 1996 and 2006, Chassang and Padro i Miquel 2009, and Jackson and Morelli 2011). However, in general this incentive analysis does not separate or highlight the role of geography or concentration of resources. A traditional preventive war incentive, due for example to an expected exogenous power shift in the future, could make a minority group decide to rebel wherever the future enrichment and strengthening of the majority group is expected to come from. Similarly, recent works on civil war rationalization and natural resources that emphasize commitment problems, like Besley and Persson (2011), Lei and Michaels (2014), and Bell and Wolford (2014), focus on the effects of changes in total amounts or values of resources rather than on distribution and concentration variables. Our focus will instead be precisely on the role of geographic concentration of groups and concentration of resources.[7]

While the effect of the total quantity and/or value of oil on conflict is ambiguous,[8] we are able to demonstrate that the unevenness of oil has a robust effect. The shadow of the future argument (related to the fear that an oil discovery could strengthen the government) is equally strong when the oil discovery is in the region controlled by the government; hence the fact that we find a strong and robust impact of the concentration of resources only in the regions mostly populated by minority groups speaks in favor of the type of war motivations that we highlight.

Our paper can also be usefully contrasted with the literature in international relations. There is a common view that preemptive as well as preventive war motivations relate to *fear*, and not to *greed*.[9] In our view, preemptive and preventive wars are much more likely if fear goes hand in hand with greed: in Esteban, Morelli and Rohner (2015) the decision by a minority group to start a war is preventive, and certainly due to fear of mass killings, but this fear comes up precisely because the group in power is made more greedy (or is afraid that the minority group will become more greedy

[7] Caselli, Morelli and Rohner (2015) find that geography, and in particular location of borders and resources, matter a lot also for interstate wars.

[8] Cotet and Tsui (2013) find that there is no robust effect of oil reserves on civil wars when controlling for country fixed effects. In contrast, Lei and Michaels (2014) find that oil discoveries lead to more civil wars, perhaps consistently with preventive war motivations.

[9] "... the spiral model contends that even a state interested in protecting the status quo can go to war (out of fear), whereas the deterrence model posits that there are status quo states and revisionist states and that only the latter are attackers." (Reiter, 1995: 8). There is a large literature in international relations on this: See e.g. Schelling (1966), Jervis (1978), and all the subsequent literature on the security dilemma and spiral of fear.

in the future) due to larger amounts of resources. In the current paper, preemptive incentives can be seen in the possibility of an attack due to the fear of being otherwise involved in a less advantageous type of conflict, but once again this fear would not materialize without the complementary greed on one side or the other. Moreover, given the existence of multiple conflict technologies that we emphasize, spirals of events or spiral of beliefs are not necessary for the existence of preemptive war motivations.[10]

In terms of relationship with other empirical studies on civil war, our findings are broadly consistent with the empirical results of Walter (2006) on the importance of group concentration; Reynal-Querol (2002), Saideman et al. (2002), and Cederman and Girardin (2007) on the importance of ethnic discrimination;[11] Gates (2002) and Buhaug, Gates and Lujala (2009) on the importance of situations where the rebelling minority group is concentrated in remote peripheral areas, where its odds of winning a local war are larger and conflict tends to be harder to avoid. As far as the empirical literature on inequality and civil conflict is concerned (see e.g. Gurr (2000), Alesina, Michalopoulos, and Papaioannou (2014) and Huber and Mayoral (2013)), our contribution is to show that between-group inequality can be expected to fuel civil conflict if such inequality is mostly in terms of control or shares of resource extraction rents, and especially when resource concentration is interacted with group concentration.

The econometric specification that we use in the analysis at the country level is related to that in Fearon and Laitin (2003), Collier and Hoeffler (2004), Montalvo and Reynal-Querol (2005), Cederman and Girardin (2007), Collier and Rohner (2008), and Esteban, Mayoral and Ray (2012).

The paper is organized as follows: In section 2 we present the model framework and characterize its equilibrium. Then we move to the empirics, with section 3 displaying the country level analysis first, followed by the ethnic group level analysis. Section 4 offers some brief concluding remarks. Supplementary material and the detailed data description are relegated to the appendix.

[10] The offense defense balance theory did already allow for the possibility that preemptive wars can occur due to technology, but in that case they talk about countries having advantages in offense technology over defense technology (see e.g. Snyder, 1984), while we focus on different technological distinctions, more related to geography and motivations.

[11] In their qualitative comparative analysis of oil-exporters Basedau and Richter (2011) find that conflict is more frequent in countries where there are groups from oil-producing regions that are excluded from power.

2 Model

2.1 Setup

Consider a country populated by two ethnic groups, i and j.[12] The country is divided in two clearly defined regions, $1, 2$. We allow for different values of extractable natural resources in regions 1 and 2, labeled R_1 and R_2, respectively, where $R_1 + R_2 \equiv R$.[13] For simplicity, natural resource extraction and export of such resources is conducted by a unique state firm. The group in power can control the sharing of the ensuing surplus.[14]

There are N_i^1 and N_i^2 members of group i in regions 1 and 2 respectively, and N_j^1, N_j^2 of group j. Assume that the two groups are to some degree concentrated in the two regions, group i in region 1 and group j in region 2, $N_i^1/N_j^1 > N_i^2/N_j^2$.

We assume that group j controls the government at the beginning of the game, which may mean that group j is a majority group ($N_j = N_j^1 + N_j^2 > N_i = N_i^1 + N_i^2$) in a democracy, or simply that j had won some conflict for power in a prior period. Group j in power can choose the shares $\alpha, (1-\alpha)$ of the total surplus R to be attributed respectively to groups i and j. There is commitment to this sharing rule conditional on peace, but j is assumed to be unable to credibly commit to always keep peace (see Powell, 1996). After the choice of α by j, each player chooses between peace and conflict.

There are two types of conflict, secessionist (s) and nationwide (n) conflict.[15] The stake in secessionist conflict is region 1, whereas in nationwide conflict both regions 1 and 2 are at stake. In particular, the winner in secessionist conflict gains R_1 (with R_2 remaining under control of group j), while the winner in nationwide conflict gains $R_1 + R_2$. We assume without loss of generality that group i only has the choice between keeping peace or initiating secessionist conflict (s), while group j only has the choice between initiating nationwide repression (n) and keeping peace.[16]

Call p_s the probability of winning for i in case they attack the local government guards by surprise for secession; call $p_n < p_s$ the probability of winning for i when surprised by an attack

[12] We call the two groups "ethnic groups" only because we need to focus on societies where the population is divided along some identifiable cleavage, but that cleavage could be religion or anything else. Hence we will never require any specific component of the concept of ethnicity in our theory.

[13] The variables R_1 and R_2 capture the total value of rents, corresponding to the product of the amounts extracted times their price. Hence, R_1 and R_2 increase when there are resource discoveries in the two regions (or when the world demand for the potentially different natural resources increases).

[14] Future work could include explicit consideration of the interaction between a government, the minority groups and the extraction companies, rather than assuming state control. For a motivating case study for this future extension, see Morelli and Pischedda (2014).

[15] See Fearon and Laitin (2003) and Ross (2012).

[16] In an extended model where group i is able to start nationwide conflict themselves, it can be easily shown that whenever group i prefers nationwide conflict to secessionist conflict, war can always be prevented. Indeed, the bargaining space is always non-empty in this case, as shown in Morelli and Rohner (2014).

for nationwide repression by j.[17] In line with the offensive advantage framework of Chassang and Padro i Miquel (2009), if only i chooses conflict, it has an offensive advantage, and vice versa. If both choose conflict the probability of winning for i is intermediate between the high probability when they attack for secession and the low probability when they are surprised by a nationwide repression attack.[18] Let p be the ex ante intermediate probability of success for i in this case. We can think of p as either a mixture of nationwide repression and secession attempts happening at the same time or, alternatively, as a probability that one or the other pure type of war will take place. In order not to carry around too many parameters, let's assume that $p \equiv p_s/2 + p_n/2$.[19]

Denote by c_{sk} the cost of a secessionist war and by c_{nk} the cost of a nationwide war, for player $k = i, j$, where $c_{sk} \equiv c_s y_k$, and $c_{nk} \equiv c_n y_k$. The variable y_k denotes the income of group k from all sources of revenue other than natural resource rents, and c_s,c_n correspond to the percentage of such income that is lost due to war.[20]

2.2 Equilibrium

We shall now characterize the subgame perfect Nash equilibrium of the game. Let's study the incentive compatibility constraints to determine the bounds of the bargaining space, i.e. the threshold levels of α for which a given group would deviate from peace to start its preferred type of war. In particular, two levels of α define the bargaining space:

The minimum α to appease i, i.e., to avoid temptations to start surprise secession attacks, is given by:[21]

$$\alpha_s = (p_s R_1 - c_{si})/(R_1 + R_2).$$

The maximum α that j is willing to concede without temptation to do nationwide repression

[17] This assumption is justified by the empirical findings that secessionist conflict generally entails a higher winning chance for group i, i.e. $p_s > p_n$ (see e.g. Gates (2002) and Buhaug, Gates and Lujala (2009)).

[18] In equilibrium, surprise attacks never occur, but the commitment problem forces the modeler to consider the most profitable unilateral deviations.

[19] In Morelli and Rohner (2014) we focused on the second interpretation of intermediate probabilities, but allowed for arbitrary weights, such that $p = p_s(1-\theta) + p_n\theta$ for any $\theta \in (0, 1)$.

[20] We are implicitly assuming here that the destruction of a war concerns exclusively (or mostly) the "other" sectors, leaving unchanged the present discounted value of natural resource extraction rents, our R variables. Esteban, Morelli and Rohner (2015) present evidence that civil wars imply mainly and disproportionately destruction of other types of wealth and income prospects (section 6.2): Typically only a small fraction of natural resources to be extracted in the future can be damaged by a war, while a substantial part of the non-resource income and accumulated physical capital is often destroyed, partly because it is a vital input for the capacity of a fighting group or state in a conflict. Oil fields and diamond mines (and especially their future productivity) "survive" fighting, while other sectors suffer considerably from political instability.

[21] This condition follows from the fact that i's expected payoff is $\alpha(R_1 + R_2)$ under peace and $p_s R_1 - c_{si}$ under secessionist conflict.

corresponds to:[22]

$$\alpha_n = p_n + c_{nj}/(R_1 + R_2).$$

Given these two incentive compatibility constraints, proposition 1 below easily follows.

Proposition 1 *The game has a unique subgame perfect Nash equilibrium outcome for every set of parameters. i) If $\alpha_s < \alpha_n$, then group j selects $\alpha = \alpha_s$, and peace prevails. ii) If $\alpha_s > \alpha_n$, then regardless of the α selected by group j in the initial stage both groups select conflict in the war-peace subgame. In other words, conflict arises in equilibrium if and only if*

$$(p_s - p_n)R_1 > p_n R_2 + c_{si} + c_{nj}. \tag{1}$$

Proof. Omitted. ■

Given that war is unavoidable for all parameter values such that $\alpha_s > \alpha_n$, we can perform some interesting comparative statics, making statements about under what economic and geographic conditions civil war is more likely to be unavoidable.[23] If R_1 is sufficiently high and R_2 and conflict costs sufficiently low, war will arise.

Figure 1[24] displays the zone where a peace equilibrium exists (labelled "Peace") and the zone where conflict can never be avoided (labelled "Conflict"). We can see that bargaining failure happens when the share of a given amount of natural resources that are situated in region 1, R_1/R, is large, and when the group out of power has a substantially larger winning probability in secessionist than in nationwide conflict, i.e. when $p_s - p_n$ is large.

So far p_s and p_n have been treated as blackboxes. We now analyze the main determinants of relative strength, and hence of the probability of winning, leading to additional testable implications.

The probability with which i wins a nationwide conflict can be expressed as:

$$p_n = \frac{N_i y_i}{N_i y_i + N_j y_j + \lambda}$$

where λ measures the extra strength of group j due to the control of the government and perhaps the army. In this contest success function (CSF) the ratio of total wealth matters. Like Jackson

[22] This follows from the fact that j's expected payoff is $(1-\alpha)(R_1+R_2)$ under peace and $(1-p_n)(R_1+R_2) - c_{nj}$ under nationwide conflict.

[23] Note that it would not help to allow j to choose two separate sharing rules, α_1, α_2, for the two regions: in fact, if the parameters are such that no α exists to make peace an equilibrium in the subsequent subgame, then trivially it is not possible to find any peace inducing pair α_1, α_2, since if this were possible then there would exist α' such that $\alpha'(R_1+R_2) = \alpha_1 R_1 + \alpha_2 R_2$ that would work, a contradiction.

[24] The following parameter values have been used in this Figure: $R_1 + R_2 = 1$, $p_n = 0.2$, $c_s y_i = 0.05$, $c_n y_j = 0.1$.

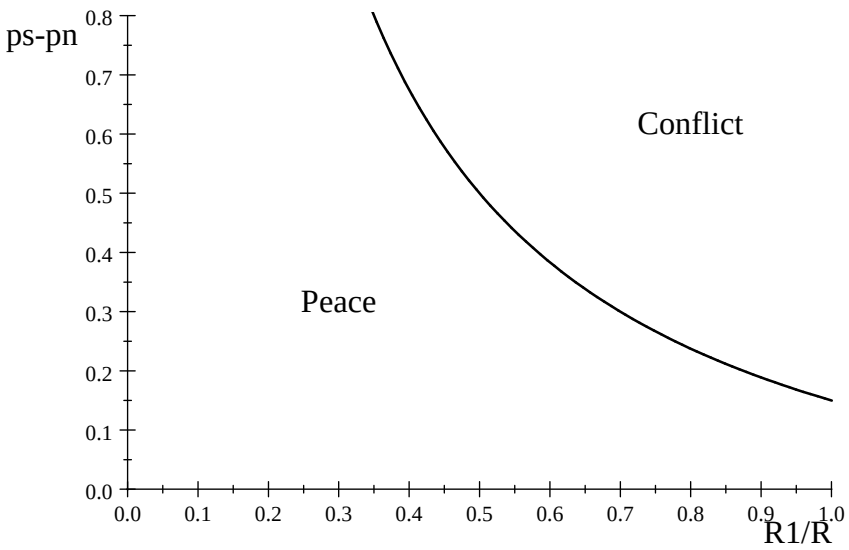

Figure 1: Zones of parameter values for peace and conflict

and Morelli (2007) and Esteban and Ray (2011) we assume a group's probability of winning to be increasing in its wealth. Finally, this expression allows for the obvious relevance of population sizes, especially in traditional warfare that is typical of many civil conflicts.

Consistently, group i's winning probability in secessionist conflict can be expressed as follows:

$$p_s = \frac{N_i^1 y_i}{N_i^1 y_i + N_j^1 y_j + \lambda/(1+\phi)}$$

where the government's advantage of controlling a standing army (λ) is discounted by $(1+\phi)$, and ϕ is a non-negative parameter that is an increasing function of the geographical distance between the secessionist homelands and the capital. This is in line with the idea that the military power of the government decays when projected at a large distance (see e.g. Boulding, 1962). It is also consistent with the literature finding that geographical remoteness and ideological recruitment provide a larger advantage for group i in secessionist than in nationwide wars (see e.g. Gates, 2002; Buhaug, Gates and Lujala, 2009). The parameter ϕ also depends on various factors that affect the scope for hide-and-ambush guerrilla insurgency. Note also that in case of a secessionist conflict the men used are only those in the conflict region.[25] Using the explicit dependence of p_s and p_n on the distribution of populations of the different groups in the territory, the condition for bargaining

[25] Gates (2002) and Buhaug, Gates and Lujala (2009) find evidence that when the minority group is concentrated the odds of winning are larger.

failure becomes

$$\left(\frac{N_i^1 y_i}{N_i^1 y_i + N_j^1 y_j + \lambda/(1+\phi)} - \frac{N_i y_i}{N_i y_i + N_j y_j + \lambda}\right) R_1 > \frac{N_i y_i}{N_i y_i + N_j y_j + \lambda} R_2 + c_s y_i + c_n y_j. \quad (2)$$

This reveals that for various possible parameter values civil war is on average more likely when the minority group i is very concentrated (large N_i^1) and the corresponding region relatively homogeneous (low N_j^1). All the relevant comparative statics are summarized as follows:

Corollary 1 *Conflict occurs for a larger set of the other parameters when*

1. *more natural resources are located in the region of the powerless group for any fixed total amount of natural resources;*
2. *the difference between the probabilities of winning in secessionist and nationwide conflict for the group out of power is larger;*
3. *war is less destructive;*
4. *the powerless group is more concentrated in a relatively homogeneous region, i.e. N_i^1/N_j^1 and N_i^1/N_i^2 are larger;*
5. *the region of the group out of power is further away from the capital, i.e. ϕ is greater.*

No clear and unambiguous comparative statics result can instead be given with respect to y_i and y_j, since for example an increase in y_i may increase the left-hand-side of (2) looking at the first term in the presence of group concentration, but affects of course the cost part as well, hence the sign of the effect depends on many parameters.

3 Empirical Analysis

The existing empirical literature on natural resources and civil war has two main weaknesses: First, it only studies the effects of the total *amount* of natural resources and not of their *geographical distribution*. Second, it studies the impact of natural resource abundance mostly on the country level rather than on the ethnic group level. On such an aggregate level of analysis there is more unobserved heterogeneity in the data, and some of our predictions relate to ethnic group characteristics, which calls for a test on this disaggregated level.

To address these concerns, we will now perform our own empirical analysis. First, we will use panel data on the country level and construct our novel oil inequality measure. Based on georeferenced petrol field and ethnic group location data, we are able to compute a Gini Index of

how unevenly petrol fields are spread between different ethnic groups in a given country. Using a standard specification of control variables and including country fixed effects, we will demonstrate that petrol unevenness has a positive and significant effect on the likelihood of civil conflict.

Afterwards, we will move to a panel on the ethnic group level. Using a variety of control variables, as well as country fixed effects / ethnic group fixed effects, it will be shown that indeed the ethnic groups out of power who are relatively oil rich with respect to the rest of the country will be significantly more likely to be involved in civil conflict. We also find that the interaction terms between oil abundance and group concentration, resp. distance from capital are positive and significant, pointing out that indeed civil wars are more likely when $R_1/(R_1 + R_2)$ and $p_s - p_n$ are both large, as predicted by our theory.

3.1 Empirical Analysis: Country Level

Our goal in this section is to analyze how the unevenness of geographical distribution of petrol fields across ethnic groups in a given country impacts on the likelihood of conflict. For this purpose we have put together a panel of 157 countries with sample period 1960-2008, and have constructed a novel variable, *Oil Gini*, which we will describe in more detail below. In our regressions we use –in addition to our main, new variable of interest– the standard battery of control variables, as well as country fixed effects and annual time dummies. Below we start by describing the data in some detail.

3.1.1 Data and Specification

Given that our theory focuses on bargaining failure to avoid conflict outbreaks, a natural choice of dependent variable is the onset of civil conflicts. The source of the civil war data is "UCDP/PRIO Armed Conflict Dataset" (UCDP, 2012), which is the most commonly used standard data source for civil wars at the country level[26]. We focus on the widely used standard definition of civil war, counting a country and year as having a civil war when at least 1000 casualties are recorded in a given year. Our conflict onset variable takes a value of 1 when a new civil war starts, is coded as missing during wars (where by definition a war cannot newly start), and is coded as 0 for peace. In a robustness check we will also show that our results are robust when focusing on war incidence as dependent variable (where ongoing wars are also coded as one).[27]

[26] This dataset has been used, among others, by Besley and Persson (2011) and Esteban, Mayoral and Ray (2012).

[27] The war onset variable relates to the question of what makes wars break out, while the war incidence variable is designed to capture the total intensity of conflict, which is not only driven by factors making wars start, but also by factors making wars last.

To construct our main novel *Oil Gini* measure, we started out from the ethnic group level. First, we constructed a panel dataset on the ethnic group and year level using the sample of ethnic groups in the "Geo-referencing of ethnic groups" (GREG) dataset (Weidmann, Rod and Cederman, 2010).[28] Relying on maps from the classical "Soviet Atlas Narodov Mira" from 1964 (Bruk and Apenchenko, 1964), which is still very extensively used for the construction of ethnolinguistic fractionalization (ELF) indices, GREG is a geo-referenced dataset containing the coordinates of the group boundaries of 1120 ethnic groups.

One major advantage of this dataset is that it provides a global coverage of ethnic groups for the whole world, containing relatively precise information on the geographical location of groups, which enables us to merge it with other geo-referenced group-level data using Geographical Information Systems (GIS). In contrast, the main alternative datasets on ethnic groups like the "Minorities at Risk" (Minorities at Risk, 2009), the "Ethnologue" (Lewis, 2009), and the lists of ethnic groups from Alesina et al. (2003), resp. Fearon (2003) lack detailed geographical information on the location of ethnic groups in all countries.[29]

The fact that the GREG data is a non-time varying snapshot from the early 1960s has both advantages and disadvantages. On the negative side, it implies that in some instances the group boundaries are not fully accurate in recent decades, although thankfully group borders generally evolve very slowly over time. The fact that the group border information is not time-varying lowers accuracy and hence adds noise to our estimations, which biases the magnitude of coefficients and the significance levels downwards, while there seems to be no other obvious bias of the results. This means that using GREG will tend to bias the results against us and this attenuation bias makes them appear *less* strong than they are in reality.

On the positive side, using ethnic group borders from the beginning of our sample –which hence pre-date the conflict observations in our regressions– has the advantage to alleviate concerns of ethnic group locations being endogenous to the wars that we want to explain. Still, there remains the possibility that past wars and existing oil fields had some impact on location patterns of ethnic groups –although such concerns are of course alleviated by the fact that ethnic group homelands

[28] Throughout the database construction we use the country borders from the time-varying, geo-referenced "CShapes" dataset (Weidmann, Kuse, and Gleditsch, 2010).

[29] Very few other datasets on ethnic groups provide geographical information. The "Ethnographic Atlas" (Gray, 1999) includes longitudes and latitudes for each group. Wucherpfennig et al. (2011) have also put together a dataset on ethnic groups with location information ("Ethnic Power Relations" (EPR)). They, however, do not include all ethnic groups, but only those that have been judged as "politically relevant", which could result in endogenous sample selection if the judgment of group "relevance" is affected by past political outcomes like wars. Controlling for past wars would not address this issue, as the concern is about the potential bias of the composition of the sample itself, with peaceful groups being less likely to be viewed as politically relevant, and hence more likely to be omitted from the sample. If these omitted groups differ in key characteristics from the included groups, the beta coefficients in the regression are biased.

are very stable. To address such remaining concerns we control for past conflict in all regressions and we include additional robustness tables in the appendix showing that our results also hold for a restricted sample that includes only observations of countries that did not have oil in 1964 when the "Soviet Atlas Narodov Mira" underlying GREG was put together.

Using GIS-Software (ArcGIS) we have then merged the GREG data with the geo-referenced petroleum dataset (PETRODATA) from Lujala, Rod and Thieme (2007), which documents where oil fields lie. Given that discovery dates of oil fields are available, this allowed us to construct a time-varying map of oil fields. Combining this information with the ethnic group data we were able to compute for every country and year the area occupied by a given ethnic group and also the area occupied by oil fields lying in the ethnic group's territory. Using this information, and the standard formula of the GINI inequality index, we were able to compute a novel, time-varying measure of the unevenness of oil field distribution across ethnic groups for a given country and year, which we call the *Oil Gini* variable.[30] Focusing on the inequality of oil fields between the homelands of ethnic groups, this measure captures well the logic of our model. For secession incentives it does not only matter who currently controls oil revenues, but even more important is the distribution of oil money between ethnic groups after a potential secession.

It is important to note the limits of the *Oil Gini* measure. Given that we lack data on the amounts of oil produced by ethnic groups, and only have data on the territory covered by oil, our novel *Oil Gini* variable is not a perfect measure of oil unevenness, but rather an imperfect proxy. However, the statistical noise of this measure does not appear biased in any particular direction. Hence, we expect our results to be conservative estimates suffering from attenuation bias, which makes our findings appear weaker that they are in reality.

Note also that for constructing the *Oil Gini* measure and the underlying ethnic group level measure of R_1/R we take oil discoveries as exogenous. While it has indeed been stressed in the literature that oil discoveries have a very important random element (see Lei and Michaels, 2014), it could still be that in productive and well-governed countries mining companies are keener to invest (which would bias downward the coefficient of variables capturing the total *amount* of oil). This concern is alleviated by the fact that we control in all specifications for GDP per capita and democracy scores. Further, given that our results on oil inequality hold when we control for the total amount of oil, any bias in the total quantity of oil discovered should not bias our results. In contrast, it would be more problematic if oil *inequality* was affected by endogenous drilling efforts. It however seems unlikely that investors or governments would deliberately miss out on opportunities for oil discoveries for strategic reasons. Still, as we show in the ethnic group level

[30] Note that for countries without oil the *Oil Gini* variable takes a value of 0.

analysis below, what drives the impact of oil inequality on conflict are oil findings in ethnic minority homelands (i.e. spikes in R_1/R) and it seems particularly unlikely that investors or a government in a conflict-prone environment would strategically favor drilling in an ethnic minority region over oil extraction in other regions. Hence, we don't expect an upward bias in our estimates.

Given that our *Oil Gini* is a novel measure, we shall now briefly describe its main summary statistics. First of all, we should point out that this new measure is only weakly correlated to other key covariates of conflict. To just mention two particularly important ones, the correlation of *Oil Gini* with the per capita oil production (from Humphreys, 2005) is -0.042, while the correlation of *Oil Gini* with GDP per capita (from Heston, Summers, and Aten, 2011) is -0.145.

In Figure 2 in the Appendix we display the country means of *Oil Gini* over the sample duration (1960-2008). There is quite some heterogeneity across countries, with the highest *Oil Gini* values found in African and Asian countries. This is intuitive: while in places with little oil (e.g. Europe) or huge amounts of oil (e.g. the Arabic peninsula) oil inequality is on average low, it is much higher in places like Africa or Asia where there are substantial oil fields, which however are scattered and punctual rather than covering the whole territory, as often the case in the Gulf states.

Similarly, in Figure 3 in the Appendix we display the country-level standard deviation over the sample period. Strikingly, the variance in *Oil Gini* is biggest in Africa (and some parts of Asia) where many of the new oil discoveries of the past decades took place. Interestingly, Africa and Asia are also the two continents where the risk for civil war onsets was the highest in the past decades (the average risk of new civil war onsets was 2.4 percent per year in Africa-Asia, while it was only 0.6 percent in the rest of the world).

In our regression analysis we use a standard battery of control variables, which results in a specification that is extremely close to the core specifications run by Fearon and Laitin (2003), Collier and Hoeffler (2004), Montalvo and Reynal-Querol (2005), Cederman and Girardin (2007), Collier and Rohner (2008), and Esteban, Mayoral and Ray (2012). Like these papers, we control for natural resource abundance (i.e., an updated version of Fearon and Laitin's (2003) "oil exporter" variable), GDP per capita, democracy, population size, whether a state was recently created, ethnic fractionalization, geography (mountainous terrain and noncontiguous states), and peace duration. In robustness checks we control for alternative war persistence measures (lagged war incidence, the number of war years in the last five year period), and alternative natural resource measures (gold producer, diamond producer, oil production per capita etc). All these variables are described in more detail in Appendix B. To account for omitted variable bias and unobserved heterogeneity we include from column 2 onwards country fixed effects and annual time dummies, and we allow

the robust standard errors to be clustered at the country level. We will as default run linear probability models, which have the advantage of providing easily interpretable coefficients and of allowing for clustered standard errors in the presence of country fixed effects. We will however show in a robustness check that the results go through when using conditional logit.

To summarize, we estimate the following main specification (see column 3 of Table 1):

$$War_{c,t} = \alpha Oil\_Gini_{c,t} + \mathbf{X}'_{c,t}\boldsymbol{\beta} + \mathbf{Y}'_{c}\boldsymbol{\gamma} + \mathbf{Z}'_{t}\boldsymbol{\delta} + u_{c,t}$$

where $\mathbf{X}'_{c,t}$ is a vector of time-varying and country-varying control variables, $\mathbf{Y}'_{c}$ is a vector of country fixed effects, $\mathbf{Z}'_{t}$ is a vector of annual time dummies, and $u_{c,t}$ is the error term. The coefficient of interest is α, which our theory predicts to be of positive sign.

3.1.2 Results

In Table 1 we display the results of the country level regressions. In the first column we run a linear probability model with just the battery of standard controls, but without our new *Oil Gini* measure. For comparability with most papers in the existing literature (e.g. the classic article of Fearon and Laitin, 2003) we focus on a specification without country and time fixed effects, where standard errors are left unclustered. Besides the "New State" variable, all other variables have the expected sign. In line with the existing literature (see e.g. Fearon and Laitin (2003), Collier and Hoeffler (2004), Montalvo and Reynal-Querol (2005), Cederman and Girardin (2007), Collier and Rohner (2008), and Esteban, Mayoral and Ray (2012)), we find that oil abundant, ethnically fractionalized and very populated states with low GDP per capita and a track record of past violence are more likely to experience civil war onsets.

In column 2 we add country fixed effects and annual time dummies, and allow for robust standard errors to be clustered at the country level, in the goal of addressing omitted variable bias and unobserved heterogeneity. As noted in the literature (Fearon and Laitin, 2003; Collier and Hoeffler, 2004; Sambanis, 2004; Collier, Hoeffler and Rohner, 2009) adding country fixed effects tends to remove the significance of most variables in such civil war regressions. In column 3 we now add our new *Oil Gini* measure, which has the expected sign and is significant at the 10% level. Its effect is quantitatively important: Moving a country from full ethnic oil equality to full oil inequality increases the risk of war onsets by 6.6 percentage points, which is more than four-fold the baseline risk of 1.5 percentage points.

In columns 4 and 5 it is shown that the effect of *Oil Gini* is robust to running logit regressions instead of OLS and to having "civil war incidence" rather than "civil war onset" as dependent

	Dependent variable: Civil War Onsets					
	(1)	(2)	(3)	(4)	(5)	(6)
Oil Gini			**0.066***	**7.955****	**0.098****	**0.065***
			(0.035)	**(3.912)**	**(0.045)**	**(0.035)**
Oil exporter (t-1)	0.016**	0.018	0.021	2.774**	0.014	0.019
	(0.007)	(0.017)	(0.019)	(1.258)	(0.014)	(0.016)
ln GDP p.c.(t-1)	-0.004*	-0.009	-0.006	-0.363	0.001	-0.006
	(0.002)	(0.008)	(0.009)	(0.589)	(0.014)	(0.009)
Democ. (t-1)	0.000	-0.001*	-0.001*	-0.040	-0.000	-0.001*
	(0.000)	(0.000)	(0.001)	(0.042)	(0.001)	(0.001)
ln Popul.(t-1)	0.004***	0.005	0.006	-2.449	0.011	0.006
	(0.001)	(0.010)	(0.011)	(1.974)	(0.014)	(0.011)
New State	-0.014***	0.007	0.008	1.799	0.009	0.009
	(0.004)	(0.006)	(0.006)	(1.273)	(0.008)	(0.006)
Ethnic Fraction.	0.014**					
	(0.007)					
Mountainous Terr.	0.003					
	(0.008)					
Noncontig. State	0.006					
	(0.005)					
Peace duration	-0.001***	-0.000	-0.000	0.025*		-0.000
	(0.000)	(0.000)	(0.000)	(0.014)		(0.000)
Civ. War Incid. (t-1)					0.575***	
					(0.040)	
Oil Exp. * Oil Gini						0.004
						(0.042)
Model	OLS	OLS	OLS	Logit	OLS	OLS
Country Fixed Eff.	No	Yes	Yes	Yes	Yes	Yes
Observations	5289	5417	5180	1093	5421	5180
R-squared	0.024	0.105	0.112	0.227	0.544	0.112

Notes: Dependent variable: Civil war onset (unless in column 5 where the dependent variable is Conflict incidence). The dependent variable is coded as 1 if a conflict causing at least 1000 fatalities is starting in a given year. Sample period: 1960-2008. Number of countries for which observations are available: 157. From column 2 onwards unreported country fixed effects and annual time dummies are included, and robust standard errors are allowed to be clustered at the country level (unless in the conditional logit regression of column 4 where clustering of standard errors is not possible). Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 1: Country Level Regressions of the Effect of Oil Gini on Conflict Onsets

variable: In both columns, *Oil Gini* has a positive sign and is significant at the 5 percent level.

In column 6 we also include an interaction term of *Oil Gini* and our dummy variable taking a value of 1 for big oil exporting countries. The coefficient of *Oil Gini* remains positive and significant, while the interaction term also has the expected positive sign, but is not statistically significant.[31]

In Table 3 in Appendix A we provide some additional results. First, columns 1-5 show that *Oil Gini* continues to be statistically significant and to have a coefficient of similar magnitude when we control for an alternative war persistence variable and for other natural resource measures (gold producer, diamond producer, oil production per capita, share of a country's area covered with oil, and oil area / population). In columns 6-8 we re-run the main regressions of columns 3, 5 and 6 of Table 1, but restricting the sample to observations from countries that did not have oil in 1964 when the "Atlas Narodov Mira" (Bruk and Apenchenko, 1964) –which is the primary source for our geo-referenced ethnic group location dataset– was put together. This restriction of the sample serves the purpose of alleviating concerns that oil abundance could have affected the location of ethnic groups. Even in these very demanding regressions run on a restricted sample our main *Oil Gini* variable has the expected sign and is statistically significant. Interestingly, also the interaction term of *Oil Gini* with Oil Exporter in column 8 is of the expected positive sign and statistically significant.

3.2 Empirical Analysis: Ethnic Group Level

To test the predictions of our theory on the impact of the relative oil richness of the ethnic groups on civil war onsets, we have put together a panel dataset *on the ethnic group level*, covering 1120 ethnic groups and spanning over the period 1960-2006. In particular, we include all ethnic groups of the "Geo-referencing of ethnic groups" (GREG) dataset (Weidmann, Rod and Cederman, 2010), and construct group level variables that capture closely the expression $R_1/(R_1+R_2)$ of the model, as explained above. There is a small number of papers in the literature that study civil war on the ethnic group level (e.g., Walter, 2006; Cederman, Buhaug and Rod, 2009), but these papers typically either ignore natural resources or use a natural resource abundance variable at the country-level, which does not allow to capture unevenness of natural resource distribution.[32] Hence, to the

[31] Note that while the focus of this paper lies on oil *inequality*, an auxiliary implication of the model is that –for a fixed resource inequality– a general increase in resources is expected to increase the conflict risk. The empirical support for this auxiliary implication of our model is rather weak. In Table 1, for example, the resource abundance variable "Oil exporter" has always the expected positive sign, but is only statistically significant in two out of six columns. One potential explanation for the weak direct effect of oil abundance is the inclusion of country fixed effects that are indispensable for a tight identification strategy, but which remove all cross-sectional variation in oil abundance.

[32] There are only few exceptions: Sorens (2011) also uses ethnic group level natural resource data to explain territorial conflict. While our sample contains all ethnic groups, he uses the selected sample of discriminated groups

best of our knowledge we are the first ones to study civil war using a panel of all ethnic groups with natural resource variables that vary for different ethnic groups.

Looking at this disaggregated level of analysis has several advantages: With more fine grained data typically unobserved heterogeneity and omitted variable problems are reduced. Further, studying group level data also allows us to better discriminate between our mechanism and others. Our theory predicts that conflicts become more likely to break out if a powerless ethnic minority group becomes very oil rich. In contrast, alternative theories may predict conflict onsets when the oil abundance of the governing ethnic group's homelands rises. So far, our country level regressions only established the link between inter-ethnic oil unevenness and civil war onsets, without being able to distinguish between our theory and alternative theories stressing for example war triggered by oil rich governing groups. In contrast, the ethnic group level regressions below will be able to distinguish between such different mechanisms, and assess whether the dangerous type of oil unevenness is indeed the one where powerless groups are oil abundant, as predicted by our theory.

3.2.1 Data and Specification

The main dependent variable is civil war onset at the group level and is taken from Cederman, Buhaug and Rod (2009). It varies on the ethnic group and year level and takes a value of 1 when in a given year a civil war newly starts that involves a given ethnic group. It is coded as missing for ongoing wars, and as 0 if the ethnic group is not involved in a civil war in a given year.

Our main independent variable is the surface of an ethnic group's territory covered with oil and gas as a percentage of the country's total surface covered with oil and gas. This proxies well $R_1/(R_1 + R_2)$. To construct this measure, we used as starting point all ethnic groups in the "Geo-referencing of ethnic groups" (GREG) dataset (Weidmann, Rod and Cederman, 2010), which allowed us to know the geographical coordinates of where a given ethnic group is located. Then we merged this with the geo-referenced petroleum dataset (PETRODATA) from Lujala, Rod and Thieme (2007), which documents where oil fields lie and when they have been discovered. Combining this information we were able to compute a variable measuring which part of the territory occupied by a given ethnic group contains oil. Expressing this in terms of the total surface containing oil in the country, we obtain a quite precise, time-varying measure of how relatively petrol-rich the homelands of a given ethnic group are.

We also include various control variables on the ethnic group level. In particular, we include

from the "Minorities at Risk" (MAR) project. Further, contrary to us, he runs pooled panel regressions which do not control for country fixed effects. Esteban, Morelli and Rohner (2015) use an ethnic-group level resource abundance measure to study massacres of civilians, but they do not include in their analysis any variable of a group's relative natural resource abundance with respect to the other groups in the country.

two time-varying demographic controls, *Group Population / Governing Groups' Population* and *Governing Groups' Population* from Cederman, Buhaug and Rod (2009). Note that while the population estimates of each ethnic group are not time varying, the composition of groups controlling the government is time varying, which makes these measures time-varying. Further, we construct different war persistence measures (*Group Peace Duration*, and lagged *Group War Incidence*). In addition, we control for several time-invariant geographical and demographic factors, i.e. *Group Diamond Mines* (constructed using raw data from Gilmore et al., 2005), *Group co-ethnics abroad* (constructed based on the GREG dataset), *Group's share of mountainous terrain* and *ln group's distance to capital* (from Cederman, Buhaug and Rod, 2009), *Group's soil quality* (constructed using the Harmonized World Soil Database from Fischer et al., 2008), and *Group population density* (constructed using the geo-referenced population density data from the Socioeconomic Data and Applications Center at Columbia University, see Sedac, 2012).

Further, we also include in some specifications interaction terms of our main independent variable $R_1/(R_1+R_2)$ with, among others, a group concentration measure constructed using raw data from Alesina and Zhuravskaya (2011), and with a group distance from capital measure (as described above). All these variables are described in more detail in Appendix B.

Finally, we also include –but do not report– all country-level standard control variables included in the country-level regressions above and annual time dummies. Like in the country-level regressions above we allow robust standard errors to be clustered at the country level. As above, we also run as main specifications linear probability models, but show that the results are robust to the use of conditional logit estimators.

To summarize, we estimate the following main specification (corresponding to column 3 of Table 2):

$$War_{e,t} = \alpha(R_1/R)_{e,t} + \mathbf{W}'_{e,t}\boldsymbol{\beta} + \mathbf{X}'_{c,t}\boldsymbol{\gamma} + \mathbf{Y}'_{e}\boldsymbol{\delta} + \mathbf{Z}'_{t}\boldsymbol{\zeta} + u_{e,t}$$

where $\mathbf{W}'_{e,t}$ is a vector of time-varying and ethnic group-varying control variables, $\mathbf{X}'_{c,t}$ is a vector of time-varying and country-varying control variables, $\mathbf{Y}'_{e}$ is a vector of ethnic group fixed effects, $\mathbf{Z}'_{t}$ is a vector of annual time dummies, and $u_{e,t}$ is the error term. The coefficient of interest is α, which our theory predicts to be of positive sign.

3.2.2 Results

The focus of our empirical analysis is to assess whether conflict indeed becomes more likely when the group out of power has an ethnic homeland that is abundant in oil, i.e. when $R_1/(R_1+R_2)$

is large, as predicted by the theory. Hence, for most of our regressions, namely in the main table 2, in the Appendix Table 4 and in the first four columns of the Appendix Table 5, we will use the sample of all ethnic groups that are out of power, which corresponds to the largest part of all ethnic groups (i.e. on average about 83% of all ethnic groups are non-governing groups).[33] As discussed in more detail below, in the columns 5-6 of the Appendix Table 5 we run as Placebo test the main regressions on a sample of governing groups.

In the first column of table 2 we regress for a pooled panel our main dependent variable, Civil War Onsets, on the main independent variable R_1/R (where $R \equiv R_1 + R_2$) and on all controls mentioned above. Our main variable R_1/R has the expected positive sign and is significant at the 5% level. The effect is sizeable: If an initially oil-less ethnic group has after a discovery all the country's oil fields on its territory, the risk of being involved in a civil war onset is increased by 2.7 percentage points, which is about 8-times larger than the relatively small baseline risk for a group to be involved in a conflict onset, which is of 0.3 percentage points.

About half of the control variables are significant. Larger ethnic groups and groups occupying valuable soils are more likely to be involved in conflict onsets, while groups with a peaceful past and ethnic groups with co-ethnics in many other countries are less likely to be involved in civil war onsets.

In column 2 we run the same specification, but now including country fixed effects. Our main variable R_1/R remains significant, with a coefficient of similar size. In column 3 we now include group fixed effects. This is a very demanding specification, as all time-invariant group characteristics are controlled for by the group fixed effects (which implies that all time-invariant variables drop from the specification) and the results are purely driven by within-group changes of relative resource abundance over time. Even in this very demanding specification our main independent variable R_1/R remains statistically significant (the p-value is 0.053), with now a somewhat larger coefficient.

In columns 4 and 5 we show that the results of columns 2 and 3 are robust to running conditional logit estimations with country, respectively group fixed effects, instead of the linear probability models used before.

In columns 6 and 7 we run the same regressions like in columns 2 and 3, but with civil war incidence rather than civil war onsets as dependent variable. While our main variable R_1/R remains significant in the presence of country fixed effects, it drops below the 10% significance threshold when group fixed effects are included. This suggest that indeed our theory is better at predicting the break-outs of wars due to bargaining failure rather than war duration, which also depends on

[33] Concretely, all ethnic groups are included that do not control the government of their country in a given year, i.e. for which the variable "ethnic group in power" (egip) of Cederman, Buhaug and Rod (2009) takes a value of zero.

additional factors.

Our theory predicts that not only R_1/R will be a main driving factor of bargaining failure to prevent war, but that also the difference between the winning prospects in secessionist versus in nationwide wars, $p_s - p_n$, matters. As shown in Corollary 1, the wedge between p_s and p_n is typically larger when a given ethnic group is very concentrated and located far from the capital.

In column 8 we hence interact our main variable of relative resource abundance, R_1/R, with a measure of group concentration ranging from 0 to 1, constructed with raw data from Alesina and Zhuravskaya (2011). While this is the most precise available data for our purpose, it only covers a subset of our ethnic groups, leading to a drop in the sample size. It is also a time-invariant variable, meaning that the baseline group concentration variable drops from the specification as we include group fixed effects, and only its interaction with R_1/R remains in the specification. We find a quantitatively strong effect of the interaction term that is significant at the 1% level, and also R_1/R and (R_1/R) * Group Concentration are jointly statistically significant at the 1% level. This highlights that indeed civil conflict outbreaks are most likely when R_1/R and $p_s - p_n$ are both large, as shown in Corollary 1 and displayed graphically in Figure 1.

In column 9 we include the interaction of our R_1/R variable with a dummy of above-median distance from the capital. In line with Corollary 1 we again expect this interaction term to be positive. Indeed, we find a quantitatively large, positive coefficient of this interaction term that is statistically significant at the 5% level. The variables R_1/R and (R_1/R)* Distance from capital are jointly significant at the 5% level. These findings are consistent with our theory's prediction that indeed civil conflict outbreaks are most likely when R_1/R and $p_s - p_c$ are both large.

In the Appendix Table 4 we perform several robustness checks. In column 1 we control for the percentage of the country's area covered by oil. The variable of interest R_1/R continues to be significant at the 5% level, while this new control variable is not statistically significantly different from zero. Similarly, in column 2 we have as additional control variable the country's oil area per capita. While this new control is not statistically significant, our variable of interest R_1/R is still statistically significant.[34]

One potential concern of our analysis could be that oil discoveries and conflict may affect the likelihood of a given ethnic group to reach power. Hence, in column 3 we replicate the baseline specification of column 3 of Table 2, but only including in the sample ethnic groups that *never* reach

[34] Note that we have run various additional robustness checks controlling for non-monotonic effects of resource abundance. In particular, in the country-level (resp., group level) analysis the coefficient of interest of the Oil Gini (resp., R_1/R) variable remains statistically significant and of similar magnitude when controlling for both the linear and square terms of the share of oil fields in the countries' total surface, or the area of oil fields divided by the countries' population.

Dep. Var.: Civil War	Onsets	Onsets	Onsets	Onsets	Onsets	Incid.	Incid.	Onsets	Onsets
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
R1 / R	**0.027****	**0.022***	**0.064***	**2.610*****	**3.189***	**0.021***	**0.048**	**0.024**	**-0.061**
	(0.012)	**(0.012)**	**(0.033)**	**(0.652)**	**(1.859)**	**(0.012)**	**(0.034)**	**(0.026)**	**(0.047)**
(R1/R)* Group. Conc.								**0.181*****	
								(0.044)	
(R1/R) * Dist.									**0.169****
									(0.068)
Gr.Pop./Gov.Pop.(t-1)	0.003*	0.006***	0.012***	0.578***	9.627**	0.007***	0.006**	0.012	0.012***
	(0.002)	(0.002)	(0.002)	(0.158)	(4.589)	(0.002)	(0.002)	(0.010)	(0.002)
Gov. Pop. (t-1)	-0.000	-0.008	0.004	-1.012	0.921	-0.009	0.007	0.042	0.004
	(0.001)	(0.009)	(0.007)	(2.365)	(3.476)	(0.006)	(0.008)	(0.054)	(0.007)
Group Peace Duration	-0.001***	-0.001	0.001	-0.062***	0.110***			0.001	0.001
	(0.000)	(0.001)	(0.001)	(0.008)	(0.018)			(0.001)	(0.001)
Group Diamond Mines	0.032	0.029		5.107*		0.043			
	(0.032)	(0.031)		(2.931)		(0.041)			
Group Co-Ethnics	-0.001**	-0.000		-0.014		-0.000			
	(0.000)	(0.000)		(0.065)		(0.000)			
Group Mount. Terr.	0.002	0.003*		0.583		0.002			
	(0.002)	(0.002)		(0.391)		(0.002)			
ln Group Dist. Cap.	0.002	0.004		1.325***		0.006**			
	(0.001)	(0.002)		(0.232)		(0.003)			
Group Soil Quality	0.005*	0.001		0.028		-0.001			
	(0.002)	(0.002)		(0.496)		(0.002)			
Group Pop. Dens.	0.009	0.014		3.058**		0.023			
	(0.012)	(0.012)		(1.227)		(0.015)			
Group War Inc.(t-1)						0.906***	0.801***		
						(0.014)	(0.030)		
Model	OLS	OLS	OLS	Logit	Logit	OLS	OLS	OLS	OLS
Fixed effects	No	Cou. FE	Gr. FE	Cou. FE	Gr. FE	Cou. FE	Gr. FE	Gr. FE	Gr. FE
Observations	28549	28594	28594	14799	2013	30164	30164	11828	28594
R-squared	0.024	0.049	0.168	0.274	0.376	0.869	0.878	0.128	0.169

Notes: Dependent variable: Civil war onsets (unless when specified differently in row 1). Sample period: 1960-2006. Number of non-governing ethnic groups for which observations are available: 930. In all columns the unreported set of country level controls of Table 1 and annual time dummies are included, and robust standard errors are allowed to be clustered at the country level (unless in the conditional logit regressions of columns 4-5 where clustering of standard errors is not possible). Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 2: Ethnic Group Level Regressions of the Effect of Oil Unevenness on Conflict Onsets

power (i.e. which are never part of the government coalition during the entire sample period). Also for this sample our main variable R_1/R remains significant.[35]

In column 4 we replicate our findings on the interaction terms of (R_1/R) with group concentration and distance from capital, respectively, but including a further set of controls variables. In particular, we include as additional controls the interaction of (R_1/R) with the share of mountainous terrain in the ethnic group's homelands, with soil quality, and with the ratio of group population over government population. The coefficients of (R_1/R) * Group Concentration and (R_1/R) * Distance from capital are still positive and significant as expected.

Finally, column 5 shows that our main variable R_1/R can also explain the presence of separatist movements in a given ethnic group. Given that this variable is ethnic group specific, but not time-varying, we are only able to include country fixed effects, but not group fixed effects.

The Appendix Table 5 provides additional results. In order to alleviate concerns of oil abundance affecting ethnic group location, we restrict in columns 1-4 the sample to groups from countries that did not have oil in 1964, which is the year when our primary ethnic group location data was put together. Column 1 replicates the main regression of column 3 of Table 2 for this restricted sample finding still a positive coefficient of comparable size that is significant at the 5% level. Also the regression of column 2 –which mirrors the war incidence regression of column 7 of Table 2– finds like before in Table 2 a positive coefficient that however is not quite statistically significant. Like before in the columns 8-9 of Table 2, we find in columns 3-4 that the interaction terms of R_1/R and group concentration, resp. distance from capital are statistically significant.

The columns 5-6 re-run the main regressions of columns 2 and 7 of Table 2, but this time for governing rather than for non-governing groups. According to our theory the expression R_2/R that captures resource abundance of the governing group should –if anything– decrease the risk of civil war. We find indeed a negative coefficient which is however imprecisely estimated and not statistically significant. This is unsurprising given the smaller sample size. Still, running such a regression on a sample of governing groups and finding a non significant coefficient of interest being of opposite, negative sign alleviates concerns that our main results could be driven by some other channel outside our model leading mechanically to more conflict in resource rich groups.

[35] Another way to address this potential concern is to check whether oil discoveries may trigger changes in the identity of the group in power. To assess this, we ran panel regressions on the ethnic group level with the sample consisting of all groups out of power. We had as dependent variable a dummy taking a value of 1 if a group that was out of power in the previous period gained power in the following period (i.e. newly entered the governing coalition), and taking a value of 0 if it remained powerless. The main explanatory variable of interest was whether there had been oil discoveries in the last year (resp. in the last five years). We looked separately at country or group level oil discoveries. In all specifications we found that oil discoveries did not statistically significantly increase the likelihood of a powerless ethnic group to newly access political power. This result was obtained in a variety of specifications, i.e. in the absence of control variables, with all standard controls of our earlier regressions, and with country or ethnic group fixed effects.

4 Conclusion

This paper employs a simple model to capture how the geographical distribution of natural resources within a country can affect the risk of civil war. We have allowed for uneven resource abundance and two different forms of conflict, secessionist and nationwide, and found that bargaining fails to prevent costly conflict if an ethnic minority group is located in a region that is particularly rich in natural resources (relative to the rest of the country) and if its winning probability for the case of secessionist conflict is substantially larger than for nationwide conflict. This leads to the presence of two relevant threat points for war, which cannot be addressed at the same time by bargaining.

This rationale for bargaining breakdown due to multiple types of civil war is not only interesting per se, but it is also shown to fit with most existing empirical findings and with the estimations on our newly constructed data sets with detailed geographic information about ethnic groups and resources.

The analysis of the conditions under which discovery of new resources in a developing country can lead to development without conflicts is still incomplete. This paper has highlighted the significance of resource concentration and group concentration for civil war, but of course the elimination of the resource curse in terms of violent incentives also requires checking that inter-state wars and other forms of violence and discriminations can be avoided. An integrated study of the role of natural resources for violent incentives of all kinds is left as next step in our research agenda.

References

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg (2003): "Fractionalization", *Journal of Economic Growth* 8: 155-194.

Alesina, Alberto, Stelios Michalopoulos and Elias Papaioannou (2014): "Ethnic Inequality," forthcoming in *Journal of Political Economy.*

Alesina, Alberto and Ekaterina Zhuravskaya (2011): "Segregation and the Quality of Government in a Cross Section of Countries", *American Economic Review* 101: 1872–1911.

Basedau, Matthias, and Thomas Richter (2011): "Why Do Some Oil Exporters Experience Civil War But Others Do Not? A Qualitative Comparative Analysis of Net Oil-Exporting Countries", mimeo, German Institute of Global and Area Studies.

Bell, Curtis, and Scott Wolford (2014): "Oil Discoveries, Shifting Power, and Civil Conflict," forthcoming in *International Studies Quarterly.*

Berman, Nicolas, Mathieu Couttenier, Dominic Rohner, and Mathias Thoenig (2014): "This mine is mine! How minerals fuel conflicts in Africa", CEPR Discussion Paper No. 10089.

Besley, Timothy, and Torsten Persson (2011): "The Logic of Political Violence", *Quarterly Journal of Economics* 126: 1411-1445.

Boulding, Kenneth (1962): *Conflict and defense: A general theory*, New York: Harper & Brothers.

Bruk, Solomon, and V. S. Apenchenko (eds) (1964): *Atlas narodov mira* [Atlas of the Peoples of the World]. Moscow: Glavnoe Upravlenie Geodezii i Kartografi.

Buhaug, Halvard, Scott Gates, and Päivi Lujala (2009): "Geography, Rebel Capability, and the Duration of Civil Conflict," *Journal of Conflict Resolution* 53: 544-69.

Caselli, Francesco and Wilbur John Coleman II (2013): "On the Theory of Ethnic Conflict," *Journal of the European Economic Association* 11: 161-192.

Caselli, Francesco, Massimo Morelli, and Dominic Rohner (2015): "The Geography of Inter State Resource Wars," *Quarterly Journal of Economics* 130: 267-315.

Cederman, Lars-Erik, Halvard Buhaug and Jan Ketil Rod (2009): "Ethno-Nationalist Dyads and Civil War : A GIS-Based Analysis," *Journal of Conflict Resolution* 53: 496-525.

Cederman, Lars-Erik and Luc Girardin (2007): "Beyond Fractionalization: Mapping Ethnicity onto Nationalist Insurgencies," *American Political Science Review* 101: 173-85.

Chassang, Sylvain, and Gerard Padro i Miquel (2009): "Economic Shocks and Civil War," *Quarterly Journal of Political Science* 4: 211-228.

Collier, Paul and Anke Hoeffler (2004): "Greed and Grievance in Civil War", *Oxford Economic Papers* 56: 563-95.

Collier, Paul, Anke Hoeffler, and Dominic Rohner (2009): "Beyond Greed and Grievance: Feasibility and Civil War", *Oxford Economic Papers* 61: 1-27.

Collier, Paul and Dominic Rohner (2008): "Democracy, Development, and Conflict", *Journal of the European Economic Association* 6: 531-40.

Cotet, Anca M., and Kevin K. Tsui (2013): "Oil and Conflict: What Does the Cross Country Evidence Really Show?", *American Economic Journal: Macroeconomics* 5: 49-80.

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

Esteban, Joan, Laura Mayoral, and Debraj Ray (2012): "Ethnicity and Conflict: An Empirical Investigation," *American Economic Review* 102: 1310-1342.

Esteban, Joan, Massimo Morelli, and Dominic Rohner (2015): "Strategic Mass Killings," forthcoming in *Journal of Political Economy*.

Esteban, Joan and Debraj Ray (2011): "A Model of Ethnic Conflict", *Journal of the European Economic Association* 9: 496–521.

Fearon, James (1995): "Rationalist Explanations for War," *International Organization* 49: 379-414.

Fearon, James (2003): "Ethnic and Cultural Diversity by Country", *Journal of Economic Growth* 8: 195-222.

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

Fearon, James (2007): "Economic Development, Insurgency, and Civil War", in Elhanan Helpman (ed), *Institutions and Economic Performance*. Cambridge, MA: Harvard University Press.

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

Fischer, Guenther, Freddy Nachtergaele, Sylvia Prieler, Harrij van Velthuizen, Luc Verelst, and David Wiberg (2008): *Global Agro-ecological Zones Assessment for Agriculture*. IIASA, Laxenburg, Austria and FAO, Rome, Italy.

Gates, Scott (2002): "Recruitment and Allegiance: The Microfoundations of Rebellion", *Journal of Conflict Resolution* 46: 111-30.

Gilmore, Elisabeth, Nils Petter Gleditsch, Päivi Lujala, and Jan Ketil Rod (2005): "Conflict Diamonds: A New Dataset", *Conflict Management and Peace Science* 22: 257–292.

Gray, J. Patrick (1999): "A Corrected Ethnographic Atlas", *World Cultures* 10: 24-85.

Grossman, Herschel and Juan Mendoza (2003): "Scarcity and appropriative competition," *European Journal of Political Economy* 19: 747-58.

Gurr, Ted (2000): *Peoples Versus States: Minorities at Risk in the New Century*, Washington DC: U.S. Institute of Peace Press.

Heston, Alan, Robert Summers, and Bettina Aten (2011): "Penn World Table Version 7.0", dataset, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, http://pwt.econ.upenn.edu/.

Hodler, Roland (2006): "The curse of natural resources in fractionalized countries," *European Economic Review* 50: 1367-86.

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

Huber, John and Laura Mayoral (2013): "Civil Conflict and Three Dimensions of Ethnic Inequality", mimeo, Columbia University.

Humphreys, Macartan (2005): "Natural Resources, Conflict, and Conflict Resolution: Uncovering the Mechanisms," *Journal of Conflict Resolution* 49: 508-37.

Jackson, Matthew O. and Massimo Morelli (2007): "Political Bias and War," *American Economic Review* 97: 1353-73.

Jackson, Matthew O. and Massimo Morelli (2011): "The Reasons for War, An Updated Survey," in *Handbook on the Political Economy of War*, edited by Chris Coyne, Cheltenham UK: Edward Elgar Publishing.

Jervis, Robert (1978): "Cooperation under the Security Dilemma," *World Politics* 30: 188-189.

Le Billon, Philippe (2001): "The political ecology of war: natural resources and armed conflicts," *Political Geography* 20: 561-584.

Lei, Yu-Hsiang, and Guy Michaels (2014): "Do Giant Oil Field Discoveries Fuel Internal Armed Conflicts?", *Journal of Development Economics* 110: 139-157.

Levy, Jack (1987): "Declining Power and the Preventive Motivation for War," *World Politics* 40: 90-92.

Lewis, M. Paul (ed.) (2009): *Ethnologue: Languages of the World* (Sixteenth edition), Dallas, Texas: SIL International.

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

Lujala, Päivi, Jan Ketil Rod and Nadja Thieme (2007): "Fighting over Oil: Introducing a New Dataset", *Conflict Management and Peace Science* 24: 239-56.

Lujala, Päivi, Nils Petter Gleditsch and Elisabeth Gilmore (2005): "A Diamond Curse? Civil War and a Lootable Resource," *Journal of Conflict Resolution* 49: 538-62.

Minorities at Risk (2009): Dataset, www.cidcm.umd.edu/mar.

Montalvo, José and Marta Reynal-Querol (2005): "Ethnic Polarization, Potential Conflict, and Civil Wars," *American Economic Review* 95: 796-815.

Morelli, Massimo and Costantino Pischedda (2014): "The Turkey-KRG Energy Partnership: Assessing Its Implications", *Middle East Policy* 21: 107-121.

Morelli, Massimo, and Dominic Rohner (2014): "Resource Concentration and Civil Wars," NBER Working Paper 20129.

Ploeg, Frederick van der and Dominic Rohner (2012): "War and Natural Resource Exploitation," *European Economic Review* 56: 1714-1729.

Polity IV, 2012, "Political Regime Characteristics and Transitions", dataset, www.systemicpeace.org/polity4.

Powell, Robert (1996): "Uncertainty, Shifting Power, and Appeasement", *American Political Science Review* 90: 749-764.

Powell, Robert (2006): "War as a Commitment Problem," *International Organization* 60: 169-203.

Ray, Debraj (2010): "Multiple Threats to Peace", mimeo, NYU.

Reiter, Dan (1995): "Exploding the Powder Keg Myth: Preemptive Wars Almost Never Happen," *International Security* 20: 5-34.

Reuveny, Rafael and John Maxwell (2001): "Conflict and Renewable Resources," *Journal of Conflict Resolution* 45: 719-42.

Reynal-Querol, Marta (2002): "Ethnicity, Political Systems, and Civil Wars," *Journal of Conflict Resolution* 46: 29-54.

Rohner, Dominic (2014): "Barrels, Books, and Bullets: How Education Can Prevent Conflict and Promote Development in Resource Rich Countries", mimeo, University of Lausanne.

Rohner, Dominic, Mathias Thoenig, and Fabrizio Zilibotti (2013): "War Signals: A Theory of Trade, Trust and Conflict," *Review of Economic Studies* 80: 1114-1147.

Ross, Michael (2004): "What Do We Know About Natural Resources and Civil War?" *Journal of Peace Research* 41: 337-56.

Ross, Michael (2004b): "How Do Natural Resources Influence Civil War? Evidence from Thirteen Cases," *International Organization* 58: 35-67.

Ross, Michael (2012): *The Oil Curse: How Petroleum Wealth Shapes the Development of Nations*, Princeton NJ: Princeton University Press.

Saideman, Stephen, David Lanoue, Michael Campenni, and Samuel Stanton (2002): "Democratization, Political Institutions, and Ethnic Conflict: A Pooled Time-Series Analysis, 1985-1998," *Comparative Political Studies* 35: 103-29.

Sambanis, Nicholas (2004): "Poverty and the Organization of Political Violence: A Review and Some Conjectures," in Carol Graham and Susan Collins (eds.), *Brookings Trade Forum: 2004: Globalization, Poverty, and Inequality*, Washington DC: Brookings Institution Press.

Schelling, Thomas C. (1966): *Arms and Influence*, New Haven and London: Yale University Press.

Sedac (2012): "Geo-referenced population density data," Dataset, Socioeconomic Data and Applications Center (Sedac), Columbia University, http://sedac.ciesin.org/.

Snyder, Jack (1984): "Civil-Military Relations and the Cult of the Offensive, 1914 and 1984," *International Security* 9: 108-146.

Sorens, Jason (2011): "Mineral production, territory, and ethnic rebellion: The role of rebel constituencies," *Journal of Peace Research* 48: 571-85.

Taylor, Alan J.P. (1954): *The Struggle for Mastery in Europe, 1848 to 1918,* Oxford: Clarendon Press.

UCDP (2012): "UCDP/PRIO Armed Conflict Dataset," dataset, www.pcr.uu.se/research/ucdp.

Walter, Barbara (2006): "Information, Uncertainty, and the Decision to Secede," *International Organization* 60: 105-35.

Walter, Barbara (2006b): "Building Reputation: Why Governments Fight Some Separatists but Not Others, " *American Journal of Political Science* 50: 313-30.

Weidmann, Nils, Doreen Kuse, and Kristian Skrede Gleditsch (2010): "The Geography of the International System: The CShapes Dataset," *International Interactions* 36: 86-106.

Weidmann, Nils, Jan Ketil Rod, and Lars-Erik Cederman (2010): "Representing Ethnic Groups in Space: A New Dataset," *Journal of Peace Research* 47: 491-499

Wimmer, Andreas (2014): "Ethnic Armed Conflict" (Version 3), dataset, http://www.princeton.edu/~awimmer/WarlistEPRVersion3.xls.

World Bank (2010): "Adjusted Net Savings Database", Dataset, http://go.worldbank.org/3AWKN2ZOY0.

World Bank (2012): "World Development Indicators," Dataset, http://data.worldbank.org/indicator.

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

Appendix A: Additional Regression Tables

The Tables 3, 4 and 5 display –as discussed in the main text– additional regressions at the country, resp. ethnic group level.

Appendix B: Data

Below we describe the data used in section 3. Table 6 provides the descriptive summary statistics for all variables. At the end of the appendix we also provide a graphical representation of the *Oil Gini* variable.

Country Level Variables

The dependent variables, civil war onset and incidence, and the main independent variable, *Oil Gini*, have been described above in the main text. In what follows we describe the control variables.

Oil exporter: Dummy variable taking a value of 1 if in a given country and year the fuel exports (in % of merchandise exports) is above 33%. Variable from Fearon and Laitin (2003), but updated

	Dependent variable: Civil War Onsets							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Oil Gini	**0.069***	**0.107****	**0.067***	**0.064***	**0.062***	**0.092***	**0.155****	**0.083***
	(0.037)	**(0.053)**	**(0.037)**	**(0.035)**	**(0.034)**	**(0.053)**	**(0.067)**	**(0.049)**
Oil exporter (t-1)	0.020	0.033				0.029	0.017	-0.008
	(0.019)	(0.021)				(0.044)	(0.026)	(0.009)
ln GDP p.c.(t-1)	-0.004	0.000	-0.017*	-0.004	-0.004	-0.005	-0.006	-0.009
	(0.009)	(0.012)	(0.010)	(0.008)	(0.008)	(0.018)	(0.029)	(0.019)
Democ. (t-1)	-0.001	-0.001	-0.001	-0.001	-0.001	-0.002***	-0.001	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
ln Popul.(t-1)	0.008	0.011	0.021	0.010	0.011	0.004	-0.007	-0.007
	(0.012)	(0.016)	(0.015)	(0.011)	(0.011)	(0.026)	(0.028)	(0.024)
New State	0.011	0.000	0.014*	0.008	0.009	-0.003	0.011	-0.002
	(0.008)	(0.006)	(0.007)	(0.006)	(0.006)	(0.013)	(0.013)	(0.013)
War yrs. last 5y.	0.017*							
	(0.009)							
Peace duration		-0.000	-0.000	-0.000	-0.000	-0.001		-0.001
		(0.001)	(0.001)	(0.000)	(0.000)	(0.001)		(0.001)
Gold Producer		-0.024						
		(0.020)						
Diamond Producer		-0.039						
		(0.027)						
Oil Prod. pc. (t-1)			0.034					
			(0.029)					
Share area with oil				0.036				
				(0.065)				
Oil area / pop. (t-1)					0.416			
					(0.476)			
Civ. War Incid. (t-1)							0.556***	
							(0.051)	
Oil Exp. * Oil Gini								1.394***
								(0.071)
Sample	All countries					All countries without oil in 1964		
Model	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Country Fixed Eff.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4949	4017	4204	5372	5335	1918	2035	1918
R-squared	0.117	0.137	0.125	0.099	0.099	0.125	0.521	0.141

Notes: Dependent variable: Civil war onset (unless in column 7 where the dependent variable is conflict incidence). The dependent variable is coded as 1 if a conflict causing at least 1000 fatalities is starting in a given year. Sample period: 1960-2008. Number of countries for which observations are available: 157. In all columns unreported country fixed effects and annual time dummies are included, and robust standard errors are allowed to be clustered at the country level. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 3: Additional country level regressions of the effect of Oil Gini on conflict onsets

Dependent variable:	CW Onset	CW Onset	CW Onset	CW Onset	Secession.
	(1)	(2)	(3)	(4)	(5)
R1 / R	**0.067****	**0.064***	**0.087***	**-0.186**	**0.126***
	(0.032)	**(0.033)**	**(0.051)**	**(0.134)**	**(0.074)**
(R1/R)* Group. Conc.				**0.192****	
				(0.081)	
(R1/R) * Dist.				**0.330*****	
				(0.086)	
Gr.Pop./Gov.Pop.(t-1)	0.012***	0.012***	0.001	0.012	0.081
	(0.002)	(0.002)	(0.001)	(0.013)	(0.050)
Gov. Pop. (t-1)	0.004	0.004	-0.001	0.042	-0.251
	(0.007)	(0.007)	(0.006)	(0.057)	(0.250)
Group Peace Duration	0.001	0.001	0.001	0.001	
	(0.001)	(0.001)	(0.001)	(0.001)	
Group Diamond Mines					-0.025
					(0.124)
Group Co-Ethnics					-0.011
					(0.015)
Group Mount. Terr.					0.029
					(0.048)
ln Group Dist. Cap.					0.089**
					(0.039)
Group Soil Quality					0.074
					(0.076)
Group Pop. Dens.					0.522
					(0.399)
Share area with oil	-0.207				
	(0.160)				
Oil area / pop. (t-1)		0.226			
		(0.270)			
(R1/R)* Mount. Terr.				0.481	
				(0.308)	
(R1/R) * Soil Quality				-0.298	
				(0.226)	
(R1/R) * Gr.Pop./Gov.Pop.				0.003	
				(0.019)	
Sample	All non-gov. groups	All non-gov. groups	All groups never in gov.	All non-gov. groups	All non-gov. groups
Fixed effects	Gr. FE	Gr. FE	Gr. FE	Gr. FE	Cou. FE
Observations	28594	28594	27294	11828	10282
R-squared	0.169	0.168	0.157	0.130	0.873

Notes: OLS regressions in all columns. Dependent variable: Civil war onsets in columns 1-4, secessionism in column 5. Sample period: 1960-2006. Number of non-governing ethnic groups for which observations are available: 930. In all columns the unreported set of country level controls of Table 1 and annual time dummies are included, and robust standard errors are allowed to be clustered at the country level. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 4: Robustness checks of the ethnic group level regressions

Dep. Var.: Civil War	Onsets	Incid.	Onsets	Onsets	Onsets	Incid.
	(1)	(2)	(3)	(4)	(5)	(6)
R1 / R	**0.096****	**0.041**	**0.032**	**-0.035**	**-0.004**	**-0.019**
	(0.044)	**(0.034)**	**(0.023)**	**(0.026)**	**(0.008)**	**(0.014)**
(R1/R)* Group. Conc.			**0.184*****			
			(0.027)			
(R1/R) * Dist.				**0.173****		
				(0.066)		
Gr.Pop./Gov.Pop.(t-1)	0.010	0.004	0.015	0.010	-0.021	-0.052***
	(0.008)	(0.006)	(0.013)	(0.008)	(0.014)	(0.010)
Gov. Pop. (t-1)	-0.006	0.049	0.067	-0.006	-0.038	-0.118***
	(0.017)	(0.056)	(0.067)	(0.017)	(0.023)	(0.041)
Group Peace Duration	0.002		0.003**	0.002	0.000	
	(0.001)		(0.001)	(0.001)	(0.001)	
Group War Inc.(t-1)		0.845***				0.722***
		(0.032)				(0.064)
Sample	Non-gov. groups in countries without oil in 1964				Governing groups	
Model	OLS	OLS	OLS	OLS	OLS	OLS
Group fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9616	10533	3431	9616	8383	8538
R-squared	0.203	0.898	0.227	0.204	0.175	0.794

Notes: Dependent variable: Civil war onsets (unless when specified differently in row 1). Sample period: 1960-2006. Number of ethnic groups for which observations are available: max. 317 for columns 1-4; 273 for columns 5-6. In all columns the unreported set of country level controls from Table 1, group fixed effects and annual time dummies are included, and robust standard errors are allowed to be clustered at the country level. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 5: Additional Ethnic Group Level Regressions of the Effect of Oil Unevenness on Conflict Onsets

with recent data of the variable "fuel exports (in % of merchandise exports)" from World Bank (2012).

ln GDP per capita: PPP adjusted GDP per capita at constant prices. From the Penn World Tables (Heston, Summers, and Aten, 2011).

Democracy: Polity scores ranging from -10 (strongly autocratic) to +10 (strongly democratic). From Polity IV (2012).

ln Population: Total population. From World Bank (2012).

New State: Coded as 1 when a state was created in the previous 10 years, coded as 0 otherwise.

Ethnic Fractionalization: Index of ethnic fractionalization. From Fearon and Laitin (2003).

Mountainous Terrain: Percentage of territory covered by mountains. From Collier, Hoeffler and Rohner (2009).

Noncontiguous State: Dummy taking a value of 1 if a state has noncontiguous territory. From Fearon and Laitin (2003).

Peace duration: Years since last civil war incidence.

Conflict Incidence (t-1): First lag of the civil war incidence variable.

War years during last 5 years: Count of how many civil war incidence years during last 5 years, ranging from 0 to 5.

Gold Production Dummy: Takes a value of 1 when there is gold production in a country year, and 0 otherwise. From World Bank (2010).

Diamonds production dummy: Takes a value of 1 when there is diamond production in a country year, and 0 otherwise. From Lujala, Gleditsch, and Gilmore (2005).

Oil production per capita: Average amount per capita of oil extracted per day in a given year, measured in millions of barrels per day. From Humphreys (2005).

*Oil exporter * Oil Gini:* Interaction term of the variables Oil exporter and *Oil Gini* defined above.

Share area with oil: Using ArcGIS we have intersected the CShapes and PETRODATA datasets (that are discussed above in Section 3.1.1), which yields the share of the countries' area with oil.

Oil area / population (t-1): This corresponds to the lagged country area with oil (obtained through the intersection of CShapes and PETRODATA datasets) divided by the lagged country population from the World Bank (2012).

Group Level Variables

The two main dependent variables, civil war onset and incidence at the group level, and the main independent variable, R_1/R, have been described above in the main text. In what follows we describe the third dependent variable and all control variables.

Separatism: Dummy taking a value of 1 when the variable Sepx from Minorities at Risk (2009) takes a value of 2 or above.

Group Population / Governing Groups' Population: From Cederman, Buhaug and Rod (2009). Note that the population estimates of each ethnic group are not time-varying, but that the composition of groups controlling the government is time varying, which makes this measure time varying.

Governing Groups' Population: In 100 Millions. From Cederman, Buhaug and Rod (2009). The population estimates of each ethnic group are not time-varying, but that the composition of groups controlling the government is time varying, which makes this measure time varying.

Group Peace Duration: Years since last civil war incidence involving the given ethnic group.

Group Diamond Mines: Total number of diamond mines on the group territory (re-scaled in 100 mines). Constructed with GIS based on the group boundaries from the "Geo-referencing of ethnic groups" (GREG) dataset (Weidmann, Rod and Cederman, 2010) and the geo-referenced DIADATA dataset on the location of diamonds (from Gilmore et al., 2005).

Group co-ethnics abroad: Number of countries in which the same ethnic group also exists. Computed with GIS based on the group boundaries from the "Geo-referencing of ethnic groups" (GREG) dataset (Weidmann, Rod and Cederman, 2010).

Group's share of mountainous terrain: From Cederman, Buhaug and Rod (2009).

ln group's distance to capital: From Cederman, Buhaug and Rod (2009).

Group's soil quality: Part of the group's territory with high-quality fertile soil. Constructed based on the Harmonized World Soil Database (Fischer et al., 2008). Their complete global grid of nutrient availability is 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"). Our dummy takes a value of 1 for categories 1 and 2, categories 3 to 6 get value 0, and category 7 is set to missing.

Group Population Density: Average population density, re-scaled in 1000, constructed with GIS based on the on the group boundaries from the "Geo-referencing of ethnic groups" (GREG) dataset (Weidmann, Rod and Cederman, 2010) and the geo-referenced population density data from the Socioeconomic Data and Applications Center at Columbia University (Sedac, 2012).

Group War Incidence (t-1): First lag of the group level civil war incidence variable.

*(R1/R) * Group Concentration:* Interaction term between the (R1/R) variable and the following group concentration measure. The group concentration variable corresponds to the ratio of the "ethnic group's population living in areas where this ethnic group is the largest group" divided by the "total population of an ethnic group" in a given country. This time-invariant variable has been constructed using raw data from Alesina and Zhuravskaya (2011).

*(R1/R) * Distance:* Interaction term between the (R1/R) variable and a dummy variable taking a value of 1 if the value of the group distance from capital variable described above is larger than its median.

*(R1/R) * Mountainous terrain:* Interaction term between the (R1/R) variable and the group's share of mountainous terrain.

*(R1/R) * Soil quality:* Interaction term between the (R1/R) variable and the group's soil quality.

*(R1/R) * (Group population / Government population):* Interaction term between the (R1/R) variable and the group's population over government population.

Graphical Representation of the Oil Gini Variable

Below the distribution of country means and standard deviations over the period 1960-2008 across the world is represented graphically in the Figures 2 and 3. The classification of all countries into five groups was made applying the Jenks natural breaks classification method.

Country Level Variables	Obs	Avg	Sd	Min	Max
Conflict Onset	8071	0.015	0.123	0	1
Conflict Incidence	8494	0.050	0.219	0	1
Oil Gini	7983	0.167	0.258	0	0.977
Oil Exp. * Oil Gini	7464	0.038	0.145	0	0.888
Oil exporter (t-1)	7775	0.139	0.346	0	1
Gold Producer	6277	0.389	0.488	0	1
Diamond Producer	7921	0.186	0.389	0	1
Share area with oil	7983	0.092	0.168	0	0.928
Oil area / pop. (t-1)	6679	0.003	0.009	0	0.103
Oil Prod. pc. (t-1)	5237	0.045	0.270	0	4.923
ln GDP p.c.(t-1)	7074	8.249	1.287	4.767	11.722
Democ. (t-1)	7662	0	7.666	-10	10
ln Popul.(t-1)	7059	15.747	1.680	11.578	21.000
New State	8494	0.126	0.332	0	1
Ethnic Fraction.	7921	0.392	0.284	0.001	0.925
Mountainous Terr.	7559	0.176	0.209	0	0.943
Noncontig. State	7921	0.169	0.375	0	1
Peace duration	8428	29.452	18.781	0	62
Conflict Incid. (t-1)	8428	0.050	0.218	0	1
Conflict yrs. last 5y.	7613	0.257	0.920	0	5

Group Level Variables	Obs	Avg	Sd	Min	Max
Group Conflict Onset	63869	0.003	0.059	0	1
Group Conflict Incid.	65934	0.035	0.183	0	1
Group Separatism	19313	0.336	0.472	0	1
R1 / R	69596	0.065	0.210	0	1
Gr.Pop./Gov.Pop.(t-1)	55884	0.165	0.441	7.80E-08	8.918
Gov. Pop. (t-1)	55884	0.686	1.564	0.001	8.673
Group Peace Duration	63869	25.610	16.330	0	59
Group Diamond Mines	69596	0.006	0.061	0	1.790
Group Co-Ethnics	69596	2.781	2.329	1	15
Group Mount. Terr.	65934	0.369	0.352	0	1
ln Group Dist. Cap.	65934	6.154	1.090	1.609	8.782
Group Soil Quality	69596	0.699	0.333	0	1
Group Pop. Dens.	69596	0.069	0.153	0	4.354
(R1/R)* Group. Conc.	28987	0.059	0.210	0	1
(R1/R) * Dist.	65934	0.018	0.102	0	1
(R1/R) * Mount. Terr.	65934	0.018	0.077	0	0.909
(R1/R) * Soil Quality	69596	0.048	0.166	0	1
(R1/R) * Gr.Pop./Gov.Pop.	55884	0.045	0.236	0	8.918
Group Conflict Inc.(t-1)	65880	0.035	0.183	0	1

Table 6: Descriptive Statistics

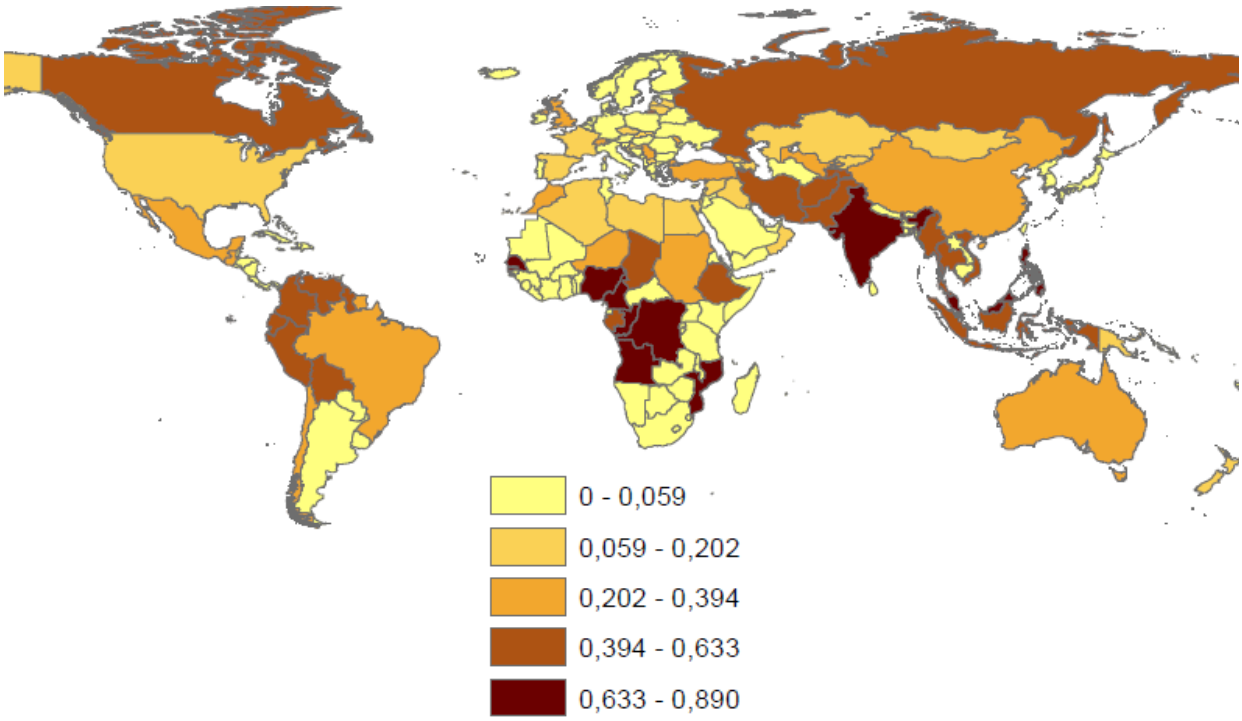

Figure 2: Oil Gini country means

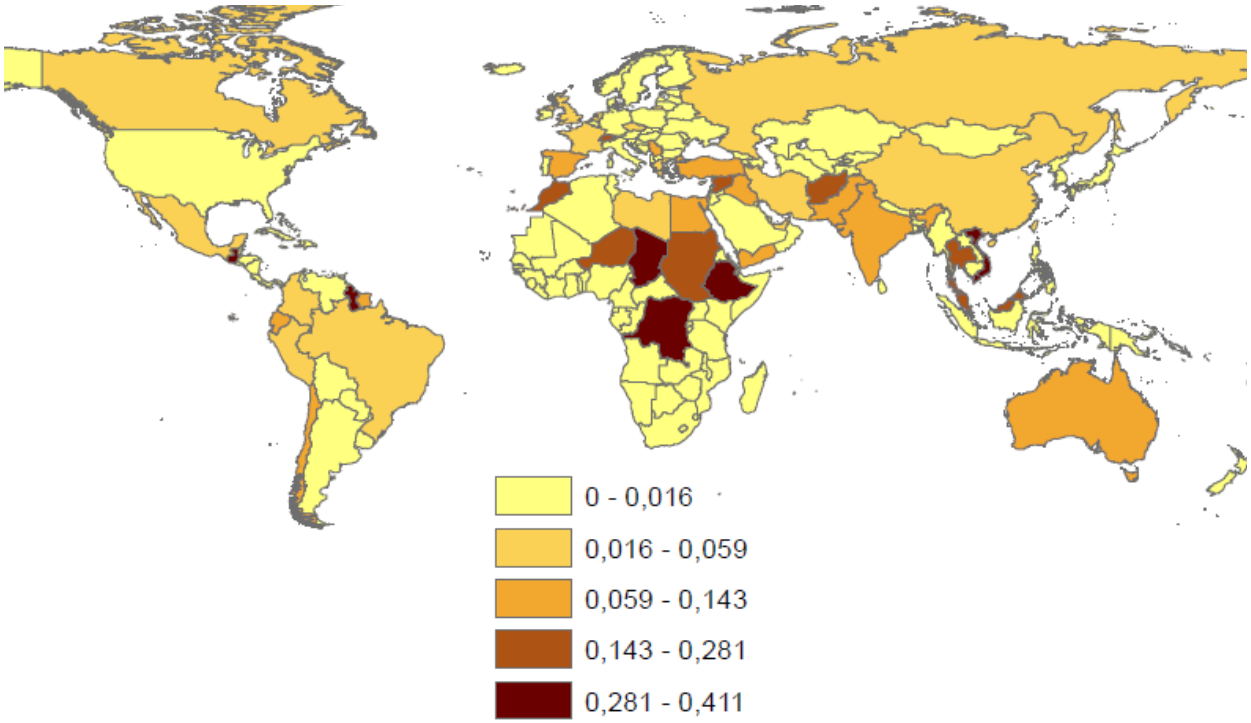

Figure 3: Oil Gini country standard deviations