**Ethno-Nationalist Dyads and Civil War: A GIS-Based Analysis**

Lars-Erik Cederman, ETH  
Halvard Buhaug, PRIO  
Jan Ketil Rød, NTNU

**Abstract (150 words)**

Previous quantitative research on ethnic civil war relies on macro-level proxies in an attempt to specify the conditions under which ethnic minorities rebel. Going beyond an exclusive focus on minorities, the present study employs Geographic Information Systems (GIS) as a way to model ethnic center-periphery dyads that confront governments with excluded groups. We construct and analyze a new dataset of georeferenced politically relevant ethnic groups, covering the entire world during the period from 1951 through 2005. Our results show that the conflict probability of marginalized groups increases with the demographic power balance compared to the group(s) in power. Further, the risk of conflict increases with the distance from the group to the capital, and the roughness of the terrain in the group’s settlement area. We also find that while the results for demographic group strength holds for all ethnic civil wars, the geographic factors apply for territorial ethnic conflicts only.
Is there a link between ethnicity and civil war? This question remains a controversial one in the quantitative literature. While recent “revisionist” research has come to the conclusion that ethnic grievances do little to explain civil wars, others attempt to establish under what conditions ethnic minorities rebel. Arguing that the controversy largely stems from over-aggregated and/or conceptually problematic research designs, we propose a model that offers a more detailed, sub-national analysis with the help of geo-coded data of ethnic groups that goes beyond the exclusive focus on minorities. Geographic Information Systems (GIS) allow us to study conflict in explicit ethnic center-periphery dyads that confront governmental actors with excluded ethno-nationalist groups.

Our analysis improves on our previous attempts to investigate ethno-political mechanisms of conflict along three dimensions. First, as opposed to previous dyadic studies of civil-war onset that have been limited to Eurasia and North Africa because of data limitations (research note by AUTHORS), this study covers the entire world. Second, whereas our previous research has relied on extensive ethnographic lists of groups regardless of their political role, the present study codes political relevance of ethnic groups explicitly. Third, we investigate to what extent our main hypotheses apply differently to territorial and governmental conflicts.

In order to extend and improve the coding of politically relevant ethnic groups and their power access, we rely on information from a new dataset on “Ethnic Power Relations” (EPR). This data source allows us to select a politically relevant subset of geo-coded groups from the Atlas Narodov Mira (ANM) with a sample period covering the entire post-WWII era. Based on this data, we show that the probability that excluded ethnic groups experience conflict increases with their relative size compared to the governmental group or coalition, the distance to their countries’ capitals, and the roughness their settlement area’s terrain. While the demographic size effect holds for all ethnic civil wars, the geographic factors apply for territorial ethnic conflicts only. Moreover, our results indicate that ethnic groups have been more likely to engage in territorial conflicts after the end of the Cold War.

The paper is organized as follows. The next section presents a review of the literature on ethnic civil wars. Then we introduce the main hypotheses together with their theoretical justification, followed by an extensive discussion of the data sources and methods. The results are presented in two sections, the first one focusing on the main hypotheses and the second one on the robustness of these findings. We devote the concluding section to the theoretical conclusions that can be drawn from this study and possible future extensions.
Theorizing ethno-nationalist actor configurations

There are several alternative ways of linking ethnicity to violence. Rather than covering an exhaustive list of explanations, this section focuses on how to theorize ethno-nationalist configurations. Further elaborating our exclusion perspective, the following section introduces dyadic mechanisms that explain the outbreak of violence.

In the 1960s and 1970s, conflict researchers paid more attention to socio-economic causes than to the ethnic balance of power, Gurr’s (1968) famous deprivation thesis being a case in point. Even his critics, who propose opportunity-based explanations, tended to favor materialist, rather than ethno-nationalist explanations of civil conflict (e.g. Tilly 1978). The result of this tendency was that at end of the Cold War, ethnicity had, with few exceptions (e.g. Horowitz 1985), almost entirely dropped from political scientists’ research agenda.

Warfare in Yugoslavia and elsewhere in the early 1990s came as a rude awakening to scholars and policy makers alike. A first scholarly reaction to these events was a wave of theorizing that views the outbreak of violence as a domestic security dilemma along the lines that had already been proposed by neorealist and rationalist IR theorists (e.g. Posen 1993). According to this perspective, mistrust among ethnic groups triggers conflict in the wake of state collapse (Hardin 1995). Framed as a critique of journalistic and nationalistic conceptions of primordial “ancient hatreds,” these approaches are right to question the historical and psychological determinism in popular accounts of the war, but the security-dilemma logic renders the actor configuration too symmetric and the outbreak of violence too automatic (Gagnon 1994/95; Brubaker 1998).

Another important body of research that emerged in the early 1990s analyzes the conditions under which ethnic minorities rebel. Coming closer to capturing the asymmetry of ethnic civil wars, these studies build on the older grievance tradition while analyzing ethnic conflicts in particular. Drawing on Gurr’s Minorities at Risk (MAR) dataset, these scholars explore conflict-inducing mechanisms relating to political discrimination (Gurr 1993; 2000; Regan and Norton 2005; Walter 2006), concentration of settlement patterns (Toft 2003), domestic diversion policies (Tir and Jasinski 2008), third-party interventions (Cetinyan 2002; Saideman 2002), and states’ reluctance or willingness to grant autonomy (Brancati 2006; Jenne 2007).

Much better suited to studying ethno-nationalist civil wars, this research tradition disaggregates the analysis to the group level and accords the state a much more active role. Moreover, MAR has the advantage of highlighting politically relevant groups as opposed to the ANM’s broader sample. There is little use to include groups that are clearly irrelevant to the causal process to be evaluated (Posner 2004). However, with the
important exception for a variable coding concentration of groups’ settlement patterns (Toft 2003), the MAR dataset does not contain any explicit geographic information, which limits the hypotheses that can be tested (Posner 2004). Moreover, conceptual problems relating to the sampling frame of the MAR data complicates matters. First, by selecting out minorities rather than considering the groups’ access to state power, this data source makes it difficult to analyze situations involving ruling minorities or complex coalitions of ruling elites. Second, the MAR dataset is not primarily structured to trace changes of minorities’ power status over time, a weakness that becomes particularly problematic where power shifts from one group to another through coups and other sudden power transitions.

More recently, several prominent political economists have tended to downplay the importance of ethnicity for conflict patterns altogether. Whether interpreted as a type of grievance (Collier and Hoeffler 2004) or primarily as demographic diversity (Fearon and Laitin 2003), this scholarship views ethnicity as a surface phenomenon that conceals more fundamental material causes of conflict (for a review, see Kalyvas 2007). Most scholars who contribute to this research tradition operationalize ethnicity by relying on the ethno-linguistic fractionalization index (ELF). It remains unclear, however, how, and whether at all, this measure is related to conflict. Based on the Herfindahl index, it measures the likelihood of two individuals in a country not belonging to the same ethnic group (Taylor and Hudson 1972). Yet, ethno-nationalist civil wars clearly are not the aggregated effects of individual-level processes. Nor can they be reduced to notions of polarization (Montalvo and Reynal-Querol 2005) or simple measures of “ethnic dominance” based on the size of the largest ethnic group (Collier and Hoeffler 2004). We argue instead that they should be explained as genuinely political conflicts over the control of the modern state that feature state elites on one side and ethno-nationalist challengers on the other.

In the following, we propose a group-based model of ethno-nationalist conflict that is based on a star-shaped actor constellation and corresponding conflict mechanisms connecting the actors with each other. We interpret ethno-nationalist conflict as the consequence of one particular path of nation-state formation. It occurs where the elites of nationalizing states were not resourceful enough to include and integrate the vast majority of the population into the imagined community of the nation. Especially in societies where no networks of civil society organizations were available, they relied on ethnic clientelism to mobilize a political following, thus excluding ethnic others from access to power and public goods (Brass 1991; Wimmer 2002).

Whereas nation-building in strong states with well developed civil societies incorporated large parts of the population, sometimes with the help of widespread ethnic cleansing
(Mann 2005), the situation in weaker states, mostly in the previously colonized parts of the world, was quite different. Exclusion along ethnic lines often led to political mobilization by counter-elites, which denounced their under-representation at the center of power and demanded inclusion (e.g. the Hutu in Rwanda), or the creation of a separate state in which their own ethnic group would represent the dominant nation (e.g. the Bengalis of former East Pakistan, now Bangladesh).

By definition, civil wars confront an incumbent government with one or more challengers. In keeping with Tilly’s (1978) “polity model,” we assume that a state consists of a government and a number of contending groups, some of which are included in the polity, and others that are excluded from executive power at the national level. Because this study focuses on ethno-nationalist conflict, we conceptualize the former as “ethnic group(s) in power” (EGIPs) and the latter “marginalized ethnic groups” (MEGs).3

Based on this specification, it is now possible to conceive of conflicts in relational terms. Rather than analyzing conflict as a general state-level condition, we introduce the notion of a center-periphery dyad that is composed of an EGIP at the center and a MEG in the periphery. For the time being, we assume that the EGIP can be composed of more than one group, but that MEGs do not form coalitions in opposition to the center.4 Whereas some countries, such as North or South Korea, are almost perfectly homogenous and thus feature no center-periphery dyads, others, as for example India and Indonesia, exhibit massive ethnic diversity and feature dozens of dyadic relations. Given this actor configuration, we now need to specify under which circumstances dyadic conflict is likely to erupt.

**Postulating dyadic conflict mechanisms**

Far from every case of political exclusion provokes an ethno-nationalist civil war. If our exclusion perspective is correct, we would expect powerful ethnic groups excluded from power to be most likely to mobilize around an ethno-nationalist program and to initiate conflict against the government, and conversely, the government to engage in repression to curb the power of such threatening contenders. Thus, we postulate that collective ethno-nationalist action requires both opportunities and willingness to challenge central state power (Most and Starr 1989; Starr 1978). Civil wars occur when peripheral contenders are powerful enough to challenge the center and sufficiently motivated to do so (Gurr 2000). In this sense, we part company from attempts to artificially prioritize either opportunities or willingness over the other. For instance, Collier and Hoeffler’s (2004) catchy formula “greed and grievance” associates greed with opportunities and grievance with motives. Discounting the causal relevance of motives, Collier and Hoeffler assert that “all societies may have groups with exaggerated grievances” (p. 564).
In their view, rebel movements’ opportunities to stage collective action determine the outbreak of civil wars: ‘‘opportunity’’ and ‘viability’ describe the common conditions sufficient for profit-seeking, or not-for-profit, rebel organizations to exist’’ (p. 565). Other researchers, building on the legacy of relative deprivation and related psychological perspectives, give priority to grievances over opportunities (e.g. Regan and Norton 2005). In the following, we elaborate on how the proposed mechanisms are constituted and how they can be measured.

For operational reasons, and without any claim to completeness, we focus on two crucial aspects of ethnic center-periphery relationships which we expect to systematically influence both the peripheral group’s opportunities and willingness to rebel: (1) the demographic balance between the government and the peripheral ethnic group, and (2) the ethno-geographic constellation – especially with respect to the distance between the capital and the peripheral group – and the roughness of the terrain within the group’s settlement area. Figure 1 illustrates how the empirical indicators, through four causal mechanisms, are connected with the key factors that are expected to generate ethno-nationalist civil wars.

[FIGURE 1 ABOUT HERE]

We first show how the demographic balance influences excluded groups’ ability and propensity to challenge the state’s power through the resource mobilization and popular representation mechanisms respectively. Then we turn to the impact of ethno-geographic factors, which is mediated by the mechanisms pertaining to the logistics of insurgency and peripheral nationalism.

Demographic mechanisms driving ethno-nationalist conflict
We are now ready to state our first operational hypothesis:

H1. The probability of conflict increases with the relative demographic size of the excluded group.

Other things being equal, we expect that larger groups will be able to stage successful collective action thanks to their superior numbers. It is not hard to find support for this claim in the general literature on rebellions and civil wars. For example, DeNardo (1985, pp. 35-36) writes that “Regardless of the political context, there always seems to be power in numbers ... Accordingly, we take the disruptiveness of protests, demonstrations, and uprisings to be first and foremost a question of number.” In a general sense, this demographic variable proxies many of the mechanisms that are central to resource mobilization theory (McCarthy and Zald 1977; Tilly 1978), which emerged as an explicit
challenge to Gurr’s original relative deprivation school and other motivation-based explanations.

This resource mobilization mechanism applies readily to ethno-nationalist collective action as a special case of the broader category of political mobilization. Quite contrary to Fearon and Laitin’s (2003) putative test of “nationalism theory” that links conflict to high degrees of fractionalization, the exclusion perspective purported here expects less conflict in highly fragmented settings. To put things in explicitly relational terms, what matters is the relationship between the state and the central groups. As the doyen of the study of ethnic politics has stated, “whether ethnic politics is more parochial or more central is mainly a function of groups size relative to state size” (Horowitz 1985, 40; see also Deutsch 1953; Hannan 1979). ⁵

It would be a mistake, however, to expect group sizes to translate automatically into successful collective action. Faced with the enforcement of centralized states, such endeavors have to overcome acute collective-action dilemmas that may thwart even large groups’ mobilization efforts (Lichbach 1995; Hechter and Okamoto 2001). However, as suggested by Kalyvas and Kocher (2007), the collective-action literature exaggerates these problems by overestimating advantages of non-participation. ⁶

At this point, our attention has already started to shift from factors related to opportunity to those linked to willingness. Clearly, numbers have an impact not only on an excluded group’s resources, but also on the motivation to take on the center. We therefore posit a complementary mechanism of popular representation. In brief, we expect that claims to power access should gain legitimacy as a function of group size. This argument is particularly evident in democracies, because of the basic norm granting equal rights to all individuals according to the formula “one man, one vote.” In addition, it is well-known that permanent exclusion of significant minorities poses a serious problem of legitimacy to any democratic system (Horowitz 1985, 86). While technically compatible with strictly majoritarian interpretations of democracy, systematic exclusion along ethnic lines is likely to provoke ethno-nationalist claims and bring forth polarization of party systems along ethnic lines, subsequent ethnic outbidding, and finally, political violence as a likely outcome.

Even in non-democratic societies, however, the political exclusion of demographically large ethnic groups poses severe problems of legitimacy. All modern nation-states, independent of political regime type, are built on the principle of ethnic representativity according to which ethnic minority rule must be perceived as a scandal (Petersen 2002, 51; Wimmer 2002, chap. 3). Even in the absence of democratic elections, norms of proportionality between group size and political power are widely held. The situation
becomes especially acute in cases of suppressed majorities, such as the black population in South Africa under Apartheid, or pluralities, such as the Shiites in Iraq before the fall of Saddam Hussein’s regime. In these cases, a pressure for democratization should increase with the size of the excluded group. At the same time, mass mobilization leading to a dramatic shift in the ethnic balance of power may provoke conflict (Deutsch 1953; Mansfield and Snyder 2005).

In sum, it can be concluded that an excluded group’s size is likely to affect both the opportunity and willingness to challenge the center’s power through the mechanisms of resource mobilization and popular representation, respectively. Because both mechanisms link ethnicity to conflict, we anticipate group size to be strongly related to our dependent variable.

Geographic mechanisms driving ethno-nationalist conflict

We have seen that demographic group size serves as an indicator of resources that enable mobilization and at the same time gives mobilization political legitimacy. However, this hypothesis needs to be complemented, because in large states with difficult terrain, even relatively small groups may be able to wage surprisingly effective campaigns against powerful political centers, as the insurgencies in Aceh and Chechnya have demonstrated. On average, then, we would expect the prospects of peripheral challenges to the central government to be the most successful in cases where the latter’s reach is the least developed. This logic yields two straight-forward hypotheses:

H2a. The probability of conflict increases with the distance between the excluded group and the capital.

H2b. The probability of conflict increases with the roughness of the terrain within the settlement area of the excluded group.

Again, we argue that remoteness in these two senses supports excluded groups’ efforts to challenge the center’s power both in terms of opportunities and willingness to stage insurgencies. The former argument, which we refer to as the \textit{logistical insurgency mechanism}, is well represented in the contemporary literature on civil wars. Based on Max Weber’s (1946) classic concept of statehood, we expect the projection of state power to decline with increasing distances from the capital. This logistical challenge has been documented as the main factor driving the shift from “indirect” to “direct” rule in European history (Tilly 1990; Hechter 2001), but it also applies to non-European cases of state formation (e.g. Herbst 2000).
Capturing this idea mathematically, Boulding’s (1962) notion of the “loss-of-strength gradient” (LSG) postulates a declining function, often expressed as exponential decay. All power diminishes as it is projected across distance. Within a military-strategic framework, this can be exemplified by the need to protect supply lines. When the state moves away from its home base, resources are increasingly diverted to non-combat operations. Accordingly, more and more investment is required to ensure a sustained military presence on the battlefield. Conventional warfare, involving armored vehicles and heavy artillery, is severely hampered by mountainous terrain and dense forests.

For this reason, superior local knowledge and taking advantage of rugged terrain constitute the core of insurgent doctrine (e.g. Guevara 1969; Laqueur 1976). Advancing a similar argument, albeit supported by a crude country-level proxy measuring percentage mountainous terrain, Fearon and Laitin (2003, p. 75) postulate that “rough terrain, poorly served by roads, at a distance from the centers of state power, should favor insurgency and civil war.” More broadly, the LSG model applies to governance techniques beyond the strictly military realm, including communications, administrative control, and resource extraction (Tilly 1990; Mann 1986; Hechter 2001; Herbst 2000).

There is good reason to believe that distance and terrain influence not only opportunities, but also the willingness to engage in collective action. We therefore postulate a final peripheral nationalism mechanism, according to which ethnic groups that are far removed from the political center and live in inaccessible territories hold more hostile attitudes towards central rule than those who have been more thoroughly socialized to tolerate central control. In this regard, distance and remoteness can be thought of as indicators of heterogeneity in preferences between the center and periphery.

Adopting such a perspective, Stein Rokkan insists that “any analysis of variations among political systems must start from notions of territory” and proposes a comprehensive analysis of center-periphery relations:

A periphery is located at some distance from the dominant centre or centres, and its transactions with the latter are fraught with costs. A periphery is also different from the central areas on one or more scores: while the degree of distinctiveness will vary, being to some extent a function of distances and dependence, there will invariably be some minimum level and sense of separate identity. Finally, a periphery is dependent upon one or more centres … in political decision-making, in cultural standardization, and in economic life (Rokkan 1999, 115).

Similarly, studies of European nationalism have established that cultural state penetration tends to be slower far away from the capital and especially in rough terrain. For example,
Eugen Weber (1976) demonstrates that it took French nation-builders an entire century to turn France’s territory truly French and that this process dragged out the longest in the geographic periphery.

The difficulties of projecting national identities are not limited to European history, however, as illustrated by Weiner’s (1978) notion of the “sons of the soil.” This conflict pattern captures situations in which members of the dominant ethnic group settle in the culturally distinct and typically geographically remote periphery of the country in question because of state inducements or other economic opportunities. This migratory pressure triggers nationalist mobilization among the ethnically distinct local populations, i.e. the “sons of the soil.” This dynamic can be exacerbated by government policies that favor the settler population. Such policies are often perceived as a threat to the identity of the newly mobilized peripheral group, with center-periphery conflict as a likely consequence (as has been the case in e.g. Bangladesh, Burma, Northeastern India and Indonesia).

The mechanism of peripheral nationalism is more general than that of Weiner, however, which depends on a particular migratory pattern and should therefore be seen as a special case of the exclusion perspective we advance here. According to Hechter (2001), peripheral nationalism emerges as a reaction to the imposition of direct rule by a state-building center on areas that have previously enjoyed wide-ranging autonomy. In such cases, repression can trigger violent bids for secession that tend to be more effective in peripheral locations.

Relevance of mechanisms according to conflict type

In order to get a clearer picture of our causal mechanisms, we consider two main types of ethno-nationalist civil wars according to rebel aim. While the challenger’s goal is to seize power from the government directly in governmental conflicts, territorial ones concern sovereignty over a part of a state’s territory. There are good reasons to believe that conflict processes operate differently depending on the type of conflict. Because governmental civil wars are harder to win for non-state organizations, and because only demographically large groups can claim national power with a reasonable sense of legitimacy, we would expect weaker contending actors to seek territorial autonomy or independence for their home territory (AUTHOR).

Following from this general principle, we would expect compensating geographic factors, such as distance and terrain, to be of special value in territorial conflicts only. The weaker the contending group relative to the state, the more it needs to exploit contextual opportunity factors that can shift the balance of power in their favor. From this follows that our ethno-geographic mechanisms (H2a,b) should apply to territorial conflicts rather
than to the universe of ethno-nationalist civil wars. In contrast, there is no reason to believe that the demographic basis of power, as stipulated by H1, would differ in these two configurations.

In short, we propose the following qualification of our geographic mechanisms:

H3. The effect of the geographic factors (H2a,b) only applies to territorial conflicts.

Beyond these geographic conditions, under what circumstances are territorial conflicts more likely to erupt? In general terms, the principle of nationalism implies that ethno-nationalist conflict follows in the wake of geopolitical upheavals (Wimmer and Min 2006). Given that nationalists oppose foreign rule, shifts from indirect to direct rule tend to prompt violent reactions (Hechter 2001). In particular, nationalizing states that attempt to boost ethno-nationalist cohesion in the wake of imperial collapse can be expected to generate secessionist conflict (Brubaker 1996).

While this effect is not limited to the post-Cold War period, we expect the collapse of ideological polarization to have powerfully increased the likelihood of secession. This effect was most strongly felt in, but not limited to, the Former Soviet Union (Zürcher 2007).

We thus introduce our last hypothesis:

H4. Territorial ethno-nationalist conflicts are more likely during the post-Cold War period.

To sum up, we argue that through the mechanisms of insurgency logistics and peripheral nationalism our ethno-geographical variables contribute to both resources and motivations that drive ethno-nationalist insurgencies. Moreover, we expect the effect of the geographical factors to be limited to territorial conflicts, where the asymmetry in power between the center and periphery is much more prominent.

Before turning to operational matters, it is important to note what the current model leaves out. As a first cut, our analysis focuses primarily on structural characteristic of the dyadic relations. For operational reasons, we have deliberately downplayed the actors’ characteristics, such as their internal cohesion and organization (e.g. Tilly 1978), degree of mobilization (e.g. Sambanis and Zinn 2005), and access to weapons and raw materials (e.g. Ross 2006). Because another paper in this special issue operationalizes Toft’s (2003) notion of settlement patterns using GIS-data, we refrain from including this factor.
in our analysis.\textsuperscript{7} Nor have we attempted to classify specific types of grievances, or for that matter, previous histories of violence (e.g. Petersen 2002). Furthermore, it can be expected that political institutions influence the likelihood of conflict, including opportunities for peaceful protest (e.g. ARTICLE in this issue), democratization processes (e.g. Mansfield and Snyder 2005), and ethnic federalism (e.g. Bunce 1999; Brancati 2006). Finally, the current study abstracts away from any transnational and international influences, such as support from kin in neighboring countries or international intervention (e.g. Weiner 1978; Brubaker 1996; Saideman 2002).\textsuperscript{8}

While believing that all the aforementioned factors are potentially important as predictors of ethno-nationalist civil wars, we have to defer their theoretical incorporation into our model to future research. For now, we focus on testing our simplified model along the demographic and ethno-geographic dimensions as specified in the following section. If our exclusion theory is strong enough, it should be statistically discernible despite the presence of confounding factors.

**Operationalizing ethno-nationalist dyads**

By measuring ethnic dyads directly and introducing an explicitly geographic dimension, the current study provides a more fine-grained picture of the mechanisms driving ethno-nationalist wars than do most other quantitative studies. Our goal is to study conflict between ethnically defined state authorities, i.e. the “ethnic group(s) in power,” and their challengers. As we will argue, the spatial dimension is crucial, because it helps us not only to identify and locate the ethnic groups, but also to estimate their demographic sizes.

To succeed in our endeavor, we thus need information on (1) the identity and location of ethnic groups, (2) demographic group sizes, (3) ethnic group(s) in power, (4) geo-coded data on distances and terrain, and (5) ethnic dyadic conflicts.

In the following, we describe our data-collection efforts before turning to a presentation of the results in the subsequent section.

(1) Geocoding ethnic groups

In the literature reviewed thus far, ethnicity typically enters the analysis as quantitative indices, such as the ELF and polarization scores. Strikingly, spatial configurations play no role in the causal arguments underpinning these measures (Posner 2004). As argued above, the MAR-based literature analyzes the influence of settlement concentration, but such information is usually narrowed down to a single, dichotomous variable (Gurr 1993; Toft 2003). To our knowledge, there is no ready-to-use dataset that systematically pins down the location of ethnic groups in a large number of comparable cases. In quantitative
studies of ethnic conflict, the exact location of the key actors thus remains undetermined and overlooked.

This raises the question from where such information could be drawn. There are several candidates: Linguists have developed detailed maps of language diffusion, most notably through various editions of the *Ethnologue* (Gordon 2005). However, its linguistic charts are typically too detailed to serve as a guide to ethnic group delimitation (for example, the *Ethnologue* catalogues 742 unique languages in Indonesia alone). Another possibility would be to infer the location of ethnic groups from census or survey data (e.g. Laitin 1998). Yet, such an approach is only viable if the data contains references to ethnicity, which is often not the case. Furthermore, it also hinges on the presence of a reasonably fine-grained provincial structure. Where federal subunits are large, the needed degree of spatial disaggregation may never be attained.

For the aforementioned reasons, we have chosen to focus on the well-known *Atlas Narodov Mira* (*ANM*) (Bruk and Apenchenko 1964). The *ANM* stems from a major project of charting ethnic groups undertaken by Soviet ethnographers in the early 1960s. Their efforts bore fruit in the extensive but still un-translated atlas, covering the entire world. It contains 57 ethnographic maps at various scales that jointly cover all inhabited parts of the world. In the maps, the ethnic groups are represented by different color shadings and unique symbols. Impressively, the colors are carefully selected so that similar shades indicate ethnic affinity. Most areas are coded as pertaining to one group only, but in some cases up to three ethnic groups are indicated as sharing the same territory. Although the *ANM* could be criticized for being outdated (and rightly so in a number of cases, which we discuss below), it is nonetheless the best available source for our purpose. Crucially, it employs a consistent classification of ethnicity with a uniform group list that is valid across state borders (e.g. Kurds in Iraq and Turkey are recognized as belonging to the same ethnic group).

Transforming the ethnic atlas into an appropriate digital format involved several non-trivial steps. First, the maps were digitized through scanning. The next task involved aligning the maps with underlying GIS country shapes for the contemporary period (1964), thereby attaching geographical coordinates to the ethnicity data. Once a map was successfully referenced, group polygons were replicated from the original map by means of screen digitizing (Longley et al. 2005). In doing so, the coder had to make sure that the polygons completely covered the country's territory, and that there were no overlapping polygons or gaps in-between. Note that some groups are represented by several polygons as they have multiple separate population clusters surrounded by populations of other ethnic origins. Finally, all legends and group names were translated from Russian into English and entered into the dataset’s attribute table with links to the corresponding
group polygons. The resulting dataset contains spatial information on the location of more than 1,600 ethnic groups identified in the *ANM*. Figure 2 compares an original *ANM* map page of the former Yugoslavia with corresponding group polygons in our GIS-based dataset.

[FIGURE 2 ABOUT HERE]

(2) Estimating group sizes
Having determined the geographic location of all ethnic groups, our next task is to construct a suitable measure of the power balance in the center-periphery dyads. Unfortunately, it is difficult to find direct indicators of group strength in such settings (though see # in this volume). As the best available option, we decided to use demographic proxies for capability and therefore needed reasonably reliable measures of group size. One possible source of such information would be the *ANM*. In addition to the map material, the *Atlas* contains group-size estimates, which have often served as a basis for the calculation of ELF values (e.g. Taylor and Hudson 1972). However, these statistics may not be ideal for our purpose because a number of states did not exist at the time of publication, such as the post-Soviet republics of Russia, Ukraine, Belarus, etc. How could one derive population statistics of ethnic groups in these states based on the Soviet data?

Fortunately, GIS provides a convenient solution to this dilemma. GIS renders possible complex computational operations involving several spatially arranged data layers. Assuming the ethnic map to be constant, we developed a spatial estimation method based on an intersection of territorial country masks, the group polygons, and population density maps. Using the boundaries in a country layer (shapefile) representing the Post-Cold War period as “cookie cutters,” we singled out the group polygons (or parts thereof) that fell within the borders of each state. We then intersected the ethnic group layer with a gridded population density layer from Columbia University (CIESIN 2005). This allowed us to measure the size of the population that fell within a given ethnic group’s “state-cropped” polygon(s) (see Figure 3). In cases of multiple groups per polygon, we divided the population figures evenly between the groups. By summarizing the population for all polygons belonging to an ethnic group and repeating this procedure for all cases we obtained population estimates for all groups in all countries in the *ANM*. In a further step, we divided these numbers by the total population for the corresponding country, which yielded a table of the relative population share for each ethnic group.

[FIGURE 3 ABOUT HERE]
To limit the computational effort, we decided to generate the spatially derived population shares for two years only; 1964, the ANM’s year of publication, and, 1994, which captures the post-Cold War situation. We then merged the two datasets, using the estimates from 1964 for countries that existed in that year and data for 1994 for younger states. Finally, by way of validating the spatially computed data, we compared our group size estimates from 1964 to the population data provided by the ANM. This simple test indicated a very high correlation, at $r = 0.99$. Validation of the 1994-based estimates posed more problems due to the difficulty of finding matching group data that can serve as a reference. Nevertheless, a comparison with the population statistics for the corresponding ethnic groups in Fearon’s (2003) dataset yielded a reasonably high correlation coefficient ($r = 0.93$).

(3) Identifying the ethnic group(s) in power
When the candidate list of the geo-referenced ethnic groups was established, two tasks remained before the sample of center-periphery dyads was complete: assessing all groups with respect to political relevance, and identifying the ethnic groups in power (EGIP) by country and year.

A number of ethnic groups mapped in the ANM, such as the Aleuts in the US and Jamaicans in the UK, bear no significant political meaning at the national level. To identify the politically relevant ethnic groups and their access to executive power, we merged the ANM group list with a new dataset on Ethnic Power Relations (EPR). The EPR dataset comprises data on politically relevant ethnic groups with a population of at least one million in all countries between 1946 and 2005. Because the EPR is based on a sustained collective coding effort that involved more than 100 expert coders, the quality of this source surpasses the largely time-invariant EGIP data used in previous research. Furthermore, unlike those studies, which are limited to Eurasia and North Africa, the EPR data covers the entire world and includes changes of governmental access during the sample period.

A group is considered politically relevant if it forms part of the national government, if it is mobilized at the national level or if it is discriminated by the government. The EPR dataset thus offers a fine-grained coding of ethnic groups’ role in national politics that goes well beyond previous attempts of this type in terms of conceptual sophistication. Because the EPR group list varies over time and deviates from the much more extensive list supplied by the ANM, a somewhat elaborate coding procedure was required when merging the data. In some cases, groups are named differently in the two datasets, in other cases, the EPR lists superset groups that comprise two or more groups identified in the ANM. For each such subgroup, we used various sources, such as Levinson (1998), Wikipedia, or geographic location to identifying the corresponding group in EPR. The
opposite case proved more difficult to code because it implies that cleavages present in EPR cannot be captured based on the ANM list. In most of these cases we coded the ANM group in accordance with the political status of the largest EPR group. Moreover, as the ANM tends to suppress religious cleavages, we decided to generate new polygons for the EPR groups for a number of Middle Eastern countries. Countries for which the EPR dataset indicated that ethnicity is irrelevant were dropped from the sample; by definition, they are exempt from risk of ethno-national conflict.

Our key independent variable – the power balance, $r$, between the EGIP and the marginalized group – is operationalized as the periphery’s share of the dyadic population. The power balance indicator is theoretically bounded between zero (the EGIP comprises the entire population) and one (everyone belongs to an excluded group). Naturally, most excluded groups are vastly inferior to the group(s) that holds national power; the mean $r$ score for the politically relevant dyads is 0.07. Only 50 excluded groups are demographically more powerful than the center in one or more years since 1946. To account for a highly right-skewed distribution and the assumption that a one percentage point change in power matters more for small than large groups, we apply a logarithmic transformation of $r$ in the empirical analysis below.

(4) Geocoding distances and terrain
All else being equal, large excluded groups are relatively more likely to be involved in conflict with the center (H1). But the context is rarely constant between cases. Above, we argued that power decays as it is projected across distance and in rough terrain. Accordingly, minority groups that reside in remote and inaccessible regions are better protected against government forces than are similar groups closer to the political core.

To capture the locational trait of the ethnic groups, we used GIS to extract their exact mid-point location and then calculated the distance from each centroid to the country capital. Groups that are present in two or more distinct areas (polygons) are represented by a population-weighted average distance measure. Some capitals shifted location during the sample period. For example, groups in former Soviet republics like Georgia are coded with the distance to Moscow for the period up to 1991, and the distance to T’Bilisi for more recent years. By default, groups belonging to the EGIP are coded with the capital city as their location. To reduce outlier influence and to account for an expected declining effect of distance with higher values, we take the natural logarithm of the distance variable.

The second geographic factor, mountainous terrain, was computed in a similar manner. Based on gridded mountain data (UNEP-WCMC 2002), we calculated the share of the two-dimensional area of each polygon covered by mountains. For groups represented by
multiple polygons, we weighted the average terrain values by the area of each polygon. The area-weighted mountain variable takes on values between zero (for groups in the plains, such as the Bretons in France) and one (for groups settled entirely in the mountains, for example the Ladakhs in India). To illustrate this procedure, Figure 4 displays an overlay of ethnic groups and mountainous terrain, as well as lines that denote the great-circle distances from the polygon centroids to the capital city.

(Figure 4 about here)

(5) Determining the onset of dyadic ethnic conflict.
Finally, we need group-level data on ethnic conflict. We selected the UCDP/PRIO Armed Conflict Dataset (Gleditsch et al. 2002) as our candidate list, and checked this dataset against earlier attempts to identifying ethnic civil wars (notably Fearon and Laitin 2003; Licklider 1995; Sambanis 2001) in order to identify likely ethnic conflicts. In addition, we conferred a number of case-specific sources to achieve further information on the characteristics of the conflicts. Note that we follow convention and code conflicts based on whether mobilization was shaped by ethnic affiliation. We do not claim that any or all of the included conflicts truly concern issues of ethnicity or national self-determination (though such questions are often part of the incompatibilities between the government and the excluded groups, an issue that we return to below).

When a list of plausible conflicts was established, the next step involved indentifying the various groups involved in each case. This is not a trivial task as information on the actors’ ethnicity is often sparse, and particularly so for conflicts early in the period. Moreover, some conflicts (e.g. in Guatemala, Indonesia, and the Philippines) effectively covered the home region of numerous small, often indigenous, groups, most of which may plausibly have contributed to shaping the mobilization structure of the rebellion without this being documented. This poses a dilemma: whether to code only those groups that positively took part on the opposition side or all groups that resided in the conflict area. The former alternative implies a bias towards excluding smaller groups whereas the latter most likely is too inclusive. In the end, we opted for a mixed strategy of selecting verified groups plus those most conceivable to have formed part of the rebellion, for example by being identified as politically active by the MAR project.

The binary dependent variable is operationalized as dyadic conflict onset, implying that only the first year is assigned a value of one. Years of ongoing conflict are dropped from the analysis as the marginalized groups by definition cannot have more than one conflict with the centre at any time. In cases where the rebellion represents several groups, an onset of conflict is coded for all implied parties. For the models where we compare the conflict types (H3), we use a three-category dependent variable that separates between
events where the rebels seek territorial self-determination (separatist conflict) and violent attempts to achieve governmental control (coup, revolution). Information on rebel objective was derived from the UCDP/PRIO incompatibility indicator.

In addition to the power balance and geographic measures, we introduce a few control variables. The first is the natural log of GDP per capita, drawn from an updated version (v. 4.1) of Gleditsch (2002), which we expect to be negatively associated with the risk of dyadic ethnic conflict. Given that these data are unavailable prior to 1951 (applying a one-year time lag), the temporal domain of the empirical analysis is 1951–2005. We also control for regime type and regime instability, derived from the Scalar Index of Polities, SIP (Gates et al. 2006). The SIP scale varies between zero (strongly authoritarian) and one (pure democratic), and is less problematic within the context of conflict onset than the Polity scale (see Vreeland 2008). The democracy dummy (one-year lag) marks off countries with a SIP score above 0.75 while the anocracies are operationalized as having SIP scores between 0.25 and 0.75. The instability proxy measures the number of years since the last significant change (>0.1) in the SIP score. In addition, we include a dummy for the post-Cold War period (1989–2005) and Beck, Katz and Tucker’s (1998) cubic splines approach to handle duration dependence. All reported models are estimated by robust logit and multinomial logit regression with standard errors clustered on countries.

**Results**

To facilitate interpretation, we limit the presentation of the results to two sets of models: a standard logit analysis of dyadic conflict onset and a multinomial analysis of territorial and governmental conflict. Both analyses are shown with two specifications – full sample and a sample limited to politically relevant dyads.

For theoretical reasons discussed above, we put most emphasis on the results obtained from the reduced models, which also constitute a more conservative evaluation of the hypotheses. Table 1 displays the results for the standard logit models. As is evident, the alternative specifications generate quite similar results, implying that our conclusions do not hinge on a particular sample selection. Importantly, we see that our proxy for relative strength of the marginalized group is a very powerful correlate of conflict risk. In marginal terms, an increase in \( r \) from the 5th to the 95th percentile value for the median politically relevant dyad is associated with a tenfold increase in the estimated risk of conflict. Hypothesis 1 is thus strongly supported.

Table 1 also suggests that relative location matters. Groups in the remote periphery and groups with their core settlement area in the mountains are more likely to fight the center than are other marginalized populations, even though some parameter estimates misses
the conventional 95% significance threshold. The relative effect of distance is also less pronounced than that for dyadic demographic balance; a 5th to 95th percentile change roughly doubles the risk of conflict, holding all other covariates at their medians. The mountainous terrain parameter has a substantial impact on the estimated risk in Model 1 but the estimate drops considerably in both size and significance in the reduced model. This is not surprising, given that rough terrain is presumably most important as an opportunity factor for the weakest groups, some of which are considered politically irrelevant in the EPR data and excluded from Model 2. Be that as it may, we conclude that Hypothesis 2a is supported while the evidence for the general expectation expressed in Hypothesis 2b is much weaker.

None of the regime controls have significant effects on the risk of dyadic ethnic conflict. As expected, the GDP per capita estimate mirrors previous country-level findings of a negative association with conflict onset, although its performance is comparatively weak and sensitive to sample selection. The post-Cold War dummy and the peace years count successfully captures important temporal patterns in the data. The dramatic change in the world order around 1989 has acted as a knock-on risk factor for subsequent years while the likelihood of dyadic conflict decreases as a function of time since independence and time since the end of the previous conflict.

Next, keeping in mind that our findings are restricted to ethnic civil wars only, we turn to the multinomial logit regression, exploring to what extent relative capacity and geography explain variations in territorial and governmental conflicts (Table 2). Again, we show results from the full as well as the sample limited to politically relevant dyads, and again, the alternative models correspond well. In both specifications, we find substantial differences between the conflict types, as anticipated by Hypothesis 3. Note the poor performance of the covariates for the governmental conflicts. Apparently, the timing and location of this particular form of conflict is not well understood. The only factor that is robustly significant across samples and conflict types is relative dyadic capacity. Demographic power is an important determinant for all military mobilization. The substantial increase in the regression estimate for governmental conflicts provides an additional insight: the strategic objective of the group is influenced by its population size: Larger groups are relatively more likely to aim for national government power.

Models 3–4 also show how geography differs in its impact on the conflict types. In line with Hypothesis 3, distance and terrain are much more important in explaining separatist conflict than conflict over state power. Indeed, the coefficient estimates for the latter conflict type are all negative, suggesting an opposite effect, if any. This supports our
notion that having a remote and inaccessible location is disadvantageous if the group seeks to topple the regime but may be a vital facilitating factor for mounting separatist rebellion. These results thus call for a specification of Hypotheses 2a and 2b to apply only to self-determination conflicts.

The controls for regime type perform poorly and add little to the models, contrasting the reputed democratic civil peace (e.g. Hegre et al. 2001). However, readers should be reminded that this analysis is conducted at the dyadic level and only covers conflicts organized along ethnic lines. Additionally, the UCDP/PRIO conflict data apply a much lower casualty threshold than comparable datasets. The evidence of a democratic civil peace is primarily derived from analyses limited to major civil wars. Regime instability, in contrast, is a statistically significant predictor of government-related ethnic unrest, and the effect is significant also in substantive terms. For the median politically relevant dyad, a change from the 5th percentile to the 95th percentile value on the instability proxy (amounting to a shift from zero to 111 years since last regime change) increases the estimated risk of governmental conflict by an order of magnitude. Interestingly, per capita income is generally unrelated to the risk of regime conflicts while reducing the risk of ethnic separatist conflict at the dyadic level.

The controls for time and conflict history further demonstrate the differences in correlates of civil war outbreak between the two conflict types. In confirmation of Hypothesis 4, the collapse of the Cold War system led to a four-fold increase in baseline probability of territorial conflict while the risk of ethnic governmental conflict has remained unaffected. The territorial conflicts also exhibit considerably more inertia – and thereby predictability – by being more likely shortly after the end of the previous violence in the dyad.

[ TABLE 2 ABOUT HERE]

Because interpretation of effects in logit models is hardly intuitive, Figure 5 visualizes how the predicted probability of conflict varies with the size of the excluded group (calculated from Model 2). Three trend lines are shown: the lower solid line represents the marginal effect of $r$, holding all other covariates at their median values, the middle, dashed plot represents a similar dyad with the distance increased to the 95th percentile value, and the upper, dotted line shows the effect of $r$ the for median dyad with both distance and mountains at the 95th percentile. Politically excluded minorities that face vastly superior EGIPs are not likely to rebel, but the risk of conflict increases markedly with the relative size of the marginalized group. Moreover, location contributes considerably to the underlying risk, implying that relatively sizable groups that reside in rough terrain and at a considerable distance from the capital (such as Kurds in Iran and
Bengalis in Pakistan pre-1971) are the most likely to challenge the center by violent means.

[FIGURE 5 ABOUT HERE]

**Robustness tests**

The results presented correspond well with expectations and add to our understanding of ethnic conflict. Yet, some additional tests might be in order as an evaluation of the sensitivity of the main findings. For example, the demographic size of the EGIP is arguably a suboptimal indicator of state capacity. In tests not shown, we generated an alternative $r$, operationalized as the group’s population relative to the CINC score of the country. This measure actually performed slightly better than the original index and it also improved the fit of the overall model. Ideally, we would have liked to test alternative measures of group strength as well, but while new data on rebel strength exists (see other contribution to the special issue), there are no comparable statistics of (latent) military capacity for non-mobilized groups.

Second, armed conflicts are not randomly distributed across space but cluster in certain geographic areas and are absent in others. This raises the question of whether certain regions have a significantly different baseline risk of armed conflict even after taking into account relevant country characteristics. An alternative Model 4 with dummies for Eastern Europe, Latin America, Sub-Saharan Africa (SSA), North Africa and the Middle East, and Asia (Western Europe, North America, and Japan being the reference group) was not particularly illuminating. The only region to display a different constant was Latin America for territorial conflict. According to the UCDP/PRIO data, this continent has not had a single separatist conflict since WWII. With the exception of SSA, the result for $r$ is also robust to the exclusion of any single region.

A conventional method to model cross-sectional idiosyncrasies is to use a fixed-effects estimator. This procedure essentially adds one dummy variable per unit and allows for a unique intercept for each ethnic dyad. To assess the robustness of our findings we re-estimated Model 2 using a conditional logit model with fixed country effects. This led to a non-trivial reduction in the sample size, down one-third from 15,362 to 9,835 dyad years. The results for our variables of primary interest – dyadic power, distance, and terrain – are nonetheless essentially similar to the original results while some of the other covariates (notably income, which now is significantly positive) behave differently. However, the fixed-effects estimator should be applied with caution, in particular in settings with a binary response, as any unit without variation on dependent variable is
excluded by design. This is often problematic as the excluded cases may have avoided conflict precisely because of the attributes on the explanatory variables.

Conclusion
In this paper, we have investigated how exclusion of ethnic groups from governmental power increases the probability of ethno-nationalist civil war. Focusing on the dyadic relationship between excluded and included groups, our theory postulates that demographic and geographic mechanisms together shape opportunities and willingness for political violence. The statistical analysis provides ample evidence for a positive effect of relative size of the excluded group on the outbreak of conflict. The analysis further shows that the relevance of geography is limited to separatist conflicts, which have been especially prevalent after the end of the Cold War. We conclude that the size of marginalized ethnic groups vis-a-vis the center is a powerful determinant of ethnonational conflict, and that this demographic power balance, in combination with contextual geographic factors, determine the objective of the rebellion.

This paper offers a number of novel methodological contributions. It presents and uses a new geo-referenced dataset on the location, size, and political relevance of all ethnic groups across the global since 1946. The applied disaggregated research design, specifically capturing center-periphery power relations, constitutes a significant improvement on existing quantitative research that relies on cruder, country-level proxies. The explicit incorporation of relative location and terrain and is another novel feature, allowing us to test commonly held beliefs about the geography of ethnic conflict in a systematic fashion.

The present research consolidates earlier findings of a connection between ethnic exclusion and organized violence. Paying particular attention to politically relevant groups, we offer a tougher test of the exclusion perspective than has so far been conducted in the literature. Supported by MAR data, Gurr’s (1993; 2000) investigations leave the governmental side of the relationship partly open, and limiting the analysis to targeted groups also entails a potential for sample selection bias. Furthermore, previous research by AUTHORS excludes large parts of the world from the analysis, including Sub-Saharan Africa and the Western Hemisphere.

Our study does not disentangle the independent effect of motivations and capacity; yet, we believe that both are necessary for violence to break out. Because our results exhibit patterns that are broadly compatible with both grievance-related and resource-based arguments, there is no reason to believe that only one or the other set of factors is responsible for the outbreak of conflict. We hope that future research will be able to
measure more directly these effects and explain how they are intertwined. The research agenda centering on mobilizational mechanisms also needs to be extended from the present focus on marginalized groups to conflicts involving ethno-nationalist civil wars that pit previous power-sharing partners against other members of the governing coalition.

Other important issues that call out for further research include complex temporal and spatial effects. While we have controlled for some basic aspects of temporal dependency, group-level analysis opens the door to more precise investigation of the impact of past conflicts on present conflict behavior. Perhaps the most important case of spatial dependency centers on the influence of kin groups across state borders. It is well known that trans-border rebel sanctuaries that elude the reach of the state can affect the balance of power and motivation within the country in question (Salehyan 2007).

While we had to leave a number of threads dangling, we hope to have convinced the reader that disaggregated, GIS-based analysis provides a more detailed picture of causal mechanisms than country-level approaches allow for, thus representing a valuable complement to data-intensive micro-level studies that is currently enjoying a boom in the civil-war literature (Tarrow 2007). Although its level of detail is less impressive than in those studies, the disaggregated GIS-based approach is able to generate new knowledge about general conflict patterns around the globe.
References


Table 1. Determinants of Dyadic Ethnic Conflict, 1951–2005

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All dyads</td>
<td>Relevant dyads</td>
</tr>
<tr>
<td>Dyadic power balance(^a)</td>
<td>0.354</td>
<td>0.328</td>
</tr>
<tr>
<td></td>
<td>(6.42)**</td>
<td>(4.96)**</td>
</tr>
<tr>
<td>Distance to capital(^b)</td>
<td>0.396</td>
<td>0.185</td>
</tr>
<tr>
<td></td>
<td>(2.46)*</td>
<td>(1.50)</td>
</tr>
<tr>
<td>Mountain</td>
<td>0.537</td>
<td>0.277</td>
</tr>
<tr>
<td></td>
<td>(1.64)</td>
<td>(0.70)</td>
</tr>
<tr>
<td>GDP capita(^{a,b})</td>
<td>-0.056</td>
<td>-0.269</td>
</tr>
<tr>
<td></td>
<td>(0.42)</td>
<td>(1.73)</td>
</tr>
<tr>
<td>Democracy(^b)</td>
<td>0.114</td>
<td>0.289</td>
</tr>
<tr>
<td></td>
<td>(0.25)</td>
<td>(0.59)</td>
</tr>
<tr>
<td>Anocracy(^b)</td>
<td>0.415</td>
<td>0.354</td>
</tr>
<tr>
<td></td>
<td>(0.77)</td>
<td>(0.84)</td>
</tr>
<tr>
<td>Time since regime change</td>
<td>0.056</td>
<td>-0.099</td>
</tr>
<tr>
<td></td>
<td>(0.37)</td>
<td>(0.73)</td>
</tr>
<tr>
<td>Post-Cold War</td>
<td>1.276</td>
<td>1.154</td>
</tr>
<tr>
<td></td>
<td>(3.81)**</td>
<td>(2.94)**</td>
</tr>
<tr>
<td>Peace years</td>
<td>-0.315</td>
<td>-0.291</td>
</tr>
<tr>
<td></td>
<td>(5.90)**</td>
<td>(6.46)**</td>
</tr>
<tr>
<td>Constant</td>
<td>-4.749</td>
<td>-1.530</td>
</tr>
<tr>
<td></td>
<td>(3.84)**</td>
<td>(1.38)</td>
</tr>
<tr>
<td>Pseudo R(^2)</td>
<td>0.150</td>
<td>0.141</td>
</tr>
<tr>
<td>Log pseudolikelihood</td>
<td>-1,190.4</td>
<td>-676.3</td>
</tr>
<tr>
<td>Observations</td>
<td>56,058</td>
<td>15,362</td>
</tr>
</tbody>
</table>

Robust z scores, clustered on countries, in parentheses. Three natural cubic splines estimated but not shown. \(^a\) = natural log; \(^b\) = one-year lag. * significant at 5%; ** significant at 1%.
Table 2. Determinants of Dyadic Ethnic Conflict by Type, 1951–2005

<table>
<thead>
<tr>
<th></th>
<th>(3) All dyads</th>
<th>(4) Relevant dyads</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Terr</td>
<td>Gov</td>
</tr>
<tr>
<td>Dyadic power balance(^a)</td>
<td>0.288</td>
<td>0.539</td>
</tr>
<tr>
<td></td>
<td>(4.43)**</td>
<td>(5.04)**</td>
</tr>
<tr>
<td>Distance to capital(^b)</td>
<td>0.686</td>
<td>-0.076</td>
</tr>
<tr>
<td></td>
<td>(4.25)**</td>
<td>(0.29)</td>
</tr>
<tr>
<td>Mountain</td>
<td>0.855</td>
<td>-0.118</td>
</tr>
<tr>
<td></td>
<td>(2.66)**</td>
<td>(0.15)</td>
</tr>
<tr>
<td>GDP capita(^b)</td>
<td>-0.158</td>
<td>0.133</td>
</tr>
<tr>
<td></td>
<td>(0.88)</td>
<td>(0.70)</td>
</tr>
<tr>
<td>Democracy(^b)</td>
<td>0.570</td>
<td>-1.172</td>
</tr>
<tr>
<td></td>
<td>(1.17)</td>
<td>(1.57)</td>
</tr>
<tr>
<td>Anocracy(^b)</td>
<td>0.924</td>
<td>-0.349</td>
</tr>
<tr>
<td></td>
<td>(1.33)</td>
<td>(0.50)</td>
</tr>
<tr>
<td>Time since regime change</td>
<td>0.253</td>
<td>-0.372</td>
</tr>
<tr>
<td></td>
<td>(1.81)</td>
<td>(1.31)</td>
</tr>
<tr>
<td>Post-Cold War</td>
<td>1.660</td>
<td>0.621</td>
</tr>
<tr>
<td></td>
<td>(4.26)**</td>
<td>(1.24)</td>
</tr>
<tr>
<td>Peace years</td>
<td>-0.346</td>
<td>-0.226</td>
</tr>
<tr>
<td></td>
<td>(5.53)**</td>
<td>(1.78)</td>
</tr>
<tr>
<td>Constant</td>
<td>-7.348</td>
<td>-2.701</td>
</tr>
<tr>
<td></td>
<td>(4.64)**</td>
<td>(1.78)</td>
</tr>
</tbody>
</table>

Robust z scores, clustered on countries, in parentheses. Three natural cubic splines estimated but not shown. \(^a\) = natural log; \(^b\) = one-year lag. * significant at 5%; ** significant at 1%.
Figure 1. Mechanisms leading to ethno-nationalist civil wars
Figure 2. Geo-coding of Ethnic Groups as Polygons Based on the Atlas Narodov Mira
Figure 3. Estimation of Population Size for an Ethnic Group Polygon
Figure 4. Geodesic Distances and Mountain Data for Ethnic Group Polygons in Georgia
Figure 5. Estimated Risk of Conflict as a Function of the Periphery’s Relative Capability ($r$)

*Note:* The figure shows the estimated probability of dyadic conflict for a typical post-Soviet republic; i.e. middle-income, post-Cold War anocracy with last regime change two years ago and five years since previous conflict.
Endnotes

1. Given the small number of conflict cases compared to mobilized groups, and EPR’s explicit sampling rules that do not depend on conflict, the risk of endogeneity is minimal. Below, we present results for both the full and restricted samples.

2. Fearon (2003) criticizes MAR for its alleged sample selection bias due to its restriction to “minorities at risk.” However, our criticism targets the focus on minorities.

3. It should be noted that this specification reifies ethnic groups which renders the explicit tracing of intra-group processes impossible. For example, some ethno-nationalist groups, such as the Assamese in India, cannot easily be classified as either EGIPs or MEGs, because parts of their peripheral elites collaborate with the central authorities of the state while others fight for secession and independence. For critiques of group-level reification, see e.g. Brubaker (1998) and Kalyvas (2008).

4. Because of the difficulties relating to inter-ethnic cooperation without state enforcement, this is a reasonable assumption, although there are many cases in which ethnic groups splinter into factions, that may or may not cooperate with each other. See Hannan (1979) for a sophisticated argument relating the extent of opposition cooperation to the level of governmental oppression. Ultimately, additional empirical research is needed to establish the effect of peripheral coalitions.

5. Recent quantitative studies have found strong support that the size of ethnic groups is associated with outbreak of violence, see e.g. Walter (2006) and Roeder (2007).

6. Another qualification relates to situations in which members of ethnic rebel groups defect to the governmental side, see Kalyvas (2008). However, so far, little systematic evidence has been presented suggesting that this constitutes a regularity. Thus, we expect group size to be a good statistical indicator for mobilizational capacity.

7. Adding such measures to our models do not change significantly the findings reported in this paper.

8. In a forthcoming, GIS-based study that is based on similar data, AUTHOR et al. show that the presence of transborder ethnic kin contribute to shifting the power balance in the group’s favor within the center-periphery dyad.

9. Since we are dealing here with intrastate power balances and the relative distribution of population, using 1964 data for some countries and 1994 data for others does not pose significant problems.

10. Some population estimates deviate significantly, such as in the Ukraine and Azerbaijan. This is most likely due to extensive post-1964 migration that renders the polygons obsolete. Even so, the results presented below do not change substantially if we substitute the more questionable estimates with Fearon’s group data.

11. Access to national power is divided into three main categories according to (1) whether group representatives enjoy full control of the executive branch at the expense of all other groups in the country, (2) whether they share power with members of other groups in a
coalition including other ethnic groups, or (3) whether they are excluded from national decision making altogether. As this study focuses on exclusion only, the cutoff point was placed between the two first main categories and the last one.

12 See Jenness (2007).
13 Indeed, an interaction between dyadic power balance and mountains improves the performance of the mountain measure somewhat.
14 For a more sophisticated way of studying historical effects based on “enduring rivalries”, see e.g. DeRouen and Bercovitch (2008).
15 Composite Index of National Capabilities, from the Correlates of War project. The index gives a joint score for military, industrial, and demographic power for each country relative to all other countries in the system in the given year.