

# Language, Religion, and Ethnic Civil War\*

Nils-Christian Bormann  
Lars-Erik Cederman  
Manuel Vogt<sup>†</sup>

Word count: 10,538

## Abstract

Are certain ethnic cleavages more conflict-prone than others? While only few scholars focus on the contents of ethnicity, most of those who do argue that political violence is more likely to occur along religious divisions than linguistic ones. We challenge this claim by proposing the path from linguistic differences to ethnic civil war along three theoretical steps: (1) the perception of grievances by group members, (2) rebel mobilization, and (3) government accommodation of rebel demands. Our argument is tested with a new dataset of ethnic cleavages that records multiple linguistic and religious segments for ethnic groups from 1946 to 2009. Adopting a relational perspective, we assess ethnic differences between potential challengers and the politically dominant group in each country. Our findings indicate that intrastate conflict is more likely within linguistic dyads than among religious ones. Moreover, we find no support for the thesis that Muslim groups are particularly conflict-prone.

---

\*We thank Michael Hechter, David Siroky, Nicholas Sambanis, Nils B. Weidmann, Patrick M. Kuhn, Seraina Rüegger, Sofia Vera, Philipp Hunziker, panelists at the 2013 Swiss Political Science Association's Annual Meeting and the 2013 Annual Meeting of American Political Science Association, seminar participants at the University of Amsterdam, Arizona State University, and the University of Konstanz, and two anonymous reviewers for helpful comments. Replication material will be available on the journal's website upon publication. We gratefully acknowledge financial support by the Swiss National Science Foundation (Grants 105511-143213 & 400240-147210) and the Swiss Agency for Development and Cooperation. All remaining mistakes are our own.

<sup>†</sup>Center for Comparative and International Studies, ETH Zurich, Switzerland. Corresponding Author: Nils-Christian Bormann, Email: bormann@icr.gess.ethz.ch

Whether or not arguing that ethnic grievances cause civil war, conflict researchers typically adopt a formal, encompassing approach to ethnicity that draws no distinction between its key dimensions, such as language, religion, caste, or race. The vast majority of scholars thus follows Barth (1969) who defines ethnicity in terms of boundaries or cleavages rather than in terms of its specific contents (cf. Brubaker, 2009). Departing from this scholarly practice, Juergensmeyer (1993), Laitin (2000) and others have singled out religion as being more conflict-prone than other ethnic dimensions (for a summary, see Gorski and Türkmen-Derivoğlu, 2013). In the aftermath of the terrorist attacks on September 11, 2001, this hypothesis has gained considerable currency beyond academia, as illustrated by Tony Blair’s (2014) recent assertion that “religious difference will fuel this century’s battles.” Rebellions by organizations such as al-Qaida or ISIS seemingly confirm fears of “civilizational clashes” between major world religions (see Huntington, 1996).

Researchers who claim that religion is the most violence-prone ethnic dimension do so for different reasons. Some argue that religious violence results from disappointment with failed secular state-building projects (Juergensmeyer, 1993; Philpott, 2002). Others claim that, since religious conflicts concern indivisible goods, religiously distinct groups are less likely to compromise and more likely to use political violence (Reynal-Querol, 2002; Hassner, 2009). Finally, some expect that religious groups are less hampered by collective action problems that impede the mobilization of rebels compared to language groups where individual group members are assumed to prefer linguistic assimilation to fighting (Laitin, 2000). In agreement with Huntington’s civilizational thesis (1996), some of these scholars also insist that among the world religions, it is Islam that is most conflict-prone (e.g., Toft, 2007).

We challenge the claim that religion is more salient in ethnic civil wars than language for two reasons. First, proponents of the religion-conflict link typically focus on isolated theoretical mechanisms without embedding them in an overarching theoretical model of civil conflict outbreak.<sup>1</sup> Second, many of these empirical studies do not compare religious differences to other ethnic dimensions. In response, we adopt

a more comprehensive theoretical framework that allows us to evaluate the conflict-fueling force of both religious and linguistic cleavages along a chain of mechanisms that link ethnic divisions with civil conflict. This theoretical framework allows us to identify a logic of mobilization for linguistically distinct ethnic groups that has been overlooked by the proponents of the religion-conflict link.

Our analysis relies on the new *Ethnic Power Relations – Ethnic Dimensions* (EPR-ED) dataset, which provides detailed information about groups' ethnic setup. Thanks to their linguistic and anthropological roots, these data offer a more exogenous starting point of inquiry than ex-post codings of ethnic mobilization. Moreover, by extending the geographic and temporal coverage, the EPR-ED dataset allows us to examine a much larger sample of countries than has hitherto been possible. In fact, the current dataset covers all politically relevant groups around the world from 1946 through 2009. Featuring several linguistic and religious segments per group, it refrains from defining members of a given group as being *either* linguistic *or* religious, thereby avoiding ex-post facto assessments of the relevant cleavage in a given ethnic conflict. In addition, coding multiple linguistic and religious segments per group enables us to capture the cross-cutting nature of cleavages between groups. Based on these data, we adopt a relational approach that explicitly assesses groups' ethnic differences in relation to the politically predominant ethnic group for each country-year.

Our empirical findings lend no support to the claim that religious differences are responsible for more outbreaks of ethnic civil war than linguistic cleavages. In fact, for the period since World War II, language is more clearly associated with ethnic conflict than religion. Nor is there any indication that Muslim identities are overrepresented in our conflict statistics. Without discounting religion as an important source of conflict in the contemporary world, we conclude that recent scholarly claims that religion is the most important conflict dimension have been overstated.

The next section introduces our theoretical framework that links cleavages to conflict, and compares the conflict-potential of both linguistic and religious divisions. We then present the EPR-ED dataset before testing our argument and discussing the im-

plications of our results for future research.

## Ethnic cleavages and civil war

Departing from the Barthian view of ethnicity as a formal and modular boundary marker, recent conflict studies focus on specific ethnic dimensions, in particular religion (Reynal-Querol, 2002; Fox, 2004; Toft, 2007). Claiming that religious cleavages are more conflict-prone than other ethnic markers, proponents of the religion-conflict link usually build their theoretical argument upon one theoretical mechanism. Laitin (2000), for example, compares religious and linguistic groups in their ability to overcome the collective action dilemma while Toft (2007) highlights the indivisibility of religious claims.

These arguments by themselves are important, but do not go far enough in explaining the link between specific ethnic markers and intrastate conflict, particularly if they are only applied to religion but not to other ethnic dimensions (Juergensmeyer, 1993; Toft, 2007; Hassner, 2009).<sup>2</sup> Therefore, we compare the conflict-inducing effect of linguistic and religious cleavages directly, by analyzing three stages in the process that leads from categorical inequalities to civil war.

More specifically, we build on classical theories of ethnic conflict that link categorical inequalities between ethnic groups to grievances among disadvantaged group members (Gurr et al., 1993; Petersen, 2002; Stewart, 2008; Cederman, Weidmann and Gleditsch, 2011; Cederman, Gleditsch and Buhaug, 2013). Grievances as a motive for conflict by themselves, however, do not go far enough in explaining the onset of civil war. As Tilly (1978) notes, individual mechanisms such as the perception of grievances by group members are insufficient to guarantee the outbreak of conflict (see also Gurr, 2000). Thus, we also focus on the organizational resources that are necessary to overcome the collective action problem among group members and to recruit rebels (Tilly, 1978). Finally, we consider the effect of repressive and accommodating government reactions on the risk of civil war outbreak .

Are religious cleavages indeed more conflict-prone than linguistic ones? We evaluate

this question by systematically examining three stages from inequalities to conflict onset for both religious and linguistic ethnic divisions (cf. Gurr et al., 1993; Cederman, Gleditsch and Buhaug, 2013). First, how do group-level inequalities along religious and linguistic dimensions come to be perceived as collective grievances? Second, how are these two types of ethnic groups able to overcome the collective action dilemma? And third, what are governments' options to accommodate these ethnic challengers? By inspecting the entire pathway to intrastate conflict onset, our theoretical approach makes it possible to reveal arguments for the connection between linguistic cleavages and civil war that have been overlooked by the proponents of the religion-conflict link, leading us to conclude that linguistic differences are at least as conflict-prone as religious ones.

## **Perception of grievances**

We begin by discussing the transformation of inter-ethnic inequalities into collective grievances that provide the motivation for civil war. Juergensmeyer (1993) was among the first to argue that religion re-emerges as a basis for collectively felt grievances due to disappointment with failed projects of "secular nationalism." Predicting a "Clash of Civilizations" between major world religions, Huntington (1993, 26) expects that "animosities stretching . . . back deep into history" such as the Western crusades will replace other motives for conflict such as ethno-linguistic desires for self-determination. Huntington and others also argue that grievances in response to a disappointing secular order are particularly acute in the Islamic world which is why Muslims are more frequently involved in intrastate conflict than members of other religions (see, e.g., Toft, 2007; Fox, 2012).

While there may be some truth to these claims in specific cases, it is important to realize that grievances are not necessarily expressed in religious terms even where religious differences exist. For example, long-lasting intrastate conflicts in the Philippines and Thailand are often interpreted as clashes between Muslim minorities and other dominant religions but are in fact fundamentally about ethno-linguistic self-determination

as Liow (2006) shows. In particular where violent conflicts are interpreted as being only about one issue, it is easy to lose sight of alternative conflict issues (see Kalyvas, 2003).<sup>3</sup>

Indeed, collective grievances are at least as likely to emerge along linguistic cleavages. The most important reason is that individuals cannot escape the pervasiveness of language in both politics and everyday life. According to Brubaker (2013, 5–6), language “is an inescapable medium of public discourse, government, administration, law, courts, education, media and public signage.” Because states simply cannot be neutral with respect to language, speakers of minority languages are often disadvantaged when dealing with state agents. Even under federal and cultural autonomy arrangements, desires for self-determination among linguistically distinct groups might persist as in Quebec or the Basque country. Moreover, in many less developed countries, jobs in the bureaucracy are one of the very few reliable sources of economic well-being, and access to state employment is frequently restricted by linguistic discrimination (Lewis, 1992; Wimmer, 1997; Bayart, 2009). This is not to say that individuals who do not speak the dominant language are incapable of learning it. However, they usually cannot master it at the same level of proficiency as native-speakers (Hyltenstam and Abrahamsson, 2000). Unlike religious affiliation, which is typically less visible in personal encounters, even minor linguistic distinctions are easily noticeable and often carry important symbolic meaning, especially where national identity is linguistically defined. In these cases, Gellner’s (1983, 58–62) fictitious example of Ruritania highlights how the perception of being second-class citizen among speakers of a minority language spurs their desire for self-determination. Likewise Horowitz (2000, 222) articulates the economic and psychological logic of linguistic grievances elegantly:

[Language] links political claims to ownership with psychological demands for the affirmation of group worth, and it ties this aggregate matter of group status to outright careerism, thereby binding elite material interests to mass concerns. . . Religion can sometimes link elite and mass concerns. But clearly it is more plausible to think that fluency in the working language

of a bureaucracy will become an issue more often than will facility in scripture. Language is so often an issue not because there is something special about linguistic differences or something mystical and therefore conflict-producing about the process of communication. The reason is more mundane: in a multilingual state, there has to be a language policy to decide trivial and important issues of linguistic choice.

In sum, collective grievances may emerge along linguistic and religious cleavages. However, the omnipresent pervasiveness of language in daily life is likely to make language the more prevalent source of perceived injustice.

## **Rebel mobilization**

Following our framework from inequalities to conflict further, grievances alone do not cause rebellions. For this to happen, rebel leaders need to overcome the collective action dilemma and mobilize fighters (Tilly, 1978; Lichbach, 1998). Adopting an instrumentalist logic, Laitin (2000) argues that individuals who belong to distinct religious groups are more likely to overcome collective action problems than members of language groups who can opt for linguistic assimilation instead of risking their lives by joining a rebellion (see also Laitin, 1998).<sup>4</sup> In contrast, it is more difficult for members of religious groups to convert to other creeds. Moreover, religious groups can use their pre-existing organizational resources to mobilize fighters (Stewart, 2009). Based on these arguments, religious elites can be expected to find it easier to set up and maintain rebel organizations than leaders of language groups.

However, we argue that linguistic differences are at least as difficult to overcome as religious ones. Darden (2013) shows that mass schooling instills lasting loyalties in linguistically defined groups that “defy the [instrumental] incentives presented in a given situation” and motivate extreme personal sacrifices (ibid., p.5). More generally, Anderson (2006) highlights how mass media, such as newspapers, create “imagined communities” that constitute the basis for nationalist mobilization. Conversely, the absence of a unifying language can block mobilization for violence. Therefore, where

religious groups are linguistically divided, their potential for collective action might well be smaller. Safran and Liu (2012, 273) argue that the very preservation of religious group identities depends on the use of a language that is distinct from the dominant language in a state. The fact that Sunni Kurdish fighters are at the forefront of fighting the Sunni Arab ISIS rebellion in Syria and Iraq is a clear example. Although many religious groups have their own sacred language, such as Latin or classical Arabic, these are rarely connected to individuals' everyday lives. In contrast, vernacular languages are more likely to link masses and elites (Lange and Dawson, 2010).

Moving from the masses to the elite-level, language-based networks often rival the role of religious organizations in violent mobilization. The functional equivalent of religious schools and organizations is the university campus, which served as an important breeding ground for ethno-nationalist movements in the former Soviet Union (Roeder, 1991), Africa (Reno, 2011, 37–75), and 19th century Europe (Hroch, 1985). Elite networks created at universities, especially if combined with mass schooling in the relevant language, offer powerful channels for spreading the message of self-determination among members of an ethnic group. Thus, on balance, there is no reason to believe that political entrepreneurs who express religious grievances enjoy a monopoly on mobilization for violent collective action.

## **Government reaction**

Even in the presence of widespread grievances and mobilization, the occurrence of intrastate war depends on the government's decision to either escalate the conflict or to accommodate the demands of the ethnically defined opposition. With regard to this last step, theorists of the religion-conflict link suggest that governments find it harder to satisfy the demands of ethno-religious challengers because these conflicts concern indivisible goods, such as the primacy of one god over another (Toft, 2007, 101). Along these lines, Hassner (2009) argues that "sacred sites" claimed by multiple religious actors, such as the Temple Mount in Jerusalem, cannot be shared and are therefore constant sources of violent conflict. Toft (2007) also suggests that the problem



of indivisibility is particularly pronounced in the Islamic world where religion and politics are not separated, and civil wars should therefore be especially likely in these countries.

While it may indeed be difficult for governments or state elites to find common ground with religious challengers, it is important to note that states can always declare religious neutrality by delegating religious matters to the individual or by recognizing multiple religions at the same time. As Goddard (2006) notes, indivisibility is itself socially constructed. In contrast, the state has to choose at least one working language, and introducing multilingualism is far more expensive since it involves the training of teachers and administrators in different languages and extensive translation of public documents. Unsurprisingly, many states refrain from offering language concessions on cost grounds. Liu's (2011, 130) global analysis of language rights in democracies reveals that there is any recognition of minority languages in only about half of all cases. Since her assessment does not distinguish between groups within states, it is likely that an even larger number of languages are legally unprotected.

The admittedly extreme example of the European Union vividly illustrates the economic costs of a multi-linguistic regime. In the European Parliament translation and interpretation services make up one third of the administrative staff and account for half of the parliamentary budget (Corbett, Jacobs and Shakleton, 2011, 43). If it is "prohibitively expensive for the EU to provide interpretation services for every meeting amongst civil servants," as Patten and Kymlicka (2003, 9) argue, it is impossible for poorer multiethnic states. Whereas issue indivisibility might hinder compromise between distinct religious groups, economic considerations make it at least as difficult for governments to accommodate the demands of language groups.

Concluding our theoretical discussion, we note that both cleavage dimensions may be linked to collective grievances, offer the potential for rebel mobilization, and pose obstacles to the accommodation of violent challengers. Therefore, we reject the idea of a specific conflict-fueling effect of religion, but expect ethnic civil conflicts along linguistic lines to be at least as likely as those between religiously distinct groups.

Empirically, most studies define conflicts as either religious or not religious (see, e.g., Fox, 2002, 2004; Svensson, 2007; Toft, 2007). Yet, this obscures the potential interaction of linguistic and religious cleavages in ethnic civil war onset (see, e.g., Selway, 2010). In the next section, we introduce a new dataset that allows us to directly compare linguistic and religious cleavages, and their interaction without pre-determining the relevance of either for any particular ethnic group or conflict.

## Concepts and data

To test our claim that linguistic distinctions are at least as conflict-prone as religious cleavages, we introduce the new *Ethnic Dimensions* (EPR-ED) dataset. EPR-ED unpacks the “cultural stuff” (Barth, 1969, 15) of the ethnic groups listed in the Ethnic Power Relations (EPR-ETH) dataset (Cederman, Wimmer and Min, 2010; Cederman, Gleditsch and Buhaug, 2013). Our data identify multiple ethnic segments on the linguistic and religious dimension for each of the more than 700 ethnic groups in over 130 states where ethnicity is politically relevant in national politics. Ethnic groups are politically relevant when group representatives make political claims on their behalf or when then state discriminates against these groups politically. In the following, we first introduce the coding of the linguistic and religious dimensions of ethnic groups and then discuss the construction of our relational cleavage variables for group-dyads.

### Multiple ethnic dimensions

EPR-ED introduces two key innovations: first, it codes religious and linguistic segments for each group rather than defining groups as either linguistic or religious. This multi-dimensional assessment enables us to avoid relying on actor claims that might be endogenous to conflict, and also makes it possible to test arguments about cross-cutting or reinforcing cleavage constellations. Second, our dataset codes multiple linguistic and religious segments per group rather than identifying only one. Thereby we measure overlap on the linguistic and religious dimensions between groups, which yields more

precise estimates than dichotomous indicators of group differences. Figure 1 illustrates the coding with regards to two Nigerian groups. While belonging to one religion, the Hausa-Fulani are divided linguistically into three segments. Conversely, a linguistically homogeneous group such as the Yoruba may count adherents from multiple religions such as Islam and Christianity among its members. We refer to these internal-group divisions as *ethnic segments*. The EPR-ED dataset identifies up to three linguistic and religious segments for each ethnic group along with their relative demographic size.<sup>5</sup> The Yoruba’s largest religious segment includes 43% of all group members.

Figure 1 about here.

The data on linguistic segments build on the well-known Ethnologue database, which uses mutual intelligibility as its main criterion for identifying languages (Lewis, 2009).<sup>6</sup> Data on religious segments derive from the Joshua Project, which connects the language groups of Ethnologue with government statistics, and a variety of sources on religious adherence, yielding a list of “people’s groups and their religious make-up” (see *Joshua Project: Unreached Peoples of the World*, 2011).

Figure 2 shows a partial example of one of the Ethnologue language families in tree-form where higher-order nodes indicate shared heritage between the language leaves. The leaves on the lowest level represent spoken languages, which correspond to our linguistic segments. The example shows four languages spoken in Nigeria that belong to the Niger-Congo language family. Yoruba is located on the bottom-left, while the two smaller Hausa-Fulani language segments, Adamawa and Nigerian Fulfulde, can be found at the bottom of the middle branch. The Hausa segment is not part of the Niger-Congo family but belongs to the Afro-Asiatic language tree. Some Ethnologue language families such as Austronesian include more than 1200 nodes and up to 16 different levels.

Figure 2 about here.

While the linguistic hierarchies from Ethnologue are available online, equivalent

trees for religious families had to be constructed. Figure 3 displays a subset of the Muslim tree including the Sunni-Muslim (Maliki) segment shared by the Yoruba and Hausa-Fulani. Resembling their linguistic counterparts in most respects, religious trees have considerably fewer levels. While the top level of each tree captures one of the major world religions, the sub-divisions describe denominations or major religious traditions. The maximum number of levels in religious trees is found in the Muslim tree (4) while the largest number of denominations is found in the Christian tree (33).

Figure 3 about here.

The dataset codes ethnic segments as time-invariant properties that do not capture multilingual or syncretic practices primarily due to limited data availability. Time invariance is a reasonable assumption because cultural change is usually a multi-generational process (see, e.g., Weber, 1976). With respect to multilingualism and syncretism, the mechanisms introduced above do not operate at the individual level but rather at the group level. Among existing studies only Laitin (2000) approaches the complexity of our data by coding linguistic and religious differences for group dyads for a far smaller number of ethnic groups. Unlike all existing efforts, our dataset is the first to code multiple sub-group segments on both ethnic dimensions, which makes it possible to assess fine-grained ethnic differences between groups that go beyond dichotomous difference indicators.

In sum, our dataset contains 629 unique languages and 67 distinct religious creeds for 793 ethnic groups (also see Figure B1). These sum to 1,147 linguistic segments and 1,535 religious segments overall. As a result, ethnic groups are more uniform in linguistic terms than along the religious dimension. While the mean of linguistic segments per ethnic group is approximately 1.5, it is 1.9 for religious segments.

## **Ethnic dyads**

While scholars who subscribe to the religious-conflict link regularly identify ethnic groups as either exclusively linguistic or religious (Fox, 2002, 2004; Svensson, 2007), the

relevance of specific ethnic segments depends on the context and cannot be determined a priori. Subscribing to the sociological literature on cleavages, we argue that the conflict-proneness of religion or language needs to be analyzed relatively between groups (Lipset and Rokkan, 1990 (1967)). Our unit of analysis is therefore the ethnic-dyad that compares the most powerful group in a state to all other groups. This setup is appropriate when analyzing civil wars that pit the incumbent government against non-state actors (Cunningham, Gleditsch and Salehyan, 2009).<sup>7</sup>

We assume that the state is not ethnically neutral but usually controlled by one specific ethnic group. This assumption holds most clearly in nation-states where one ethno-nationalist group dominates. But even in multi-ethnic states one group usually is more closely associated with the state than any other, and therefore the relevant opponent in violent conflicts. Although the pre-war Yugoslav government included Croats, Bosniaks, Macedonians, and Montenegrins next to Serbs, Serbs were clearly the most powerful group in that coalition, and the relevant conflict actor on the side of the government.<sup>8</sup>

We identify such a “reference group” for each country-year in two stages. First, using the relative power ranking coded in EPR-ETH, we identify the most powerful group(s). If one group is coded as dominant, as holding monopoly power, or as the only senior partner, it is selected as the reference group. Second, in the few cases in which EPR-ETH codes a power-sharing arrangement with more than one “senior partner”, the demographically largest of the most powerful groups is selected. In Yugoslavia, for example, all ethnic groups in power (EGIPs) hold the rank of senior partner until 1991. Here, Serbs are selected as the most powerful group, as they constituted a demographic plurality.<sup>9</sup>

Figure 4 illustrates our dyadic setup for Nigeria in 2006. The government coalition consists of two included groups: the Yoruba as senior partner and the Hausa-Fulani as junior partner. Two additional groups, the Ijaw and the Ogoni, are excluded from the executive. In our dyadic setup all other included and excluded groups are potential challengers of the Yoruba. In other words, we consider both challenges within the

government coalition (infighting) and rebellions by excluded groups. As EPR codes power shifts, the reference group sometimes changes. Prior to 1999, EPR identifies the Hausa-Fulani as dominant, and therefore all group-dyads include the Hausa-Fulani as the reference group during that period.

Figure 4 about here.

## Cleavage variables

We are now ready to introduce our cleavage variables, which combine the data on linguistic and religious segments with the dyadic setup introduced above. For each group dyad and ethnic dimension, we compare the segments, and note whether a difference exists or not.<sup>10</sup> Recall the example of the Yoruba and Hausa-Fulani. On the linguistic dimension, the two groups differ completely since none of the three Hausa-Fulani languages is spoken by the Yoruba. In contrast, the Yoruba's Sunni Muslim (Maliki) religious segments corresponds to the Hausa-Fulani's only religious affiliation. We therefore code a difference for two of the three religious segments of the Yoruba but not for the Sunni Muslim (Maliki) segment. We repeat this exercise for all ethnic group-dyads and country-years on both the linguistic and religious dimensions.

A second step incorporates the information on segment size in order to determine whether linguistic and religious differences exist. We first compute linguistic and religious dummy variables that take the value of one whenever more than 50% of all group members differ on the given dimension and zero otherwise. In case of the Yoruba/Hausa-Fulani dyad, this dummy takes the value of one on both ethnic dimensions as only a minority of the Yoruba shares the Sunni Muslim (Maliki) affiliation with the Hausa-Fulani. To reach a higher level of precision, we also compute the probability that two randomly selected individuals from each group speak the same language or adhere to the same religious creed. With regards to the Nigerian example, randomly selecting a member of each group would yield a 57% chance of selecting two individuals with different religious beliefs (see Figure 4).<sup>11</sup>

Our main explanatory variables thus capture three important aspects of cleavages: their relational character, their multi-dimensional nature, and the possible overlap of group boundaries. First, by calculating cleavages with respect to a particular reference group, we take into account the dyadic character of civil war. Second, the coding of both linguistic and religious cleavages recognizes that it cannot be determined a priori which of the two markers is more relevant for a particular group, or whether both matter at the same time. Third, our mean difference variable reflects the fuzziness of ethnic differences by recognizing that linguistic and religious segments can cross-cut group boundaries. Where more members of an ethnic group dyad share one cleavage dimension, the salience of that marker to highlight group differences is likely to be lower than where no shared ethnic dimensions exist. While capturing some aspects of a constructivist understanding of ethnic cleavages, our measures are limited in that they do not identify the political meaning of religion and language at a given point in time. In the presence of both religious and linguistic cleavages, our data do not reveal which dimension serves as the actual basis of identification.

Despite this restriction, we believe that our dataset offers improvements over existing data on ethnic groups and cleavages. Compared to existing fractionalization indices (cf. Roeder, 2001; Alesina et al., 2003) our measures capture the ethnic-dyad level rather than country-level variation, and clearly identify the potential combatants in a civil war rather than comparing all groups with one another. Additionally, the explicit coding of linguistic and religious cleavages allows us to estimate both individual cleavage dimension effects and their interaction, which country-level indices of cleavage constellations cannot (see Selway, 2010).

Rather than comparing religious and linguistic dimensions most existing studies on the conflict-inducing effect of ethnic cleavages identify groups as either linguistic or religious (Svensson, 2007; Toft, 2007; Fox, 2012). This practice runs the risk of miscoding a particular conflict or overlooking the importance of another salient conflict dimension (cf. Kalyvas, 2003). By coding both linguistic and religious cleavages, we do not fall into the trap of falsely attributing conflict issues to either one dimension. In the

Israeli-Palestinian conflict, for example, linguistic nationalism arguably played a more decisive role in the first Palestinian uprising in 1988 and religious motives became more relevant in the Second Intifada (Litvak, 1998). Our data do not detect this change in salient cleavages but they also do not bias the results in favor of either language or religion. Moreover, following actor claims in order to identify the master cleavage of particular civil wars, makes those studies, which only identify conflict dyads but omit non-conflict dyads from their samples, vulnerable to selection bias. In contrast, our data coded for all peaceful and conflictual ethnic dyads alike are less susceptible to reverse causality.

Using Minorities at Risk (MAR) data by Gurr et al. (1993), Laitin (2000) considers potential conflict dyads and compares linguistic and religious cleavages. MAR provides a constructivist coding of policies that discriminate against linguistic and religious identities. Selection bias is a concern here since the ethnic groups in the MAR data are selected exactly because they are “at risk,” due to, for example, linguistic or religious discrimination (Hug, 2003). Moreover, the MAR data cover less than half the number of groups included in EPR and is only available for a few years at the beginning of the 1990s and 2000s.

## **Other variables**

Our dependent variable, ethnic civil war onset, is based on the civil war coding of the UCDP/PRIO Armed Conflict Database, which uses a minimum of 25 battle deaths as its lower threshold (Gleditsch et al., 2002; Harbom and Wallensteen, 2010). The ACD2EPR Dataset connects civil wars to particular ethnic groups through the rebel groups in the Non-State Actors (NSA) dataset (Cunningham, Gleditsch and Salehyan, 2009; Wucherpfennig et al., 2012). A civil war is defined as ethnic if a rebel organization fighting in it claims to do so on behalf of and recruits from a particular ethnic group.

Introducing a number of standard control variables from the literature, we assume that inequalities and grievances increase the likelihood of civil war onset. They are measured by two variables, one capturing the exclusion of political actors who rep-



resent an ethnic group from executive power, and the other a recent loss of relative power within the previous two years (Petersen, 2002). The former reflects structural disadvantages an ethnic group faces in the political arena while the latter reflects more dynamic shocks to its political status. A group's relative share of the politically relevant population captures larger fighting capacity. Larger groups are more likely to fight than smaller groups as, on average, they have a larger pool of potential recruits and supporters if recruitment occurs along ethnic lines. Both the grievance and the size variables come from the EPR-ETH dataset by Cederman, Wimmer and Min (2010); Cederman, Gleditsch and Buhaug (2013). Moving to the country-level, low GDP per capita and large population size are expected to increase the likelihood of civil war (Hegre and Sambanis, 2006).<sup>12</sup> Finally, we control for temporal effects by including the number of previous civil wars fought by a group, a linear term for peace-years, and three cubic splines (Beck, Katz and Tucker, 1998).

## Analysis

This section tests our main theoretical claim that linguistic differences are at least as conflict-prone as religious cleavages.<sup>13</sup> First, we discuss the bivariate relationships of civil war onset and the two ethnic cleavages. While civil wars are about nine times more frequent in the presence of linguistic differences than in their absence, the corresponding ratio for religious differences is merely two (also see Table C1).<sup>14</sup> Although religious differences are significantly related to civil war onset in this simple setup their effect is far weaker than that of linguistic differences. These results should be interpreted with caution, however, because other factors may affect the association. In order to gauge their individual effects we include both cleavage types in a number of logit regressions with robust country-clustered standard errors.

Table 1 presents four regression models. Model 1 introduces the mean difference variables that measure the degree of difference for linguistic and religious cleavages for both languages and religions in addition to a range of standard controls as described in the previous section. Model 2 replaces the mean difference variables with their

majority equivalents that record a “1” whenever more than 50% of group members differ on one of the dimensions. In Models 1 and 2, the cleavage variables overlap. This means that for those dyads that differ along both ethnic dimensions, the two cleavage variables each indicate the respective difference values. In contrast, Model 3 adds a third dummy variable that identifies those ethnic dyads that differ on both cleavage dimensions. Thus, the three dummy variables in this model indicate whether a given ethnic dyad is only linguistically distinct, only religiously distinct, and both linguistically and religiously distinct. This allows us to assess the conflict risk of ethnic dyads characterized by reinforcing cleavages compared to those where at least one ethnic dimension cross-cuts group boundaries (cf. Chandra, 2005). Finally, Model 4 includes all cases with both religious and linguistic cleavages in the religious category and compares them to those dyads that differ only on the linguistic dimension. Hence, this is an easy test for the religion-conflict thesis and a hard test for our counter-claim.

Table 1 about here.

Across all specifications, the effect of linguistic differences on the risk of ethnic conflict is positive and statistically significant. In contrast, the estimated coefficients for religious differences fall short of conventional levels of statistical significance in Models 1 to 3: in the first two models, they are essentially zero. Moreover, Wald tests of equality reject the null hypothesis that religious cleavages have the same effect on civil war onset as linguistic differences ( $p=.045$  in Model 1,  $p=.007$  in Model 2, and  $p=.054$  in Model 3). Only in Model 4, where we interpret all mixed dyads as religious ones, does the estimated effect of religion reach conventional levels of significance. Yet, even under these unlikely assumptions, the residual cases with only linguistic differences exert a substantially larger effect on the risk of ethnic conflict than religious differences. This finding lends support to the theoretical mechanisms we proposed above with respect to linguistic cleavages, while casting doubts on claims about the conflict-inducing effect of religious cleavages. Finally, in Model 3, the effect of language is not statistically different from those cases where both cleavages reinforce one another

( $p=.174$ ). Thus, our results deviate from previous findings that report a lower risk of civil war for cases with cross-cutting cleavages (Dunning and Harrison, 2010; Selway, 2010).<sup>15</sup>

The results of the control variables are in accordance with previous findings. Exclusion from executive power or a recent downgrading of a group's power status substantially and significantly increases the likelihood of conflict onset. Additionally, larger groups and past conflict likewise increase the probability of ethnic civil war while higher economic development at the country level makes civil war less likely. In sum, even after the introduction of important control variables, linguistic differences are very strongly linked to ethnic conflict probability while religious differences exert a much weaker effect in the post-World War II period.<sup>16</sup>

## Spatio-temporal variations

Our main results indicate that civil wars are more likely to be fought along linguistic cleavages than religious ones. In the following, we subject this finding to a variety of sensitivity analyses, foremost because the results are likely to exhibit considerable regional and temporal variation. Moreover, it is necessary to test alternative operationalizations of ethnic cleavages. The results presented below, however, demonstrate that such considerations do not undermine our general finding.

The relative importance of language and religion may vary across world regions due to distinct political and economic conditions. According to Gellner (1983), linguistic differences should be especially consequential in societies that are thoroughly industrialized and where mass literacy has made major advances. This implies that Eastern Europe and Asia should be more likely to host ethnic conflict along linguistic cleavages than Sub-Saharan Africa. In contrast, the apparent importance of Islam and the lower frequency of linguistic cleavages due to the widely shared Arabic language should make religious differences particularly conflict-prone in the Middle East.

Figure 5 shows the estimated size and accuracy of the language and religion coefficients for five world regions (also see Table C1 in our Online Appendix).<sup>17</sup> While

highly relevant for conflict in Eastern Europe and Asia, linguistic cleavages are an even stronger covariate of conflict than religious differences in the Middle East. This is illustrated by cases of linguistically distinct ethnic groups, such as the Kurds in Turkey and Iraq, or the Baloch and Arabs in Iran, that have engaged in multiple civil wars with their respective governments (cf. Gurr et al., 1993, 317). The only region where religious differences show a slightly stronger relation to conflict than linguistic differences is Sub-Saharan Africa although the estimated difference is not statistically significant. This is not surprising given that our theoretical argument emphasizes reduced individual life chances due to linguistic discrimination as one source of collective grievances. This mechanism should be less pronounced in African states that do not feature Gellner's ideal type of modernizing, industrialized society because of their lack of economic development and relatively low literacy rates. In sum, with the exception of this part of the world, linguistic differences explain at least an equal amount, if not more, of the variation in the likelihood of ethnic civil war onset as religious cleavages.

Figure 5 about here.

Having considered spatial differences, we now turn to temporal variation. Huntington (1993) argued that civilizational differences among world religions would become the focal point of conflict after the Cold War. In Table 2, we therefore focus on the time period since 1990.<sup>18</sup> Yet, contrary to Huntington's expectations, religious differences do not exhibit a stronger effect on ethnic conflict in the post-Cold War period regardless of the model specification. In fact, the effect becomes even weaker. In contrast, the effect of linguistic differences becomes somewhat stronger during this period, although there is no statistically significant difference between the two periods before and after the Cold War. However, our operationalization of religious differences does not completely match Huntington's idea of civilizations, which is based on the highest level of aggregation and thus predicts a clash between world religions. Accordingly, religious differences should only matter for conflicts between Christians and Muslims, Hindus and Sikhs, or Buddhists and Animists, but not for disputes within major reli-

gions. To test this more specific claim, Model 7 replaces our main variable of religious differences with an indicator that measures civilizational differences between major world religions. Again, this expectation of Huntington is not borne out by our data as the coefficient for civilizational differences turns negative for the period after 1989.<sup>19</sup>

Table 2 about here.

In addition to postulating religious cleavages in general as a source of violent backlashes, some scholars have singled out cleavages associated with Islam as particularly susceptible to violence (Huntington, 1996; Fox, 2012). Similarly, Toft (2007) suggests that Muslim involvement in civil wars is disproportionately frequent. In Figure 6, we compare dyads with at least one Muslim group to dyads encompassing groups of all other major world religions (see Table C2). Instead of our relational cleavage variable, we include major religion dummies that indicate the majority religions of the groups in the dyad. For example, the Muslim dummy variable is coded as “1” for every dyad, in which more than 50% of the members of at least one group are Muslims. We chose dyads that include Christian groups as our baseline category since most of them are situated in the West – the region with the least amount of civil wars. We find that dyads that include Muslim groups are indeed more likely to experience ethnic civil war than dyads that include Christian groups. Yet, with respect to all major religions except for groups that are predominantly atheist, the t-test results reveal that the differences are not statistically significant. In view of these findings, we reject the claim that Islam is a particularly conflict-prone religion.

Figure 6 about here.

In contrast to Toft’s approach, which is reflected in Figure 9, our empirical strategy is relational. Following this line, we construct a new dummy variable that identifies all dyads that include either one Muslim group and one non-Muslim group *or* two groups of different Muslim denominations (for example, Sunnis and Shi’a). Thus, cases of infighting between moderate and extremist Muslims as, for instance, the intrastate

conflict between the Pakistani government and the Pakistani Taliban, are not considered here. Model 8 in Table 2 uses this new Muslim dummy variable along with a residual religious difference dummy that identifies the remaining dyads that are religiously distinct. The results reveal that dyads consisting of Muslims and non-Muslims or of two different Muslim denominations are not more likely to experience ethnic civil war than other ethnic dyads. Yet, the effect of the residual religious difference variable now becomes negative, which means that without Muslims, religious differences might actually decrease conflict risk. In other words, Toft might be right in that Muslim differences are more often associated with ethnic conflict - but this is only true in relation to other religious differences, not in relation to other cleavage dimensions (also see Table D2).

## **Additional robustness tests**

Following up the evaluation of spatio-temporal variations, the current section exposes our findings to a number of additional sensitivity tests (see Section D in our Online Appendix for details). This analysis shows that the main result exposing the weakness of religious cleavages as an explanation of civil war holds up regardless of whether one controls for temporal heterogeneity (Table D1 and Figures D1 and D2) or introduces alternative specifications of the cleavages variables (Table D2 and D3) and the ethnic dyad (Table D4). By comparing each ethnic group to all ethnic groups in power (EGIPs), we also account for conflicts that occur among two peripheral groups – one of which is a member of the government coalition – and later escalate to civil wars (Lacina, 2013).

In addition, a series of models explores possible sensitivity to uncontrolled variance and endogeneity (see Tables D5-D11). Using multi-way clustering, hierarchical models, and fixed effects at the level of countries and years, this analysis further bolsters the robustness of the main results. Although the underlying EPR data were selected with respect to political relevance, there is also no reason to believe that this selection criterion introduces systematic bias in favor of any particular cleavage dimension or

cases that involved conflict. It is correct that our data contain far more linguistic than religious segments (see Figure B1), but this imbalance matters little for our analysis as the difference in the frequency of linguistic and religious differences is much smaller (see Table C1).<sup>20</sup>

As opposed to claims-based codings of cleavage dimensions, the EPR-ED data themselves can be considered relatively immune to endogeneity with respect to conflict thanks to their origin in ethnographic data that were collected by researchers who have no particular interest in conflict. As argued above, our cleavage variables only change very slowly over time on the aggregate group-level (Weber, 1976), and are thus unlikely to be endogenous to civil war. We therefore refrain from any formal instrumentation of our cleavage variables, which is in line with the standard practise in economic studies that investigate the effects of structural diversity on public goods provision or conflict (see, e.g., Alesina, Baqir and Easterly, 1999; Esteban and Ray, 2011; Spolaore and Wacziarg, 2013).<sup>21</sup>

## Conclusion

In this paper we have revisited the relationship between language, religion, and ethnic civil war. We find that linguistic cleavages are overrepresented as a potential source of such violence compared to religious differences. These findings cast doubts on the claim that religion is a more divisive cleavage than language. Moreover, our findings lend little support to recent studies that attribute an especially destabilizing potential to cleavages involving Muslim groups compared to other major world religions.

Stressing the political logic of grievances, rebel mobilization, and the accommodation of rebel demands by governments, this study compares the conflict potential of linguistic and religious differences by considering how they influence the main mechanisms that link horizontal inequalities to civil war. While our empirical results corroborate the link between linguistic cleavages and ethnic civil war, future research needs to examine the individual mechanisms of our theoretical framework more closely.

Although we have found much less evidence for the religion-conflict link than the

existing literature, we do not deny the fact that the political activation of religious cleavages has triggered ethnic conflict in many instances. When both linguistic and religious differences are present, our data do not reveal which of these cleavages is more important in a conflict. While this would be desirable, it is also important to bear in mind that it is often difficult to identify the most salient cleavage in a particular conflict (Kalyvas, 2003). Not all rebels of the Islamic State of Iraq and Syria (ISIS) organization, for example, are motivated by religious ideology. Instead, many Sunni Arabs have joined the rebel group in response to discriminatory state policies by the Shi'a dominated government. As we demonstrate, the relative salience of cleavages also depends on regional context with significantly higher levels of linguistic conflict in Eastern Europe and Asia compared to the rest of the world while temporal variation seems to be less pronounced. Our analysis may also underestimate the effect of religion because it does not focus on conflicts between moderates and extremists within a given religious group. In order to more clearly understand the precise role of both religious and linguistic divisions in the emergence of ethnic civil war, future research needs to collect data on actors' claims and government policies with respect to language and religion.

Overall, however, our results suggest that much of the literature focusing on religion and violence has tended to be one-sided and too narrow. The present study prompts us to rethink overly alarmist renderings that religious cleavages are becoming more violent, especially those that involve Muslim groups. Through the EPR-ED dataset we now have a new and powerful tool at our disposal by which we can more thoroughly assess the differential impact of specific cleavage types by considering the structural preconditions around the world without limiting the sample to conflict cases. Thanks to its wide coverage and relatively exogenous nature, it is becoming easier to detect patterns linking the contents of ethnicity to conflict processes.



## Notes

<sup>1</sup>In the remainder of this study, we will interchangeably refer to ethnic civil war onset as civil war, intrastate conflict, civil conflict, rebellion, or ethnic conflict. Our main research focus is thus on civil wars between a government and ethnically defined challengers. We consider neither non-ethnic civil wars nor communal conflicts without government involvement.

<sup>2</sup>Laitin (2000) is an exception here.

<sup>3</sup>Kalyvas also suggests that ISIS, although rooted in fundamentalist religious ideology, is far better understood as a revolutionary actor that could as easily draw on nationalist and other ethnic identities: <http://www.washingtonpost.com/blogs/monkey-cage/wp/2014/07/07/the-logic-of-violence-in-islamic-states-war/>

<sup>4</sup>Other conflict researchers question the importance of the collective action dilemma as a fundamental obstacle to mobilization (Staniland, 2014). Kalyvas and Kocher (2007), for example, argue that defecting is likely to be more dangerous than fighting.

<sup>5</sup>In principle, a politically relevant ethnic group in EPR-ETH could consist of more than three linguistic or religious segments but in the vast majority of cases they do not. In exceptional cases such as the Indigenous Peoples in Brazil, we only coded the three largest segments. Note that a segment was only recorded if it accounted for at least 10% of group members.

<sup>6</sup>There are some exceptions, however, which we discuss in Section A of our Online Appendix.

<sup>7</sup>When analyzing communal conflicts between non-state actors, it would be more appropriate to consider all ethnic group dyads in a state.

<sup>8</sup>Slovenes did not fight Croats or Montenegrins. Also see Horowitz (2002) who argues that ethnic identities in Sub-Saharan Africa tend to polarize into dyadic blocs.

<sup>9</sup>Overall, we applied the second coding rule only 22 times in 17 different states.

<sup>10</sup>With respect to the tree structure above, this simply means that we check whether two groups share the same node in a tree but not whether they share any parent nodes. Note that a new reference group implies different group dyads, and thus, changes in our cleavage variables.

<sup>11</sup>Formally, assume a reference group  $A$  and another group  $B$ , each which have ethnic segments  $a$  and  $b$  of relative sizes  $s$ , then the mean difference between those two groups is:

$$d_{mean,AB} = \sum_a \sum_b \mathbb{1}_{ab} * s_a * s_b$$

where  $\mathbb{1}$  is an indicator function that takes the value of 1 if the ethnic segments  $a$  and  $b$  are different and 0 if they are the same. Summary statistics and a correlation matrix of our main variables can be found in Section B of our Online Appendix.

<sup>12</sup>The data are taken from Hunziker and Bormann (2013).

<sup>13</sup>Our Online Appendix contains additional, detailed information on our data (Sections A & B), regression models that we display only graphically (Section C), and an extensive number of robustness checks (Section D). Tables and Figures from the appendix are referenced by these letters.

<sup>14</sup>As mentioned above, a dyad is coded as different when more than 50% of all members differ on the given dimension. Recall that we do not define the ethnic dimension of a dyad a priori. Hence linguistic dyads can also differ on religious grounds and vice versa.

<sup>15</sup>In Table D2 we interact the language and religion variables from Model 2 but find no support for the importance of an interaction effect.

<sup>16</sup>See Figure C for actual effect estimates.

<sup>17</sup>In this graph, we do not display the results for Latin America, the Caribbean or

Oceania for a lack of ethnic civil war in these regions. These regions are included in all other analyses.

<sup>18</sup>The corresponding regressions for the Cold War period can be found in Table D1.

<sup>19</sup>Additional inconclusive models with a continuous measure of time are included in Table D1 and Figures D1 and D2. Our results are in line with previous empirical research concerned with either intra- (Fox, 2002, 2005) or interstate conflicts (Gartzke and Gleditsch, 2006).

<sup>20</sup>As opposed to the Soviet Atlas Narodov Mira (Bruk and Apenchenko, 1964), which is known to suffer from ideologically driven anti-religious bias, EPR was coded by a large number of experts who were entirely free to highlight any politically salient cleavages. Selection on the dependent variable is also unlikely as confirmed by Birnir et al.'s (2011) use of the much more inclusive dataset A-MAR to replicate previous EPR-based studies covering the link between exclusion and civil war.

<sup>21</sup>We recognize that our political power access variables are potentially endogenous to conflict. Yet research by one of the authors instruments for ethnic exclusion based on variation in colonial strategies and finds that naive studies underestimate the effect of exclusion on civil war onset.

## References

- Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat and Romain Wacziarg. 2003. "Fractionalization." *Journal of Economic Growth* 8(2):155–194.
- Alesina, Alberto, Reza Baqir and William Easterly. 1999. "Public Goods and Ethnic Divisions." *Quarterly Journal of Economics* 114(4):1243–1284.
- Anderson, Benedict. 2006. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. New York, NY: Verso.
- Barth, Fredrik. 1969. "Ethnic Groups and Boundaries - Introduction." In *Fredrik Barth (Ed.) Ethnic Groups and Boundaries: The Organization of Cultural Differences*. Little, Brown: Boston pp. 9–37.
- Bayart, Jean-François. 2009. *The State in Africa: The Politics of the Belly*. 2nd ed. Cambridge, UK: Polity Press.
- Beck, Nathaniel, Jonathan N. Katz and Richard Tucker. 1998. "Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable." *American Journal of Political Science* 42(4):1260–1288.
- Birner, Johanna K., Jonathan Wilkenfeld, Dawn Brancati, James D. Fearon, Ted R. Gurr, David D. Laitin, Amy Pate and Stephen Saidemann. 2011. A-MAR (All-Minorities at Risk): Addressing the Selection Bias Issue. Technical report University of Maryland.  
**URL:** [http://gvpt.umd.edu/birner/Johanna\\_Birner/Research\\_files/A-MAR%208-11.pdf](http://gvpt.umd.edu/birner/Johanna_Birner/Research_files/A-MAR%208-11.pdf)
- Blair, Tony. 2014. "Religious Difference, Not Ideology, Will Fuel This Century's Epic Battles." *The Observer*, 25 January 2014. Available: <http://www.theguardian.com/commentisfree/2014/jan/25/religious-difference-ideology-conflicts-middle-east-tony-blair> [Last accessed: 11 February 2014].
- Brubaker, Rogers. 2009. "Ethnicity, Race, and Nationalism." *Annual Review of Sociology* 35:21–42.
- Brubaker, Rogers. 2013. "Language, Religion and the Politics of Difference." *Nations and Nationalism* 19(1):1–20.
- Bruk, Solomon I. and V. S. Apenchenko, eds. 1964. *Atlas Narodov Mira (Atlas of the Peoples of the World)*. Moscow, RU: Glavnoe Upravlenie Geodezii i Kartografi.
- Cederman, Lars-Erik, Andreas Wimmer and Brian Min. 2010. "Why Do Ethnic Groups Rebel? New Data and Analysis." *World Politics* 62(1):87–119.
- Cederman, Lars-Erik, Kristian S. Gleditsch and Halvard Buhaug. 2013. *Inequality, Grievances, and Civil War*. New York, NY: Cambridge University Press.
- Cederman, Lars-Erik, Nils B. Weidmann and Kristian S. Gleditsch. 2011. "Horizontal Inequalities and Ethno-Nationalist Civil War: A Global Comparison." *American Political Science Review* 105(2):478–495.

- Chandra, Kanchan. 2005. "Ethnic Parties and Democratic Stability." *Perspectives on Politics* 3(02):235–252.
- Corbett, Richard, Francis Jacobs and Michael Shakleton. 2011. *The European Parliament*. 8 ed. London, UK: John Harper Publishing.
- Cunningham, David E., Kristian S. Gleditsch and Idean Salehyan. 2009. "It Takes Two: A Dyadic Analysis of Civil War Duration and Outcome." *Journal of Conflict Resolution* 53(4):570–597.
- Darden, Keith. 2013. *Resisting Occupation: Mass Schooling and the Creation of Durable National Loyalties*. New York, NY: Cambridge University Press.
- Dunning, Thad and Lauren Harrison. 2010. "Cross-Cutting Cleavages and Ethnic Voting: An Experimental Study of Cousinage in Mali." *American Political Science Review* 104(01):21–39.
- Esteban, Joan and Debraj Ray. 2011. "A Model of Ethnic Conflict." *Journal of the European Economic Association* 9(3):496–521.
- Fox, Jonathan. 2002. "Ethnic Minorities and the Clash of Civilizations: A Quantitative Analysis of Huntington's Thesis." *British Journal of Political Science* 32(03):415–434.
- Fox, Jonathan. 2004. "The Rise of Religious Nationalism and Conflict: Ethnic Conflict and Revolutionary Wars, 1945-2001." *Journal of Peace Research* 41(6):715–731.
- Fox, Jonathan. 2005. "Paradigm Lost: Huntington's Unfulfilled Clash of Civilizations Prediction into the 21st Century." *International Politics* 42(4):428–457.
- Fox, Jonathan. 2012. "The Religious Wave: Religion and Domestic Conflict from 1960 to 2009." *Civil Wars* 14(2):141–158.
- Gartzke, Erik and Kristian S. Gleditsch. 2006. "Identity and Conflict: Ties that Bind and Differences that Divide." *European Journal of International Relations* 12(1):53–87.
- Gellner, Ernest. 1983. *Nations and Nationalism*. Ithaca, NY: Cornell University Press.
- Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg and Havard Strand. 2002. "Armed Conflict 1946-2001: A New Dataset." *Journal of Peace Research* 39(5):615–637.
- Goddard, Stacie E. 2006. "Uncommon Ground: Indivisible Territory and the Politics of Legitimacy." *International Organization* 60(1):35–68.
- Gorski, Philip S and Gülay Türkmen-Derviřođlu. 2013. "Religion, Nationalism, and Violence: An Integrated Approach." *Sociology* 39:193–210.
- Gurr, Ted R. 2000. *Peoples versus States: Minorities At Risk in the New Century*. Washington, DC: United States Institute of Peace Press.
- Gurr, Ted .R., Barbara Harff, M.G. Marshall and J.R. Scarritt. 1993. *Minorities at Risk: A Global View of Ethnopolitical Conflicts*. Washington, DC: United States Institute of Peace Press.

- Harbom, Lotta. and Peter. Wallensteen. 2010. "Armed Conflicts, 1946–2009." *Journal of Peace Research* 47(4):501–509.
- Hassner, Ron E. 2009. *War on Sacred Grounds*. Ithaca, NY: Cornell University Press.
- Hegre, Havard and Nicholas Sambanis. 2006. "Sensitivity Analysis of Empirical Results on Civil War Onset." *Journal of Conflict Resolution* 50(4):508–535.
- Horowitz, Donald L. 2000. *Ethnic Groups in Conflict*. Berkeley, CA: University of California Press.
- Horowitz, Donald L. 2002. Constitutional Design: Proposals versus Processes. In *The Architecture of Democracy: Constitutional Design, Conflict Management, and Democracy*, ed. Andrew Reynolds. New York: Oxford University Press pp. 15–36.
- Hroch, Miroslav. 1985. *Social Pre-Conditions of National Revival in Europe. A Comparative Analysis of the Social Composition of Patriotic Groups among the Smaller European Nations*. Cambridge, UK: Cambridge University Press.
- Hug, Simon. 2003. "Selection Bias in Comparative Research: The Case of Incomplete Data Sets." *Political Analysis* 11(3):255–74.
- Huntington, Samuel P. 1993. "The Clash of Civilizations?" *Foreign Affairs* 72(3):22–49.
- Huntington, Samuel P. 1996. *The Clash of Civilizations and the Remaking of World Order*. New York, NY: Simon & Schuster.
- Hunziker, Philipp and Nils-Christian Bormann. 2013. "Size and Wealth in the International System: Population and GDP per capita Data for Political Science." Typescript: ETH Zurich.
- Hyltenstam, Kenneth and Niclas Abrahamsson. 2000. "Who Can Become Native-Like in a Second Language? All, Some, or None?" *Studia Linguistica* 54(2):150–166.
- Joshua Project: Unreached Peoples of the World*. 2011.  
**URL:** <http://www.joshuaproject.net/>
- Juergensmeyer, Mark. 1993. *The New Cold War?: Religious Nationalism Confronts the Secular State*. Berkeley, CA: University of California Press.
- Kalyvas, Stathis N. 2003. "The Ontology of "Political Violence": Action and Identity in Civil Wars." *Perspectives on Politics* 1(3):475–494.
- Kalyvas, Stathis N. and Matthew Adam Kocher. 2007. "How "Free" is Free Riding in Civil Wars?: Violence, Insurgency, and the Collective Action Problem." *World Politics* 59(2):177–216.
- Lacina, Bethany. 2013. "How Governments Shape the Risk of Civil Violence: India's Federal Reorganization, 1950–56." *American Journal of Political Science* Forthcoming:1–19.  
**URL:** <http://dx.doi.org/10.1111/ajps.12074>
- Laitin, David D. 1998. *Identity in Formation: The Russian-Speaking Populations in the Near Abroad*. Ithaca, NY: Cornell University Press.

- Laitin, David D. 2000. "Language Conflict and Violence: The Straw that Strengthens the Camel's Back." *European Journal of Sociology* 41(1):97–137.
- Lange, Matthew and Andrew Dawson. 2010. "Education and Ethnic Violence: A Cross-National Time-Series Analysis." *Nationalism and Ethnic Politics* 16(2):216–239.
- Lewis, M. Paul, ed. 2009. *Ethnologue: Languages of the world*. Vol. 16 SIL International Dallas, TX.
- Lewis, Peter. 1992. "Political Transition and the Dilemma of Civil Society in Africa." *Journal of International Affairs* 46(1):31–54.
- Lichbach, Mark Irving. 1998. *The Rebel's Dilemma*. Ann Arbor, MI: University of Michigan Press.
- Liow, Joseph Chinyong. 2006. "Muslim Resistance in Southern Thailand and Southern Philippines: Religion, Ideology, and Politics." *East-West Center Policy Studies* 24:1–84.
- Lipset, Seymour Martin and Stein Rokkan. 1990 (1967). Cleavage Structures, Party Systems, and Voter Alignments. In *The West European Party System*, ed. Peter Mair. Oxford, UK: Oxford University Press pp. 91–138.
- Litvak, Meir. 1998. "The Islamization of the Palestinian-Israeli Conflict: The Case of Hamas." *Middle Eastern Studies* 34(1):148–163.
- Liu, Amy H. 2011. "The Linguistic Effects of Political Institutions." *Journal of Politics* 73(1):125–39.
- Patten, Alan and Will Kymlicka. 2003. Introduction: Context, Issues, and Approaches. In *Language Rights and Political Theory*, ed. W. Kymlicka & A. Patten. Oxford, UK: Oxford University Press.
- Petersen, Roger D. 2002. *Understanding Ethnic Violence: Fear, Hatred, and Resentment in Twentieth-Century Eastern Europe*. New York, NY: Cambridge University Press.
- Philpott, Daniel. 2002. "The Challenge of September 11 to Secularism in International Relations." *World Politics* 55(01):66–95.
- Reno, William. 2011. *Warfare in Independent Africa*. New York, NY: Cambridge University Press.
- Reynal-Querol, Marta. 2002. "Ethnicity, Political Systems, and Civil Wars." *Journal of Conflict Resolution* 46(1):29.
- Roeder, Philip G. 1991. "Soviet Federalism and Ethnic Mobilization." *World Politics* 43(2):196–232.
- Roeder, Philip G. 2001. "Ethnolinguistic Fractionalization (ELF) Indices, 1961 and 1985." Online.  
**URL:** <http://weber.ucsd.edu/~proeder/elf.htm>

- Safran, William and Amy H Liu. 2012. "Nation-Building, Collective Identity, and Language Choices: Between Instrumental and Value Rationalities." *Nationalism and Ethnic Politics* 18(3):269–292.
- Selway, Joel S. 2010. "Cross-Cuttingness, Cleavage Structures and Civil War Onset." *British Journal of Political Science* 41(1):1–28.
- Spolaore, Enrico and Romain Wacziarg. 2013. "How Deep Are the Roots of Economic Development?" *Journal of Economics Literature* 51(2):1–45.
- Staniland, Paul. 2014. *Networks of Rebellion: Explaining Insurgent Cohesion and Collapse*. Ithaca, NY: Cornell University Press.
- Stewart, Frances. 2008. *Horizontal Inequalities and Conflict: Understanding Group Violence in Multiethnic Societies*. London, UK: Palgrave Macmillan.
- Stewart, Frances. 2009. "Religion versus Ethnicity as a Source of Mobilisation: Are There Differences?" *MICROCON Research Working Paper* 18.
- Svensson, Isak. 2007. "Fighting with Faith." *Journal of Conflict Resolution* 51(6):930–949.
- Tilly, Charles. 1978. *From Mobilization to Revolution*. Reading, MA: Addison-Wesley.
- Toft, Monica Duffy. 2007. "Getting Religion? The Puzzling Case of Islam and Civil War." *International Security* 31(4):97–131.
- Weber, Eugene. 1976. *Peasants into Frenchmen: The Modernization of Rural France, 1870-1914*. Redwood City, CA: Stanford University Press.
- Wimmer, Andreas. 1997. "Who Owns the State? Understanding Ethnic Conflict in Post-Colonial Societies." *Nations and Nationalism* 3(4):631–666.
- Wucherpfennig, Julian, Nils Metternich, Lars-Erik Cederman and Kristian S. Gleditsch. 2012. "Ethnicity, the State and the Duration of Civil War." *World Politics* 64(1):79–115.



## Figures

Figure 1: Linguistic and religious segments for the Hausa-Fulani and Yoruba from Nigeria.

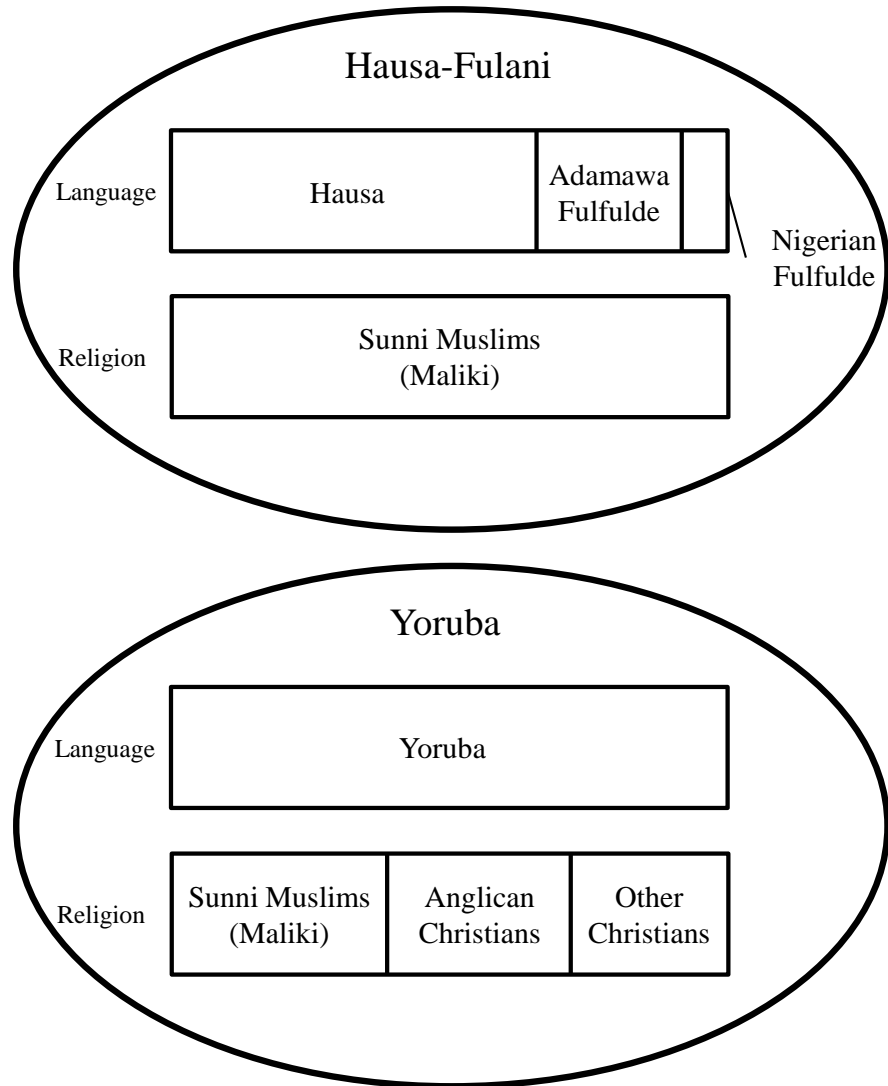


Figure 2: Ethnologue language tree example.

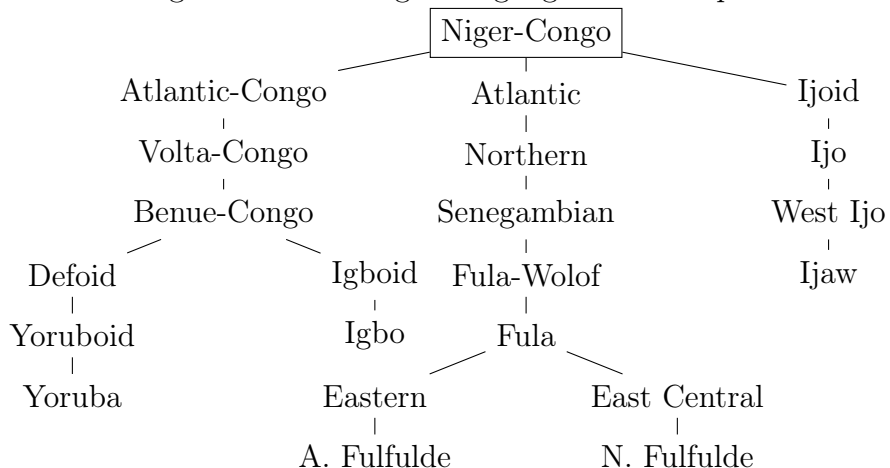


Figure 3: Religion tree example.

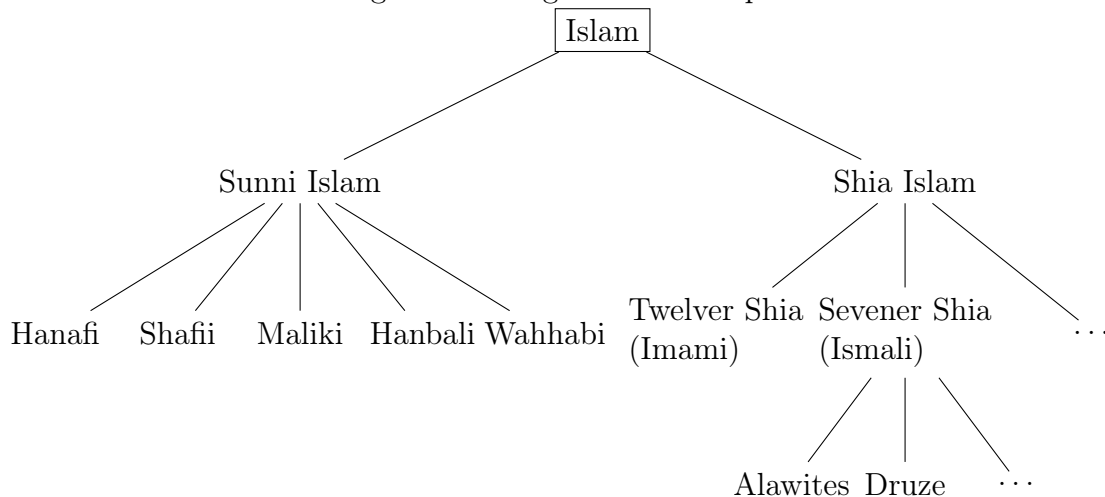


Figure 4: Ethnic power constellation and ethnic cleavages in Nigeria in 2006 (LD = Linguistic Differences; RD = Religious Differences).

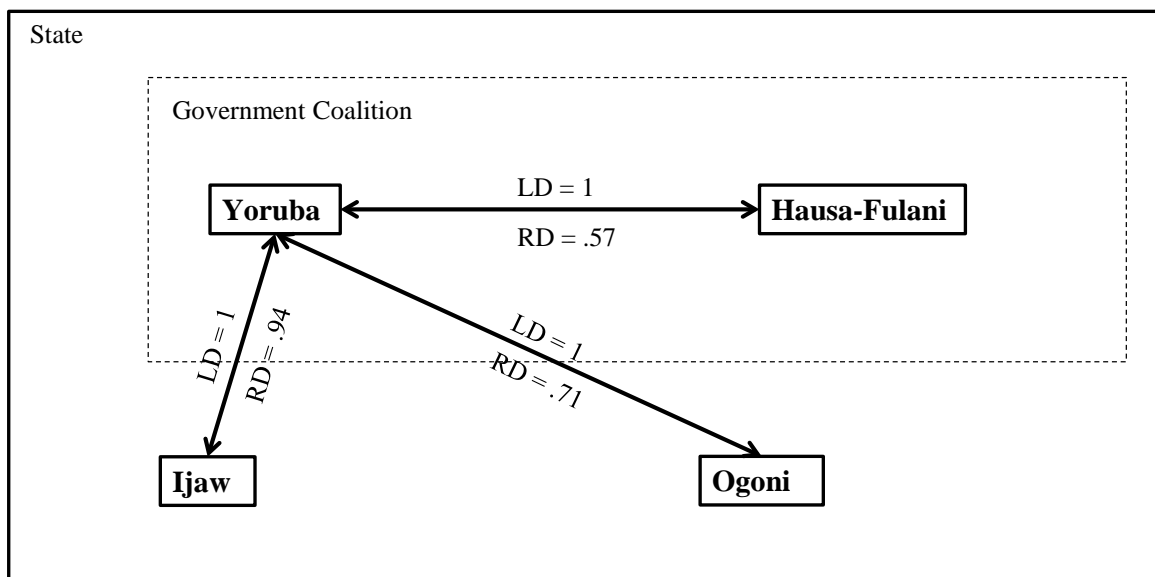


Figure 5: Estimated effects of ethnic cleavages in different world regions. Grey lines represent 95% confidence intervals. Full regression models in Table C1.

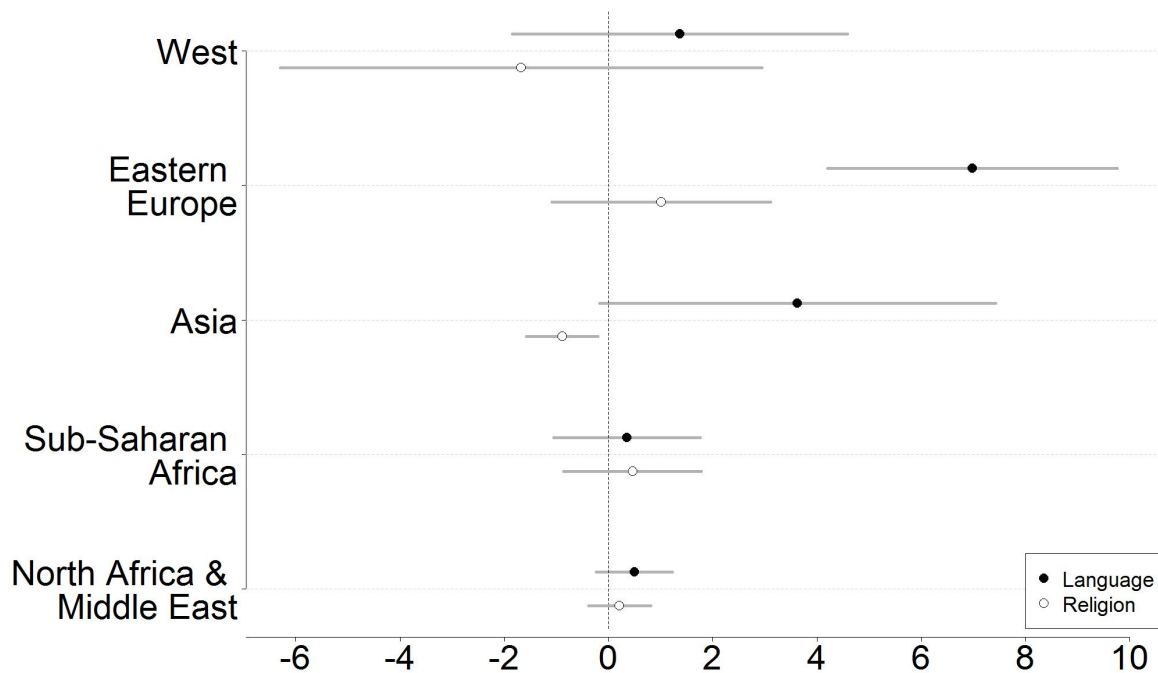
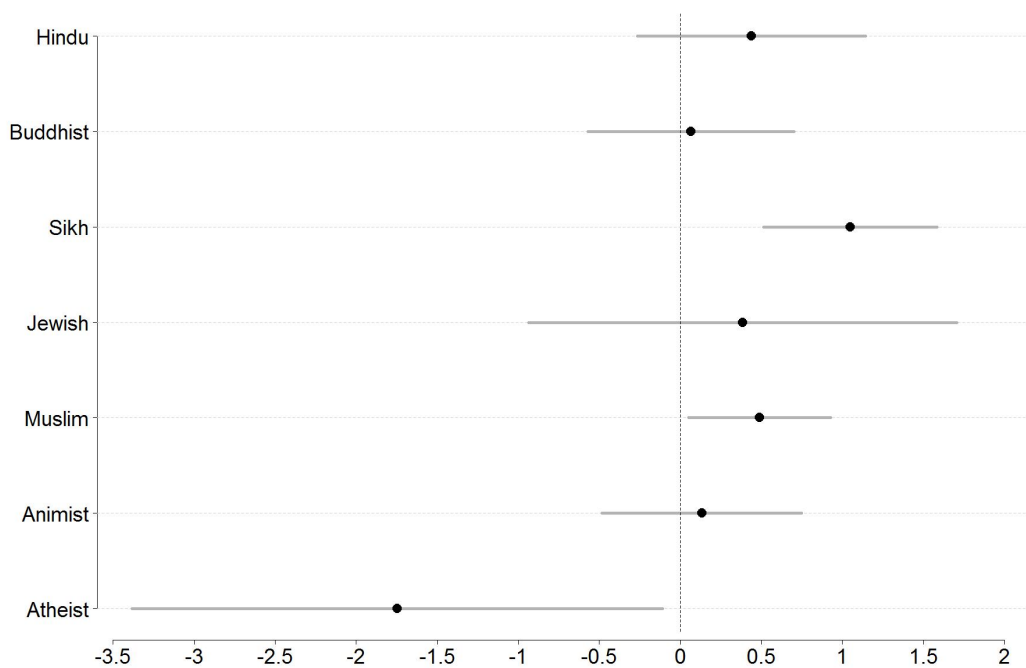


Figure 6: Estimated effects of various major religious identities in an ethnic group dyad on civil war onset. Grey lines represent 95% confidence intervals. Full regression model in Table C2.



## Tables

Table 1: Logit models of ethnic differences and civil war onset, 1946-2009.

DV: Ethnic Civil War Onset	(1)	(2)	(3)	(4)
Linguistic Difference (Mean)	1.150** (0.393)			
Religious Difference (Mean)	0.021 (0.276)			
Linguistic Difference (Dummy)		1.066*** (0.324)		
Religious Difference (Dummy)		-0.102 (0.201)		1.477*** (0.439)
Only Ling. Diff. (Dummy)			1.794*** (0.455)	1.754*** (0.448)
Only Rel. Diff. (Dummy)			1.020 (0.574)	
Ling. + Rel. Difference (Dummy)			1.583*** (0.448)	
Excluded	0.907*** (0.219)	0.935*** (0.219)	0.841*** (0.201)	0.833*** (0.201)
Downgraded	1.491*** (0.292)	1.491*** (0.293)	1.499*** (0.291)	1.509*** (0.292)
Groupsize	1.862*** (0.422)	1.791*** (0.425)	1.911*** (0.444)	1.769*** (0.435)
War History	0.710*** (0.079)	0.713*** (0.079)	0.719*** (0.078)	0.729*** (0.080)
Ln(GDP p.c.)	-0.253* (0.105)	-0.256* (0.105)	-0.263* (0.105)	-0.279** (0.103)
Ln(Population)	0.038 (0.103)	0.037 (0.103)	0.039 (0.102)	0.033 (0.104)
Peace Years	-0.081 (0.049)	-0.081 (0.049)	-0.080 (0.049)	-0.079 (0.050)
Constant	-5.237** (1.863)	-5.055** (1.872)	-5.638** (1.921)	-5.359** (1.938)
Cubic Splines	Yes	Yes	Yes	Yes
Observations	28375	28375	28375	28375
$\ell$	-1079.917	-1080.050	-1077.327	-1079.203
$\chi^2$	677.134	654.110	640.582	663.867

Country-clustered standard errors in parentheses.

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Table 2: Logit models of ethnic differences and civil war onset after the Cold War.

DV: Ethnic Civil War Onset	(5)	(6)	(7)	(8)
Linguistic Difference (Dummy)	1.123* (0.459)		1.124** (0.432)	1.129* (0.450)
Religious Difference (Dummy)	-0.237 (0.296)	1.461* (0.652)		
Only Ling. Diff. (Dummy)		1.861** (0.632)		
Civilizational Difference			-0.303 (0.323)	
Residual Rel. Difference				-0.757* (0.330)
Muslim Difference				0.050 (0.389)
Excluded	0.887** (0.294)	0.730** (0.266)	0.885** (0.287)	0.947** (0.302)
Downgraded	1.519*** (0.363)	1.542*** (0.363)	1.506*** (0.360)	1.546*** (0.359)
Groupsize	2.028** (0.640)	1.842** (0.572)	1.981** (0.625)	1.981** (0.716)
War History	0.341** (0.122)	0.343** (0.124)	0.345** (0.114)	0.323** (0.116)
Ln(GDP p.c.)	-0.295** (0.109)	-0.306** (0.106)	-0.307** (0.114)	-0.318** (0.111)
Ln(Population)	0.215 (0.115)	0.209 (0.115)	0.228* (0.112)	0.224 (0.118)
Peace Years	-0.018 (0.052)	-0.014 (0.052)	-0.017 (0.053)	-0.015 (0.048)
Constant	-6.882** (2.117)	-7.237*** (2.085)	-7.047*** (2.039)	-6.853** (2.195)
Cubic Splines	Yes	Yes	Yes	Yes
Observations	11187	11187	11187	11187
$\ell$	-534.229	-533.061	-533.875	-528.425
$\chi^2$	160.661	146.133	167.378	180.567

Country-clustered standard errors in parentheses.

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$