



<http://www.diva-portal.org>

## Postprint

This is the accepted version of a paper published in *Journal of Peace Research*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Bara, C. (2014)

Incentives and Opportunities: A Complexity-oriented Explanation of Violent Ethnic Conflict.

*Journal of Peace Research*, 51(6): 696-710

<https://doi.org/10.1177/0022343314534458>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-331091>

**Incentives and opportunities:**  
**A complexity-oriented explanation of violent ethnic conflict**

Corinne Bara

Center for Security Studies, ETH Zurich

bara@sipo.gess.ethz.ch

CORINNE BARA, b. 1981. PhD candidate in political science, ETH Zurich (2011–); current research interests: ethnic conflict, conflict diffusion, transnational rebels, causal complexity.

**Acknowledgments**

I thank Seraina Rüegger and Philipp Hunziker for sharing their data with me early on in the production of this paper. Philipp also provided essential help with the predictions. I am indebted to Gudrun Østby and three anonymous referees, whose in-depth engagement with previous drafts have helped me take this paper to the next level. Thanks also to Andreas Wenger and the PhD group at the Center for Security Studies, and to Alrik Thiem, Sebastian Schutte, Carsten Q. Schneider, and the participants at the 2013 QCA Expert Seminar in Budapest.

## **Abstract**

*Existing research on the causes of violent ethnic conflict is characterized by an enduring debate on whether these conflicts are the result of deeply felt grievances, or rather the product of an opportunity structure in which rebellion is an attractive and/or viable option. This article argues that the question of whether incentive- or opportunity-based explanations of conflict have more explanatory power is fundamentally misguided, as conflict is more likely the result of a complex interaction of both. Fact is, however, that there is little generalized knowledge about these interactions. This study aims to fill this gap and applies crisp-set Qualitative Comparative Analysis (QCA) in order to identify constellations of risk factors that are conducive to ethnic conflict. The results demonstrate the explanatory leverage gained by taking causal complexity in the form of risk patterns into account. It takes no more than four different configurations of totally eight conditions to reliably explain almost two thirds of all ethnic conflict onsets between 1990 and 2009. Moreover, these four configurations are quasi-sufficient for onset, leading to conflict in 88% of all cases covered. The QCA model generated in this paper also fares well in predicting conflicts in-sample and out-of-sample, with the in-sample predictions being more precise than those generated by a simple binary logistic regression.*

## **Introduction**

The question of whether violent conflicts are the result of grievances, or rather the product of an environment in which rebellion is an attractive and/or viable option has divided scholars of intra-state conflict for decades. The debate was reignited by Collier & Hoeffler (2004), who claimed that rebellion cannot be explained by grievances resulting from ethnic animosities and economic and political inequalities, because situations in which people want to rebel are ubiquitous. Opportunity structures in which people are able to rebel, on the other hand, are considered sufficiently rare to constitute the explanation.

In this article on the causes of violent ethnic conflict, I argue that the competition between incentive- and opportunity-oriented explanations is misplaced altogether, because conflict is likely the result of both. This argument is not entirely new. Reflecting a growing unease with the either-or framing of the debate so far, it is now in vogue to state that conflict is the result of a complex interaction of incentives and opportunities (Ballentine & Sherman, 2003: 6; Korf, 2005: 201-202; Østby, 2008: 145; Sambanis, 2005: 329). Yet apart from this basic finding, we have little systematic knowledge about how they interact.

To fill this gap, this study aims to identify the various constellations of explanatory factors that are particularly conducive to the outbreak of ethnic conflict. I analyze global data on 500 five-year periods of ethnic conflict onset and non-onset on the level of ethnic groups. To identify recurring causal patterns and build a configurational model of ethnic conflict, I apply crisp-set Qualitative Comparative Analysis (csQCA), a method well-suited to detect complex causal relationships.

The most striking finding is that it takes no more than four different patterns to explain almost two thirds of all ethnic conflict onsets between 1990 and 2009. Using some of the field's well-known catchphrases I label them 'conflict trap,' 'bad neighborhood,' 'ousted rulers,' and 'resource curse.' The combinations of explanatory factors in these four patterns are quasi-sufficient for conflict, i.e., they lead to conflict in 88% of all cases covered. From a theoretical perspective, the results are largely in line with the importance accorded to the risk from previous conflict, neighborhood effects, political exclusion, or natural resources in the recent scholarship on conflict onset. What this study demonstrates powerfully, however, is the explanatory leverage we can gain if we take different combinations of those risk factors into account. The model generated in this paper also performs well in predicting conflicts, with in-sample predictions that are more precise than those generated by a simple binary logistic regression. Out-of-sample, both models predict very well and with comparable precision.

The article proceeds as follows: I first briefly recapitulate the debate on incentive- and opportunity-oriented explanations of conflict. Next, I describe my own analytical framework and introduce the explanatory conditions examined, followed by the research design. I then present the QCA results and offer a more substantive interpretation of each risk pattern before I conclude with an assessment of the predictive power of the model.

### **Incentives, opportunities, and conflict**

In the past three decades, ethnic conflict has become the prevalent type of civil war (Fearon & Laitin, 2011: 199). Reflecting the world's shock and outrage at the slaughtering of innocent men, women, and children in Bosnia, Rwanda, and elsewhere, these conflicts have drawn and continue to draw enormous scholarly interest. Scholars disagree, however, on whether ethnic identities are really an incentive for violence in such conflicts (as in the 'ancient hatreds' thesis, see Kaplan, 1993), or whether these identities are merely (re)created and instrumentalized by extremist leaders who sense an opportunity to come to — or hold on to — power (Gagnon, 2004; Snyder, 2000). A similar disagreement concerns the causal role of poverty and economic inequality in what is known as the 'greed-grievance' debate. In their famous article, Collier & Hoeffler (2004) rejected the popular argument that economic grievances are a powerful incentive for rebellion. They argued instead that rebel leaders merely employ a discourse of popular grievances to justify their violent strategy in order to profit from the war by looting or eventually controlling the resources of the state. Rejecting both ethnic antagonisms and economic inequalities as meaningful explanations of conflict, Fearon & Laitin (2003: 4) similarly argued that grievances 'fail to postdict civil war onset,' while measures of an opportunity structure that favors insurgency (like rough terrain or weak states) did fairly well.

These (stylized) debates are but two examples of a controversy that on a meta-level has run through ethnic conflict and civil war research like a golden thread: the incentive-opportunity debate. Ever since the exchanges in the 1970s between relative deprivation theorists (Davies, 1962; Gurr, 1970) and the resource mobilization school (Snyder & Tilly, 1972; Tilly, 1978), the controversy circles around the question whether conflict can really be understood by looking at the incentives for collective action, or whether we should rather examine the opportunity structure that makes collective action by ethnic groups possible. The morally charged phrasing in terms of 'greed' and 'grievance' in the latest manifestation of this controversy has certainly contributed to the unease with the either-or framing of the debate (Korf, 2005: 201-202). Increasingly, scholars claim that conflict is more likely the result of a complex interaction of both incentives and opportunities (Ballentine & Sherman, 2003: 6; Østby, 2008:

145; Sambanis, 2005: 329). Case studies usually highlight this complex interplay of risk factors in leading up to conflict, but even the occasional interaction terms tested in quantitative models of conflict (for example Brown, 2009; Østby et al., 2011) offer some preliminary evidence of complex causal relationships. Most research, however, has not systematically explored how incentives and opportunities interact, at least not in a comparative manner. This study aims to fill this gap.

### **Analytical framework and explanatory factors**

This paper undertakes a comprehensive analysis of risk patterns. The general assumption behind the configurational approach adopted here is that both incentives and opportunities are necessary for a conflict to start, because it seems common sense that a group has to be both willing and able to rebel (see also Starr, 1978: 375). Empirically, this assumption cannot be tested, lest we can claim that we know all possible incentives and opportunities for rebellion and have included them in our models. Moreover, the distinction between incentives and opportunities is far from clear-cut. For many variables commonly found to have an influence on conflict onset, it is not obvious whether they do so via an incentive or an opportunity mechanism. Natural resources, for example, may be a source of grievances if the population feels that the wealth from 'their' resources is siphoned away from the region while the population faces the negative externalities of the extraction process, such as environmental damage and displacement (Humphreys, 2005: 512; Tadjoeeddin, Suharyo & Mishra, 2001). They could, however, also offer an opportunity to finance a rebellion, because many resources can be either looted or used for extortion (Collier & Hoeffler, 2004: 565, 588). To complicate things, groups are not unitary actors: While some group members fight against an unfair distribution of resource wealth, others in the group may see a rebellion as an opportunity to accumulate private wealth from the control of resources (Lujala, Rod & Thieme, 2007: 240). With these ambiguities in mind, I have nevertheless aimed to select some 'typical' incentive and opportunity variables for the empirical analysis due to their theoretical and empirical importance in the ethnic conflict and civil war literatures (see Table I).

Table I. Explanatory conditions included in the analysis

<b>Condition</b>	<b>Assumptions on causal mechanism</b>
Political exclusion	Incentive
Ousted from rule	Incentive, possibly also opportunity
Ruling group	Absence of incentive
Oil and gas	Ambiguous, both incentive and opportunity
Previous conflict	Ambiguous, both incentive and opportunity
Tiny group	Lack of opportunity
Territorial concentration	Opportunity
Political instability	Opportunity
Extreme state poverty	Opportunity, possibly also incentive
Neighboring ethnic kin	Opportunity
Kin in conflict	Opportunity, may alter incentives

The *political exclusion* of an ethnic group from national-level decision-making is a typical incentive variable and has received much attention and empirical support within research on horizontal inequalities and conflict (Brown, 2009; Cederman, Wimmer & Min, 2010). Groups who have no say in government are lacking an important means to redress grievances and may not consider the state to be their legitimate representative. Exclusion is even more explosive if a group is *suddenly ousted from a position of power* (Cederman, Wimmer & Min, 2010). Such groups obviously have an incentive to regain the privileges once held. At the same time, their inside knowledge of the state, professional networks, and even state resources they may still partially control (Roessler, 2011) may offer them formidable opportunities to launch a rebellion against the new rulers. The condition *ruling group* was included to control for the fact that the group who holds most power in a state may not have an incentive to rebel at all, no matter what other risk factors are present at the same time. Unfortunately, no high-quality data was available on economic inequalities at the level of ethnic groups, hence this incentive remains unaccounted for in this study.<sup>1</sup>

The strong academic interest in the link between natural resources — especially oil — and conflict (Collier & Hoeffler, 2004; Humphreys, 2005; Lujala, Rod & Thieme, 2007; Ross, 2003) warrants the inclusion of the condition *oil or gas*, despite the above-mentioned ambiguousness with regard to its exact causal effect. *Previous conflict* is also usually included in quantitative models of conflict onset — if only to control for the temporal dependence of observations — but it is again not entirely clear how it

<sup>1</sup> The economic inequality measure available for EPR-ETH groups has a number of shortcomings that deter me from using this data: It is only available for spatially concentrated groups; the estimates for economic performance are influenced by local natural resources, thus underestimating inequality for groups resource-rich areas; and data quality varies considerably across countries (Cederman, Weidmann & Gleditsch, 2011: 483).

facilitates renewed onset. Previous conflict may have caused hurt, loss, and feelings of revenge, thus contributing to the incentives for renewed conflict; or it may have left a legacy of weapons and trained rebels that facilitate the organization of a rebellion (Collier & Hoeffler, 2004; Walter, 2004).

*Group size* and *territorial concentration* were included because they are typical for what Gurr (2000: 70) calls group capacity — and thus, opportunity — variables. Group size influences the resources a group can mobilize (Cederman, Wimmer & Min, 2010: 96), and tiny groups may not be able to gather enough financial and personal resources to challenge the state. The same applies to dispersed groups, who face coordination problems in organizing collective action, while territorial concentration positively influences a group's capacity for mobilization (Weidmann, 2009).

*Political instability* and *extreme state poverty* are included because they are two key aspects of state strength — a typical opportunity concept. Political instability as a temporary weakness signals to potential rebels that there is a vulnerability of the state to be exploited (Fearon & Laitin, 2003: 16). Extreme state poverty in terms of GDP per capita is both a cause and a result of bad administrative quality and weak state institutions, and reflects a chronic weakness of the state (Hendrix, 2010). Again, there are alternative mechanisms by which poverty could lead to conflict, such as being an incentive in itself, or via an opportunity cost mechanism: If income from regular employment is absent or low, joining the rebels may be an option to make a living (Collier & Hoeffler, 2004: 569).

The last two factors — *having ethnic kin* in a neighboring state, and having *neighboring kin that are in conflict* — are included to account for the international dimensions of 'internal' conflict. While there are various ways in which the neighborhood can influence the chance of war in another country, there is evidence that links stemming from transnational ethnic groups are particularly important (Buhaug & Gleditsch, 2008; Cederman, Girardin & Gleditsch, 2009: 409). Such groups may provide safe havens for rebels, and —especially if they are in a conflict themselves — can be a source of both inspiration and support in the form of weapons, finances, fighters, and even rebel leaders (Salehyan, 2009).

## **Method of analysis and research design**

With the 11 conditions described above, a process of pattern-finding — explained in this section — is employed in order to identify the multiple configurations of incentive and opportunity conditions that likely lead to ethnic conflict.



## Qualitative Comparative Analysis (QCA)

This study employs crisp-set QCA (csQCA), which was developed to permit valid generalization on complex causal relationships even with small to intermediate case numbers (Ragin, 2000, 2008).<sup>2</sup> In this study, however, the choice of QCA is not guided by the number of cases available for study, which is sufficient for using standard statistical techniques. Instead, QCA is applied because it can handle two aspects of causal complexity that are of core theoretical interest in this paper: conjunctural causation and equifinality. Conjunctural causation is a situation in which the effect of one explanatory factor depends on the presence or absence of other variables (Braumoeller, 2003: 4). Equifinality refers to the fact that there may be multiple paths to the same outcome, i.e., that conflict may be the result of different configurations of explanatory factors. In assuming conjunctural causation and equifinality, csQCA differs fundamentally from binary logistic regression, which is frequently used in onset studies but is founded on the assumption that each variable has an independent ('net') effect on the risk of conflict (Ragin, 2000: 95). While there are efforts to incorporate individual aspects of causal complexity into statistical methods such as regression analysis (see Schneider & Wagemann, 2012: 88), most of these attempts can handle just one aspect of complexity at a time.

CsQCA starts from an assumption of maximum complexity and lists *all* logically possible combinations of the conditions examined in a truth table, indicating for each row what proportion of the cases with this combination also have the outcome. The analyst selects only those configurations for further analysis that are sufficient for the outcome, i.e., in which all or at least most of the cases have a conflict. QCA then employs the Quine-McCluskey algorithm to discard all redundant information from the selected truth table rows (Schneider & Wagemann, 2012: 104). The result is a logically minimized solution, or in other words, causal complexity reduced to its most simple, valid expression.<sup>3</sup> Given the large-N character of this study, the aim is to identify quasi-sufficient rather than perfectly sufficient causal combinations, which means that most, but not necessarily all cases with a certain combination also have conflict (Ragin, 2000: 109-115). It is the task of the analyst to set a *consistency* threshold at which the proportion of cases that have the conditions *and* the outcome is considered high enough to warrant a statement of sufficiency. Consistency is not only relevant in the selection of truth table rows, but is also one of two important parameters of fit of the final QCA solution. The second parameter of fit is *coverage*, which reports how many onset cases are explained by the final QCA solution or the individual paths, because

---

<sup>2</sup> The choice of the crisp-set rather than the more sophisticated fuzzy-set variant of QCA is determined by the binary coding of the outcome. In QCA, this requires the dichotomization of all explanatory conditions, with any disadvantages that may entail.

<sup>3</sup> Readers not familiar with csQCA are referred to the introductory texts by Rihoux & De Meur (2009) or Grofman & Schneider (2009). The latter also offers a comparison of csQCA with binary logistic regression.

QCA does not make any statements about cases that do not exhibit one of the sufficient paths to conflict. In this sense, consistency fulfills a similar function (but is not the same!) as the parameters of significance in a regression analysis — indicating whether it is worth interpreting a causal relationship. Coverage, in turn, resembles the parameters of strength, i.e., correlation coefficients and total variance explained (Ragin, 2008: 45).

### **Outcome and sample population**

The outcome to be explained in this study is why some politically relevant ethnic groups experience the onset of ethnic conflict within a five-year period, while others do not. I use the *onset\_do\_flag* variable downloaded from the GROW<sup>up</sup> data portal to identify ethnic conflict onsets (Cederman, Wimmer & Min, 2010; Cunningham, Gleditsch & Salehyan, 2009; Gleditsch et al., 2002; Wucherpfennig et al., 2012).<sup>4</sup> An ethnic conflict is a conflict in which at least one rebel organization in an internal conflict (as defined by UCDP/PRI0, 2011: 9) explicitly or implicitly claims to represent this group in the conflict *AND* predominantly recruits fighters from the respective ethnic group (Wucherpfennig et al., 2012: 95). This strict definition of ‘ethnic’ has a caveat worth mentioning: Internal conflicts that do not meet the double requirement of ethnic claim *and* recruitment are coded as zero and are thus treated like cases that had no conflict at all. Note that this problem applies to most analyses using the EPR-ETH group-level data. The consequence is that we need to be careful in the interpretation of deviant cases — they may be deviant because they really had no conflict, or because they had conflict that was not coded as ethnic.

The coding for the *onset\_do\_flag* variable applies a two-year rule to collapse renewed episodes of a conflict that are within two years of the last episode into one single onset, assuming that a mere suspension of hostilities for a year does not mean that a conflict has ended in between. The analysis is limited to the time period 1990-2009, which accounts for the fact that the end of the Cold War fundamentally altered a number of conditions that states and potential rebels were facing on local, national, and regional levels, such as the dissolution of multiethnic empires, a proliferation of cheap weapons, or withdrawal of superpower support, to name just a few (Kalyvas & Balcells, 2010: 416).

With these rules, the dataset contains 102 onsets of ethnic conflict between 1990 and 2009.<sup>5</sup> To these onset periods, a random sample of just below 400 non-onsets was added, yielding a dataset of 500 observations. Random sampling allowed me to focus data collection and coding efforts on more

---

<sup>4</sup> <http://growup.ethz.ch/pfe/>.

<sup>5</sup> The ten onsets concerning the Caucasus Emirate in Russia in 2007 were merged into one conflict (peoples of the Caucasus against the government of Russia), as the ten groups are all rather small in (relative) size and are represented by the same rebel group.

interesting variables rather than more observations (King & Zeng, 2001: 137).<sup>6</sup> From a statisticians' viewpoint, both the random sampling and the study of five-year periods instead of group-years amounts to 'wasting data.' However, the number of conflict onsets to be studied is finite and analyzing group-years merely inflates datasets with zeroes (non-events), while most of the knowledge we gain about conflict is obviously gained from events. Moreover, studying five-year periods allows for flexible 'incubation periods,' i.e., the time it takes for a change in a condition to exhibit its conflict-triggering effect.

### **Measuring the explanatory conditions**

The following paragraphs detail the measurement of the explanatory conditions.<sup>7</sup> *Political exclusion (polx)* is measured with the *StatusID* variable from the EPR-ETH dataset (Cederman, Wimmer & Min, 2010) and takes on the value 1 if an ethnic group is excluded from central executive power in the majority of period-years.<sup>8</sup> A group was considered *ousted from rule (oust)* if it was excluded from central executive power in the course of a period.<sup>9</sup> Groups who retained their status as senior partners in government during all period-years were considered to be the *ruling group (ruler)*.

The condition *territorial concentration (conc)* takes on the value 1 if a group has a defined settlement pattern in the geo-coded version of EPR-ETH (GeoEPR-ETH) rather than being dispersed, migrant, or predominantly urban (Cederman, Wimmer & Min, 2010; Wucherpennig et al., 2011). The condition *petrol* measures whether there is at least one giant oil or gas field in a group's settlement area. It was obtained by combining GeoEPR-ETH data on groups' settlement areas with a georeferenced petroleum dataset by Lujala, Rod & Thieme (2007).<sup>10</sup> *Previous conflict (precon)* indicates whether a group already had an ethnic conflict as defined above within the past ten years. The condition *tiny group (tiny)* was coded 1 if a group makes up less than 1% of the total country population as reported in the EPR-ETH

---

<sup>6</sup> Given the explorative nature of this study, I do not test for statistical significance as would theoretically be possible in QCA, although rarely done (Ragin, 2000: 109-115). Technically, this does not allow any inferences beyond the sample studied. As Ward, Greenhill & Bakke (2010) write, however, statistical significance may not always be the best way to evaluate the 'real world' usefulness of a model, and the out-of-sample predictions reported further below offer an alternative heuristic to evaluate the model generated.

<sup>7</sup> The Online Appendix contains more details about the coding rules.

<sup>8</sup> Following Cederman, Weidmann & Gleditsch (2011: 484), ethnic groups that have absolute political power (EPR status 1=monopoly and 2=dominant) were dropped from the dataset. They may launch coups from within the government (see Roessler, 2011: 325), but these do not meet the conflict definition adopted here. Also, one may argue that groups with regional autonomy may be satisfied with their political status, especially in decentralized political systems. However, the results reported in the next section do not substantially change if these groups are coded as politically included (see Online Appendix).

<sup>9</sup> To avoid endogeneity, a qualitative check was performed to ensure that the group was ousted temporally before conflict onset, and not as a result of rebellion.

<sup>10</sup> Philipp Hunziker from the International Conflict Research group (ETH Zurich) kindly shared this data.

*GroupSize* variable (Cederman, Wimmer & Min, 2010) and has an absolute group population of less than one million people (CIA, 2013a).

On the country level, the condition *political instability (instab)* denotes whether there was a regime change, i.e., a substantial shift from democracy to autocracy or the other way within a group-period, as measured in the *Polity* variable of the Polity IV dataset (Marshall, Jaggers & Gurr, 2011).<sup>11</sup> *Extreme state poverty (xpoor)* indicates whether a country is among the lowest 10% of all countries with regard to real per capita GDP, with GDP data from the Penn World Tables' *rgdpch* variable (Heston, Summers & Aten, 2011), extrapolated for missing years using World Bank growth rates (World Bank, 2013).

Data on transnational ethnic kin (TEK) is from the International Conflict Research group at ETH Zurich (Cederman et al., 2013). The condition *havtek* is coded 1 if an ethnic group has a kin group in a country that is connected to its host country by a land border (CIA, 2013b). The condition *tekcon* is coded 1 if such a group has an ongoing ethnic conflict in any year of the group-period and if the settlement areas of the two groups are adjoining.

## Results

This section presents the results of the comparative analysis, followed by a discussion of the risk patterns identified.

### A configurational model of ethnic conflict

The first step in a QCA is the analysis of necessary conditions. We should be satisfied to find a quasi-necessary rather than fully necessary condition. To this end, I set a consistency threshold of 0.95, which means that at least 95% of all conflict cases should exhibit a necessary condition, with five deviant cases allowed. Only two single conditions fulfill this criterion:  $\sim$ *tiny* and  $\sim$ *ruler*, with the tilde indicating the absence of this condition. This means that with very few exceptions, ethnic conflict should only happen if an ethnic group is not tiny and not the senior partner in government, and indeed, only five very small groups have staged a rebellion, and only three ruling groups have done so. However, with very low coverage scores these are trivial necessary conditions, i.e., they are too common in the sample to be of much substantive interest (Schneider & Wagemann, 2012: 144-147, 233-237).

The sufficiency test moves away from looking at single conditions, and aims to identify configurations of conditions that are quasi-sufficient for conflict onset. The consistency threshold to include a truth table

---

<sup>11</sup> Again, given endogeneity concerns with the Polity IV dataset in the context of civil war research (see Vreeland, 2008), a qualitative check was performed to ensure that changes in the polity score were not a *result* of rebellion.

row into the minimization process is set at 0.7, which is a bit lower than the 0.75 consistency threshold often recommended (Ragin, 2008: 46; Schneider & Wagemann, 2012: 212), but seems justified given the large number of cases and the fact that some problems connected to lower thresholds apply only to fuzzy-set QCA (Schneider & Wagemann, 2012: 238-244).

As was to be expected, using all 11 conditions discussed in the previous sections in one single model yields a complex, unwieldy solution that is difficult to interpret. It explains 62% of all onsets with a very high consistency of 0.97, but it does so by identifying many different paths, most of which only explain very few conflicts.<sup>12</sup> Conditions were subsequently dropped from this full model to find a model that was both parsimonious, i.e., had a limited number of paths that could explain a *group* of onsets each and at the same time had acceptable consistency and coverage scores. Dropped were the conditions territorial concentration (*conc*), extreme state poverty (*xpoor*), and having transnational ethnic kin (*havtek*). These may still be important risk factors, but they did not contribute to a better explanation in terms of consistency and coverage and split the solution up into many paths that rendered a meaningful interpretation difficult.

Consisting of only eight conditions, the solution presented in Table II offers the best combination of consistency, coverage, and parsimony.<sup>13</sup> The solution consistency is 0.88, and with a coverage of 0.60 it explains almost two thirds of all conflicts (61 out of 102) in an elegant solution of only four paths.

---

<sup>12</sup> The R code to replicate that solution is provided in the supplementary materials to this paper.

<sup>13</sup> Most QCA software packages report three solutions (conservative, parsimonious, and intermediate), which differ with regard to the assumptions they allow about logical remainders, i.e., combinations of conditions for which no cases exist. I prefer the intermediate solution, which allows for the inclusion of easy counterfactuals (Schneider & Wagemann, 2012: 165-177). The R code in the supplementary materials to this paper permits a replication of all three solutions.

Table II. Preferred QCA solution (intermediate solution term)

	Solution/ configuration consistency	Solution coverage	Configuration raw coverage*	Configuration unique coverage*
<b>Model:</b>				
polx * oust * ruler * petrol * precon * tiny * instab * tekcon → onset				
(frequency cutoff: 1.00 / consistency cutoff: 0.70)				
<b>Model parameters:</b>	<b>0.88</b>	<b>0.60</b>		
~tiny*precon*polx*~ruler	0.93		0.40	0.37
tekcon*instab*~tiny*~petrol*~ruler	0.77		0.10	0.10
instab*~tiny*oust*~polx*~ruler	0.83		0.05	0.05
~tekcon*instab*~tiny*petrol*polx*~ruler	0.89		0.08	0.05

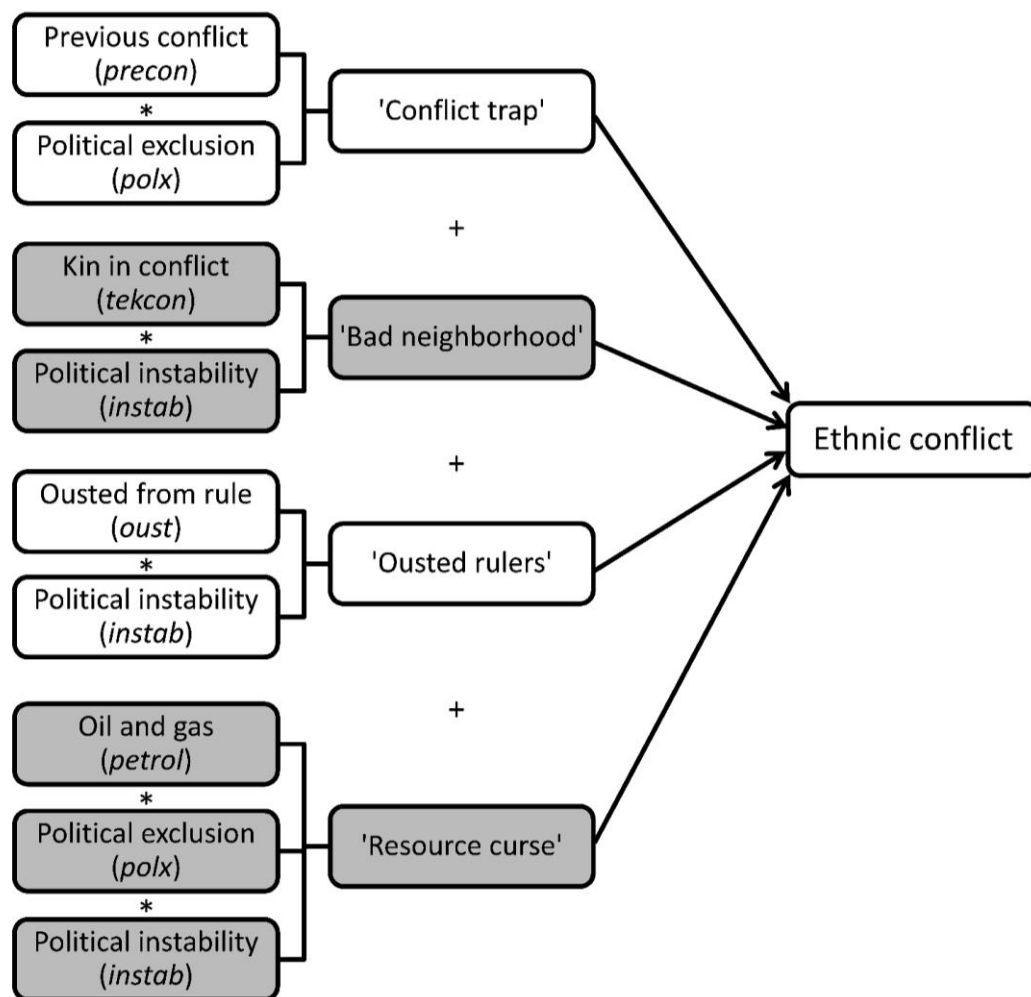
\* Raw coverage includes cases explained by more than one configuration, while unique coverage includes only cases exclusively covered by that configuration.

Not surprisingly, all four quasi-sufficient configurations contain the two trivial necessary conditions identified above. This does not make them any less trivial *per se*, but in those four paths they are important prerequisites for the other conditions to have their strong joint effect.

Some of the paths in Table II contain conditions that should contribute to conflict in their absence, namely *~petrol* (no oil and gas in a group's settlement area) in the second path, *~polx* (no political exclusion) in the third path, and *~tekcon* (no transnational ethnic kin in conflict) in the fourth path. This is not unusual in QCA solutions given that the effect of any condition is assumed to be dependent on the presence or absence of other conditions. In this case, however, these conditions do not seem to make sense theoretically, and a brief analysis confirms that they are not needed for a valid interpretation of the solution, because the respective paths are quasi-sufficient without them.<sup>14</sup>

Ignoring these 'absent conditions' and the two trivial necessary conditions from the QCA solution, Figure 1 captures the structure of the argument contained in Table II.

<sup>14</sup> For a brief explanation, see Online Appendix.



**Figure 1.** Configurational model of ethnic conflict, 1990-2009. Adapted from Goertz and Mahoney (2005). \* = logical AND (conjunction of conditions); + = logical OR (substitutable/equifinal paths); → = quasi-sufficient relationship.

There are four quasi-sufficient paths to conflict, which I labelled using some of the field’s well-known catchphrases: ‘Conflict trap’ for the conjunction of previous conflict and political exclusion, ‘bad neighborhood’ for a combination of ethnic kin in conflict and political instability at home, ‘ousted rulers’ for groups that are ousted from a position of power in a situation of political instability, and ‘resource curse’ for the conjunction of oil and gas reserves, political exclusion, and instability. The following paragraphs explain and discuss these four paths in more detail, with Table III listing all the conflicts that are explained by either of them as well as deviant cases.

Table III. Cases explained per quasi-sufficient configuration

Onsets explained by configuration	Deviant cases, no onset
<b>Configuration 1: Conflict trap</b>	
Catholics in Northern Ireland, UK (1995-98)	Basques in Spain (1997-2001)
Basques in Spain (1988-91)	Afar in Ethiopia (2000-04)
Chechens in Russia (1997-99)	Papua in Indonesia (1990-94)
South Ossetians in Georgia (2000-04; 2005-08)	
Armenians in Azerbaijan (1991; 2001-05)	
Tuareg in Niger (1995-97)	
Lari/Bakongo in the Rep. Congo (2000-02)	
Hutu in Rwanda (2005-09)	
Somali (Ogaden) in Ethiopia (1990-94; 1997-99)	
Bakongo in Angola (1992-94*; 1999-2002)	
Cabindan Mayombe in Angola (1992-94*; 1999-2002; 2005-07)	
Kurds in Iran (1991-93; 1994-96; 2001-05)	
Kurds in Iraq (1993-95; 2000-04)	
Shi'a Arabs in Iraq (1988-1991)	
Palestinian Arabs in Israel (1997-2000)	
Bodo in India (1991-93; 2005-09)	
Indigenous Tripuri in India (1989-92)	
Manipuri in India (1989-92; 2001-03)	
Naga in India (1998-2000; 2001-05)	
Mohajirs in Pakistan (1991-95)	
Kayin (Karens) in Myanmar (1993-95)	
Mons in Myanmar (1994-96)	
Muslim Arakanese in Myanmar (1987-91)	
Shan in Myanmar (1989-93; 2003-05)	
Moro in the Philippines (1991-93)	
Achinese in Indonesia (1995-99)*	
East Timorese in Indonesia (1990-92; 1993-97)	
<b>Configuration 2: Bad neighborhood</b>	
Serbs in Croatia (1991-92)	Croats in Slovenia (1991-95)
Serbs in Bosnia and Herzegovina (1992)	Banyarwanda in Uganda (1990-94)
Croats in Bosnia (1992-93)	Baloch in Afghanistan (2004-08)
Tuareg in Niger (1987-91)	
Bakongo in the DRC (2003-07)	
Tutsi-Banyamulenge in the DRC (1992-96)	
Hutu in Rwanda (1993-97)	
Afar in Djibouti (1998-99)	
Afar in Ethiopia (1992-96)	
Baloch in Iran (2002-06)	
<b>Configuration 3: Ousted rulers</b>	
Lari/Bakongo in Congo (1994-98)	Russians in Kazakhstan (1991-95)
Sunni Arabs in Iraq (2000-04)	
Tajiks in Afghanistan (1992-96)	
Uzbeks in Afghanistan (1992-96)	
Mohajirs in Pakistan (1986-90)	



#### **Configuration 4: Resource curse**

Azeri in Russia (1986-90)

Indigenous peoples of the Amazon in Peru (1993-97)

Chechens in Russia (1991-94)

Bakongo in Angola (1987-91; 1992-94\*)

Cabindan Mayombe in Angola (1987-91; 1992-94\*)

Pashtuns in Afghanistan (1991-95)

Achinese in Indonesia (1995-99)\*

---

\* This case is not uniquely covered by that specific configuration, i.e., two configurations explain this case.

#### **Conflict trap**

The first path to conflict is via the recurrence of a previous conflict: If a (non-tiny, non-ruling) ethnic group already had an ethnic conflict in the past ten years *and* is still politically excluded, conflict breaks out with a high consistency (0.93). With a raw coverage of 0.40, this combination explains the highest number of conflicts covered by the total model (41 onsets).<sup>15</sup> This finding is congruent with the central argument of a recent book on conflict recurrence by Call (2012: 4), who argues that political exclusion is the crucial variable in explaining most cases of civil war recurrence. It also corroborates Walter's (2004: 372, 385) finding that besides the improvement of basic living conditions, access to central political decision-making significantly decreases the risk of conflict recurrence. The result as such cannot answer the question of whether the high risk of renewed conflict is due to a grievance or an opportunity effect of the previous conflict. What it shows, however, is that the combination of a clear *current* grievance (political exclusion) with a situation in which a previous conflict may have left both emotional scars and a legacy of conflict-specific capital (opportunity) poses a threat. The conflict that started in 2005 between the Kurdish PJAK (The Free Life Party of Kurdistan) and the Iranian government is a case in point. Since 1946, but in particular since the Iranian Revolution in 1979, Kurdish opposition forces have repeatedly challenged the state in order to create an autonomous Kurdistan and put an end to the discriminatory and assimilatory policies of the regime. The conflict was last active in 1996, after which Mohammad Khatami's presidency introduced at least some cultural and political freedoms for the Iranian Kurds, although they were still discriminated and politically excluded. With the election of Mahmoud Ahmadinejad as president in 2005, Kurdish hopes for reform were crushed, giving way to renewed conflict (Stansfield, Lowe & Ahmadzadeh, 2007: 6-7).

---

<sup>15</sup> A robustness test was conducted with onsets only included in the dataset after three and four years of peace, respectively. Although the results above are confirmed, the coverage of the QCA solution does decrease, indicating that in some cases we may be studying the continuation of an existing conflict rather than a 'real' conflict recurrence. The supplementary materials for this paper contain the two datasets to replicate this test.

### **Bad neighborhood**

The second configuration can be summarized as a situation of instability both at home and in the neighborhood. Ethnic groups who have warring ethnic kin across the border are likely to rebel themselves *if* the government at home is at the same time vulnerable because of regime change (consistency 0.77). With a coverage of 0.10, this configuration explains ten onsets in my sample. The finding is fully in line with recent research that has demonstrated an increased risk both for neighbors of a country in conflict (Buhaug & Gleditsch, 2008), and for ethnic groups who have kin groups across the border (Cederman, Girardin & Gleditsch, 2009; Salehyan, 2009). The fact that this condition is not equally dangerous for stable governments supports Buhaug & Gleditsch (2008: 230) who find that the risk of conflicts spilling over is the highest when the 'host' state already has a high baseline risk for conflict due to domestic characteristics. Braithwaite (2010) similarly finds evidence that state capacity modifies the risk of conflict contagion. Typical for this configuration is the rebellion by the Tutsi-Banyamulenge in the Democratic Republic of the Congo (DRC). The genocide of Tutsi and the subsequent change in power relations in neighboring Rwanda had a tremendous impact on the Tutsi in the DRC — a country that was already at the brink of anarchy when president Mobutu lost crucial support from his Western allies by 1996 (Prunier, 2009: 78-79). This is the only configuration that includes no unambiguous incentive condition, but in which an extraordinary opportunity structure seems to be sufficient for onset. This does not imply that there were no mass grievances, but they are not captured by the model.

### **Ousted rulers**

The third configuration describes rebellions by groups who were recently excluded from central government power in a situation accompanied by political instability. Typical are the Sunni Arabs in Iraq in 2004, who lost the political advantages they enjoyed under Saddam Hussein's regime when he was ousted in a US-led invasion of Iraq. With a coverage of 0.05, this configuration explains five conflict onsets in my sample at a consistency of 0.83. It supports research by Cederman, Wimmer & Min (2010: 104) who find that groups whose power status decreased during the previous two years are much more likely to rebel. They argue that anger and resentment is especially strong after a group loses power and prestige, especially when this anger can be directed at the ethnic group that is considered guilty of the ousting. Gurr (2000: 108) also posits that advantaged groups are at a special risk of being the target of reprisals and revenge once displaced from power, giving them an incentive to fight back. Again, in this configuration incentives (loss of power and privileges) and opportunities (political instability, resources still available to the ousted group) coincide to increase the risk of conflict onset.

### **Resource curse**

The fourth path to conflict is when oil-rich but politically excluded groups can make use of the window of opportunity offered by political instability at the center. With a raw coverage of 0.08 this path explains eight conflicts at a consistency of 0.89. The finding supports the view that at the heart of both ethnic and non-ethnic conflicts is frequently a dispute about the control over natural resources. Oil and gas in particular are relatively unlootable commodities and as such only offer a benefit to those who have direct control over it, i.e., the extraction firm and the government (Ross, 2003: 55-56). This makes the political exclusion condition in this path so salient. However, this lack of control over resources affects many groups who still do not rebel violently against the exploitation of their lands. It needs an extraordinary opportunity offered by the rupture of regime change to make the combination of natural resources and political exclusion quasi-sufficient for conflict. The movement by the Bakongo and Cabindan Mayombe for the independence of the Angolan enclave Cabinda is a typical case: Cabinda accounts for more than half of Angola's oil production, yet neither the political power nor the economic welfare of the two groups have been positively influenced by these riches (le Billon, 2001). At the same time, the presence of oil reserves may have influenced the strategic calculus of the rebels in Angola, fueling beliefs that 'going it alone' could be feasible and an independent Cabinda potentially prosperous (Humphreys, 2005: 511). When the instability caused by the country's transition to multi-party democracy offered a window of opportunity in the early 1990s, the simmering conflict escalated.

### **Predicting conflict with QCA?**

The results reported above and the rich theoretical interpretations they give rise to lend some legitimacy to the ontological assumption made at the outset of this paper, namely that the causal relationships involved in ethnic conflict onset are complex. Nevertheless, it is the nature of ontological assumptions that we cannot empirically test whether they are right or wrong (Hay, 2008: 87-88). We can, however, assess the degree to which a model based on such assumptions is *useful* for policy purposes by testing its capacity to actually *predict* conflicts. Predictive capacity can also serve as a criterion for comparing the usefulness of competing models across methodological divides. To this end, I compare the predictive power of the QCA model developed in this paper with a binary logistic regression model that is founded on very different ontological assumptions.

Equation (1) shows the specification of the binary logit model used for this comparison, where  $p$  is the predicted probability of onset contingent on the values of the independent variables.<sup>16</sup>

$p(\textit{onset}) =$

$$\textit{Logit}(\beta_0 + \beta_1*\textit{polx} + \beta_2*\textit{oust} + \beta_3*\textit{tiny} + \beta_4*\textit{ruler} + \beta_5*\textit{precon} + \beta_6*\textit{petrol} + \beta_7*\textit{instab} + \beta_8*\textit{tekcon}) \quad (1)$$

This is a very naïve model with a linear predictor that is purely additive in the regressors, and more complex causal relationships could be modelled using interaction terms. The purpose here, however, was not to build a statistical model that mimics the type of causal relationships analyzed with QCA, but to compare two model specifications that are on the extreme ends of causal complexity as defined further above.

I use the two models to predict conflicts both in-sample and out-of-sample, using the same data. Out-of-sample prediction is the ability to predict conflicts outside the dataset that was used to fit a model in the first instance, often in a different time period. Because most data used in this paper are not yet updated for the time period 2010-2013, I instead refit both the QCA and regression analyses for the time period 1990-2004 in order to assess how well the resulting models predict the onsets that happened in the last period, 2005-2009. The QCA solution for 1990-2004 is almost identical to the one for the full period: The same four paths are quasi-sufficient for conflict onset, with comparable consistency and coverage levels.<sup>17</sup>

A good model in terms of predictive power correctly predicts as many onsets as possible (true positives) and at the same time makes few mistakes in the form of false positives, i.e., predicting conflict where there was none (Ward, Greenhill & Bakke, 2010). Accordingly, two measures of predictive power are reported in Table IV: *Sensitivity* is the fraction of all onsets that are correctly anticipated, while *precision* takes false positives into account by reporting the percentage of onset predictions that are correct.

Making predictions based on the QCA model is straightforward: Conflict is predicted for all cases that exhibit any of the quasi-sufficient paths to conflict.<sup>18</sup> In order to make point predictions based on the

---

<sup>16</sup> The logit parameter estimates are of no substantive interest here, but are reported in the Online Appendix. Note that in order to permit a fair comparison of the two models and because the goal is not interpretation but prediction, one key assumption in binary logistic regression — that observations are independent of each other — is violated given that the model does not adjust for spatial and temporal autocorrelation.

<sup>17</sup> Solution reported in the Online Appendix.

<sup>18</sup> Note, however, that because QCA assumes asymmetric causation in the form of necessary and sufficient conditions, the QCA analyst would not predict all remaining cases to have peace. To predict non-onset, the procedure in QCA is to conduct a separate analysis of the absence of the outcome, reflecting the fact that in

logit model, however, we have to define a threshold above which the predicted probabilities (which run from zero to one) are deemed high enough to predict an onset. Because this choice of threshold is arbitrary and because predicted probabilities are difficult to compare across different models, scholars comparing the predictive capacities of statistical models prefer to make use of Receiver Operator Characteristic (ROC) curves, which plot the true positive rate against the false positive rate for all possible thresholds (Ward, Greenhill & Bakke, 2010: 4). QCA results do not easily lend themselves to comparison by means of a ROC curve, hence the two models' precision is instead compared at two different sensitivity scores: The first, termed QCA1 and Logit1, is determined by the consistency threshold set to achieve a good QCA result. Because this may not be the sensitivity level at which the logit model predicts 'best,' the second comparison, termed QCA2 and Logit2, is made at the logit model's optimum true to false positive ratio, assuming that we value an additional true positive at equal value as an additional false positive.<sup>19</sup> To then achieve the same true positive rate for QCA2, I chose from the QCA truth table just those rows for minimization with the highest row consistencies until the number of onsets that Logit2 correctly predicts were covered by the solution.

Table IV shows that the QCA model has a considerably better in-sample predictive capacity than the logit model, even at the sensitivity level at which the logit model performs best. The QCA solution correctly predicts 61 out of total 102 onsets. At this true positive rate, QCA makes eight mistakes, i.e., predicts conflict for eight cases that did not experience an onset of ethnic conflict, while the logit model wrongly predicts 16 conflicts that did not happen. Even at the logit model's optimum sensitivity, at which both models correctly identify 47 out of 102 onsets, the logit model still produces twice as many false positives as the QCA solution (six instead of only three).

---

asymmetric causal relationships, the causes of an outcome and its opposite are not mirror images (Schneider & Wagemann, 2012: 112). The Online Appendix reports the results of this 'peace' analysis.

<sup>19</sup> This optimum threshold is where a diagonal of slope  $m=1$  touches the upper-left corner of the ROC curve.

Table IV. Predictive capacities of QCA and binary logit models

<b><i>In-sample prediction (1990-2009)</i></b>			
<i>N=500 (102 onsets, 398 non-onsets)</i>			
<b>Model:</b>	<b>Sensitivity (true positive rate)</b>	<b>False positives:</b>	<b>Precision:</b>
QCA1	61/102 (0.60)	8	0.88
Logit1	62/102 (0.61)	16	0.79
Logit2	47/102 (0.46)	6	0.89
QCA2	47/102 (0.46)	3	0.94
<b><i>Out-of-sample predictions for 2005-2009 based on 1990-2004</i></b>			
<i>N=106 (14 onsets, 92 non-onsets)</i>			
<b>Model:</b>	<b>Sensitivity (true positive rate)</b>	<b>False positives:</b>	<b>Precision:</b>
QCA1	10/14 (0.71)	1	0.91
Logit1	11/14 (0.79)	2	0.85
Logit2	9/14 (0.64)	0	1.0
QCA2	8/14 (0.57)	0	1.0

While in-sample predictive power is a useful indicator for the validity of a model, out-of-sample prediction is an even more powerful evaluative tool, especially for policy purposes (Ward, Greenhill & Bakke, 2010). Both the QCA and logit models perform extremely well in the out-of-sample test, with their predictive capacities not differing much. The standard QCA solution (QCA1) correctly identifies 10 out of 14 onsets, and yields only one false positive. Because the logit model identifies two onsets at the same threshold of predicted probabilities, it correctly predicts 11 onsets, but produces an additional false positive. As I have chosen to give the same weight to true and false positives, the predictive capacity of both models is almost the same here. At its optimum, the logit model can predict nine out of 14 onsets without making a single mistake. Achieving this 100% precision with the QCA model requires a consistency cut-off of 0.75 in the estimation data truth table, resulting in a five-path solution that correctly predicts eight onsets in the period 2005-2009.

To sum up, the QCA model fares equally well as the logit model in the out-of-sample prediction, and considerably better in the in-sample prediction.<sup>20</sup> Given the low number of onsets in the test sample, the out-of-sample results have to be treated with caution and should be re-tested once data on further time

<sup>20</sup> The inclusion of certain interaction terms improves the predictive capacity of the logit model, but does not surpass the predictive capacity of the QCA model reported here. The best logit model in terms of predictions was — not surprisingly — the one that included the key interactions of all four paths identified with QCA (see Online Appendix).

periods become available, but these preliminary results suggest that the assumptions of causal complexity at the heart of this paper are warranted.

## Conclusion

This article set out to enrich — and hopefully overcome — the incentive-opportunity debate by exploring how incentives and opportunities combine to give way to ethnic conflict. The patterns that were identified suggest that it may be time to abandon the either-or framing of the debate in favor of a more inclusive approach. The ‘resource curse’ pattern in particular is a textbook example of incentives and opportunities coinciding at a certain point in time to facilitate violent uprising: The ethnic groups in question had a reason to rebel (grievances induced by political exclusion and possibly by the oil and gas resources on their territory), and did so when a window of opportunity opened up through political instability at the center. At the same time, there are clear limits to the interpretation of risk patterns in terms of incentives and opportunities in a macro-level study like the current one, and assessing the causal mechanisms by which explanatory factors really contribute to conflict risk would require more in-depth case analyses. What this study has undoubtedly demonstrated, however, is that a complexity-oriented approach to the explanation of ethnic conflict is fruitful both for explaining and predicting conflict onset.

QCA is an outcome-oriented method, i.e., it is targeted at finding explanations for outcomes rather than identifying average effects of causes (Ragin, 2000: 32-33, 39). For policy purposes, this feature of QCA has the advantage that results directly correspond to actual outcomes of individual cases, which permits the scholar to easily communicate research findings to policy-makers. The real added value of QCA for conflict studies, however, is the ability to identify *multiple* paths to conflict, for even if conventional statistical models can incorporate more complex relationships using interaction terms, they do not help us identify these relationships in the first place, and certainly do they not easily lend themselves to the identification of substitutable (equifinal) paths to conflict.

Policy-relevant is also the quasi-sufficiency of the risk patterns identified: That a specific combination of risk factors leads to conflict most of the time is powerful knowledge. The price of this confidence about the consistent effect of *some* causal patterns is that we can say nothing at all about the 41 conflict onsets not covered by the model.<sup>21</sup> Omitted explanatory factors are most probably responsible for this lack of coverage, especially on the incentive side. 16 out of the 41 non-covered cases are coded as politically included. A brief look at them suffices, however, to see that their political inclusion is either just a ‘token

---

<sup>21</sup> Onsets not explained by the QCA solution are listed in the Online Appendix.

inclusion,' for example in the transition to multiparty democracy, or that even the full inclusion of some political leaders of an ethnic group may not be able to offset the pervasive feeling of economic disadvantages and cultural discrimination of the masses — both conditions not accounted for in this study. Future research should expand this configurational model of ethnic conflict and try to cover more conflicts not currently explained.

More generally, however, further research should capitalize on the added value of a complexity-oriented approach. Methodological avenues to be explored are those that are suitable for the type of pattern-seeking employed here, such as cluster analysis (Cooper & Glaesser, 2011), or methods that avoid strong parametric assumptions, such as Kernel Regularized Least Squares (KRLS) (Hainmueller & Hazlett, 2014) or neural network models (Beck, King & Zeng, 2000). After all, this paper has demonstrated how evaluating the predictive capacity of different models may be a way to compare our empirical results even if we do not use the same methodological approaches, thus facilitating communication across methodological boundaries.

### **Statistics package**

All results in this article were generated using the R software, and in particular the QCA package for R, v1.0-5 (Thiem & Duşa, 2013).

### **Data replication**

Datasets, code, and supplementary documentation are available from <http://www.prio.no/jpr/datasets>.



## References

Ballentine, Karen & Jake Sherman (2003) Introduction. In: Karen Ballentine & Jake Sherman (eds) *The Political Economy of Armed Conflict: Beyond Greed and Grievance*. Boulder, CO: Lynne Rienner Publishers, 1-15.

Beck, Nathaniel; Gary King & Langche Zeng (2000) Improving quantitative studies of international conflict: A conjecture. *American Political Science Review* 94(1): 21-35.

Braithwaite, Alex (2010) Resisting infection: How state capacity conditions conflict contagion. *Journal of Peace Research* 47(3): 311-319.

Braumoeller, Bear F (2003) Causal complexity and the study of politics. *Political Analysis* 11(3): 209-233.

Brown, Graham K (2009) Regional autonomy, spatial disparity and ethnoregional protest in contemporary democracies: A panel data analysis, 1985–2003. *Ethnopolitics* 8(1): 47-66.

Buhaug, Halvard & Kristian Skrede Gleditsch (2008) Contagion or confusion? Why conflicts cluster in space. *International Studies Quarterly* 52(2): 215-233.

Call, Charles T (2012) *Why Peace Fails: The Causes and Prevention of Civil War Recurrence*. Washington, DC: Georgetown University Press.

Cederman, Lars-Erik; Luc Girardin & Kristian Skrede Gleditsch (2009) Ethnonationalist triads: Assessing the influence of kin groups on civil wars. *World Politics* 61(3): 403-437.

Cederman, Lars-Erik; Kristian Skrede Gleditsch, Idean Salehyan & Julian Wucherpfennig (2013) Transborder ethnic kin and civil war. *International Organization* 67(02): 389-410.

Cederman, Lars-Erik; Nils B Weidmann & Kristian Skrede Gleditsch (2011) Horizontal inequalities and ethno-nationalist civil war: A global comparison. *American Political Science Review* 105(3): 478-495.

Cederman, Lars-Erik; Andreas Wimmer & Brian Min (2010) Why do ethnic groups rebel? New data and analysis. *World Politics* 62(1): 87-119.

Central Intelligence Agency (CIA) (2013a) *CIA World Factbook Country Comparison: Population* (<https://www.cia.gov/library/publications/the-world-factbook/rankorder/2119rank.html>).

Central Intelligence Agency (CIA) (2013b) *CIA World Factbook Field Listing: Land Boundaries* (<https://www.cia.gov/library/publications/the-world-factbook/fields/2096.html>).

Collier, Paul & Anke Hoeffler (2004) Greed and grievance in civil war. *Oxford Economic Papers* 56(4): 563-595.

Cooper, Barry & Judith Glaesser (2011) Using case-based approaches to analyse large datasets: A comparison of Ragin's fsQCA and Fuzzy Cluster Analysis. *International Journal of Social Research Methodology* 14(1): 31-48.

Cunningham, David E; Kristian Skrede Gleditsch & Idean Salehyan (2009) It takes two: A dyadic analysis of civil war duration and outcome. *Journal of Conflict Resolution* 53(4): 570-597.

Davies, James Chowning (1962) Toward a theory of revolution. *American Sociological Review* 27(1): 5-19.

Fearon, James D & David D Laitin (2003) Ethnicity, insurgency, and civil war. *American Political Science Review* 97(1): 1-16.

Fearon, James D & David D Laitin (2011) Sons of the soil, migrants, and civil war. *World Development* 39(2): 199-211.

Gagnon, Valère Philip Jr. (2004) *The Myth of Ethnic War: Serbia and Croatia in the 1990s*. Ithaca, NY: Cornell University Press.

Gleditsch, Nils Petter; Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg & Havard Strand (2002) Armed conflict 1946-2001: A new dataset. *Journal of Peace Research* 39(5): 615-637.

Goertz, Gary & James Mahoney (2005) Two-level theories and fuzzy set analysis. *Sociological Methods & Research* 33(4): 497-538.

Grofman, Bernard & Carsten Q Schneider (2009) An introduction to crisp set QCA, with a comparison to binary logistic regression. *Political Research Quarterly* 62(4): 662-672.

Gurr, Ted Robert (1970) *Why Men Rebel*. Princeton, NJ: Princeton University Press.

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

Hainmueller, Jens & Chad Hazlett (2014) Kernel regularized least squares: Reducing misspecification bias with a flexible and interpretable machine learning approach. *Political Analysis* 22(2): 143-168.

Hay, Colin (2008) Political ontology. In: Robert E. Goodin & Charles Tilly (eds) *The Oxford Handbook of Contextual Political Analysis*. Oxford: Oxford University Press, 78-96.

Hendrix, Cullen S (2010) Measuring state capacity: Theoretical and empirical implications for the study of civil conflict. *Journal of Peace Research* 47(3): 273-285.

Heston, Alan; Robert Summers & Bettina Aten (2011) *Penn World Table Version 7.0*. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania ([https://pwt.sas.upenn.edu/php\\_site/pwt\\_index.php](https://pwt.sas.upenn.edu/php_site/pwt_index.php)).

Humphreys, Macartan (2005) Natural resources, conflict, and conflict resolution: Uncovering the mechanisms. *Journal of Conflict Resolution* 49(4): 508-537.

Kalyvas, Stathis N & Laia Balcells (2010) International system and technologies of rebellion: How the Cold War shaped internal conflict. *American Political Science Review* 104(3): 415-429.

Kaplan, Robert D (1993) *Balkan Ghosts: A Journey Through History*. New York: St. Martin's Press.

King, Gary & Langche Zeng (2001) Logistic regression in rare events data. *Political Analysis* 9(2): 137-163.

Korf, Benedikt (2005) Rethinking the greed–grievance nexus: Property rights and the political economy of war in Sri Lanka. *Journal of Peace Research* 42(2): 201-217.

le Billon, Philippe (2001) Angola's political economy of war: The role of oil and diamonds, 1975–2000. *African Affairs* 100(398): 55-80.

Lujala, Päivi; Jan Ketil Rod & Nadja Thieme (2007) Fighting over oil: Introducing a new dataset. *Conflict Management and Peace Science* 24(3): 239-256.

Marshall, Monty G; Keith Jagers & Ted Gurr (2011) *Polity IV Project*. Center for Systemic Peace (<http://www.systemicpeace.org/inscr/inscr.htm>).

Østby, Gudrun (2008) Polarization, horizontal inequalities and violent civil conflict. *Journal of Peace Research* 45(2): 143-162.

Østby, Gudrun; Henrik Urdal, Zulfan Mohammed Tadjoeeddin, S Mansoob Murshed & Havard Strand (2011) Population pressure, horizontal inequality and political violence: A disaggregated study of Indonesian provinces, 1990–2003. *Journal of Development Studies* 47(3): 377-398.

Prunier, Gérard (2009) *Africa's World War: Congo, the Rwandan Genocide, and the Making of a Continental Catastrophe*. Oxford: Oxford University Press.

Ragin, Charles (2000) *Fuzzy-Set Social Science*. Chicago, IL: University of Chicago Press.

Ragin, Charles (2008) *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago, IL: University of Chicago Press.

Rihoux, Benoît & Gisèle De Meur (2009) Crisp-set Qualitative Comparative Analysis (csQCA). In: Benoît Rihoux & Charles Ragin (eds) *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. Thousand Oaks, CA: SAGE Publications, 33-68.

Roessler, Philip (2011) The enemy within: Personal rule, coups, and civil war in Africa. *World Politics* 63(2): 300-346.

Ross, Michael L (2003) Oil, drugs, and diamonds: The varying roles of natural resources in civil war. In: Karen Ballentine & Jake Sherman (eds) *The Political Economy of Armed Conflict: Beyond Greed and Grievance*. Boulder, CO: Lynne Rienner Publishers, 47-70.

Salehyan, Idean (2009) *Rebels Without Borders: Transnational Insurgencies in World Politics*. Ithaca, NY: Cornell University Press.

Sambanis, Nicholas (2005) Conclusion: Using case studies to refine and expand the theory of civil war. In: Paul Collier & Nicholas Sambanis (eds) *Understanding Civil War: Evidence and Analysis*. Washington, DC: The World Bank, 303-334.

Schneider, Carsten Q & Claudius Wagemann (2012) *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis*. Cambridge: Cambridge University Press.

Snyder, David & Charles Tilly (1972) Hardship and collective violence in France: 1830 to 1960. *American Sociological Review* 37(5): 520-532.

Snyder, Jack L (2000) *From Voting to Violence*. New York: W.W. Norton.

Stansfield, Gareth; Robert Lowe & Hashem Ahmadzadeh (2007) The Kurdish policy imperative. *Chatham House Briefing Paper* (December 2007)  
(<http://www.chathamhouse.org/publications/papers/view/108733#node-108733>).

Starr, Harvey (1978) 'Opportunity' and 'willingness' as ordering concepts in the study of war. *International Interactions* 4(4): 363-387.

Tadjoeddin, Mohammad Zulfan; Widjajanti I Suharyo & Satish Mishra (2001) Regional disparity and vertical conflict in Indonesia. *Journal of the Asia Pacific Economy* 6(3): 283-304.

Thiem, Alrik & Adrian Duşa (2013) *Qualitative Comparative Analysis with R: A User's Guide*. New York: Springer.

Tilly, Charles (1978) *From Mobilization to Revolution*. Reading, MA: Addison-Wesley.

UCDP/PRIO (2011) *UCDP/PRIO Armed Conflict Dataset Codebook, Version 4-2011*. Uppsala Conflict Data Program (UCDP), Centre for the Study of Civil Wars, International Peace Research Institute (PRIO) (<http://www.pcr.uu.se/research/ucdp/datasets/>).

Vreeland, James Raymond (2008) The effect of political regime on civil war. *Journal of Conflict Resolution* 52(3): 401-425.

Walter, Barbara (2004) Does conflict beget conflict? Explaining recurring civil war. *Journal of Peace Research* 41(3): 371-388.

Ward, Michael D; Brian D Greenhill & Kristin M Bakke (2010) The perils of policy by p-value: Predicting civil conflicts. *Journal of Peace Research* 47(4): 363-375.

Weidmann, Nils B (2009) Geography as motivation and opportunity: Group concentration and ethnic conflict. *Journal of Conflict Resolution* 53(4): 526-543.

World Bank (2013) *GDP per capita growth (annual %)* (<http://data.worldbank.org/indicator/NY.GDP.PCAP.KD.ZG>).

Wucherpfennig, Julian; Nils W Metternich, Lars-Erik Cederman & Kristian Skrede Gleditsch (2012) Ethnicity, the state and the duration of civil war. *World Politics* 64(1): 79-115.

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