

The Failure to Ameliorate Grievances: Measuring Bargaining to Explain Intrastate Conflict

Matthew Ruppert
Undergraduate Thesis
Department of Political Science, Columbia University
April 3, 2024

Abstract

This paper explores the causes of civil war onset at the ethnic group level. I argue that economic inequality between ethnic groups that fuels grievances and that the ability of potential rebels to reach a bargain with the state to address these grievances without resorting to violence are key explanations for onset. While these conditions will impact the likelihood of civil war onset independently, I theorize that that they provide the most powerful explanation together because easy bargaining moderates the effect of grievances. Ultimately, I show that economic grievances and the moderating effect of bargaining on these grievances hold poor explanatory power. Instead, conditions that make it easier to reach a bargain, most importantly an improved ability for the state to collect information about domestic happenings, reduce the likelihood of intrastate conflict.

I would like to thank my thesis advisor, Page Fortna, for her guidance and thoughtful comments during this process, Martin Devaux for his crucial solutions to my issues with R and helpful feedback, and John Huber for asking difficult questions and providing structured support. I would also like to thank Sarah Daly and Gregory Wawro for their assistance with brainstorming and clarification at critical moments during this project.

The causes of civil war onset have been widely studied by political scientists. Civil wars have major ramifications for political systems and state behavior, cause widespread death and destruction, and make up the vast majority of political conflicts involving states since World War II (Davies, Pettersson, and Öberg 2023, 695). Understanding what causes this violence will allow policy makers, from the international to local levels, to make informed decisions to prevent it. Like many previous scholars, I take up the question of what conditions increase the likelihood that an ethnic group, or people claiming to represent an ethnic group, will engage in organized, politically motivated violence against the state?

I argue that economic grievances and the ability of potential rebels to reach a bargain with the state to address these grievances without resorting to violence are key explanations for onset. Economic grievances caused by a difference between actual and expected economic well-being drives individuals to feel anger towards the state, making them more likely to engage in violence to appropriate the state's power in order to improve the well-being of themselves and their ethnic group. Bargaining matters because the costs of war should incentivize both sides to strike a deal that will sufficiently ameliorate the grievances of potential rebels and avoid these costs, so onset will only occur when bargaining fails. This is most likely to occur after rebels misrepresent their potential fighting power in search of the best possible deal and the state is unable to verify their true power. This leaves the sides unable to reach a bargain that is mutually preferable to war. While these conditions will impact the likelihood of civil war onset independently, I theorize that they provide the most powerful explanation together. I argue that bargaining moderates the impact of economic grievances so that increased grievances matter little when bargaining is easy (because bargaining will prevent escalation to war) but that increased economic grievances will more dramatically increase the likelihood of conflict when

bargaining is difficult. While economic grievances and bargaining have received attention independently, combining the theories in this way is new. It is an important step for understanding what is actually being bargained over.

This paper also contributes a novel way of measuring economic grievances and the first the large-N quantitative test of the bargaining theory in the context of civil war. I measure economic grievances as income inequality between a given ethnic group and the ethnic group with the most political power in the state, providing a real reference group against whom potential rebels can be aggrieved and a clear path for anger to be directed against the state. Bargaining conditions are tested with two different measures of the state's general ability to collect information about its citizens, whether the state formally collects data on ethnicity, a measure of the ethnic group's ability to organize internationally, and a proxy for the internal cohesion of an ethnic group. Each measure is designed to capture conditions that make either the state or a potential rebel group more uncertain about the potential rebel group's capabilities or resolve; more uncertainty means bargaining is more likely to fail. Ideas about bargaining have been used to explain a variety of aspects of conflict, so this empirical test is a useful starting point to begin quantifying its effects. Using a worldwide dataset of all politically relevant ethnic groups from 1985 to 2010, I use a series of regressions to test the association between grievances, bargaining conditions, and the interaction between them and the likelihood that an ethnic group will experience civil war onset.

Ultimately, I find support for only certain aspects of my theory. Economic grievances, by my measure, have an insignificant impact on the likelihood of civil war. Additionally, the moderation of economic grievances by bargaining conditions has a very unreliable relationship with civil war onset, but may exist as theorized. In contrast, this study strongly supports the

theory that certain conditions that make bargaining easier by reducing uncertainty, namely an improved ability of the state to collect information about domestic happenings, make civil war less likely. These measures of bargaining uncertainty have a robust association with the likelihood that an ethnic group will experience civil war onset: less uncertainty is correlated with less likelihood. When successful bargaining is more likely, civil war is more likely to be avoided.

This paper proceeds as follows. First, I review the existing literature and develop my theory, constructing four formal hypotheses. Second, I provide explanations of the data and key variables. Third, I test my theory at the group-year level and analyze these results. Fourth, I collapse my data to test and analyze my theory at the group level, not tied to years. I then provide a brief conclusion.

1. Literature Review and Theory

In this section I review relevant literature and, using these past studies, construct my own theory and hypotheses. I focus on the literature about civil war onset, provide a theory relating economic grievances by itself to civil war onset, and then turn to the literature and my theory about how certain conditions lead to bargaining failure and cause onset. This section concludes by laying out how economic grievances and bargaining interact to cause intrastate war.

Explaining Civil War Onset. Studies seeking to understand the onset of civil war abound. One group of scholars argue that civil war onset is best explained by conditions that make fighting easier or more personally beneficial. Collier and Hoeffler (2004) are among the strongest proponents of such viability theories, arguing that civil war onset is more likely when (1) rebel groups have greater access to funds for their startup costs, (2) low earnings in the formal economy make fighting cheap, and (3) rebels have greater military parity with the state. Fearon and Laitin (2003) find that more available weapons, recent political shocks, a large

population, greater mountainous terrain, and lower per capita income in a country are correlated with civil war onset because they make insurgency easier. Low income is especially important because it implies a low state capacity, making it easier for rebels to hide and states worse at fighting them. Interestingly, Koren and Sarbahi (2018) contextualize these country-level findings with subnational analysis indicating that civil wars are most likely to begin in high-capacity areas of a country.

Viability theories' emphasis on state capacity begs for further disaggregation of this concept. For example, Fearon and Laitin's (2003) measure of low per capita income simultaneously captures states without financial resources to run the government, provide welfare, surveil their citizens, or raise a strong military. All of these are distinct aspects of a state's "capacity" and they do not necessarily vary together. In their work on terrorism, Hendrix and Young (2014b) make an important distinction between military capacity and administrative capacity, which impact political violence differently. Existing studies seem to imply that military capacity is what matters, but I explicitly pick up on this separation and instead argue that administrative capacity has a unique role in explaining onset.

Another group of scholarship can be categorized as supporting grievance theories of civil war onset. The viability studies discussed above either assume that grievances are widespread or that they do not matter. Grievance theories instead propose that people fight when there is cause, and that varying levels of grievance explain the likelihood of conflict. The theory of relative deprivation from Gurr (1970)—individuals who feel they are getting less than they deserve are more likely to rebel—underlies much of this theory.

General models of *political* grievance hold that anger against the state based on poor governance or exclusion from power leads to peaceful mobilization which escalates to civil war

if it is repressed violently by the state (Goodwin 2001; Rost 2011). A key insight here is that civil war breaks out only after less extreme measures have been attempted. Studies also link grievance to ethnicity in the causal pathway towards civil war. Greater exclusion of an ethnic group from political power is correlated with a greater likelihood of civil war onset, especially if that group lost power recently, can effectively mobilize, and has experienced past conflict (Cederman, Wimmer, and Min 2010). These explanations, however, are missing a clear and testable explanation for why a compromise isn't reached before grievances escalate fully to civil war.

The ethnicity delineated grievances mechanism has also been applied specifically to *economic* grievances. Buhaug, Cederman, and Gleditsch (2014) find that countries are more likely to experience civil war onset the greater the gap between the most economically deprived and politically excluded ethnic group and the rest of society. Most relevant to this study, Cederman, Weidmann, and Gleditsch (2011) analyze horizontal inequalities at the group level, finding that the richest and the poorest ethnic groups in a country are both more likely to fight a civil war than middle income groups. Government spending on welfare programs, though, can strongly reduce the risk of civil war onset (Taydas and Peksen 2012). The nexus of ethnically defined economic grievances and civil war is at the heart of this paper.

The importance of ethnicity, though, should be qualified appropriately. Fearon and Laitin (1996) astutely observe that different ethnic groups generally live peacefully together due to self-reinforcing mechanisms that reduce intergroup conflict. The literature does not support the idea that ethnic groups fight each other. Instead, they fight the state. Discontent, anger, political demands, and violence are directed at agents of the state and generally not at members of other ethnic groups not affiliated with the state (Goodwin 2001). Denny and Walter (2014) find that

ethnic groups are more likely than any other type of group—including class, geographic, ideological, and religious groups—to begin a civil war, because they have greater grievances, greater ability to mobilize, and greater bargaining problems. This strongly justifies this paper’s focus on ethnic groups as the unit of analysis.

The Economic Grievance Theory of Civil War. My argument picks up from past studies that argue and find support for the theory that higher levels of economic grievances independently increase the likelihood that civil war will occur. In this subsection, I lay out a theory for why this is the case. Grievances are inherently comparative; to be aggrieved, one must feel worse off than someone else or than one should be (Gurr 1970). In the economic sphere, these feelings can arise from a variety of sources, including poverty, poor infrastructure, low education, reduced welfare, or the transfer of wealth. Possible sources of comparison include other nearby individuals, the ‘average’ fellow citizen, the highest income few, citizens of other countries, or other social groups. Most relevant to civil war onset are horizontal inequalities: collective differences between politically relevant social groupings (Cederman, Weidman, and Gleditsch 2011, 480). Social groupings, more commonly referred to as ethnic groups, are identity-based groups of people with a “subjectively experienced sense of commonality based on the belief in common ancestry and shared culture” (Vogt et al. 2015, 1329; Weber 1976). This definition encompasses racial, religious, national, tribal, and traditional “ethnic” groupings, whichever captures the relevant cleavages in a country’s politics. It does not, however, include class or ideology.

I argue that greater economic inequality between a given ethnic group and the country’s most politically powerful ethnic group increases the likelihood that the given ethnic group will engage in war against the state. The relevant comparison is to the most well-off ethnic group

because it provides the clearest measure of what an ethnic group could expect, the psychological yardstick for relative deprivation. Further, the comparison group's hold on political power ensures that individuals are aware of this group's existence and position, providing a real group to reference for comparison (Siroky et al. 2020, 699). This method of comparison also accounts for second-tier ethnic groups in multi-ethnic states who feel they should be first.¹ Below, I will explain how inequality between ethnic groups affects individual fighters and potential rebel leaders, as well as the structural factors of ethnic groups that make civil war more likely. I address each in turn.

Members of an ethnic group that is economically worse off than the most politically powerful group in their state are likely to believe, as a result of this real comparison, that their own group does not—and that they themselves individually do not—have what they are due. This feeling, termed “relative deprivation,” will be more intense as the gap grows. More intensity, in turn, raises the potential for “collective violence” by creating anger and discontent, such that people are willing to engage in high-risk aggressive behavior in order to get what they are “due” (Gurr 2016, 320). The greater the inequality, the more willing potential rebels will be to fight. Economic standing has an immense impact on individuals' everyday life, adding to the willingness to take risks to rectify the situation for oneself and one's family. Therefore, a situation of high economic grievances provides a fertile ground for recruiting rebel soldiers. Further, the standard of comparison being another ethnic group within the same state clearly marks the level of economic well-being that could reasonably be expected, demonstrating that the policies of the incumbent government, from overt favoritism of their own ethnic group to

¹ Traditional measures of horizontal inequality—comparing ethnic groups to the country mean—fail to account for groups that are better off than the country average but are still “second-tier” compared to the group in power and believe they should be first.

failing to rectify historical differences, are responsible for this inequality (Goodwin 2001; Gurr 2016, 319). It also demonstrates what level of economic well-being the individual could hope to obtain if their group captures control of the state. This directs the anger and violence of potential rebel fighters against the ethnic group in power and the state that they control.

Potential rebel leaders—motivated themselves by the anger of relative deprivation—are therefore able to draw from and stoke these anti-state sentiments. Discontent from relative deprivation becomes “political violence” when the aggrieved view violence as justified and view the incumbent government as responsible (Gurr 2016, 319). Leaders can frame the issue, incite violence as the solution, and reinforce in- and out-group sentiments to gain the backing of their co-ethnics to engage in organized violence (Granzow, Hasenclever, and Sändig, 2015). They can therefore recruit from the population of aggrieved co-ethnics to mobilize a fighting force and demand an amelioration of these economic grievances, either via a bargain or through war.

The structure of horizontal inequality also makes civil war more likely. Because it pits groups against groups, it is easier to overcome barriers to collective action. Members of ethnic groups engage with each other frequently in everyday life, creating shared understandings, norms, and knowledge about each other (Weinstein 2007, 99). This allows them to more easily craft selective incentives, create expectations of reciprocity, use pre-existing trust, and monitor compliance to avoid free-riding in order to mobilize a larger and more cohesive fighting force (Weinstein 2007, 98-101; Cederman, Wimmer, and Min 2010, 96). Greater information and stronger relationships allow ethnic rebel groups to overcome the collective action problem, communicate more clearly to avoid coordination problems, and trust one another to follow through on promises to navigate the commitment problem. Structurally, therefore, ethnic groups are well positioned to translate economic discontent into an organized political group capable of

pursuing a civil war. With economic grievances emerging due to inequality, group members will have the motivation as well. Formally, this can be written as:

H1: Greater economic grievances will make representatives of an ethnic group more likely to become engaged in an ethnic civil war.

Although grievances create a motivation for war, they do not rationally justify a war for its own sake. Why, therefore, do aggrieved ethnic groups and the government not negotiate sufficient redress to avoid war? This leads us to the theory of bargaining failure.

Bargaining Failure. The majority of scholarly literature on bargaining focuses on interstate war. Most prominently, Fearon (1995) argues that states could rationally enter into war against each other when the states could not find a mutually preferable agreement that avoided the costs of war and that fairly reflected the outcome if a war had actually been fought. Such a bargaining failure, Fearon suggests, could occur due to (1) informational asymmetries, (2) commitment problems, or (3) indivisible stakes.

First, in order to accurately predict the outcome of a war to make a fair bargain, both sides need information about the other's military strength and fighting resolve. In order to maximize their gains from negotiation states should overstate their strength or resolve and in order to increase the viability of a first strike states may hide their capabilities, giving both sides incentives to misrepresent their relative strength in various ways. Therefore, uncertainty over relative capabilities can make finding a mutually preferable agreement impossible, resulting in war because the sides' disagreement over relative strength can then only be resolved by fighting.

Second, states will fight a war when they cannot reliably hold their potential opponent to the agreed upon bargain. First strike advantage, military buildup, and benefits gained from negotiation create a logic where states cannot trust that their potential opponent will not renege

after a bargain has been reached and demand re-negotiation under more favorable terms. This can also result in failed negotiation and war now. Third, when the aims of the disputing sides cannot be split—such as ideology, ethnic homelands, or central government control—only war can resolve the dispute (Hassner 2003).

Bargaining in *intrastate* conflicts has received less explicit attention and testing. Walter (2009) argues that information asymmetries regarding potential rebels' funding and states' willingness to fight are the primary causes of bargaining failure in intrastate wars. Arena and Hardt (2014) have found that governments are less likely to offer concessions and instead go to war when the potential rebellion they are facing will not enjoy widespread support. Disputes over ethnic homelands are also more likely to result in war because it is indivisible for both actors: the ethnic group has deep ties to the land and multinational states want to avoid setting a precedent for other ethnic groups who may want secession or autonomy (Toft 2010). Cunningham (2013) also finds that greater division within potential rebel groups worsens information and commitment problems, leading to a greater likelihood that bargaining will fail and civil war will begin. I build on this literature by theorizing about and testing how the informational asymmetry mechanism interacts with the object of dispute in order to understand why some economic grievances cause war and others are resolved.

The Bargaining Theory of Civil War Onset. Because relative deprivation does not promote war for its own sake, potential rebels (leaders, rank and file, and groups as a whole) motivated by economic grievances are only motivated to engage in war insofar as these grievances are not ameliorated through bargaining. In other words, the costs of war ensure that there is a collection of negotiated outcomes that should be preferred over war by the rebels (Fearon 1995, 380). This is also true for the incumbent government: it would be cheaper to pay

the monetary costs to reduce grievances to a tolerable level than to fight. Economic grievances, in particular, should be relatively simple to reach a bargain over.

Bargains to ameliorate economic grievances can take any number of forms, including increased public investment in infrastructure or education, increased local control of resources, lower taxes for the potential rebels' ethnic group, and increased welfare. "Bargains" can also occur in any timeframe; they can be long term measures designed to keep grievances below a tolerable threshold or short-term payoffs to rebels and their communities on the eve of potential conflict. Since these issues are based explicitly around money, bargains are simply a matter of negotiation over funding. However, the incumbent government has incentives to not address economic grievances. They want to avoid providing too much amelioration because then they are overpaying and wasting resources; resources that could be kept for the political leaders themselves or their constituencies. Over paying one group of potential rebels also risks other potential rebel group seeking similar pay offs.

The importance and inherent divisibility of money as the object of bargaining suggests that indivisible stakes are not a barrier to bargains over economic grievances. Commitment problems also hold little relevance. Certain deals, like completed infrastructure and granting local control of resources, are hard to renege on. More importantly, because rebel strength is based heavily on manpower and local knowledge (Fearon and Laitin 2003, 76) it does not diminish and because renegeing on a deal would only increase a group's resolve, the potential for the government to renege does not pose a major risk to rebels. This leaves Fearon's information asymmetry mechanism as the most likely reason for bargaining failure in intrastate disputes over economic grievances. In order for a bargain to be reached—requiring that the government does not hold out in order to pay less and that the rebels are sufficiently satisfied so as not to engage in

conflict—both sides must have accurate information about relative rebel strength, rebel resolve, and the threshold at which potential rebels will no longer be motivated to fight.

As implied by Fearon's theory, though, the rebels have motivations to misrepresent information to the incumbent about these factors. Before war begins, the extent of popular support for rebel groups in the process of forming is nebulous. The level of armaments available may also unclear. Plus, rebels want to misrepresent this information in order to protect their funding sources, signal a higher resolve to their potential supporters, and convince the state that they are strong in order to get a better deal (Walter 2009, 250). Rebels will not reliably and accurately reveal their true strength to the state during negotiations. If the state is uncertain about the rebel's strength, they will struggle to offer a mutually agreeable pay-off to avert war. However, the potential rebels' (mis)representations are not the only way for the state to gain an understanding of their capabilities. Therefore, where potentially rebellious populations are not effectively monitored or tracked by the state (the alternative way to determine rebel strength), bargaining is more difficult. In other words, states with weak administrative capacities and poor ability to "collect reliable and consistent information about groups operating within their borders" will be uncertain about rebel strength and resolve, making successful bargaining unlikely (Walter 2009, 249-250). This is the first uncertainty mechanism for bargaining failure.

The second mechanism involves rebel groups who may organize in or receive funding from other states. Rebel groups have been known to receive money from international diasporas and organize their operations in foreign states (Walter 2009, 248; Salehyan 2007). The difficulty of monitoring training, funding, armaments, and other rebel activities across borders means that potential rebels having sympathetic supporters in a foreign state will introduce uncertainty that makes bargaining harder (Salehyan 2007, 226-227).

Under certain conditions, potential rebels will also be uncertain about their own strength, resolve, and thresholds. A misrepresentation of known information is not what causes this uncertainty, but instead rebels' overly optimistic assessment of their capabilities (which the government does not share) will eliminate the range of mutually agreeable bargains. Potential rebels do not know the full extent of their funding sources, the willingness of the population to join them, or how effective their military strategy will be. Because they are about to launch a rebellion, though, rebels must be optimistic about these factors and are therefore likely to overestimate. Potential rebels may also be uncertain about the population they (claim to) represent's true threshold for an acceptable bargain (Cunningham 2013, 663). Because those who are the angriest/most discontent are more likely to join a rebel group early, views of this threshold may be skewed. If rebel leaders view themselves and their early members as representative of the group as the whole, then even the supportive population are likely to accept a worse deal to avoid war than rebel leaders think they will. This will squeeze the bargaining range. Most of these uncertainties can be rectified by rebel groups' having knowledge of their co-ethnics, which will generally be the case because of intra-ethnic ties (Weinstein 2007). However, uncertainties will arise when multiple ethnic groups coalesce into one politically relevant ethnic entity within which ties are less strong.² When rebel leaders are unsure about the nature of their own position and disagree with the state, they will be unable to reach a deal with the state that satisfies all parties. Thus, the third bargaining failure mechanism is rebel uncertainty about their own strength, resolve, and/or demands.³ This subsection can be summarized formally as:

² The classic example of this is Black people in South Africa during Apartheid (Vogt et al. 2015).

³ It is also possible for one or both sides to be uncertain about the strength or resolve of the incumbent government. Uncertainty over government strength is likely to be a limited factor. First, states know about their own capabilities well because they track taxes, expenses, military funding, and more. Additionally, states have limited reasons to

H2: Conditions that make the state more uncertain about rebel capabilities will increase the likelihood of ethnic civil war onset.

H3: Conditions that make potential rebels more uncertain about their own capabilities will increase the likelihood of ethnic civil war onset.

The Combined Economic Grievances and Bargaining Failure Theory. Economic grievances and bargaining failure, while both independently increasing the risk of civil war, are most powerful as explanations for conflict when combined. Economic grievances will result in civil war only when an appropriate bargain cannot be reached, making it such that the only way to determine a “fair” bargain is to fight. Bargaining conditions are only relevant when there is something that must be negotiated over, such as economic inequalities. Therefore, circumstances where an ethnic group’s economic grievances are high (due to horizontal inequalities blamed on the state increasing the motivation to engage in civil war) *and* where conditions make bargaining difficult (due to uncertainty about the rebel’s capabilities) create the highest risk for civil war onset. Easier conditions for bargaining mitigate the heightened risk of civil war when grievances are higher whereas difficult bargaining conditions will exacerbate this risk. This leaves us with:

H4: The impact of increased economic grievances on the likelihood of ethnic civil war onset increases as bargaining becomes more difficult.

My argument, though, has important scope conditions. It cannot explain the presence or lack of civil wars in countries that do not have multiple politically relevant ethnic groups. This includes ethnically homogenous nations such as Portugal, Barbados, Sweden, and South Korea

misrepresent their strength because their vastly superior conventional military power is a known fact for both sides. The incumbent’s resolve to maintain control of the state or to prevent secession is more relevant. This study, though, takes the ethnic group as its level of analysis and thus will explore the impact of variations in state resolve on bargaining and conflict. Interested readers could consult Walter’s (2009) description of the importance of state resolve for civil war onset (pages 249-250) or scholarship on state resolve in interstate disputes, including Clare and Danilovic (2012), Trager (2013), or Dafoe, Zwetsloot, and Cebul (2021), for further discussion.

(Vogt et al. 2015). It also cannot explain civil wars that lack any ethnic dimension, defined loosely. Importantly, this is not the same as saying my theory only applies to conflicts traditionally defined as ethnic civil wars (for example, Sambanis 2001, 261-262.). Rebel groups may claim to represent an ethnic group and recruit from this ethnic group despite limited or minimal support from members of that ethnic group. This (mis)representation may help rebels gain legitimacy, increase local support, or recruit additional fighters. However, an ethnic rebel group's true level of ethnic support is not relevant to my theory. So long as the rebels recruit fighters from an ethnic group and make some sort of claim to fight for said group, the above theory applies. Even if only some members of an ethnic group feel sufficient relative deprivation, blame the state, and justify violence, these select few members are still motivated by their ethnic group's relative economic position. It is when there is truly no relationship between the rebel group and any ethnic group that a case falls outside the scope of my theory.

2. Group-Year Level Data

I will test these four hypotheses with a quantitative analysis of all politically relevant ethnic groups from 1985 to 2010.⁴ An ethnic group is “politically relevant” and enters my dataset in a given year if “at least one political organization claims to represent it in national politics or if its members are subjected to state-led political discrimination” during that year (Vogt et al. 2015; Bormann et al 2021, 25). This excludes nations without any politically relevant ethnic cleavages. Further excluded from my dataset are ethnic groups that are primarily settled in urban environments or scattered around the country, as well as those that migrate regularly (Wucherpfennig et al. 2011). The exclusion of these groups, amounting to 2,293 group-years and

⁴ While many of the original datasets include data that extends both before and after this time frame, this study can only focus on the years 1985 to 2010 because the economic data on ethnic groups is time limited.

three civil war onsets, is unfortunate and a result of data collection strategies that rely on defined, mappable settlement areas to determine economic and population data. The remaining groups live in one region, in a combination of cities and regions, as a combination of smaller ethnic groups settled in their areas, or statewide (Wucherpfennig et al. 2011). As a result, this analysis can only test civil war onset among ethnic groups settled in these ways. The final dataset includes 15,364 group years.

Dependent Variable. The dependent variable in all models is civil war *onset*, measured at the group-year level. An ethnic group experiences onset in a given year when it is the first year of fighting within a state between the government and a rebel organization linked to that ethnic group that results in 25 battle-related deaths (Wucherpfennig et al. 2012, 98).⁵ A rebel organization is linked to an ethnic group when that organization makes some claim, explicit or implicit, to represent that ethnic group and recruits “significantly” from that ethnic group (Vogt et al. 2015; Bormann et al. 2021, 27). There are 123 incidents of civil war onset in the dataset.

Independent Variable: Economic Grievances. The first explanatory variable is gross domestic product (GDP) per capita compared to the most politically well-off ethnic group—*grievance score*. This variable is constructed by dividing the GDP per capita of the ethnic group in same country-year with the most political power by the relevant group’s GDP per capita. In cases where multiple groups share the top political power, all groups are compared to the average of these groups. Additionally, groups that are economically better off than the most politically powerful ethnic group are, for the purpose of this variable, capped at an equivalent GDP per

⁵ I employ a lag time of five years in my main models, and ten years in my robustness checks, before a conflict can be coded as a new onset. In other words, fighting must be paused for five (ten) years before a resumption of fighting between the same groups/over the same issues would be coded as a new civil war onset. Additionally, it is possible for the same civil war to involve multiple ethnic groups. In such a case, each ethnic groups’ initial involvement would be coded as a unique instance of onset.

capita because groups cannot have negative grievances; based on my theory, if a group is equivalent to or economically stronger than the group with political power, they simply have no grievances. This construction creates scores ranging from 1 to 57.4.⁶ A higher grievance score means the group is economically worse off and thus more aggrieved and a grievance score of 1 means the group has no economic grievances. While these scores are created with objective measures of economic well-being and having a grievance is a psychological state, the argument I am making is probabilistic rather than deterministic; greater objective inequality compared to a reference group should correlate with more intense and more common feelings of inequality, relative deprivation, and grievances against this group and their state. Thus, my grievance score variable is able to accurately capture grievance levels.

This formula for measuring economic grievances is distinct from past studies, most importantly the measure used by Cederman, Weidman, and Gleditsch (2011). These scholars, rather than comparing each ethnic group's GDP per capita to some reference group as I do, compare them to "the average GDP per capita of all groups in the country" (486). This lack of reference group is theoretically suspect because grievance theories of onset require "intergroup comparisons" (Cederman, Weidman, and Gleditsch 2011, 481). It is hard to link resentment, anger, and violence to comparisons between one's own ethnic group and an amorphous conception of average income. It is even harder to link these feelings to violence against the state, as both our definitions of civil war require (Cederman, Weidman, and Gleditsch 2011, 484). A lower economic well-being than average is not as directly or clearly attributable to decisions made by the state nor does it necessarily indicate that taking control of the state will alleviate this inequality. Comparing oneself to the ethnic group holding state power provides a

⁶ Given the skew of this data, the log of grievance score is used in almost all analyses. Although less interpretable, the functional form is significantly more normal. The log of grievance score ranges from 0 to 4.05.

direct pathway for animosity and frustration to rise against a well-defined comparison group and the state. Another aspect of comparing to country average is that Cederman, Weidman, and Gleditsch either treat ethnic groups above and below the country mean as the same—without a clear explanation why well-off groups would develop grievances, in line with their causal pathway (and mine) from horizontal inequality, to grievances, to collective action—by only measuring absolute distance from the mean or they collapse all groups above the *mean* into the same horizontal inequality score, ignoring the possibility that a group above the mean may develop grievances against another group still higher than themselves.

Generating my measure of grievance relied on economic data about each ethnic group, which was collected by overlaying worldwide economic production data from 1990, 1995, 2000, and 2005 on geospatial estimates of the settlement area of each group the first year they were included in the dataset (Nordhaus et al. 2006; Wucherpfennig et al. 2011). Population data, for moving from raw economic output to per capita, was generated using the same method for years 1990, 2000, and 2010 (CIESIN 2011). To fill in all missing years, linear growth was assumed between the years with known data. This method does not provide precise estimates of group GDP per capita for every given year. However, grievance scores are relational and this method provides enough accuracy to compare ethnic groups to each other; it is sufficient to capture whether an ethnic group's GDP per capita is above, equal, below, or significantly below the relevant comparison group. Because all “statewide” groups in the same country-year receive the same exact estimates of GDP per capita, I remove all statewide group-years with another statewide group in the same country-year after the comparisons necessary for calculating grievance scores are made. While this is unfortunate, there are only 434 such group-years. For determining which group was the appropriate comparison, the most politically powerful ethnic

group was selected using the *Ethnic Power Relations* dataset's ordinal coding of access to state power, based on information from country experts, ranging from monopoly on power to discriminated against (Vogt et al. 2015).

Independent Variables: Bargaining. The second group of explanatory variables capture conditions that impact uncertainty about rebel capabilities and thus the ability of potential rebels and the incumbent government to reach a bargain. One variable is the *legibility score* (Lee and Zhang 2020). This calculates the accuracy of citizen's ages collected in the most recent census (within the prior 15 years) in order to measure how much the state knows about its population. If the state has a strong bureaucracy that regularly interacts with citizens (a bureaucracy capable of regularly collecting information about happenings within state borders), people will have birth certificates, passports, or other documents that record their age in ways that are intelligible to census recorders and other government administrators. In instances where individuals don't have these documents, the estimates of their age on the census will cluster around numbers ending in five and zero (Lee and Zhang 2017, 122). The less precise census age collection is, the lower the legibility score, and the less the state knows about its citizens—their ages but also their activities, local practices, and living situations—and the less it interacts with them (Lee and Zhang 2017).⁷ In bargaining situations, this low capacity will leave the state with poor knowledge about potential rebels and without the systems to collect further information. As a result, the state will be more uncertain about potential rebels' strength, funding, and threshold for backing down. In line with *H2* and *H4*, a lower legibility score is theorized to correspond to greater likelihood of

⁷ Note that in my data I have inverted the Lee and Zhang's original measure to make it more easily interpretable. In their dataset and paper, less accuracy corresponds to a higher score. In this paper, less accuracy corresponds to a lower score.

civil war and an increased impact of economic grievances on onset. Because this is a country level variable, it is the same for all ethnic groups in the same country-year.

The second variable in this group is *administrative capacity*, which is a different operationalization of the first uncertainty mechanism described in section one, measuring how much a state knows about the people, ethnic groups, and potential rebels within its borders (Hendrix and Young 2014a). This variable is generated for each country-year by a factor analysis of the International Country Risk Guide's rule of law and bureaucratic quality data, which are both originally hand coded ordinal variables (Hendrix and Young 2014b, 340-341). It is explicitly distinguished from military capacity, which captures opportunity theories of civil war onset, and exclusively measures "the state's ability to collect and manage information on potential dissidents" (Hendrix and Young 2014b, 336). As a result, it captures a country-level measure of uncertainty: as administrative capacity decreases, the state's uncertainty about potential rebels will increase, making bargains more difficult to reach and civil war more likely.

A third variable for this group is a binary indicator of whether the government collected data on their citizens' ethnicity in a census in the prior 15 years: *ethnicity collection* (Lieberman and Singh 2016). Accurate ethnic data provides the government key information—including populations of young men, relative economic standing, and geographic concentration—for bargaining with representatives of an ethnic group. Without this information and with potential rebels incentivized to misrepresent it during bargaining, state assessments of rebel strength and demands may be murky, leading to bargaining uncertainty. For otherwise high-capacity states that have an ideological opposition to codifying ethnicity, like France, the lack of census data on ethnicity is unlikely to impact the government's ability to estimate this information. For many states, however, without collecting information on ethnicity in their census, they may have little

accurate information about ethnic groups in the country and their capacity for rebellion. This variable is also designed to measure state knowledge about the people within its borders, but it does so quite distinctly from legibility score or administrative capacity. It captures state knowledge about only a very specific set of demographic data and it proxies for a previous instance of information collection rather than an ongoing ability to do so. Just like legibility score and administrative capacity, though, ethnicity collection is a country level variable and therefore cannot account for variations in capacity within states.

A fourth variable from this group is *transnational kin*, which records the number of ethnic groups in foreign nations who are ethnically linked to this group (Cederman et al. 2013). For example, Basques in France and Spain are linked to each other. Inspired by the second uncertainty mechanism discussed in section one, this variable captures the ability of the state to accurately understand the level of foreign support that potential rebels associated with this ethnic group have. If a group has more transnational kin, tracking potential avenues of support is more difficult and therefore state uncertainty about rebel capabilities is greater. It is concerning that this variable treats every foreign kin group as equal, since a kin group that controls a neighboring state will have a wildly different relationship with the domestic group than a diasporic minority contingent on another continent. The focus on state uncertainty regarding activities external to its territory distinguishes transnational kin from the previous three bargaining variables.

A fifth and final bargaining variable, which captures an aspect of rebel uncertainty about their own strength, is a dummy variable indicating whether the relevant group is in fact an *aggregate group* (Wucherpfennig et al. 2011; Vogt et al. 2015). The ethnic cleavages that are politically relevant in a nation are not constant. Ethnic groups can form alliances, become amalgamated into larger connected groups, or the criteria on which ethnicity is sorted can

Table 1: Group-Year Level Descriptive Statistics						
Variable	N	Mean	S. Dev.	Min.	Median	Max.
Group Population (mil.)	15317	8.9	53.3	0.000424	0.971	1274
Settlement Type	15364					
... Aggregate	187	1.2%				
... Regional & urban	2568	16.7%				
... Regionally based	11293	73.5%				
... Statewide	1316	8.6%				
Onset	14357	0.00857	0.0922	0	0	1
Grievance Score	15213	1.28	1.19	1	1	57.4
Grievance Score [log]	15213	0.167	0.316	0	0	4.05
Legibility Score	9630	0.378	0.593	0.0222	0.203	5.75
Admin. Capacity	10892	-0.173	0.703	-1.88	-0.244	1.39
Ethnicity Collection	13665	0.651	0.477	0	1	1
Transnational Kin	15364	3.21	6.13	0	1	33
Aggregate Group	15364	0.0122	0.11	0	0	1

Table 1: Descriptive statistics of key group-year variables, including basic information about groups and all outcome and explanatory variables. These data include groups that are later removed for analysis due to the five- and ten-year lags applied for re-entering the dataset.

change. This leads certain politically relevant ethnic groups to actually be an aggregation of multiple ethnic groups that are not, by themselves, politically relevant. When this is the case, the relationships, understandings, and regular interactions that usually give members of an ethnic group good information about their co-ethnics are not present (Weinstein 2007). As a result, leaders of rebel groups linked to an aggregate ethnic group will be more uncertain about the resolve and minimum demands of the populace, hindering their ability to strike a bargain—the third uncertainty mechanism from section one. Unfortunately, there are only 187 such groups which could easily lead to imprecise estimates. Descriptive statistics for this and all other explanatory and outcome variables can be seen in Table 1.

Control Variables. My main set of controls at the country level are as follows: total population and geographic size to account for differences between countries (Collier and

Hoeffler 2002; CIESIN 2011, Schvitz et al. 2021); level of democracy to ensure existence or lack of democratic participation isn't facilitating bargaining or fueling grievances (Sambanis 2001; Marshall and Gurr 2020); the presence of another conflict because that lowers the barrier for new conflicts to begin (Gleditsch et al. 2002); being a Cold War year as that could impact international funding and support (Collier and Hoeffler 2002); and political instability and the state being newly independent because these fuel grievances and make building rebel group strength easier (Fearon and Laitin 2003; Marshall and Gurr 2020). My main ethnic group level controls are: group size as a percentage of the population because rebel groups connected to large ethnic groups will find it easier to recruit sufficient fighters (Vogt et al. 2015); group GDP per capita to exclude the effects of absolute and isolate relative economic deprivation (Nordhaus et al. 2006; CIESIN 2011; Wucherpfennig et al. 2011); level of political power as this is a political grievance that fuels conflict (Cederman, Wimmer, and Min 2010; Vogt et al. 2015); and number of previous civil wars because of its strong connection to new civil war onsets (Gleditsch et al. 2002; Wucherpfennig et al. 2012).

I also have a set of control variables that are included only in certain models. First, country GDP per capita because of its strong association in the literature with general state capacity, which I am attempting to isolate specific aspects of with some of my bargaining proxies (Fearon and Laitin 2003; Koubi and Bohmelt 2013; Nordhaus et al. 2006; CIESIN 2011, Schvitz et al. 2021). As a result, this control is not included in the main multivariate models that include legibility score or administrative capacity. Second, the "extra controls" that systematically reduce N. State military expenditures per capita (Hendrix and Young 2014a), number of politically relevant ethnic groups in the country (Collier and Hoeffler 2004; Vogt et al. 2015), mountainous terrain in the ethnic group's territory (Fearon and Laitin 2003; Tollefsen et

al. 2012; Blyth et al. 2002), the presence of oil and alluvial diamond resources in the ethnic group's territory (Asal et al. 2016; Paine 2019; Tollefsen et al. 2012; Lujala, Rod, and Thieme 2005; Gilmore et al. 2005; Lujala, Gleditsch, and Gilmore 2005), and distance from the edge of a group's territory to the nearest border of another state (Koren and Sarbahi 2018; Tollefsen et al. 2012; Weidmann, Kuse, and Gleditsch 2010) have systematically limited data availability that risks skewing which group-years are considered by my models.⁸ As will become apparent, robustness tests indicate that the inclusion or exclusion of these controls has little impact on the results of my tests. Appendix A contains a table of descriptive statistics for all control variables.

3. Group-Year Level Tests and Analysis

I will test my theory with regressions of onset on the explanatory variables described above. I will test *H1* by regressing onset on grievance scores, *H2* and *H3* by regressing onset on the proxies for uncertainty about rebel capabilities, and *H4* by regressing onset on the interaction between these things. For the majority of my analysis, I use a binomial generalized linear model (GLM) to conduct the regressions, with standard errors clustered by country.

Scholars (e.g., King and Zheng 2001; Leitgöb 2013) have suggested that standard logistic regression models such as GLM tend to be biased toward a null result when the dependent variable is dichotomous and there are many less occurrences of the outcome of interest than non-occurrences. Because onset is such a “rare event” in my data, occurring in only 0.86% of group-years, the GLM models will return conservative estimates of the odds of civil war onset. As a

⁸ Two other control variables considered in the literature that I do not include in any of my models are ethnolinguistic or religious fractionalization and rates of male secondary schooling. Fractionalization is discarded because it is better captured by a simple count of the total politically relevant groups in the country, which more clearly measures how ethnically (defined to include religious groups where relevant) differentiated a country is. Schooling, although popular with Collier and Hoeffler (2002; 2004), cannot be included because it is not a measure of rebel capacity but rather a form of (economic) grievance. Low rates of school attendance can be caused by poverty, lack of infrastructure, the need to start working young, and other economic conditions.

result, I also employ a penalized maximum likelihood estimation (PMLE) regression in certain models, originally proposed by Firth (1993), to correct for this rare events bias (Leitgöb 2013).⁹ These results, on the whole, differ little from the GLM results.

In all models, the coefficients are reported as odds ratios. Coefficients less than one mean that an increase in the predictor variable decreases the likelihood of civil war onset, and coefficients greater than one mean that an increase in the predictor variable increases the odds of onset.¹⁰ What follows in this section are six tables and three figures displaying the results of regression models that use a five-year lag after any previous conflict concludes for groups to re-enter the dataset. There is one regression table for grievance score alone and one table for each of the proxies for bargaining uncertainty described in section two.

Economic Grievances. Economic grievance has an insignificant relationship with civil war onset; grievance score is in the expected direction but not statistically significant in any of the models shown in Table 2. Model 2, the most theoretically sound, is a multivariate model employing the GLM method to regress onset on the log of grievance score and my main set of control variables outlined in section two. The odds ratio of 1.42 means that increasing the log of grievance score by 1 is associated with a 42% increase in the probability of civil war onset, holding all other variables constant. The independent effect of one standard deviation (0.316) increase in grievance score is thus a group-year being associated with a 13% greater likelihood

⁹ The reason that I do not employ PMLE models for all regressions is that the software cannot correctly cluster standard errors. Therefore, all PMLE models report non-clustered standard errors. While this is unfortunate, the combination of clustered standard error GLMs and non-clustered PMLEs returning similar results suggests that my results are robust to different types of models.

¹⁰ Odds ratios report the percentage change in the likelihood that conflict occurs associated with a one unit change in the predictor variable. Let's consider a dummy variable that takes values either 0 or 1 with an odds ratio of 2.0, meaning that conflict is 100% more likely when the variable takes the value 1 than when it is 0 (holding other variables constant). If the probability of onset when this variable is 0 were 0.15, then the probability of onset is 0.30 when it is 1—twice as likely. The odds ratio of 2.0 does not mean that there is a 200% chance of conflict onset.

Table 2: Grievance Score Regression Results				
	Dependent Variable: Onset			
	(1)	(2)	(3)	(4)
	GLM	GLM	PMLE	GLM
<i>Predictors</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>
Grievance Score [log]	1.20 (0.42)	1.42 (0.87)	1.60 (0.91)	1.17 (0.79)
Controls Included	No	Yes	Yes	Yes
Extra Controls	No	No	No	Yes
(Intercept)	0.01*** (0.00)	2.29 (3.78)	2.53 (3.98)	3.06 (8.46)
Observations	13512	12883	12883	8321
R ² Tjur	0.000	0.035	0.037	0.033
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01			

Table 2: Models of the regression of civil war onset on grievance score. Model 1 is a bivariate regression, Models 2 and 3 are multivariate with best set of control variables available, and Model 4 adds extra controls beyond that for robustness. Models 1, 2, and 4 show standard errors clustered by country. See Appendix B, Table 1 to view the full results with all controls.

of experiencing onset.¹¹ To put these percentages into more substantive context, a group with half the GDP per capita of the most powerful group is 29% more likely to experience conflict than a group with GDP per capita equal to that of the most powerful group, holding other variables constant.¹² Models 1, 3, and 4 have similar results. There is thus some evidence in support of *HI*. However, the lack of statistical significance means that these changes in likelihood cannot be statistically distinguished from no change, and thus these results should be treated with extreme caution. Grievance score’s lack of statistical significance carries through all other models—PMLE models, alterations to functional form, and a 10-year instead of 5-year lag

¹¹ The math for these percentages is as follows. Looking at the odds ratio of 1.42 in Model 2, this ratio means that a one unit increase in grievance score is associated with a group being 142% as likely (42% more likely) to experience onset. The odds ratio for a score increase of 0.316 (one standard deviation) is $1 - [(1 - 1.42)(0.316)] = 1.13$.

¹² This odds ratio is calculated as follows: $1 - [(1 - 1.42) * \log(2)] = 1.29$.

for groups to re-enter (Appendix B, Tables 2-4).¹³ There is thus no solid evidence to be found in support of *H1*.

This finding, that the relationship between economic grievances and civil war onset is probably null, is at odds with many past studies, most notably that of Cederman, Weidmann, and Gleditsch (2011). This is likely the result of some combination of two factors. First, the economic data is noisy. The calculation of ethnic group economic data based on rough outlines of where they live is not precise nor specific enough to provide evidence for the theory of economic grievances. Although Cederman, Weidman, and Gleditsch (2011) use the same data, they adjust for this imprecision in most of their models by removing groups with less than 500,000 members on the grounds that the imprecision is greatest for these groups (487). I do not make this adjustment because it systematically alters the composition of groups by only considering groups of a certain size, who may be more likely to experience conflict onset. The fact that my results are so consistently in the expected direction but lack statistical significance is evidence that precision is an issue.

Second, the difference in results is also explicable by the difference in our measures of economic grievance. I measure grievance by comparing groups' GDP per capita to that of the most politically powerful group in the country, whereas Cederman, Weidman, and Gleditsch compare groups to the average GDP per capita of groups in the country. Their main finding, that groups with income higher and lower than the country mean both fight more often than groups at the mean, is fundamentally different from my findings based on a measure of horizontal inequality that, in line with theory, only accounts for less well-off groups fighting. For the

¹³ Grievance score is in the unexpected direction in only one model: a bivariate regression where grievance score is not logged. In that model, grievance score has a coefficient of 0.99 and $p > 0.95$ so it lacks substantive and statistical significance by any measure.

reasons outlined in section two, I posit that my null findings are based on a test of a theoretically stronger measure of horizontal inequality at the group-year level. However, the Cederman, Weidman, and Gleditsch (2011) findings provide an important explanation for my results: even if low-income groups are more likely to experience civil war onset as a result of economic grievances, this association will be obscured in the data by high-income groups without economic grievances also being more likely to experience onset (for some other reason) compared to those in the middle. An improved theory of the causal pathway that leads high-income groups to fight, distinct from a pathway of economic grievances, could help future studies disentangle this relationship and improve upon both of our tests.

Legibility Score. Table 3 displays models using the legibility score proxy for government uncertainty about rebels during bargaining. Model 5 is the bivariate regression of onset on legibility score, Model 6 adds controls to make it multivariate, Model 7 then adds grievance score, and finally Models 8 (GLM) and 9 (PMLE) add the interaction of grievance and legibility scores. In each model, legibility score is both statistically significant at $p < 0.05$ and substantively significant in the expected direction: a greater score is associated with a lower likelihood of conflict onset. In Model 6, the coefficient of 0.37 means that for every one-point increase in the log of legibility score, a group-year is 63% less likely to experience civil war onset. Every standard deviation (1.17) increase correlates with a 74% reduction in the likelihood of conflict. The results from Model 7, with grievance score included, are very similar. This is a large drop, providing evidence consistent with *H2* that increased state knowledge of its citizenry makes bargaining easier and thus civil war less likely.

Table 3: Legibility Score Regression Results					
	Dependent Variable: Onset				
	(5) GLM	(6) GLM	(7) GLM	(8) GLM	(9) PMLE
<i>Predictors</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>
Legibility Score [log]	0.53*** (0.08)	0.37*** (0.07)	0.39*** (0.08)	0.56** (0.13)	0.57*** (0.12)
Grievance Score [log]			0.73 (0.55)	0.00*** (0.01)	0.01*** (0.01)
Grievance Score [log] x Legibility Score [log]				0.09*** (0.07)	0.10*** (0.07)
Controls Included	No	Yes	Yes	Yes	Yes
(Intercept)	0.00*** (0.00)	0.02 (0.08)	0.03 (0.10)	0.02 (0.06)	0.02 (0.06)
Observations	8516	8389	8389	8389	8389
R ² Tjur	0.003	0.061	0.061	0.099	0.110
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01				

Table 3: Model 5 is the bivariate regression of onset on legibility score and Model 6 is the multivariate version of the same regression. Model 7 regresses onset on both legibility score and grievance score, with controls. Model 8 and Model 9 show the regression of onset on the interaction between grievance score and legibility score, with controls. Models 5, 6, 7, and 8 use the GLM method with standard errors (in parentheses) clustered by country and Model 9 uses PMLE with non-clustered standard errors. See Appendix C, Table 1 for full results with all controls.

Model 8 serves to test *H4* by including the interaction of grievance score and legibility score. In this model, both variables and their interaction are statistically significant at $p < 0.05$.¹⁴ The interaction effect is best understood through Figure 1, which displays a predicted probability plot of the interaction in Model 8. The log of grievance score is plotted against probabilities of onset at the mean (-1.6) and one standard deviation above (-0.43) and below (-2.77) the mean of

¹⁴ The fact that legibility score is independently significant in Model 8 with a substantively significant coefficient of 0.56 means that an increase in state knowledge (ease of bargaining) is correlated with a decrease in the likelihood of conflict even when the log of grievance score is 0 (no economic grievances). This suggests that there are additional factors, beyond economic grievances, that are the subject of bargaining and affect civil war onset. Future research might test whether political power and group autonomy are objects of bargaining.

Figure 1: Predicted Probabilities of Civil War Onset
The Interaction of Grievance Score and Legibility Score

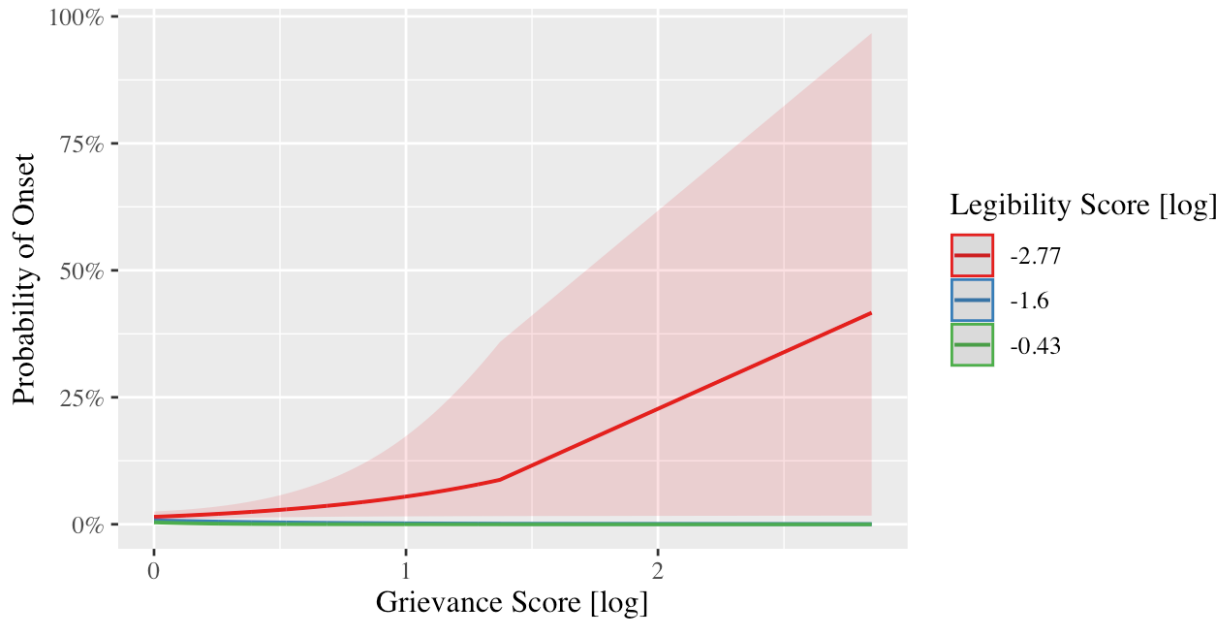


Figure 1: A plot of the regression of onset on the interaction of grievance score and legibility score from Model 8. Probability of onset is on the y-axis, the log of grievance score is on the x-axis, and the lines show the relationship at varied levels of the log of legibility score. -2.77 is one standard deviation below the mean, -1.6 is the mean, and -0.43 is one standard deviation above the mean. All other variables are held constant at their median. The 90% confidence interval is shown for clarity.

the log of legibility score, with all other variables held constant at their medians.¹⁵ Although difficult to see at this scale, at higher (less negative) legibility scores, increasing grievances is actually correlated with a (marginally) lower chance of conflict. With legibility score at one standard deviation above the mean, the probability of onset is always essentially zero. At the mean of legibility score, probability of onset decreases from 1% to nearly 0% as the log of grievance increases from 0 to 0.5. This change in grievance score is the same as going from a GDP per capita equal to that of the most powerful ethnic group to 61% of that GDP per capita. However, Figure 1 clearly illustrates that at lower legibility scores, higher grievances are correlated with higher likelihoods of conflict. When log of legibility score is one standard

¹⁵ Figure 1 shows the 90% confidence interval for the sake of clarity. When the 95% confidence interval is used, the gap between the confidence interval for the log legibility score set at -2.77 and the interval when it is set at -1.6 is not perceptible without seriously zooming in.

deviation below the mean, a 1% probability of conflict at a log grievance score of 0 rises to a 3% probability at a log grievance score of 0.5 and to an over 30% probability at a log grievance score of 2.5 (equivalent to a group having 8% of the GDP per capita of the most powerful group). Model 9, employing the PMLE method, differs very little from Model 8. The results shown in Table 3 are also robust to other alterations in model specifications, including the addition of control variables that reduce N, a ten-year lag for data re-entry after conflict, adjustments to functional form, and use of the PMLE model (Appendix C tables 2–5).¹⁶

Thus, these models suggest that when the state’s good knowledge of its citizens and potential rebels makes bargaining easy, increasing grievances make little impact on the likelihood of conflict. However, when the state’s knowledge is poor and thus bargaining is difficult, increased economic grievances can dramatically elevate the likelihood of conflict. These results are aligned with *H4*, lending support for the theory that bargaining conditions alter the impact of grievances. Model 8 and Figure 1 also suggest an explanation for the null results that were found when testing the effect of grievance score on onset in Models 1 through 4. Because increasing economic grievances has divergent effects at different levels of state knowledge, the failure to differentiate between low and high knowledge states in those models would obscure the relationship between grievances and onset.

Administrative Capacity. Table 4 displays the models that use administrative capacity as the proxy for bargaining uncertainty. Mimicking Table 3, administrative capacity is included in bivariate (Model 10), multivariate (Model 11), multivariate with grievance score (Model 12),

¹⁶ There were two variations in significance or direction of the important coefficients. First, when the functional form of legibility and grievance scores was altered by not taking the log in the model that included the interaction effect, the p-score decreased to $p=0.105$, just outside the range of statistical significance (Appendix C, Table 4). The other coefficients, though, retained their significance. Second, when extra controls were added to the GLM model that included legibility score and grievance score but not their interaction, grievance scored flipped to be in the expected direction but continued to not reach statistical significance (Appendix C, Table 2). Both of these changes, while interesting, do not affect my conclusions.

Table 4: Administrative Capacity Regression Results					
	Dependent Variable: Onset				
	(10)	(11)	(12)	(13)	(14)
	GLM	GLM	GLM	GLM	PMLE
<i>Predictors</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>
Admin. Capacity	0.28*** (0.06)	0.23*** (0.07)	0.22*** (0.07)	0.28*** (0.10)	0.29*** (0.10)
Grievance Score [log]			1.37 (0.80)	0.73 (0.55)	0.80 (0.54)
Grievance Score [log] x Admin. Capacity				0.24 (0.25)	0.25 (0.22)
Controls Included	No	Yes	Yes	Yes	Yes
(Intercept)	0.00*** (0.00)	0.00** (0.01)	0.00** (0.01)	0.00** (0.01)	0.00** (0.01)
Observations	9716	9462	9462	9462	9462
R ² Tjur	0.004	0.056	0.056	0.060	0.063
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01				

Table 4: Model 10 is a bivariate regression of onset on administrative capacity and Model 11 is the multivariate version of the same regression. Model 12 regresses onset on both administrative capacity and grievance score, with controls. Model 13 and Model 14 show the regression of onset on the interaction between grievance score and administrative capacity, with controls. Models 10, 11, 12, and 13 use the GLM method with standard errors clustered by country and Model 14 uses PMLE with non-clustered standard errors. See Appendix D, Table 1 to view the full results with all controls.

and multivariate interacted with grievance score (Models 13 and 14) models. In each model, the coefficient for administrative capacity is in the expected direction and statistically significant at the $p < 0.01$ level, lending strong support to the theory in *H2* that increased state knowledge of their citizens reduces the likelihood that an ethnic group will experience civil war onset.

Substantively, the models without an interaction effect suggest that a one standard deviation (0.703) increase in capacity is correlated with a group being between 51% (Model 10) and 55% (Model 12) less likely to experience civil war onset.¹⁷ This is a major reduction in the odds of

¹⁷ Odds ratios for these are calculated as follows: $1 - [(1 - 0.28) * (0.703)] = 0.49$ and $1 - [(1 - 0.22) * (0.703)] = 0.45$.

conflict and—especially in combination with the results from Models 5, 6, and 7—provides strong evidence that reducing state uncertainty about its citizens and potential rebels is associated with a lower likelihood of conflict. Easier conditions for bargaining appear to make war less likely. These results for the independent effect of administrative capacity on the likelihood of civil war onset are perfectly robust across all model specifications (Appendix D, Models 2-4).

This is the extent of statistically significant findings from Table 4. Neither grievance score nor the interaction of grievance score with administrative capacity are statistically significant in the GLM models or the PMLE model. Altering the model specifications does not change the nulls result displayed here (Appendix D, Models 2-4). Interestingly, despite its lack of statistical significance, Figure 2 shows that this interaction is similar in shape to the interaction of grievance score and legibility score: at lower levels of state administrative

Figure 2: Predicted Probabilities of Civil War Onset
The Interaction of Grievance Score and Administrative Capacity

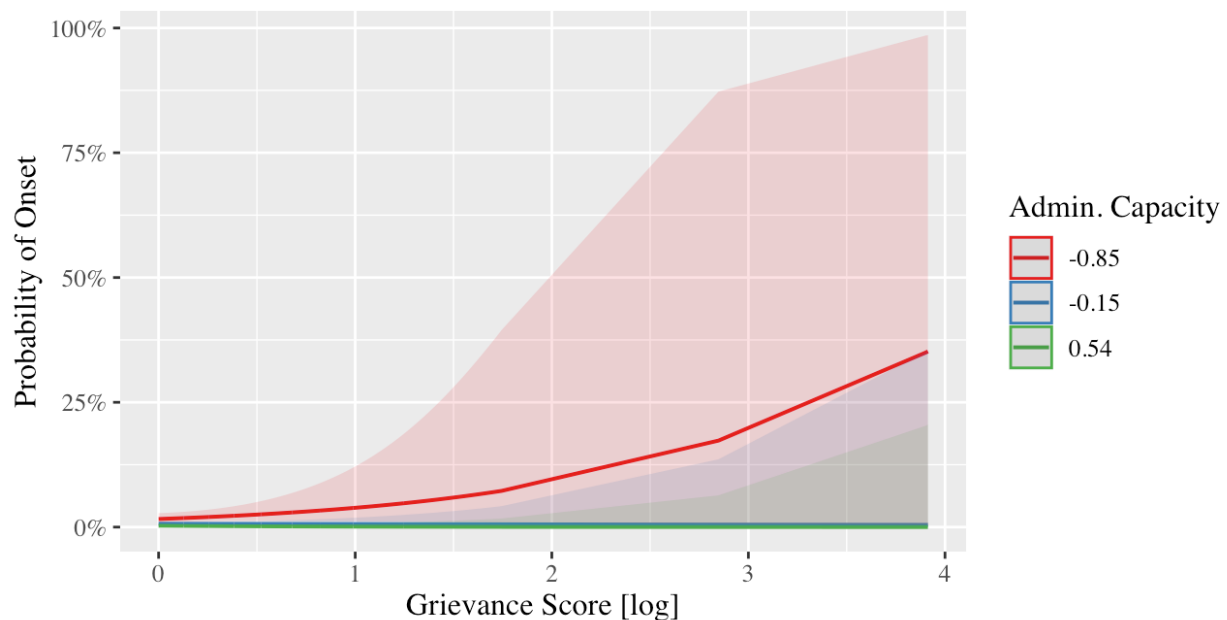


Figure 2: A plot of the regression of onset on the interaction of grievance score and administrative capacity from Model 13. Probability of onset is on the y-axis, the log of grievance score is on the x-axis, and the lines show the relationship at varied levels of administrative capacity. -0.85 is one standard deviation below the mean, -0.15 is the mean, and 0.54 is one standard deviation above the mean. All other variables are held constant at their median. The 95% confidence interval is shown.

capacity, difficult conditions for successful bargaining mean that increased economic grievances are associated with an increased probability of civil war but at higher levels of administrative capacity this relationship does not materialize. Nonetheless, the lack of statistical significance of the interaction in Table 4 reduces support for *H4*.

Ethnicity Collection. Table 5 shows models that test the ethnicity collection proxy for bargaining uncertainty. In all models, the collection of ethnicity information in a recent census is statistically significant but in the unexpected direction. Against my prediction, state collection of ethnicity data is actually correlated with an increase in the likelihood of conflict onset at the

Table 5: Ethnicity Collection Regression Results					
	Dependent Variable: Onset				
	(15)	(16)	(17)	(18)	(19)
	GLM	GLM	GLM	GLM	PMLE
<i>Predictors</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>
Ethnicity Collection	2.60*** (0.83)	4.83*** (1.84)	5.12*** (2.00)	3.24*** (1.35)	3.09*** (1.20)
Grievance Score [log]			2.04 (1.30)	0.23 (0.32)	0.32 (0.38)
Grievance Score [log] x Ethnicity Collection				11.30* (14.67)	9.36* (10.80)
Controls Included	No	Yes	Yes	Yes	Yes
(Intercept)	0.00*** (0.00)	0.02* (0.05)	0.03* (0.06)	0.02* (0.04)	0.02* (0.04)
Observations	12028	11758	11758	11758	11758
R ² Tjur	0.001	0.038	0.037	0.040	0.038
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01				

Table 5: Model 15 is a bivariate regression of onset on ethnicity collection and Model 16 is the multivariate version of the same regression. Model 17 regresses onset on both ethnicity collection and grievance score, with controls. Model 18 and Model 19 show the regression of onset on the interaction between grievance score and ethnicity collection, with controls. Models 15, 16, 17, and 18 use the GLM method with standard errors clustered by country and Model 19 uses PMLE with non-clustered standard errors. See Appendix E, Table 1 to view the full results with all controls.

group level, robust across all model specifications (Appendix E). The impact ranges quite widely—from conflict onset being 412% more likely (Model 17) to onset being 160% more likely (Model 15) when ethnicity has been collected compared to when it hasn't—but the result is substantively large in all models. On its face, this result undermines *H2*; greater government knowledge about its citizens, and thus less uncertainty about potential rebels, appears to actually make civil war *more* likely. However, given the strength of support for *H2* from the models testing legibility score and administrative capacity, I suspect that ethnic data collection's impact does not come from the bargaining space. Static demographic measures of ethnicity appear to not be critical aspects of uncertainty or key sticking points during bargaining. Instead, states' ability to generate more nuanced, up to date, and varied information about potential rebels is important. One explanation for why ethnicity collection points so strongly in the unexpected direction is that the government's decision to collect ethnic data on the census may respond to existing ethnic tensions that simultaneously drive conflict and make ethnicity salient enough that the government chooses to collect data on it. This is consistent with Bhavani and Miodownik's (2009) argument that when ethnicity is "fixed" as salient for all individuals in an ethnic group, conflict is more likely. The result is therefore unlikely to be an indictment of the bargaining theory and instead of the choice of proxy.

In Models 18 and 19, the interaction of grievance score with ethnicity collection is statistically significant, but only at the $p < 0.1$ level. Although grievance score doesn't reach statistical significance on its own and so the results should be treated with caution, this interaction is worth considering. Since it clusters standard errors by country and differs only in magnitude, I will focus on Model 18, which is shown plotted in Figure 3 with all control variables held at their medians. When ethnicity data has not been collected in a recent census

Figure 3: Predicted Probabilities of Civil War Onset
The Interaction of Grievance Score and Ethnicity Collection

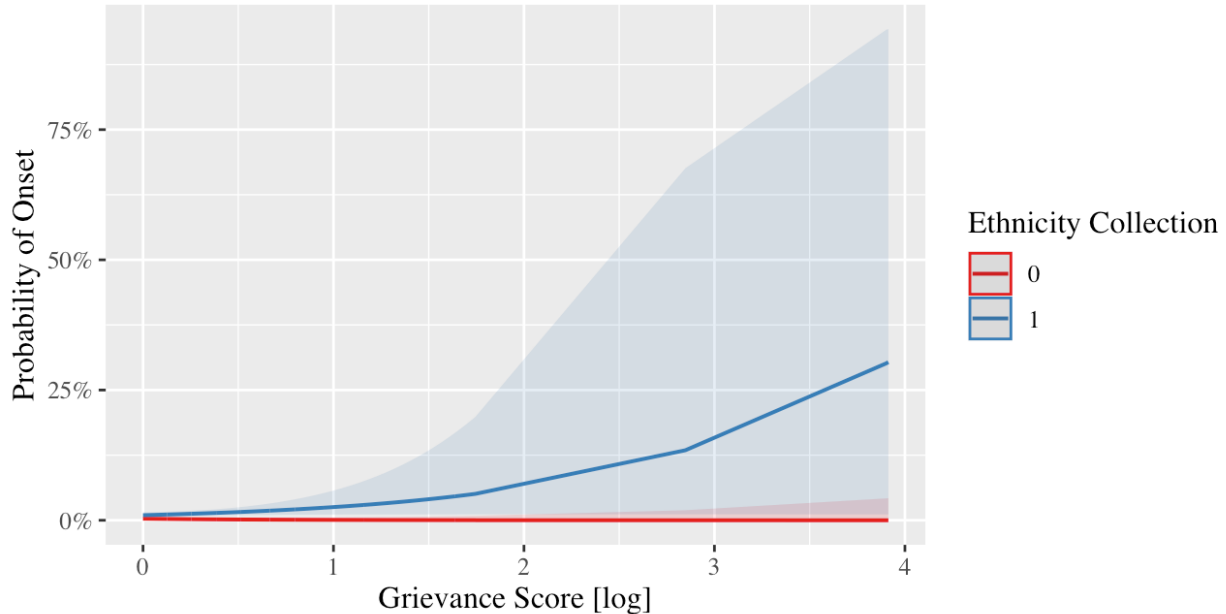


Figure 3: A plot of the regression of onset on the interaction of grievance score and ethnicity collection from Model 18. Probability of onset is on the y-axis, the log of grievance score is on the x-axis, and the lines show the relationship when ethnicity data was collected on a recent census (1) and when it was not collected (0). All other variables are held constant at their median. The 90% confidence interval is shown.

(“0”), the probability of onset remains near zero at all levels of grievance score. However, when ethnicity was collected in a recent census (“1”), an increase in the log of grievance score from 0 to 2.5 increases the likelihood of conflict onset for a group-year from about 1% to over 10%.

This means that an ethnic group with 8% of the GDP per capita of the most politically powerful group in their country is much more likely to experience conflict, but only if ethnicity information was collected on a recent census. More modestly, a group having half the GDP per capita of the most powerful group has a 2% probability of experiencing conflict when ethnicity was collected—double the likelihood of when there are no grievances but still not tremendously likely.

If collecting ethnicity data in a recent census does in fact proxy for easier bargaining, then this result suggests that easier bargaining is associated with grievances being positively associated with conflict rather than mitigating this association—directly contradicting *H4*.

However, and more likely, if ethnicity collection actually proxies for ethnic salience, then the results indicate that increased economic grievances are associated with an increased likelihood of conflict only when ethnicity is sufficiently salient. While this was not explicitly hypothesized, it aligns with my theory's emphasis on conflict arising from the feeling that one's *ethnic group* does not have what they are due, making ethnic salience a critical mechanism for turning group level inequalities into feelings of grievance. All that said, the significance of this interaction effect is sensitive to different model specifications (it isn't significant when a 10-year lag is used or when the extra controls are added) and so this result should be treated with caution and subjected to more explicit testing in future studies (Appendix E, Models 2-3).

Transnational Kin. The lack of any statistical significance in Table 6 provides evidence that increases in the number of transnational kin groups have no independent effect on the likelihood of civil war onset. Under any model specification, including with extra controls and the 10-year lag, transnational kin never reaches statistical significance (Appendix F, Tables 2-4). The transnational kin proxy provides no support for the bargaining theory of civil war onset, *H2*. Further, this result suggests that conditions that add uncertainty about activities that might be occurring outside a state's territory, such as fundraising and training, do not impact the difficulty of bargaining. The lack of effect could be a result of states' variable ability to know what's going on *inside* its territory, which does not vary with transnational kin, making up for their uncertainty about conditions outside their borders. The number of transnational kin an ethnic group has is extremely poorly correlated with legibility score (-0.03) and administrative capacity (0.05), providing evidence for this explanation.

Table 6: Transnational Kin Regression Results					
	Dependent Variable: Onset				
	(20)	(21)	(22)	(23)	(24)
	GLM	GLM	GLM	GLM	PMLE
<i>Predictors</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>
Transnational Kin	1.00 (0.02)	0.99 (0.03)	0.99 (0.03)	1.00 (0.03)	1.00 (0.02)
Grievance Score [log]			1.41 (0.86)	1.78 (1.20)	2.06 (1.26)
Grievance Score [log] x Transnational Kin				0.84 (0.20)	0.85 (0.18)
Controls Included	No	Yes	Yes	Yes	Yes
(Intercept)	0.01*** (0.00)	1.80 (2.92)	2.24 (3.70)	2.48 (4.13)	2.76 (4.33)
Observations	13644	12883	12883	12883	12883
R ² Tjur	0.000	0.035	0.035	0.036	0.038
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01				

Table 6: Model 20 is a bivariate regression of onset on transnational kin and Model 21 is the multivariate version of the same regression. Model 22 regresses onset on both transnational kin and grievance score, with controls. Models 23 and 24 show the regression of onset on the interaction between grievance score and transnational kin, with controls. Models 20, 21, 22, and 23 use the GLM method with standard errors clustered by country and Model 24 uses PMLE with non-clustered standard errors. See Appendix F, Table 1 to view the full results with all controls.

Aggregate Group. Finally, Table 7 displays results for the impact of being an aggregate group on onset, the sole measure of *H3* capturing the impact of rebel uncertainty about their own strength, resolve, and threshold for demands. These models provide no evidence for this hypothesis. Models with extra controls or other changes to their specification also failed to return significant results (Appendix G, Tables 2-4). Therefore, this data cannot provide evidence that being an aggregate group has an independent effect on the likelihood of conflict onset. The small number of aggregate groups in the dataset, just 187, is a potential explanation for the lack of significance. Another explanation is that rebel leaders are more adept at understanding the group

Table 7: Aggregate Group Regression Results					
	Dependent Variable: Onset				
	(25)	(26)	(27)	(28)	(29)
	GLM	GLM	GLM	GLM	PMLE
<i>Predictors</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>
Aggregate Group	1.06 (1.07)	0.74 (0.77)	0.70 (0.73)	0.00 (0.00)	0.29 (0.51)
Grievance Score [log]			1.44 (0.89)	1.18 (0.77)	1.31 (0.79)
Grievance Score [log] x Aggregate Group				501218.26 (7076362.62)	24.01 (59.05)
Controls Included	No	Yes	Yes	Yes	Yes
(Intercept)	0.01*** (0.00)	1.87 (3.05)	2.38 (3.93)	2.05 (3.42)	2.25 (3.56)
Observations	13644	12883	12883	12883	12883
R ² Tjur	0.000	0.035	0.035	0.036	0.038
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01				

Table 7: Model 25 is a bivariate regression of onset on aggregate group and Model 26 is the multivariate version of the same regression. Model 27 regresses onset on both aggregate group and grievance score, with controls. Models 28 and 29 show the regression of onset on the interaction between grievance score and aggregate group, with controls. Models 25, 26, 27, and 28 use the GLM method with standard errors clustered by country and Model 29 uses PMLE with non-clustered standard errors. See Appendix G, Table 1 to view the full results with all controls.

they claim to represent than theorized. This allows leaders to realistically account for the fact that they and their early supporters likely hold more extreme views than the general population when the leaders appraise their chances at war.

Assessing the Bargaining Theory. The proxies for bargaining returned varied support for *H2* and *H3*. There is strong support for the generalized measures of state information collection and knowledge about its citizens behaving as expected. Legibility score and administrative capacity were both significant and in the expected direction in all models. As states gain

increased knowledge about their citizens, they are more capable of reaching a mutually beneficial bargain and ethnic groups are therefore less likely to experience civil war onset.

This is an interesting finding in relation to existing theories that increases in non-disaggregated measures of state capacity reduce conflict by reducing the opportunity to rebel (e.g. Fearon and Laitin 2003; Fearon 2005). While state knowledge does not affect the military opportunity of rebellion, it does impact bargaining, suggesting that at least part of the impact of generalized state capacity is a result of easier bargaining. Adding a control for military expenditure (the other aspect of state capacity) and country GDP per capita (the standard measure of general state capacity) to the models did not meaningfully alter the substantive or statistical significance of legibility score (Appendix C, table 2) or administrative capacity (Appendix D, table 2).

The findings for ethnicity collection, while similarly robust, were resoundingly in the unexpected direction—collecting ethnic data in a recent census is associated with onset being more likely. As described above, however, a correlation between collecting ethnic data and high ethnic salience best explains the observed results. Transnational kin and being an aggregate group were not at all significant predictors of civil war onset; conditions that increase state uncertainty about activities beyond its borders appear to have no effect and potential rebels' uncertainty about themselves does not impact their ability to effectively bargaining with the state. As a result, there is no evidence for *H3*. However, the group-year level provides strong evidence for a qualified version of *H2*. Increases in certain types of state knowledge about its citizens—generalized, actively collected domestic information—are associated with a reduction in the likelihood of armed conflict as a result of easier bargaining.

Evidence for *H4* exists but is limited. The interaction of economic grievances with bargaining proxies was only rarely significant and in the expected direction. The interaction between legibility score and grievance score suggests that increased economic grievances increase the likelihood of an ethnic group experiencing civil war onset when low state knowledge about its citizenry makes bargaining difficult, directly supporting *H4*. This is supported by the interaction effect of grievance score and administrative capacity taking the same shape. However, the lack of significance when grievance score was interacted with administrative capacity—which, like legibility score, proxies for state knowledge about its citizens—suggests that bargaining’s moderating effect on the relationship between economic grievances and civil war onset is tenuous. Instead, as discussed above, conditions that make bargaining easier as a result of increased state knowledge have their most consistent impact on the likelihood of onset as an independent effect. This is likely because economic grievances are not the sole object of bargaining.

4. Group Level Data, Tests, and Analysis

All the tests above were conducted at the level of group-year. This is of greatest interest in order to pinpoint, with maximum accuracy, the factors that lead civil wars to begin when they do. However, many of the factors impacting civil war onset are very static. Ethnic group wide economic situations are unlikely to change rapidly. Ethnic data collection can only change when there is a census. Administrative capacity, the number of relevant transnational kin, and the settlement patterns of groups are slow moving phenomena. From the theoretical point of view, economic grievances and the bargaining space are conditions that impact the likelihood of conflict, not sparks for onset. This suggests another way of testing the theory of economic grievances, bargaining uncertainty, and their interaction: across the 26 years for which I have

data, do varying levels of these conditions alter the likelihood that an ethnic group experiences conflict at some point in that time? This section pursues this question.

The data for this analysis is the same as that for the group-year level examined above, just collapsed into the ethnic group level without any years. After removing group-years that had been involved in a conflict in the previous five years because their data is conflict affected, I collapse all remaining group-years into timeless data by taking the mean of continuous or ordinal variables and the maximum of binary variables. For onset, therefore, this captures whether a group experiences civil war onset at any point during the period from 1985 to 2010.¹⁸ There are 688 groups in the dataset, 75 of which experienced at least one civil war onset. Table 8 displays descriptive statistics for the group level versions of the dependent and explanatory variables, and Appendix H Table 1 includes all controls as well.

Table 8: Group Level Descriptive Statistics						
Variable	N	Mean	S. Dev.	Min.	Median	Max.
Onset	686	0.109	0.312	0	0	1
Avg. Grievance Score	682	1.35	1.95	1	1.01	49.8
Avg. Legibility Score	483	0.361	0.499	0.0259	0.228	4.46
Avg. Administrative Capacity	573	-0.177	0.597	-1.58	-0.252	1.39
Ethnicity Collection Ever	623	0.822	0.383	0	1	1
Avg. Transnational Kin	688	3.14	5.92	0	1	32.3
Aggregate Group Ever	688	0.0247	0.155	0	0	1

Table 8: Descriptive statistics for dependent and independent group level variables.

¹⁸ This is similar for the other binary variables, such as aggregate group and ethnicity collection. The group level versions capture whether the group was *ever* an aggregate group or was *ever* in a country that collected ethnicity data on the census. The continuous variables, like grievance score, legibility score, administrative capacity, transnational kin, and many control variables looks different. These variables, now reported with “Avg.” in front of their old variable names, report their average level across 26 years. Grievance score, for example, reports the average yearly difference between the given group and the most politically powerful ethnic group that year during the 1985-to-2010 timeframe. The one exception to my rules is that I took the minimum of each groups’ war history count, a continuous control variable, to capture the number of civil war wars they had been involved in prior to this period rather than including civil wars they experienced during the period of study.

Like the models shown in the previous section, in this section I will explore the regression of onset on average grievance score, on each of the proxies for bargaining uncertainty, and on the interactions between them.¹⁹ In these models, much is the same as the group-year level: grievance score stays in the expected direction but not statistically significant; the bargaining proxies that were significant at the group-year level stay significant and in the same direction; and the proxies that weren't significant stay insignificant. Unlike the group-year level, the interaction effects at the group level uniformly failed to achieve statistical significance.

Table 9 shows results for the regression of onset on grievance score, which is not statistically significant under any variation of control variables, is it significant in any differently

Table 9: Grievance Score Group Level Regression Results			
Dependent Variable: Onset			
	(30) GLM	(31) GLM	(32) GLM
<i>Predictors</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>
Avg. Grievance Score [log]	1.02 (0.36)	1.24 (0.70)	2.11 (1.32)
Controls Included	No	Yes	Yes
Extra Controls	No	No	Yes
(Intercept)	0.12*** (0.02)	60.45** (103.91)	1.36 (4.31)
Observations	680	649	582
R ² Tjur	0.000	0.088	0.086
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01		

Table 9: Model 30 shows the bivariate GLM regression of onset on the log of average grievance score at the group level. Model 31 adds standard controls to make this multivariate and Model 32 adds extra controls. See Appendix H, Table 2 for the full models with all controls shown. Additionally, all standard errors are clustered by country.

¹⁹ Unlike the group-year level models, the group level results are obtained exclusively through GLM models. The fact that over 10% of groups experience onset during the time period of this data means that the rare events correction offered by PMLE models is not necessary. Onset is no longer so rare that one would expect the standard GLM function to produce conservative estimates.

specified model (Appendix I, Table 3). This reinforces my findings against *HI*. In both the short and long term, economic grievance does not independently affect the likelihood of conflict. Of note, though, is that average grievance score is in the expected direction in Models 30, 31, and 32. Collapsing the measure of horizontal inequalities from yearly estimates to a 26-year time frame likely obscures variation in grievance scores and results in muddled data, contributing to larger standard errors. That said, the consistent lack of significance across both group level and group-year level models strongly suggest that the explanation truly is the lack of a relationship.

Table 10: Legibility Score Group Level Regression Results				
	Dependent Variable: Onset			
	(33) GLM	(34) GLM	(35) GLM	(36) GLM
<i>Predictors</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>
Avg. Legibility Score [log]	0.50*** (0.08)	0.67** (0.13)	0.69* (0.14)	0.70 (0.18)
Avg. Grievance Score [log]			0.30* (0.19)	0.24 (0.43)
Avg. Grievance Score [log] x Avg. Legibility Score [log]				0.92 (0.57)
Controls Included	No	Yes	Yes	Yes
(Intercept)	0.03*** (0.01)	30.81 (86.53)	36.42 (103.02)	36.52 (103.31)
Observations	483	475	475	475
R ² Tjur	0.053	0.138	0.157	0.158
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01			

Table 10: Model 33 shows the bivariate GLM regression of onset on average legibility score at the group level. Model 34 adds standard controls to make this multivariate and Model 35 adds average grievance score as well. In Model 36, onset is regressed on the interaction of average grievance score and average legibility score. See Appendix H, Table 3 for the full models with all controls shown. Additionally, all standard errors are clustered by country.

Tables 10 and 11 show regressions for legibility score and administrative capacity, respectively. Legibility score is statistically significant at least at $p < 0.1$ in the models without an interaction effect and is always in the expected direction. Administrative capacity is statistically significant at the $p < 0.01$ level and in the expected direction in all models. Together, these results strongly suggest that increased state knowledge of their population correlates with lower likelihood for an ethnic group to experience onset. The impact of a one standard deviation increase in the log of average legibility score (1.16) is that conflict onset during this 26-year timeframe becomes between 58% (Model 33) and 36% (Model 35) less likely, holding all other

Table 11: Administrative Capacity Group Level Regression Results				
	Dependent Variable: Onset			
	(37)	(38)	(39)	(40)
	GLM	GLM	GLM	GLM
<i>Predictors</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>
Administrative Capacity	0.19*** (0.06)	0.22*** (0.09)	0.22*** (0.09)	0.25*** (0.11)
Avg. Grievance Score [log]			0.82 (0.42)	0.59 (0.48)
Avg. Grievance Score [log] x Administrative Capacity				0.54 (0.59)
Controls Included	No	Yes	Yes	Yes
(Intercept)	0.06*** (0.01)	0.20 (0.51)	0.21 (0.53)	0.19 (0.48)
Observations	572	557	557	557
R ² Tjur	0.081	0.150	0.150	0.151
<i>Note:</i>	* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$			

Table 11: Model 37 shows the bivariate GLM regression of onset on average administrative capacity at the group level. Model 38 adds standard controls to make this multivariate and Model 39 adds average grievance score as well. In Model 40, onset is regressed on the interaction of average grievance score and average administrative capacity. See Appendix H, Table 4 for the full models with all controls shown. All standard errors are clustered by country.

variables constant.²⁰ Setting aside the model that includes the interaction, an increase in average administrative capacity by one standard deviation (0.597) is associated with conflict being 48% (Model 37) or 47% (Models 38 and 39) less likely.²¹ These findings are robust to extra control variables that systematically reduce the number of observations (Appendix I, Table 4 and Table 6). However, when the functional form of average legibility score is altered to not take its log, it loses its statistical significance in models with controls (Appendix I, Table 5). This means that the group level findings in support of *H2* should be treated with a little extra caution.

The models shown in Tables 10 and 11 provide evidence that increased general state knowledge about domestic happenings contributes to a reduced likelihood of conflict over the long term. This adds to the robustness of the group-year level findings in support of a qualified version of *H2* and demonstrates that this state knowledge is a condition that effects both short-term and long-term prospects of civil war. This effect can be attributed to, at least in part, a greater ability of states and potential rebels to consistently reach bargains in order to avoid the mutually harmful costs of war. The lack of significance of the interaction effects in Models 36 and 40 suggests that while conditions that effect bargaining may be relatively static, their moderating effect on the impact of grievances fluctuates more quickly, as grievances likely do as well. Especially given the shakiness of the support for *H4* from the group-year level, the group level casts additional doubt on that hypothesis.

Table 12 displays models that regress onset on a binary variable indicating whether ethnicity was ever collected on a census immediately prior to or during the period analyzed, 1985-2010. Ethnicity collection is tenuously significant in the bivariate and interaction effect included models (Models 41 and 44) and, more importantly, is significant at $p < 0.05$ in the

²⁰ These odds ratios are calculated as follows: $1 - [(1 - 0.5) * (1.16)] = 0.42$ and $1 - [(1 - 0.69) * (1.16)] = 0.64$.

²¹ These odds ratios are calculated as follows: $1 - [(1 - 0.19) * (0.597)] = 0.52$ and $1 - [(1 - 0.22) * (0.587)] = 0.53$.

Table 12: Ethnicity Collection Group Level Regression Results				
	Dependent Variable: Onset			
	(41)	(42)	(43)	(44)
	GLM	GLM	GLM	GLM
<i>Predictors</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>
Ethnicity Collection Ever	2.18* (0.90)	2.70** (1.26)	2.73** (1.27)	2.53* (1.33)
Avg. Grievance Score [log]			1.24 (0.72)	0.82 (1.28)
Avg. Grievance Score [log] x Ethnicity Collection Ever				1.56 (2.40)
Controls Included	No	Yes	Yes	Yes
(Intercept)	0.07*** (0.03)	15.01 (26.53)	15.79 (27.99)	17.35 (31.21)
Observations	621	610	610	610
R ² Tjur	0.006	0.110	0.111	0.111
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01			

Table 12: Model 41 shows the bivariate GLM regression of onset on the binary variable capturing whether ethnicity data was ever collected on a recent census. Model 42 adds standard controls to make this multivariate and Model 43 adds average grievance score as well. In Model 44, onset is regressed on the interaction of average grievance score and ethnicity collection ever. See Appendix H, Table 5 for the full models with all controls shown. All standard errors are clustered by country.

models with controls (Models 42 and 43)—all in the direction contradicting my original hypothesis. This is similar to the group-year level regressions analyzing the impact of ethnic data collection on a census and maintains significance at $p < 0.1$ with the addition of extra controls (Appendix I, Table 7). In all models, collecting ethnicity data is associated with conflict during this long-term period being more than twice as likely. The contradiction between this result and the results from the measures of general state information gathering ability, I again argue, is likely a result of ethnic data collection covarying with the salience of ethnic identities. If this is

correct, it suggests that even over the long term, the salience of ethnic identities in a country is an important explainer of which ethnic groups are likely to experience conflict onset.

Just as in the group-year level models, average transnational kin and ever being an aggregate group fail to reach statistical significance in any group level model (Appendix I, Tables 1 and 2). State uncertainty about activity beyond its borders and rebel uncertainty about their own capabilities do not appear to impact the likelihood of civil war onset over the long or near term. As a whole, the group level models for civil war onset largely confirm the group-year level findings. The general measures of bargaining uncertainty—average legibility score and average administrative capacity—retain their robust significance; ethnicity collection ever still points in the unexpected direction; average grievance score, average transnational kin, and aggregate group ever still have no independent effect on the likelihood of civil war onset. The bargaining conditions that matter for likelihood in any given year also matter for the likelihood of onset across 26 years. There is no support for *H1*, strong support for the qualified version of *H2*, no support for *H3*, and no support for *H4* in the group level, long term models.

5. Conclusion

The analysis of economic grievances, bargaining, and civil war onset presented in this paper has two major conclusions. First, economic grievances arising from horizontal inequalities are not effective predictors of which ethnic groups experience civil war onset. Across various model specifications, the GDP per capita of a group compared to that of the most politically well-off group in their country does not have a statistically significant relationship with onset, although greater difference generally points toward greater onset likelihood. While I maintain that the comparison to a specific reference group at the top of society is an improvement over other studies' conceptions of horizontal inequality, this study unfortunately joins the ranks of papers

questioning the ability of grievances to explain civil war onset (e.g., Collier and Hoeffler 2002; Fearon and Laitin 2003). However, like many of those studies, tests of economic grievances continue to be plagued by data issues. Measures of income or wealth at the ethnic group level are difficult to construct precisely and there is a gap between horizontal inequalities, which are factual, and grievances spawned by relative deprivation, which are psychological (Siroky et al. 2020). Whether it is expert coding of ethnic grievances (Siroky et al. 2020) or measures of inequality that more precisely capture the frustrations of individuals' everyday lived experiences (like inequalities in living condition or infrastructure), future studies should continue to refine our measures and tests of the relationship between economic grievances and civil war onset.

The second major conclusion is that increasing the incumbent government's knowledge of its citizens and ability to collect information about happenings within its territory is strongly associated with a reduced risk that ethnic groups within this state will experience civil war onset. This is most likely due to the impact that this knowledge has on bargaining uncertainty: rebel groups are incentivized to misrepresent information about their capabilities, resolve, and their supporters' true demands, so states' ability to gather information about these factors in other ways is critical for reducing uncertainty during bargaining, reaching a deal, and avoiding the escalation to conflict. The temporally different tests shown in this paper provide evidence that bargaining can be over both the short and long terms, but future research should examine more precisely how states strategically provide compensation to ethnic groups in a way that helps to avoid war. Contrary to what was originally theorized, this relationship between certain aspects of bargaining and civil war appears to be mostly an independent effect and does not consistently affect onset by moderating the impact of economic grievances. While there was some evidence that greater economic grievances correlate with greater likelihood of conflict when bargaining is

difficult, this was not a robust finding. This result could stem from the imprecision of my measure of economic grievances, but it also suggests that states and rebels may primarily bargain over other factors that can be explored in future studies, most notably political grievances. Also contrary to original theorizing, certain aspects of uncertainty do not appear relevant to bargaining; I find no support for formal data on ethnic demographics from censuses, uncertainty about events beyond a state's border, or rebel uncertainty about their own strength impacting the ease of bargaining.

Beginning with Fearon's "Rationalist Explanations for War" (1995), many prominent analyses have been written in support of bargaining as an explanation for war onset, both inter- and intra-state. The current paper adds to this line of scholarship with the first generalized, large-N test of bargaining theory in the context of civil war onset and contributes further support for the general principles of the theory's information asymmetries mechanism. For those interested in policy solutions for reducing civil war, this study suggests a feasible, although resource intensive, step for reducing the likelihood of onset: states should collect better information about the lived experiences of their citizens so that they are prepared to offer selective inducements to ethnic groups that effectively ameliorate grievances and alleviate motivations for conflict.

References

- Arena, Philip, and Brian Hardt. 2014. "Incentives to Rebel, Bargaining, and Civil War." *International Interactions* 40 (1): 127–41.
<https://doi.org/10.1080/03050629.2013.863193>.
- Asal, Victor, Michael Findley, James A. Piazza, and James Igoe Walsh. 2016. "Political Exclusion, Oil, and Ethnic Armed Conflict." *Journal of Conflict Resolution* 60 (8): 1343–67. <https://doi.org/10.1177/0022002714567948>.
- Bhavnani, Ravi, and Dan Miodownik. 2009. "Ethnic Polarization, Ethnic Salience, and Civil War." *Journal of Conflict Resolution* 53 (1): 30–49.
<https://doi.org/10.1177/0022002708325945>.
- Blyth, Simon, Brian Groombridge, Igor Lysenko, Lera Miles, and Adrian Newton. 2002. "Mountain Watch: environmental change & sustainable development in mountains." UNEP-WCMC Biodiversity Series 12.
- Bormann, Nils-Christian, Luc Girardin, Philipp Hunziker, and Manuel Vogt. 2021. "GROWup Research Front-End Documentation RFE Release 2021." <http://growup.ethz.ch/rfe>.
- Buhaug, Halvard, Lars-Erik Cederman, and Kristian Skrede Gleditsch. 2014. "Square Pegs in Round Holes: Inequalities, Grievances, and Civil War." *International Studies Quarterly* 58 (2): 418–31. <https://doi.org/10.1111/isqu.12068>.
- Cederman, Lars-Erik, Kristian Skrede Gleditsch, Idean Salehyan, and Julian Wucherpfennig. 2013. "Transborder Ethnic Kin and Civil War." *International Organization* 67 (2): 389–410. <https://doi.org/10.1017/S0020818313000064>.

- Cederman, Lars-Erik, Nils B. Weidmann, and Kristian Skrede Gleditsch. 2011. "Horizontal Inequalities and Ethnonationalist Civil War: A Global Comparison." *American Political Science Review* 105 (3): 478–95. <https://doi.org/10.1017/S0003055411000207>.
- Cederman, Lars-Erik, Andreas Wimmer, and Brian Min. 2010. "Why Do Ethnic Groups Rebel? New Data and Analysis." *World Politics* 62 (1): 87–119. <https://doi.org/10.1017/S0043887109990219>.
- Center for International Earth Science Information Network (CIESIN) Columbia University, International Food Policy Research Institute, The World Bank, and Centro Internacional de Agricultura Tropical. 2011. Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Population Count Grid. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC).
- Clare, Joe, and Vesna Danilovic. 2012. "Reputation for Resolve, Interests, and Conflict." *Conflict Management and Peace Science* 29 (1): 3–27. <https://doi.org/10.1177/0738894211430190>.
- Clark, David, and Patrick Regan. 2016. "Mass Mobilization Protest Data." Harvard Dataverse. <https://doi.org/10.7910/DVN/HTTWYL>.
- Collier, Paul, and Anke Hoeffler. 1999. "Justice-Seeking and Loot-Seeking in Civil War." World Bank Working Paper.
- . 2002. "Greed and Grievance in Civil War." *Center for the Study of African Economies*, March. <https://econwpa.ub.uni-muenchen.de/econ-wp/dev/papers/0409/0409007.pdf>.
- . 2004. "Greed and Grievance in Civil War." *Oxford Economic Papers* 56 (4): 563–95. <https://doi.org/10.1093/oep/gpf064>.

- Cunningham, Kathleen Gallagher. 2013. "Actor Fragmentation and Civil War Bargaining: How Internal Divisions Generate Civil Conflict." *American Journal of Political Science* 57 (3): 659–72. <https://doi.org/10.1111/ajps.12003>.
- Dafoe, Allan, Remco Zwetsloot, and Matthew Cebul. 2021. "Reputations for Resolve and Higher-Order Beliefs in Crisis Bargaining." *Journal of Conflict Resolution* 65 (7–8): 1378–1404. <https://doi.org/10.1177/0022002721995549>.
- Davies, Shawn, Therése Pettersson, and Magnus Öberg. 2023. "Organized Violence 1989–2022, and the Return of Conflict between States." *Journal of Peace Research* 60 (4): 691–708. <https://doi.org/10.1177/00223433231185169>.
- Denny, Elaine K, and Barbara F Walter. 2014. "Ethnicity and Civil War." *Journal of Peace Research* 51 (2): 199–212. <https://doi.org/10.1177/0022343313512853>.
- Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49 (3): 379–414. <https://doi.org/10.1017/S0020818300033324>.
- . 2005. "Primary Commodity Exports and Civil War." *Journal of Conflict Resolution* 49 (4): 483–507. <https://doi.org/10.1177/0022002705277544>.
- Fearon, James D., and David D. Laitin. 1996. "Explaining Interethnic Cooperation." *American Political Science Review* 90 (4): 715–35. <https://doi.org/10.2307/2945838>.
- . 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97 (01): 75–90. <https://doi.org/10.1017/S0003055403000534>.
- Firth, David. 1993. "Bias Reduction of Maximum Likelihood Estimates." *Biometrika* 80 (1): 27–38. <https://doi.org/10.1093/biomet/80.1.27>.
- Gilmore, Elisabeth, Nils Petter Gleditsch, Päivi Lujala, and Jan Ketil Rød. 2005. "Conflict Diamonds: A New Dataset." *Conflict Management and Peace Science*, 22 (3): 257–292.

- Gleditsch, Kristian S., and Michael D. Ward. 1999. "A Revised List of Independent States since the Congress of Vienna." *International Interactions* 25 (4): 393–413.
<https://doi.org/10.1080/03050629908434958>.
- Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Håvard Strand. 2002. "Armed Conflict 1946-2001: A New Dataset." *Journal of Peace Research* 39 (5): 615–37. <https://doi.org/10.1177/0022343302039005007>.
- Goodwin, Jeff. 2001. *No Other Way out: States and Revolutionary Movements, 1945-1991*. Cambridge Studies in Comparative Politics. Cambridge; New York: Cambridge University Press.
- Granzow, Tanja, Andreas Hasenclever, and Jan Sändig. 2015. "Introduction: Framing Political Violence – A Micro-Approach to Civil War Studies." *Civil Wars* 17 (2): 113–19.
<https://doi.org/10.1080/13698249.2015.1070448>.
- Gurr, Ted Robert. 1970. *Why Men Rebel*. Princeton, NJ: Princeton University Press.
- . 2016. *Why Men Rebel*. 40th anniversary ed. London: Routledge.
- Hassner, Ron E. 2003. "'To Halve and to Hold': Conflicts over Sacred Space and the Problem of Indivisibility." *Security Studies* 12 (4): 1–33.
<https://doi.org/10.1080/09636410390447617>.
- Heinze, Georg, Meinhard Ploner, Daniela Dunkler, Harry Southworth, Lena Jiricka, and Gregor Steiner. 2023. "Package 'Logistf.'" R. <https://cemsii.meduniwien.ac.at/en/kb/science-research/software/statistical-software/firth-correction/>.
- Hendrix, Cullen S., and Joseph K. Young. 2014a. "Replication Data for: State Capacity and Terrorism: A Two-Dimensional Approach." <https://doi.org/10.7910/DVN/26475>.

- . 2014b. “State Capacity and Terrorism: A Two-Dimensional Approach.” *Security Studies* 23 (2): 329–63. <https://doi.org/10.1080/09636412.2014.905358>.
- King, Gary, and Langche Zeng. 2001. “Logistic Regression in Rare Events Data.” *Society for Political Methodology*, February. <https://gking.harvard.edu/files/0s.pdf>.
- Koren, Ore, and Anoop K Sarbahi. 2018. “State Capacity, Insurgency, and Civil War: A Disaggregated Analysis.” *International Studies Quarterly*, February. <https://doi.org/10.1093/isq/sqx076>.
- Koubi, Vally, and Tobias Böhmelt. 2014. “Grievances, Economic Wealth, and Civil Conflict.” *Journal of Peace Research* 51 (1): 19–33. <https://doi.org/10.1177/0022343313500501>.
- Lee, Melissa M., and Nan Zhang. 2017. “Legibility and the Informational Foundations of State Capacity.” *The Journal of Politics* 79 (1): 118–32. <https://doi.org/10.1086/688053>.
- . 2020. State Capacity Scores. <http://statecapacityscores.org>.
- Leitgöb, Heinz. 2013. “The Problem of Modeling Rare Events in ML-Based Logistic Regression - Assessing Potential Remedies via MC Simulations.” In. Ljubljana: ResearchGate. https://www.researchgate.net/publication/269708531_The_Problem_of_Modeling_Rare_Events_in_ML-based_Logistic_Regression_-_Assessing_Potential_Remedies_via_MC_Simulations.
- Lieberman, Evan, and Prerna Singh. 2016. “Replication Data for Lieberman, Evan and Prerna Singh, ‘Census Enumeration and Group Conflict: A Global Analysis of the Consequences of Counting,’ World Politics, January 2017, 1-53.” Harvard Dataverse. <https://doi.org/10.7910/DVN/KJV6OJ>.
- Lujala, Päivi, Jan Ketil Rød, and Nadia Thieme. 2007. "Fighting over Oil: Introducing A New Dataset." *Conflict Management and Peace Science*, 24 (3): 239–256.

- Lujala, Päivi, Nils Petter Gleditsch, and Elisabeth Gilmore. 2005. "A Diamond Curse? Civil War and a Lootable Resource." *Journal of Conflict Resolution*, 49 (4): 538–562.
- Marshall, Monty G., and Gurr, Ted. 2020. POLITY5: Political Regime Characteristics and Transitions, 1800-2018 (p5v2018) [dataset]. Center for Systemic Peace.
<https://www.systemicpeace.org/inscrdata.html>.
- NOAA. DMSP Data Collection. National Geophysical Data Center and the US Air Force Weather Agency.
- Nordhaus, William, Qazi Azam, David Corderi, Kyle Hood, Nadejda Makarova Victor, Mukhtar Mohammed, and Jyldyz Weiss. 2006. "The G-Econ database on gridded output: methods and data." Yale University, New Haven.
- Paine, Jack. 2019. "Economic Grievances and Civil War: An Application to the Resource Curse." *International Studies Quarterly* 63 (2): 244–58.
<https://doi.org/10.1093/isq/sqy065>.
- Posen, Barry R. 1993. "The Security Dilemma and Ethnic Conflict." *Survival* 35 (1): 27–47.
<https://doi.org/10.1080/00396339308442672>.
- Rost, Nicolas. 2011. "Human Rights Violations, Weak States, and Civil War." *Human Rights Review* 12 (4): 417–40. <https://doi.org/10.1007/s12142-011-0196-9>.
- Salehyan, Idean. 2007. "Transnational Rebels: Neighboring States as Sanctuary for Rebel Groups." *World Politics* 59 (2): 217–42. <https://doi.org/10.1353/wp.2007.0024>.
- Sambanis, Nicholas. 2001. "Do Ethnic and Nonethnic Civil Wars Have the Same Causes?: A Theoretical and Empirical Inquiry (Part 1)." *Journal of Conflict Resolution* 45 (3): 259–82. <https://doi.org/10.1177/0022002701045003001>.

- Schvitz, Guy, Luc Girardin, Seraina Rüegger, Nils B. Weidmann, Lars-Erik Cederman, and Kristian Skrede Gleditsch. 2021. "Mapping the International System, 1886-2019: The CShapes 2.0 Dataset." *Journal of Conflict Resolution* 66 (1): 144–61.
<https://doi.org/10.1177/00220027211013563>.
- Siroky, David, Carolyn M. Warner, Gabrielle Filip-Crawford, Anna Berlin, and Steven L. Neuberg. 2020. "Grievances and Rebellion: Comparing Relative Deprivation and Horizontal Inequality." *Conflict Management and Peace Science* 37 (6): 694–715.
<https://doi.org/10.1177/0738894220906372>.
- Taydas, Zeynep, and Dursun Peksen. 2012. "Can States Buy Peace? Social Welfare Spending and Civil Conflicts." *Journal of Peace Research* 49 (2): 273–87.
<https://doi.org/10.1177/0022343311431286>.
- Toft, Monica Duffy. 2010. *The Geography of Ethnic Violence: Identity, Interests, and the Indivisibility of Territory*. Princeton University Press.
<https://doi.org/10.1515/9781400835744>.
- Tollefsen, Andreas Forø, Håvard Strand, and Halvard Buhaug. 2012. "PRIO-GRID: A Unified Spatial Data Structure." *Journal of Peace Research* 49 (2): 363–74.
<https://doi.org/10.1177/0022343311431287>.
- Trager, Robert F. 2013. "How the Scope of a Demand Conveys Resolve." *International Theory* 5 (3): 414–45. <https://doi.org/10.1017/S1752971913000250>.
- Vogt, Manuel, Nils-Christian Bormann, Seraina Rüegger, Lars-Erik Cederman, Philipp Hunziker, and Luc Girardin. 2015. "Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Data Set Family." *Journal of Conflict Resolution* 59 (7): 1327–42. <https://doi.org/10.1177/0022002715591215>.

- Walter, Barbara F. 2009. "Bargaining Failures and Civil War." *Annual Review of Political Science* 12 (1): 243–61. <https://doi.org/10.1146/annurev.polisci.10.101405.135301>.
- Weber, Max. 1976. *Wirtschaft und Gesellschaft: Grundriss der verstehenden Soziologie*. 5th ed. Tübingen, Germany: J.C.B. Mohr.
- Weinstein, Jeremy M. 2007. *Inside Rebellion: The Politics of Insurgent Violence*. Cambridge Studies in Comparative Politics. Cambridge; New York: Cambridge University Press.
- Wucherpfennig, Julian, Nils W. Metternich, Lars-Erik Cederman, and Kristian Skrede Gleditsch. 2012. "Ethnicity, the State, and the Duration of Civil War." *World Politics* 64 (1): 79–115. <https://doi.org/10.1017/S004388711100030X>.
- Wucherpfennig, Julian, Nils B. Weidmann, Luc Girardin, Lars-Erik Cederman, and Andreas Wimmer. 2011. "Politically Relevant Ethnic Groups across Space and Time: Introducing the GeoEPR Dataset." *Conflict Management and Peace Science* 28 (5): 423–37. <https://doi.org/10.1177/0738894210393217>.

Appendix A: Group-Year Level Descriptive Statistics

Appendix A, Table 1: Group-Year Level Descriptive Statistics						
Variable	N	Mean	S. Dev.	Min.	Median	Max.
Group Size	15364	0.2	0.268	0.0001	0.0773	0.981
Group Population	15317	8901648	53269772	424	970639	1274127910
Onset	14357	0.00857	0.0922	0	0	1
Territorial Onset	14790	0.00527	0.0724	0	0	1
Government Onset	14931	0.00308	0.0554	0	0	1
Grievance Score	15213	1.28	1.19	1	1	57.4
Grievance Score [log]	15213	0.167	0.316	0	0	4.05
Administrative Capacity	10892	-0.173	0.703	-1.88	-0.244	1.39
Legibility Score	9630	0.378	0.593	0.0222	0.203	5.75
Ethnicity Collection	13665	0.651	0.477	0	1	1
Transnational Kin	15364	3.21	6.13	0	1	33
Aggregate Group	15364	0.0122	0.11	0	0	1
Country GDP/capita	15044	6468	8291	8.11	3034	67438
Country Population	15147	150251319	329964980	156338	19943271	1333778366
Military Expenditures [log]	12676	14.5	4.67	-17	14.5	26.9
Country Size	15346	2448183	4779673	629	578322	22065955
Total Ethnic Groups	15364	10.3	11.7	2	5	53
Democracy Level	14886	1.16	6.73	-10	3	10
Ongoing War	15364	0.276	0.447	0	0	1
Post Cold War	15364	0.828	0.378	0	1	1
New Country	15166	0.0162	0.126	0	0	1
Instability	14886	0.152	0.359	0	0	1
Group GDP/capita	15213	6430	11130	1.84	2680	218780
Group Political Power	15364	3.24	1.71	1	2	7
Previous Wars	15364	0.233	0.649	0	0	6
Years at Peace	15364	35.5	19.3	0	41	64
Mountainous Terrain	13975	0.361	0.335	0	0.259	1
Oil Presence	10058	0.492	0.5	0	0	1
Alluvial Diamonds	10622	0.176	0.381	0	0	1

Distance to Border	14001	39.2	108	0.00338	3.01	1263
Settlement Type	15364					
... Aggregate	187	1.2%				
... Regional & urban	2568	16.7%				
... Regionally based	11293	73.5%				
... Statewide	1316	8.6%				

Appendix A, Table 1. A full table of descriptive statistics for all variables at the group-year level.

Appendix B: Grievance Score Regression Models

Appendix B, Table 1: Grievance Score Regression Results								
Dependent Variable: Onset								
<i>Predictors</i>	(1) GLM		(2) GLM		(3) PMLE		(4) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Grievance Score [log]	1.20 (0.42)	0.605	1.42 (0.87)	0.567	1.60 (0.91)	0.451	1.17 (0.79)	0.811
Country Pop. [log]			0.87 (0.10)	0.212	0.87 (0.09)	0.195	0.90 (0.15)	0.538
Country Size [log]			1.05 (0.12)	0.645	1.06 (0.11)	0.619	1.13 (0.17)	0.434
Democracy Level			0.99 (0.02)	0.549	0.99 (0.02)	0.545	0.99 (0.03)	0.683
Ongoing War			0.00 (0.00)	0.971	0.01 (0.01)	<0.001	0.00 (0.00)	0.976
Post Cold War			1.71 (0.65)	0.158	1.63 (0.59)	0.168	2.35 (1.01)	0.046
New Country			14.68 (6.91)	<0.001	15.27 (6.77)	<0.001	13.65 (7.55)	<0.001
Instability			2.79 (0.79)	<0.001	2.79 (0.75)	<0.001	2.63 (0.86)	0.003
Group Pop. [log]			1.10 (0.09)	0.232	1.09 (0.08)	0.239	0.98 (0.10)	0.875
Group GDP/capita [log]			0.95 (0.29)	0.877	1.02 (0.28)	0.942	0.75 (0.22)	0.332
Group Political Power			0.59 (0.06)	<0.001	0.60 (0.06)	<0.001	0.64 (0.08)	<0.001
Past Wars			1.87 (0.26)	<0.001	1.88 (0.25)	<0.001	1.73 (0.36)	0.009

Grievance Score	0.56	0.073	0.53	0.076	0.63	0.146		
Outlier	(0.18)		(0.15)		(0.20)			
Country	0.00	0.999	0.89	0.961	0.00	0.999		
GDP/capita [log]	(0.00)		(2.12)		(0.00)			
Country	0.01	1.000	4.39	0.542				
GDP/capita	(147)		(10.07)					
Outlier								
Military Expenditures [log]					1.05	0.414		
					(0.06)			
Total Ethnic Groups					0.98	0.396		
					(0.03)			
Mountainous Terrain					1.16	0.748		
					(0.55)			
Oil Presence					1.19	0.591		
					(0.39)			
Alluvial Diamonds					1.30	0.496		
					(0.50)			
Distance to Border [log]					0.89	0.090		
					(0.06)			
(Intercept)	0.01	<0.001	2.29	0.616	2.53	0.572	3.06	0.686
	(0.00)		(3.78)		(3.98)		(8.46)	
Observations	13512		12883		12883		8321	
R ² Tjur	0.000		0.035		0.037		0.033	

Appendix B, Table 1: Models of the regression of civil war onset on grievance score. Model 1 is a bivariate regression, Models 2 and 3 are multivariate with best set of control variables available, and Model 4 adds extra controls beyond that for robustness. Models 1, 2, and 4 show standard errors clustered by country. Corresponds to Table 2 in the body of the paper.

Appendix B, Table 2: 10-Year Lag Grievance Score Regressions								
Dependent Variable: Onset								
	(1) GLM		(2) GLM		(3) PMLE		(4) GLM	
<i>Predictors</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>

Grievance Score [log]	1.34 (0.47)	0.400	1.65 (1.00)	0.409	1.87 (1.05)	0.315	1.28 (0.86)	0.719
Country Pop. [log]			0.83 (0.09)	0.097	0.83 (0.09)	0.089	0.89 (0.16)	0.494
Country Size [log]			1.08 (0.13)	0.488	1.09 (0.12)	0.465	1.15 (0.18)	0.388
Democracy Level			0.99 (0.02)	0.690	0.99 (0.02)	0.681	0.98 (0.03)	0.482
Ongoing War			0.00 (0.00)	0.973	0.01 (0.02)	<0.001	0.00 (0.00)	0.977
Post Cold War			1.65 (0.63)	0.192	1.57 (0.57)	0.206	2.34 (1.01)	0.048
New Country			15.22 (7.24)	<0.001	15.79 (7.05)	<0.001	13.96 (7.83)	<0.001
Instability			2.52 (0.76)	0.002	2.53 (0.73)	0.003	2.49 (0.85)	0.008
Group Pop. [log]			1.13 (0.09)	0.135	1.12 (0.09)	0.142	0.97 (0.11)	0.775
Group GDP/capita [log]			0.92 (0.28)	0.794	0.99 (0.28)	0.978	0.76 (0.22)	0.347
Group Political Power			0.59 (0.06)	<0.001	0.60 (0.06)	<0.001	0.64 (0.08)	0.001
Past Wars			1.59 (0.40)	0.068	1.65 (0.37)	0.074	1.49 (0.46)	0.195
Grievance Score Outlier			0.56 (0.18)	0.077	0.52 (0.15)	0.079	0.63 (0.20)	0.153
Country GDP/capita [log]			0.00 (0.00)	0.999	0.40 (0.96)	0.711	0.00 (0.00)	0.999
Country GDP/capita Outlier			0.02 (186.9)	1.000	5.52 (13.1)	0.503		
Military Expenditures [log]							1.02 (0.06)	0.741

Total Ethnic Groups							0.98 (0.03)	0.468
Mountainous Terrain							1.35 (0.66)	0.533
Oil Presence							1.32 (0.44)	0.405
Alluvial Diamonds							1.40 (0.54)	0.388
Distance to Border [log]							0.86 (0.06)	0.039
(Intercept)	0.00 (0.00)	<0.001	3.71 (6.39)	0.447	4.15 (6.80)	0.407	4.36 (12.75)	0.614
Observations	13199		12597		12597		8155	
R ² Tjur	0.000		0.020		0.016		0.028	

Appendix B, Table 2: Robustness checks for the Models in Table 2 of the main paper (and Appendix B, Table 1). The same regressions were run, but with a 10-year lag for groups to re-enter the dataset instead of a 5-year lag.

Appendix B, Table 3: Grievance Score Functional Form Robustness								
Dependent Variable: Onset								
<i>Predictors</i>	(1) GLM		(2) GLM		(3) PMLE		(4) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Grievance Score	0.99 (0.11)	0.951	1.09 (0.26)	0.727	1.22 (0.15)	0.263	1.07 (0.30)	0.797
Country Pop. [log]			0.88 (0.10)	0.231	0.87 (0.09)	0.209	0.90 (0.15)	0.538
Country Size [log]			1.06 (0.12)	0.613	1.06 (0.11)	0.613	1.13 (0.17)	0.431
Democracy Level			0.99 (0.02)	0.527	0.99 (0.02)	0.529	0.99 (0.03)	0.679
Ongoing War			0.00 (0.00)	0.971	0.01 (0.01)	<0.001	0.00 (0.00)	0.976

Post Cold War	1.73 (0.66)	0.148	1.65 (0.59)	0.159	2.36 (1.01)	0.044
New Country	14.60 (6.87)	<0.001	15.20 (6.75)	<0.001	13.61 (7.52)	<0.001
Instability	2.78 (0.78)	<0.001	2.78 (0.75)	<0.001	2.63 (0.86)	0.003
Group Pop. [log]	1.09 (0.08)	0.256	1.09 (0.08)	0.250	0.98 (0.10)	0.865
Group GDP/capita [log]	0.90 (0.26)	0.710	0.98 (0.24)	0.946	0.75 (0.21)	0.289
Group Political Power	0.59 (0.06)	<0.001	0.60 (0.06)	<0.001	0.64 (0.08)	<0.001
Past Wars	1.87 (0.26)	<0.001	1.88 (0.25)	<0.001	1.73 (0.36)	0.009
Grievance Score Outlier	0.60 (0.18)	0.088	0.55 (0.14)	0.056	0.63 (0.19)	0.133
Country GDP/capita [log]	0.00 (0.00)	0.999	0.00 (0.00)	0.304	0.00 (0.00)	0.999
Country GDP/capita Outlier	0.02 (185)	1.000	5.15 (11.6)	0.492		
Military Expenditures [log]					1.05 (0.06)	0.408
Total Ethnic Groups					0.98 (0.03)	0.396
Mountainous Terrain					1.17 (0.55)	0.737
Oil Presence					1.19 (0.39)	0.590
Alluvial Diamonds					1.29 (0.49)	0.498
Distance to Border [log]					0.89 (0.06)	0.087

(Intercept)	0.01 (0.00)	<0.001	1.91 (3.13)	0.691	1.98 (3.09)	0.674	2.84 (7.96)	0.709
Observations	13512		12883		12883		8321	
R ² Tjur	0.000		0.035		0.036		0.033	

Appendix B, Table 3: Robustness checks for the Models in Table 2 of the main paper (and Appendix B, Table 1). The same regressions were run, but grievance score was not logged.

Appendix B, Table 4: PMLE Grievance Score Robustness				
<i>Predictors</i>	Dependent Variable: Onset			
	(1) PMLE		(2) PMLE	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Grievance Score	1.28 (0.43)	0.499	1.27 (0.77)	0.724
Country Pop. [log]			0.88 (0.14)	0.454
Country Size [log]			1.14 (0.16)	0.395
Democracy Level			0.99 (0.03)	0.722
Ongoing War			0.01 (0.02)	<0.001
Post Cold War			2.17 (0.84)	0.044
New Country			13.66 (6.87)	<0.001
Instability			2.61 (0.79)	0.004
Group Pop. [log]			0.98 (0.10)	0.873
Group GDP/capita [log]			0.78 (0.20)	0.366

Group Political Power			0.65 (0.07)	<0.001
Past Wars			1.76 (0.33)	0.014
Country GDP/capita [log]			0.60 (0.17)	0.134
Military Expenditures [log]			1.06 (0.05)	0.325
Total Ethnic Groups			0.98 (0.02)	0.397
Mountainous Terrain			1.16 (0.50)	0.757
Oil Presence			1.20 (0.36)	0.580
Alluvial Diamonds			1.30 (0.46)	0.493
Distance to Border [log]			0.89 (0.06)	0.091
Grievance Score Outlier			0.60 (1.56)	0.851
(Intercept)	0.01 (0.00)	<0.001	4.16 (10.57)	0.601
Observations	13512		8321	
R ²	0.000		0.035	

Appendix B, Table 4: Robustness checks for the Models in Table 2 of the main paper (and Appendix B, Table 1). The same regressions as Model 1 and Model 4 were run except using the PMLE method instead of GLM.

Appendix C: Legibility Score Regression Models

Appendix C, Table 1: Legibility Score Regression Results										
Dependent Variable: Onset										
<i>Predictors</i>	(1) GLM		(2) GLM		(3) GLM		(4) GLM		(5) PMLE	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Legibility Score [log]	0.53 (0.08)	<0.001	0.37 (0.07)	<0.001	0.39 (0.08)	<0.001	0.56 (0.13)	0.010	0.57 (0.12)	0.008
Grievance Score [log]					0.73 (0.55)	0.677	0.00 (0.01)	0.004	0.01 (0.01)	0.003
Grievance Score [log] x Legibility Score [log]							0.09 (0.07)	0.002	0.10 (0.07)	0.001
Country Pop. [log]			0.90 (0.12)	0.428	0.90 (0.12)	0.451	1.04 (0.16)	0.769	1.05 (0.14)	0.747
Country Size [log]			1.44 (0.28)	0.060	1.46 (0.28)	0.054	1.40 (0.28)	0.096	1.37 (0.25)	0.107
Democracy Level			1.03 (0.03)	0.438	1.03 (0.03)	0.433	1.02 (0.03)	0.583	1.02 (0.03)	0.566
Ongoing War			0.00 (0.00)	0.984	0.00 (0.00)	0.984	0.00 (0.00)	0.980	0.00 (0.00)	<0.001

Post Cold War			1.90 (0.97)	0.207	1.95 (1.00)	0.192	1.99 (1.04)	0.186	1.77 (0.83)	0.227
New Country			0.00 (0.00)	0.999	0.00 (0.00)	0.999	0.00 (0.00)	0.999	13.80 (21.22)	0.207
Instability			1.83 (0.71)	0.121	1.82 (0.71)	0.124	1.96 (0.77)	0.085	1.97 (0.72)	0.085
Group Pop. [log]			0.92 (0.10)	0.452	0.92 (0.10)	0.440	0.90 (0.09)	0.290	0.89 (0.09)	0.283
Group GDP/capita [log]			0.59 (0.14)	0.030	0.57 (0.15)	0.031	0.55 (0.14)	0.018	0.56 (0.13)	0.020
Group Political Power			0.42 (0.08)	<0.001	0.42 (0.08)	<0.001	0.45 (0.08)	<0.001	0.46 (0.08)	<0.001
Past Wars			1.88 (0.34)	<0.001	1.87 (0.34)	0.001	1.91 (0.34)	<0.001	1.93 (0.32)	0.001
Legibility Score Outlier			0.00 (0.02)	0.998	0.00 (0.02)	0.998	0.00 (0.01)	0.997	24.88 (42.13)	0.157
(Intercept)	0.00 (0.00)	<0.001	0.02 (0.08)	0.235	0.03 (0.10)	0.278	0.02 (0.06)	0.194	0.02 (0.06)	0.204
Observations	8516		8389		8389		8389		8389	
R ² Tjur	0.003		0.061		0.061		0.099		0.110	

Appendix C, Table 1: Model 1 is the bivariate regression of onset on legibility score and Model 2 is the multivariate version of the same regression. Model 3 regresses onset on both legibility score and grievance score, with controls. Model 4 and Model 4 show the regression of onset on the interaction between grievance score and legibility score, with controls. Models 1, 2, 3, and 4 use the GLM method with standard errors clustered by country and Model 5 uses PMLE. Corresponds to Table 3 in the body of the paper.

Appendix C, Table 2: Legibility Score Extra Controls Regressions

Dependent Variable: Onset												
<i>Predictors</i>	(1) GLM		(2) GLM		(3) GLM		(4) GLM		(5) GLM		(6) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Legibility Score [log]	0.43 (0.09)	<0.001	0.26 (0.08)	<0.001	0.42 (0.09)	<0.001	0.25 (0.08)	<0.001	0.58 (0.13)	0.016	0.40 (0.14)	0.007
Grievance Score [log]					1.97 (2.12)	0.529	4.81 (5.98)	0.206	0.01 (0.02)	0.032	0.00 (0.01)	0.045
Grievance Score [log] x Legibility Score [log]									0.09 (0.07)	0.003	0.06 (0.06)	0.003
Country Pop. [log]	0.90 (0.12)	0.432	0.63 (0.18)	0.108	0.89 (0.12)	0.400	0.63 (0.18)	0.109	1.04 (0.16)	0.817	0.75 (0.22)	0.320
Country Size [log]	1.46 (0.28)	0.050	1.37 (0.33)	0.199	1.43 (0.28)	0.070	1.31 (0.32)	0.270	1.38 (0.28)	0.114	1.24 (0.31)	0.394
Democracy Level	1.03 (0.03)	0.404	1.08 (0.05)	0.052	1.03 (0.03)	0.390	1.09 (0.05)	0.050	1.02 (0.03)	0.529	1.08 (0.05)	0.075
Ongoing War	0.00 (0.00)	0.984	0.00 (0.00)	0.986	0.00 (0.00)	0.984	0.00 (0.00)	0.986	0.00 (0.00)	0.980	0.00 (0.00)	0.982
Post Cold War	2.02 (1.04)	0.171	2.07 (1.14)	0.188	1.97 (1.02)	0.190	2.02 (1.11)	0.200	2.03 (1.07)	0.178	1.94 (1.10)	0.241

New Country	0.00 (0.00)	0.999	0.00 (0.00)	0.999	0.00 (0.00)	0.999	0.00 (0.00)	0.999	0.00 (0.00)	0.999	0.00 (0.00)	0.999
Instability	1.79 (0.70)	0.136	1.52 (0.67)	0.337	1.79 (0.70)	0.138	1.51 (0.66)	0.349	1.95 (0.76)	0.090	1.69 (0.76)	0.239
Group Pop. [log]	0.92 (0.10)	0.450	0.79 (0.11)	0.083	0.93 (0.10)	0.466	0.79 (0.11)	0.078	0.90 (0.09)	0.310	0.77 (0.10)	0.052
Group GDP/capita [log]	0.92 (0.43)	0.861	0.74 (0.46)	0.626	1.30 (0.91)	0.711	1.76 (1.58)	0.529	0.94 (0.68)	0.933	1.01 (0.92)	0.991
Group Political Power	0.42 (0.08)	<0.001	0.47 (0.10)	<0.001	0.43 (0.08)	<0.001	0.49 (0.10)	0.001	0.45 (0.08)	<0.001	0.52 (0.11)	0.002
Past Wars	1.89 (0.34)	<0.001	2.72 (0.76)	<0.001	1.93 (0.35)	<0.001	2.82 (0.79)	<0.001	1.94 (0.35)	<0.001	2.81 (0.76)	<0.001
Country GDP/ capita [log]	0.55 (0.29)	0.262	0.24 (0.16)	0.033	0.38 (0.29)	0.212	0.10 (0.09)	0.016	0.54 (0.42)	0.431	0.18 (0.18)	0.080
Legibility Score Outlier	0.00 (0.01)	0.998	0.00 (0.16)	0.999	0.00 (0.02)	0.998	0.00 (0.21)	0.999	0.00 (0.01)	0.997	0.00 (0.05)	0.998
Military Expen- ditures [log]			1.53 (0.18)	<0.001			1.53 (0.18)	<0.001			1.49 (0.18)	0.001
Total Ethnic Groups			0.98 (0.04)	0.692			0.97 (0.04)	0.572			0.98 (0.05)	0.698
Mountainous Territory			1.27 (0.93)	0.744			1.51 (1.14)	0.582			1.57 (1.16)	0.542

Oil Presence			0.24 (0.14)	0.017			0.24 (0.15)	0.019			0.27 (0.16)	0.025
Alluvial Diamonds			1.92 (1.08)	0.242			1.93 (1.09)	0.246			2.05 (1.16)	0.204
Distance to Border [log]			0.96 (0.09)	0.684			0.95 (0.09)	0.584			0.96 (0.10)	0.698
(Intercept)	0.09 (0.30)	0.465	2958.97 (15041)	0.116	0.13 (0.43)	0.539	5429.8 (27759)	0.093	0.04 (0.13)	0.317	1835.5 (9063)	0.128
Observations	8389		5283		8389		5283		8389		5283	
R ² Tjur	0.062		0.128		0.062		0.133		0.100		0.188	

Appendix C, Table 2: Models 1 and 2 in this table correspond to Model 6 in the main paper. Models 3 and 4 correspond to Model 7 in the main paper. Models 5 and 6 correspond to Model 8 in the main paper. In each instance, the models in this table add additional controls to the models in the main paper, either country GDP per capita or a series of extra controls that reduce N.

Appendix C, Table 3: Legibility Score 10-Year Lag Regressions

Dependent Variable: Onset								
<i>Predictors</i>	(1) GLM		(2) GLM		(3) GLM		(4) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Legibility Score [log]	0.43 (0.07)	<0.001	0.30 (0.06)	<0.001	0.31 (0.07)	<0.001	0.48 (0.12)	0.003
Grievance Score [log]					0.80 (0.60)	0.769	0.00 (0.00)	0.004
Grievance Score [log] x Legibility Score [log]							0.07 (0.06)	0.002
Country Pop. [log]			0.88 (0.13)	0.381	0.88 (0.13)	0.395	1.02 (0.16)	0.899
Country Size [log]			1.52 (0.31)	0.045	1.53 (0.32)	0.042	1.47 (0.32)	0.072
Democracy Level			1.04 (0.04)	0.310	1.04 (0.04)	0.305	1.03 (0.04)	0.462
Ongoing War			0.00 (0.00)	0.984	0.00 (0.00)	0.984	0.00 (0.00)	0.980
Post Cold War			2.05 (1.06)	0.166	2.09 (1.09)	0.159	2.10 (1.12)	0.166
New Country			0.00 (0.00)	0.999	0.00 (0.00)	0.999	0.00 (0.00)	0.999
Instability			1.26 (0.55)	0.597	1.25 (0.54)	0.603	1.37 (0.60)	0.467
Group Pop. [log]			0.90 (0.10)	0.353	0.90 (0.10)	0.348	0.88 (0.10)	0.257
Group GDP/capita [log]			0.47 (0.13)	0.005	0.45 (0.13)	0.007	0.43 (0.13)	0.004

Group Political Power			0.42 (0.08)	< 0.001	0.41 (0.08)	< 0.001	0.45 (0.09)	< 0.001
Past Wars			1.61 (0.47)	0.106	1.60 (0.47)	0.113	1.63 (0.48)	0.098
Legibility Score Outlier			0.00 (0.06)	0.998	0.00 (0.06)	0.998	0.00 (0.02)	0.997
(Intercept)	0.00 (0.00)	< 0.001	0.08 (0.26)	0.444	0.10 (0.34)	0.502	0.07 (0.25)	0.438
Observations	8343		8220		8220		8220	
R ² Tjur	0.004		0.058		0.058		0.107	

Appendix C, Table 3: Regression results corresponding to Models 5, 6, 7 and 8 in the main paper but with a 10-year lag for groups to re-enter the dataset instead of a 5-year lag.

Appendix C, Table 4: Legibility Score Functional Form Robustness Checks								
Dependent Variable: Onset								
<i>Predictors</i>	(1) GLM		(2) GLM		(3) GLM		(4) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Legibility Score	0.07 (0.06)	0.002	0.01 (0.01)	< 0.001	0.01 (0.02)	0.002	0.11 (0.15)	0.105
Grievance Score					0.59 (0.26)	0.225	0.02 (0.02)	0.001
Grievance Score x Legibility Score							0.06 (0.04)	< 0.001
Country Pop. [log]			0.91 (0.13)	0.520	0.92 (0.13)	0.559	1.06 (0.16)	0.715
Country Size [log]			1.39 (0.28)	0.099	1.45 (0.29)	0.067	1.35 (0.27)	0.141
Democracy Level			1.03 (0.03)	0.327	1.03 (0.03)	0.310	1.03 (0.03)	0.437

Ongoing War			0.00 (0.00)	0.985	0.00 (0.00)	0.985	0.00 (0.00)	0.983
Post Cold War			1.70 (0.86)	0.297	1.89 (0.98)	0.219	2.06 (1.09)	0.171
New Country			0.00 (0.00)	0.999	0.00 (0.00)	0.999	0.00 (0.00)	0.999
Instability			2.01 (0.79)	0.073	1.98 (0.77)	0.079	2.19 (0.86)	0.046
Group Pop. [log]			0.96 (0.10)	0.688	0.95 (0.10)	0.621	0.92 (0.10)	0.434
Group GDP /capita [log]			0.57 (0.13)	0.017	0.51 (0.13)	0.008	0.49 (0.12)	0.005
Group Political Power			0.42 (0.07)	<0.001	0.40 (0.07)	<0.001	0.43 (0.08)	<0.001
Past Wars			1.94 (0.34)	<0.001	1.88 (0.33)	<0.001	1.89 (0.33)	<0.001
Legibility Score Outlier			328.77 (1152219)	0.999	35.71 (127506)	0.999	0.01 (23.10)	0.999
(Intercept)	0.01 (0.00)	<0.001	0.58 (1.67)	0.849	1.49 (4.48)	0.894	9.84 (29.66)	0.448
Observations	8516		8389		8389		8389	
R ² Tjur	0.002		0.053		0.057		0.093	

Appendix C, Table 4: Regression results corresponding to Models 5, 6, 7 and 8 in the main paper but with functional form alterations. Legibility Score and Grievance Score are not logged in models shown in this table.

Appendix C, Table 5: Legibility Score PMLE Regression Results

Dependent Variable: Onset						
<i>Predictors</i>	(1) PMLE		(2) PMLE		(3) PMLE	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Legibility Score [log]	0.54 (0.08)	<0.001	0.38 (0.07)	<0.001	0.40 (0.08)	<0.001
Grievance Score [log]					0.82 (0.54)	0.780
Country Pop. [log]			0.90 (0.11)	0.453	0.91 (0.11)	0.463
Country Size [log]			1.41 (0.25)	0.065	1.42 (0.25)	0.061
Democracy Level			1.03 (0.03)	0.438	1.03 (0.03)	0.432
Ongoing War			0.01 (0.01)	<0.001	0.01 (0.01)	<0.001
Post Cold War			1.72 (0.80)	0.243	1.75 (0.81)	0.234
New Country			16.82 (24.74)	0.181	15.54 (23.02)	0.191
Instability			1.84 (0.67)	0.119	1.84 (0.67)	0.121
Group Pop. [log]			0.92 (0.09)	0.434	0.92 (0.09)	0.426
Group GDP/capita [log]			0.61 (0.14)	0.036	0.58 (0.14)	0.036
Group Political Power			0.43 (0.07)	<0.001	0.43 (0.07)	<0.001
Past Wars			1.91 (0.32)	0.001	1.90 (0.32)	0.002

Legibility Score Outlier			78.89 (131.45)	0.076	73.04 (122.49)	0.081
(Intercept)	0.00 (0.00)	<0.001	0.03 (0.09)	0.246	0.04 (0.12)	0.299
Observations	8516		8389		8389	
R ²	0.003		0.059		0.060	

Appendix C, Table 5: Regression results corresponding to Models 5, 6, and 7 in the main paper but using the PMLE method instead of GLM.

Appendix D: Administrative Capacity Regression Results

Appendix D, Table 1: Administrative Capacity Regression Results										
Dependent Variable: Onset										
<i>Predictors</i>	(10) GLM		(11) GLM		(12) GLM		(13) GLM		(14) PMLE	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Admin. Capacity	0.28 (0.06)	< 0.001	0.23 (0.07)	< 0.001	0.22 (0.07)	< 0.001	0.28 (0.10)	0.001	0.29 (0.10)	< 0.001
Grievance Score [log]					1.37 (0.80)	0.587	0.73 (0.55)	0.675	0.80 (0.54)	0.762
Grievance Score [log] x Admin. Capacity							0.24 (0.25)	0.162	0.25 (0.22)	0.160
Country Pop. [log]			1.05 (0.16)	0.724	1.05 (0.16)	0.770	1.08 (0.17)	0.601	1.07 (0.15)	0.643
Country Size [log]			1.07 (0.16)	0.675	1.05 (0.16)	0.760	1.06 (0.17)	0.719	1.06 (0.15)	0.698
Democracy Level			0.99 (0.03)	0.678	0.99 (0.03)	0.671	0.98 (0.03)	0.593	0.99 (0.03)	0.596
Ongoing War			0.00 (0.00)	0.974	0.00 (0.00)	0.974	0.00 (0.00)	0.973	0.01 (0.01)	< 0.001
Post Cold War			3.85 (1.87)	0.005	3.90 (1.90)	0.005	4.27 (2.13)	0.004	3.89 (1.77)	0.002

New Country	0.00 (0.00)	0.995	0.00 (0.00)	0.995	0.00 (0.00)	0.995	1.32 (1.83)	0.856
Instability	2.19 (0.73)	0.019	2.19 (0.73)	0.019	2.24 (0.75)	0.016	2.24 (0.71)	0.018
Group Pop. [log]	1.01 (0.09)	0.952	1.02 (0.09)	0.848	1.00 (0.09)	0.985	1.00 (0.08)	0.995
Group GDP/capita [log]	0.85 (0.13)	0.291	0.86 (0.14)	0.343	0.83 (0.13)	0.256	0.84 (0.12)	0.267
Group Political Power	0.54 (0.08)	<0.001	0.55 (0.08)	<0.001	0.55 (0.08)	<0.001	0.56 (0.07)	<0.001
Past Wars	1.96 (0.30)	<0.001	1.96 (0.30)	<0.001	1.98 (0.30)	<0.001	1.98 (0.28)	<0.001
Grievance Score Outlier			0.00 (0.00)	0.999	0.00 (0.05)	0.999	438.23 (1756.62)	0.160
(Intercept)	0.00 (0.00)	<0.001	0.00 (0.01)	0.019	0.00 (0.01)	0.020	0.00 (0.01)	0.016
Observations	9716	9462	9462	9462	9462	9462	9462	
R ² Tjur	0.004	0.056	0.056	0.056	0.060	0.060	0.063	

Appendix D, Table 1: Model 10 is a bivariate regression of onset on administrative capacity and Model 11 is the multivariate version of the same regression. Model 12 regresses onset on both administrative capacity and grievance score, with controls. Model 13 and Model 14 show the regression of onset on the interaction between grievance score and administrative capacity, with controls. Models 10, 11, 12, and 13 use the GLM method with standard errors clustered by country and Model 14 uses PMLE with non-clustered standard errors. Corresponds directly to Table 4 in the main paper.

Appendix D, Table 2: Administrative Capacity Extra Controls

Dependent Variable: Onset												
<i>Predictors</i>	(1) GLM		(2) GLM		(3) GLM		(4) GLM		(5) GLM		(6) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Admin. Capacity	0.24 (0.08)	< 0.001	0.16 (0.06)	< 0.001	0.22 (0.07)	< 0.001	0.15 (0.06)	< 0.001	0.28 (0.10)	0.001	0.18 (0.07)	< 0.001
Grievance Score [log]					1.99 (1.43)	0.336	1.66 (1.28)	0.511	1.04 (0.92)	0.969	1.14 (1.10)	0.890
Grievance Score [log] x Admin. Capacity									0.26 (0.26)	0.182	0.47 (0.54)	0.507
Country Pop. [log]	1.06 (0.16)	0.705	1.06 (0.20)	0.777	1.05 (0.16)	0.751	1.07 (0.21)	0.711	1.09 (0.17)	0.591	1.09 (0.21)	0.649
Country Size [log]	1.07 (0.16)	0.655	1.26 (0.24)	0.208	1.04 (0.16)	0.793	1.25 (0.24)	0.233	1.05 (0.16)	0.750	1.26 (0.24)	0.214
Democracy Level	0.99 (0.03)	0.707	1.02 (0.03)	0.612	0.99 (0.03)	0.735	1.02 (0.03)	0.634	0.99 (0.03)	0.648	1.01 (0.03)	0.678
Ongoing War	0.00 (0.00)	0.974	0.00 (0.00)	0.976	0.00 (0.00)	0.974	0.00 (0.00)	0.975	0.00 (0.00)	0.973	0.00 (0.00)	0.974
Post Cold War	3.80 (1.84)	0.006	4.96 (2.67)	0.003	3.80 (1.85)	0.006	4.87 (2.63)	0.003	4.17 (2.07)	0.004	5.07 (2.78)	0.003

New Country	0.00 (0.00)	0.995	0.00 (0.00)	0.995	0.00 (0.00)	0.995	0.00 (0.00)	0.995	0.00 (0.00)	0.995	0.00 (0.00)	0.995
Instability	2.18 (0.73)	0.020	2.15 (0.80)	0.041	2.19 (0.73)	0.019	2.14 (0.80)	0.042	2.24 (0.75)	0.016	2.16 (0.81)	0.039
Group Pop. [log]	1.00 (0.09)	0.995	0.95 (0.10)	0.637	1.02 (0.09)	0.852	0.95 (0.11)	0.669	1.00 (0.09)	0.986	0.94 (0.11)	0.610
Group GDP/ capita [log]	0.94 (0.30)	0.849	0.81 (0.26)	0.503	1.17 (0.44)	0.677	0.92 (0.35)	0.831	1.09 (0.42)	0.819	0.90 (0.34)	0.780
Group Political Power	0.54 (0.08)	<0.001	0.62 (0.10)	0.002	0.55 (0.08)	<0.001	0.63 (0.10)	0.003	0.56 (0.08)	<0.001	0.64 (0.10)	0.004
Past Wars	1.95 (0.30)	<0.001	1.74 (0.41)	0.019	1.97 (0.30)	<0.001	1.76 (0.42)	0.017	1.98 (0.30)	<0.001	1.78 (0.42)	0.014
Country GDP/ capita [log]	0.88 (0.30)	0.718	0.77 (0.28)	0.487	0.71 (0.27)	0.369	0.68 (0.27)	0.330	0.74 (0.29)	0.442	0.69 (0.28)	0.356
Country GDP /capita Outlier	0.10 (1070.46)	1.000			0.05 (515.41)	1.000			0.07 (701.46)	1.000		
Military Expe- nditures [log]			1.14 (0.08)	0.064			1.13 (0.08)	0.074			1.13 (0.08)	0.084
Total Ethnic Groups			0.98 (0.03)	0.447			0.98 (0.03)	0.366			0.97 (0.03)	0.361
Mountainous Territory			0.69 (0.38)	0.495			0.68 (0.37)	0.477			0.71 (0.40)	0.545
Oil Presence			0.63 (0.25)	0.245			0.63 (0.25)	0.248			0.63 (0.25)	0.247

Alluvial Diamonds			0.60 (0.27)	0.256			0.61 (0.28)	0.275			0.63 (0.29)	0.307
Distance to Border [log]			0.84 (0.06)	0.018			0.84 (0.06)	0.022			0.84 (0.06)	0.021
Grievance Score Outlier					0.00 (0.00)	0.999	0.00 (0.00)	0.999	0.00 (0.03)	0.999	0.00 (0.00)	0.999
(Intercept)	0.00 (0.01)	0.024	0.00 (0.00)	0.069	0.00 (0.01)	0.029	0.00 (0.00)	0.062	0.00 (0.01)	0.021	0.00 (0.00)	0.056
Observations	9462		7400		9462		7400		9462		7400	
R ² Tjur	0.056		0.066		0.057		0.066		0.060		0.067	

Appendix D, Table 2: Models 1 and 2 in this table correspond to Model 11 in the main paper. Models 3 and 4 correspond to Model 12 in the main paper. Models 5 and 6 correspond to Model 13 in the main paper. In each instance, the models in this table add additional controls to the models in the main paper, either country GDP per capita or a series of extra controls that reduce

Appendix D, Table 3: Administrative Capacity 10-Year Lag Regressions

Dependent Variable: Onset								
<i>Predictors</i>	(1) GLM		(2) GLM		(3) GLM		(4) GLM	
	<i>Odds Ratios</i>	<i>P</i>	<i>Odds Ratios</i>	<i>P</i>	<i>Odds Ratios</i>	<i>P</i>	<i>Odds Ratios</i>	<i>P</i>
Admin. Capacity	0.25 (0.06)	< 0.001	0.20 (0.07)	< 0.001	0.18 (0.06)	< 0.001	0.23 (0.09)	< 0.001
Grievance Score [log]					1.70 (0.98)	0.356	0.81 (0.66)	0.798
Grievance Score [log] x Admin. Capacity							0.24 (0.25)	0.173
Country Pop. [log]			1.00 (0.16)	0.985	0.99 (0.16)	0.934	1.02 (0.17)	0.890
Country Size [log]			1.13 (0.18)	0.443	1.10 (0.18)	0.550	1.12 (0.19)	0.501
Democracy Level			0.99 (0.03)	0.747	0.99 (0.03)	0.706	0.99 (0.03)	0.649
Ongoing War			0.00 (0.00)	0.975	0.00 (0.00)	0.975	0.00 (0.00)	0.983
Post Cold War			3.90 (1.93)	0.006	4.06 (2.04)	0.005	4.41 (2.26)	0.004
New Country			0.00 (0.00)	0.995	0.00 (0.00)	0.995	0.00 (0.00)	0.997
Instability			1.89 (0.69)	0.079	1.91 (0.70)	0.076	1.96 (0.72)	0.065
Group Pop. [log]			1.03 (0.09)	0.723	1.05 (0.10)	0.575	1.03 (0.10)	0.719
Group GDP/capita [log]			0.84 (0.14)	0.273	0.86 (0.15)	0.361	0.83 (0.14)	0.267

Group Political Power		0.53 (0.08)	< 0.001	0.54 (0.08)	< 0.001	0.54 (0.08)	< 0.001	
Past Wars		1.62 (0.51)	0.128	1.63 (0.51)	0.119	1.70 (0.53)	0.092	
Grievance Score Outlier				0.00 (0.00)	0.999	0.00 (0.02)	0.999	
(Intercept)	0.00 (0.00)	< 0.001	0.00 (0.01)	0.020	0.00 (0.01)	0.020	0.00 (0.00)	0.016
Observations	9517		9267		9267		9267	
R ² Tjur	0.004		0.032		0.033		0.037	

Appendix D, Table 3: Regression results corresponding to Models 10, 11, 12 and 13 in the main paper but with a 10-year lag for groups to re-enter the dataset instead of a 5-year lag.

Appendix D, Table 4: Administrative Capacity PMLE Regression Results						
Dependent Variable: Onset						
<i>Predictors</i>	(1) PMLE		(2) PMLE		(3) PMLE	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Admin. Capacity	0.28 (0.06)	< 0.001	0.24 (0.07)	< 0.001	0.23 (0.07)	< 0.001
Grievance Score [log]					1.45 (0.76)	0.527
Country Pop. [log]			1.05 (0.15)	0.757	1.03 (0.15)	0.818
Country Size [log]			1.07 (0.15)	0.663	1.05 (0.15)	0.742
Democracy Level			0.99 (0.03)	0.686	0.99 (0.03)	0.681
Ongoing War			0.01 (0.01)	< 0.001	0.01 (0.01)	< 0.001
Post Cold War			3.55 (1.60)	0.003	3.56 (1.60)	0.003

New Country			1.31 (1.85)	0.858	1.32 (1.84)	0.856
Instability			2.18 (0.69)	0.022	2.18 (0.69)	0.022
Group Pop. [log]			1.00 (0.08)	0.979	1.02 (0.08)	0.861
Group GDP/capita [log]			0.85 (0.12)	0.303	0.86 (0.13)	0.357
Group Political Power			0.55 (0.07)	<0.001	0.55 (0.07)	<0.001
Past Wars			1.96 (0.28)	<0.001	1.97 (0.28)	<0.001
Grievance Score Outlier					4.77 (12.30)	0.572
(Intercept)	0.00 (0.00)	<0.001	0.00 (0.01)	0.020	0.00 (0.01)	0.021
Observations	9716			9462		9462
R ²	0.004			0.057		0.056

Appendix D, Table 4: Regression results corresponding to Models 10, 11, and 12 in the main paper but using PMLE instead of GLM.

Appendix E: Ethnicity Collection Regression Models

Appendix E, Table 1: Ethnicity Collection Regression Results										
Dependent Variable: Onset										
<i>Predictors</i>	(15) GLM		(16) GLM		(17) GLM		(18) GLM		(19) PMLE	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Ethnicity Collection	2.60 (0.83)	0.003	4.83 (1.84)	<0.001	5.12 (2.00)	<0.001	3.24 (1.35)	0.005	3.09 (1.20)	0.002
Grievance Score [log]					2.04 (1.30)	0.263	0.23 (0.32)	0.282	0.32 (0.38)	0.362
Grievance Score [log] x Ethnicity Collection							11.30 (14.67)	0.062	9.36 (10.80)	0.099
Country Pop. [log]			1.13 (0.14)	0.321	1.13 (0.14)	0.358	1.20 (0.16)	0.178	1.19 (0.15)	0.184
Country Size [log]			0.93 (0.12)	0.595	0.90 (0.12)	0.455	0.90 (0.12)	0.456	0.90 (0.12)	0.450
Democracy Level			0.97 (0.02)	0.181	0.97 (0.03)	0.196	0.96 (0.02)	0.110	0.96 (0.02)	0.108
Ongoing War			0.00 (0.00)	0.972	0.00 (0.00)	0.972	0.00 (0.00)	0.972	0.01 (0.01)	<0.001

Post Cold War			4.37 (2.13)	0.002	4.30 (2.09)	0.003	5.50 (2.82)	0.001	5.04 (2.39)	<0.001
New Country			6.27 (4.97)	0.020	6.21 (4.93)	0.021	5.20 (4.18)	0.040	6.15 (4.37)	0.038
Instability			1.83 (0.56)	0.049	1.85 (0.57)	0.045	1.83 (0.56)	0.050	1.84 (0.54)	0.052
Group Pop. [log]			1.04 (0.08)	0.585	1.06 (0.09)	0.492	1.03 (0.08)	0.703	1.03 (0.08)	0.734
Group GDP/capita [log]			1.07 (0.30)	0.798	1.33 (0.44)	0.395	1.19 (0.40)	0.599	1.28 (0.38)	0.480
Group Political Power			0.60 (0.07)	<0.001	0.61 (0.07)	<0.001	0.63 (0.07)	<0.001	0.64 (0.07)	<0.001
Past Wars			1.74 (0.24)	<0.001	1.76 (0.25)	<0.001	1.75 (0.25)	<0.001	1.76 (0.23)	<0.001
Country GDP/capita [log]			0.55 (0.16)	0.035	0.44 (0.15)	0.015	0.49 (0.17)	0.034	0.46 (0.14)	0.041
Country GDP/capita outlier			0.04 (391.62)	1.000	0.02 (181.01)	1.000	0.03 (281.66)	1.000	7.53 (17.37)	0.415
Grievance Score Outlier					0.00 (0.00)	0.999	0.00 (0.00)	0.998	0.40 (0.96)	0.709
(Intercept)	0.00 (0.00)	<0.001	0.02 (0.05)	0.068	0.03 (0.06)	0.098	0.02 (0.04)	0.057	0.02 (0.04)	0.068
Observations	12028		11758		11758		11758		11758	

R ² Tjur	0.001	0.038	0.037	0.040	0.038
---------------------	-------	-------	-------	-------	-------

Appendix E, Table 1: Model 1 is a bivariate regression of onset on ethnicity collection and Model 2 is the multivariate version of the same regression. Model 3 regresses onset on both ethnicity collection and grievance score, with controls. Model 4 and Model 5 show the regression of onset on the interaction between grievance score and ethnicity collection, with controls. Models 1, 2, 3, and 4 use the GLM method with standard errors clustered by country and Model 5 uses PMLE with non-clustered standard errors. Corresponds to Table 5 in the main paper

Appendix E, Table 2: Ethnicity Collection Extra Controls

<i>Predictors</i>	Dependent Variable: Onset					
	(1) GLM		(2) GLM		(3) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Ethnicity Collection	17.10 (10.11)	<0.001	17.73 (10.65)	<0.001	15.58 (12.13)	<0.001
Grievance Score [log]			1.67 (1.10)	0.441	1.10 (2.00)	0.959
Grievance Score [log] x Ethnicity Collection					1.54 (2.69)	0.805
Country Pop. [log]	1.10 (0.20)	0.601	1.11 (0.20)	0.581	1.11 (0.21)	0.562
Country Size [log]	0.95 (0.16)	0.774	0.93 (0.16)	0.685	0.93 (0.16)	0.675
Democracy Level	0.98 (0.03)	0.465	0.98 (0.03)	0.485	0.98 (0.03)	0.472
Ongoing War	0.00 (0.00)	0.976	0.00 (0.00)	0.976	0.00 (0.00)	0.976
Post Cold War	8.18 (4.89)	<0.001	7.91 (4.73)	0.001	8.17 (5.02)	0.001
New Country	1.47 (1.63)	0.728	1.50 (1.66)	0.713	1.47 (1.63)	0.728
Instability	1.94 (0.69)	0.062	1.92 (0.68)	0.066	1.92 (0.68)	0.067
Group Pop. [log]	0.94 (0.09)	0.539	0.94 (0.10)	0.572	0.94 (0.10)	0.563
Group GDP/capita [log]	0.97 (0.27)	0.911	1.09 (0.38)	0.796	1.08 (0.38)	0.830
Group Political Power	0.65 (0.08)	0.001	0.66 (0.09)	0.001	0.66 (0.09)	0.001

Past Wars	1.52 (0.32)	0.045	1.54 (0.32)	0.038	1.55 (0.33)	0.037
Country GDP/capita [log]	0.47 (0.15)	0.017	0.41 (0.15)	0.014	0.42 (0.15)	0.018
Military Expenditures [log]	1.12 (0.07)	0.064	1.12 (0.07)	0.071	1.12 (0.07)	0.078
Total Ethnic Groups	1.00 (0.02)	0.937	1.00 (0.03)	0.957	1.00 (0.03)	0.986
Mountainous Territory	0.79 (0.39)	0.627	0.80 (0.40)	0.647	0.81 (0.41)	0.679
Oil Presence	1.06 (0.40)	0.885	1.08 (0.41)	0.833	1.07 (0.41)	0.852
Alluvial Diamonds	0.75 (0.32)	0.494	0.77 (0.33)	0.548	0.78 (0.34)	0.570
Distance to Border [log]	0.85 (0.06)	0.032	0.85 (0.07)	0.041	0.86 (0.07)	0.042
Grievance Score Outlier			0.00 (0.00)	0.999	0.00 (0.00)	0.999
(Intercept)	0.04 (0.14)	0.336	0.05 (0.16)	0.360	0.05 (0.16)	0.365
Observations	7608		7608		7608	
R ² Tjur	0.051		0.050		0.050	

Appendix E, Table 2: Model 1 this table corresponds to Model 16 in the main paper. Model 2 corresponds to Model 17 in the main paper. Model 3 corresponds to Model 18 in the main paper. In each instance, the models in this table add a series of extra controls that reduce N.

Appendix E, Table 3: Ethnicity Collection 10-Year Lag Regressions								
Dependent Variable: Onset								
<i>Predictors</i>	(1) GLM		(2) GLM		(3) GLM		(4) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Ethnicity Collection	2.70 (0.94)	0.004	5.33 (2.25)	<0.001	5.63 (2.42)	<0.001	3.84 (1.78)	0.004

Grievance Score [log]			2.13 (1.36)	0.236	0.39 (0.53)	0.489
Grievance Score [log] x Ethnicity Collection					6.53 (8.59)	0.153
Country Pop. [log]	1.09 (0.15)	0.514	1.08 (0.15)	0.571	1.13 (0.16)	0.378
Country Size [log]	0.97 (0.13)	0.800	0.93 (0.13)	0.638	0.93 (0.13)	0.636
Democracy Level	0.97 (0.03)	0.252	0.97 (0.03)	0.267	0.96 (0.03)	0.179
Ongoing War	0.00 (0.00)	0.973	0.00 (0.00)	0.973	0.00 (0.00)	0.973
Post Cold War	4.30 (2.12)	0.003	4.24 (2.09)	0.003	5.11 (2.63)	0.002
New Country	6.33 (5.06)	0.021	6.25 (5.02)	0.023	5.48 (4.45)	0.036
Instability	1.61 (0.53)	0.154	1.62 (0.54)	0.147	1.61 (0.53)	0.153
Group Pop. [log]	1.07 (0.09)	0.403	1.09 (0.09)	0.330	1.06 (0.09)	0.468
Group GDP/capita [log]	1.00 (0.29)	0.990	1.27 (0.44)	0.503	1.17 (0.40)	0.652
Group Political Power	0.59 (0.07)	<0.001	0.60 (0.07)	<0.001	0.61 (0.07)	<0.001
Past Wars	1.51 (0.38)	0.101	1.53 (0.38)	0.086	1.56 (0.39)	0.079
Country GDP/capita [log]	0.57 (0.17)	0.066	0.45 (0.16)	0.026	0.49 (0.17)	0.043
Country GDP/ capita outlier	0.05 (667.2)	1.000	0.02 (282)	1.000	0.03 (377)	1.000
Grievance Score Outlier			0.00 (0.00)	0.999	0.00 (0.00)	0.998

(Intercept)	0.00 (0.00)	<0.001	0.02 (0.05)	0.083	0.03 (0.07)	0.119	0.02 (0.05)	0.086
Observations	11735		11492		11492		11492	
R ² Tjur	0.001		0.022		0.022		0.023	

Appendix E, Table 3: Regression results corresponding to Models 15, 16, 17, and 18 in the main paper but with a 10-year lag for groups to re-enter the dataset instead of a 5-year lag.

Appendix E, Table 4: Ethnicity Collection PMLE Regressions						
Dependent Variable: Onset						
<i>Predictors</i>	(1) PMLE		(2) PMLE		(3) PMLE	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Ethnicity Collection	2.52 (0.79)	0.001	4.55 (1.63)	<0.001	4.83 (1.75)	<0.001
Grievance Score [log]					2.29 (1.33)	0.209
Country Pop. [log]			1.13 (0.14)	0.325	1.12 (0.14)	0.378
Country Size [log]	\		0.93 (0.12)	0.593	0.90 (0.12)	0.453
Democracy Level			0.97 (0.02)	0.181	0.97 (0.02)	0.197
Ongoing War			0.01 (0.01)	<0.001	0.01 (0.01)	<0.001
Post Cold War			4.02 (1.83)	0.001	3.92 (1.77)	0.001
New Country			7.54 (5.27)	0.022	7.43 (5.19)	0.024
Instability			1.84 (0.54)	0.052	1.87 (0.55)	0.047
Group Pop. [log]			1.04 (0.08)	0.616	1.05 (0.08)	0.510

Group GDP/capita [log]			1.10 (0.29)	0.744	1.43 (0.42)	0.315
Group Political Power			0.61 (0.06)	<0.001	0.62 (0.07)	<0.001
Past Wars			1.74 (0.23)	<0.001	1.76 (0.23)	<0.001
Country GDP/capita [log]			0.54 (0.14)	0.042	0.41 (0.12)	0.022
Country GDP/capita outlier			12.23 (27.83)	0.314	4.83 (11.13)	0.520
Grievance Score Outlier					1.01 (2.46)	0.996
(Intercept)	0.00 (0.00)	<0.001	0.03 (0.06)	0.079	0.04 (0.08)	0.118
Observations	12028			11758		11758
R ²	0.001			0.036		0.034

Appendix E, Table 4: Regression results corresponding to Models 15, 16, and 17 in the main paper but using PMLE instead of GLM.

Appendix F: Transnational Kin Regression Results

Appendix F, Table 1: Transnational Kin Regression Results										
Dependent Variable: Onset										
<i>Predictors</i>	(20) GLM		(21) GLM		(22) GLM		(23) GLM		(24) PMLE	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Transnational Kin	1.00 (0.02)	0.880	0.99 (0.03)	0.676	0.99 (0.03)	0.694	1.00 (0.03)	0.911	1.00 (0.02)	0.881
Grievance Score [log]					1.41 (0.86)	0.579	1.78 (1.20)	0.389	2.06 (1.26)	0.297
Grievance Score [log] x Transnational Kin							0.84 (0.20)	0.452	0.85 (0.18)	0.461
Country Pop. [log]			0.88 (0.10)	0.225	0.87 (0.10)	0.198	0.86 (0.10)	0.167	0.85 (0.09)	0.153
Country Size [log]			1.07 (0.12)	0.560	1.05 (0.12)	0.636	1.05 (0.12)	0.660	1.05 (0.11)	0.633
Democracy Level			0.98 (0.02)	0.464	0.99 (0.02)	0.522	0.99 (0.02)	0.530	0.99 (0.02)	0.529
Ongoing War			0.00 (0.00)	0.971	0.00 (0.00)	0.971	0.00 (0.00)	0.971	0.01 (0.01)	<0.001
Post Cold War			1.75 (0.66)	0.141	1.72 (0.66)	0.152	1.75 (0.67)	0.142	1.66 (0.59)	0.153

New Country	14.73 (6.95)	<0.001	14.81 (6.99)	<0.001	14.85 (7.01)	<0.001	15.34 (6.76)	<0.001
Instability	2.79 (0.79)	<0.001	2.78 (0.79)	<0.001	2.79 (0.79)	<0.001	2.80 (0.75)	<0.001
Group Pop. [log]	1.09 (0.08)	0.255	1.10 (0.09)	0.213	1.12 (0.09)	0.172	1.11 (0.09)	0.179
Group GDP/capita [log]	0.90 (0.21)	0.646	0.96 (0.29)	0.893	0.95 (0.29)	0.868	1.02 (0.28)	0.954
Group Political Power	0.59 (0.06)	<0.001	0.59 (0.06)	<0.001	0.58 (0.06)	<0.001	0.59 (0.06)	<0.001
Past Wars	1.86 (0.26)	<0.001	1.87 (0.26)	<0.001	1.90 (0.27)	<0.001	1.90 (0.25)	<0.001
Country GDP/capita [log]	0.61 (0.16)	0.055	0.56 (0.18)	0.073	0.57 (0.18)	0.079	0.53 (0.15)	0.080
Country GDP/capita outlier	0.02 (204.98)	1.000	0.01 (151.79)	1.000	0.02 (175.28)	1.000	5.18 (11.80)	0.498
Grievance Score Outlier			0.00 (0.00)	0.999	0.00 (0.00)	0.999	0.31 (0.82)	0.673
(Intercept)	0.01 (0.00)	<0.001	1.80 (2.92)	0.719	2.24 (3.70)	0.626	2.48 (4.13)	0.539
Observations	13644	12883	12883	12883	12883	12883	12883	12883
R ² Tjur	0.000	0.035	0.035	0.035	0.036	0.036	0.038	0.038

Appendix F, Table 1: Model 20 is a bivariate regression of onset on transnational kin and Model 21 is the multivariate version of the same regression. Model 22 regresses onset on both transnational kin and grievance score, with controls. Models 23 and 24 show the

regression of onset on the interaction between grievance score and transnational kin, with controls. Models 20, 21, 22, and 23 use the GLM method with standard errors clustered by country and Model 24 uses PMLE with non-clustered standard errors. Corresponds to Table 6 in the main paper.

Appendix F, Table 2: Transnational Kin Extra Controls

Dependent Variable: Onset						
<i>Predictors</i>	(1) GLM		(2) GLM		(3) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Transnational Kin	0.99 (0.03)	0.767	0.99 (0.03)	0.764	1.00 (0.03)	0.973
Grievance Score [log]			1.17 (0.78)	0.817	2.02 (1.54)	0.358
Grievance Score [log] x Transnational Kin					0.71 (0.20)	0.215
Country Pop. [log]	0.90 (0.15)	0.551	0.90 (0.15)	0.527	0.90 (0.15)	0.518
Country Size [log]	1.14 (0.17)	0.375	1.13 (0.17)	0.428	1.13 (0.17)	0.436
Democracy Level	0.98 (0.03)	0.572	0.99 (0.03)	0.659	0.99 (0.03)	0.650
Ongoing War	0.00 (0.00)	0.976	0.00 (0.00)	0.976	0.00 (0.00)	0.976
Post Cold War	2.40 (1.03)	0.040	2.36 (1.01)	0.045	2.44 (1.04)	0.038
New Country	13.76 (7.61)	<0.001	13.65 (7.56)	<0.001	13.31 (7.38)	<0.001
Instability	2.67 (0.87)	0.003	2.62 (0.86)	0.003	2.63 (0.86)	0.003
Group Pop. [log]	0.98 (0.10)	0.827	0.99 (0.10)	0.896	0.99 (0.11)	0.931
Group GDP/capita [log]	0.77 (0.18)	0.253	0.75 (0.22)	0.333	0.74 (0.22)	0.295
Group Political Power	0.64 (0.08)	<0.001	0.64 (0.08)	<0.001	0.63 (0.08)	<0.001

Past Wars	1.72 (0.36)	0.009	1.73 (0.36)	0.009	1.77 (0.37)	0.006
Country GDP/capita [log]	0.63 (0.18)	0.096	0.62 (0.20)	0.143	0.63 (0.20)	0.157
Military Expenditures [log]	1.04 (0.06)	0.448	1.05 (0.06)	0.403	1.04 (0.06)	0.427
Total Ethnic Groups	0.98 (0.02)	0.390	0.98 (0.03)	0.379	0.97 (0.03)	0.297
Mountainous Territory	1.17 (0.55)	0.737	1.16 (0.55)	0.756	1.16 (0.55)	0.756
Oil Presence	1.18 (0.39)	0.609	1.20 (0.39)	0.573	1.23 (0.41)	0.528
Alluvial Diamonds	1.28 (0.49)	0.515	1.29 (0.49)	0.507	1.30 (0.50)	0.498
Distance to Border [log]	0.88 (0.06)	0.068	0.89 (0.06)	0.086	0.87 (0.06)	0.050
Grievance Score Outlier			0.00 (0.00)	0.999	0.00 (0.00)	0.999
(Intercept)	2.31 (6.30)	0.758	3.03 (8.37)	0.688	3.47 (9.60)	0.652
Observations	8321		8321		8321	
R ² Tjur	0.033		0.034		0.037	

Appendix F, Table 2: Model 1 this table corresponds to Model 21 in the main paper. Model 2 corresponds to Model 22 in the main paper. Model 3 corresponds to Model 23 in the main paper. In each instance, the models in this table add a series of extra controls that reduce N.

Appendix F, Table 3: Transnational Kin 10-Year Lag Regressions								
Dependent Variable: Onset								
Predictors	(1) GLM		(2) GLM		(3) GLM		(4) GLM	
	Odds Ratios	<i>p</i>	Odds Ratios	<i>p</i>	Odds Ratios	<i>p</i>	Odds Ratios	<i>p</i>
Transnational Kin	1.00 (0.02)	0.896	0.99 (0.03)	0.828	1.00 (0.03)	0.856	1.00 (0.03)	0.922

Grievance Score [log]			1.64 (1.00)	0.414	2.11 (1.39)	0.259
Grievance Score [log] x Transnational Kin					0.83 (0.19)	0.417
Country Pop. [log]	0.84 (0.09)	0.119	0.83 (0.09)	0.095	0.81 (0.09)	0.076
Country Size [log]	1.10 (0.13)	0.396	1.08 (0.13)	0.483	1.08 (0.13)	0.506
Democracy Level	0.99 (0.02)	0.592	0.99 (0.02)	0.677	0.99 (0.02)	0.691
Ongoing War	0.00 (0.00)	0.973	0.00 (0.00)	0.973	0.00 (0.00)	0.972
Post Cold War	1.68 (0.64)	0.175	1.65 (0.63)	0.189	1.69 (0.65)	0.172
New Country	15.14 (7.20)	<0.001	15.26 (7.27)	<0.001	15.27 (7.27)	<0.001
Instability	2.54 (0.76)	0.002	2.52 (0.76)	0.002	2.52 (0.76)	0.002
Group Pop. [log]	1.11 (0.09)	0.179	1.13 (0.09)	0.133	1.15 (0.10)	0.104
Group GDP/capita [log]	0.84 (0.20)	0.461	0.93 (0.29)	0.801	0.91 (0.28)	0.771
Group Political Power	0.59 (0.06)	<0.001	0.59 (0.06)	<0.001	0.58 (0.06)	<0.001
Past Wars	1.58 (0.40)	0.070	1.59 (0.40)	0.067	1.63 (0.42)	0.054
Country GDP/capita [log]	0.63 (0.17)	0.086	0.56 (0.18)	0.076	0.56 (0.19)	0.083
Country GDP/capita outlier	0.02 (298.2)	1.000	0.02 (189.5)	1.000	0.02 (216.4)	1.000

Grievance Score Outlier					0.00 (0.00)	0.999	0.00 (0.00)	0.999
(Intercept)	0.01 (0.00)	<0.001	2.64 (4.49)	0.567	3.67 (6.34)	0.451	4.28 (7.44)	0.403
Observations	13331		12597		12597		12597	
R ² Tjur	0.000		0.019		0.020		0.021	

Appendix F, Table 3: Regression results corresponding to Models 20, 21, 22, and 23 in the main paper but with a 10-year lag for groups to re-enter the dataset instead of a 5-year lag. Like in all GLM models, standard errors are clustered by country.

Appendix F, Table 4: Transnational Kin PMLE Regression Results						
Dependent Variable: Onset						
<i>Predictors</i>	(1) PMLE		(2) PMLE		(3) PMLE	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Transnational Kin	1.00 (0.02)	0.970	0.99 (0.02)	0.825	1.00 (0.02)	0.848
Grievance Score [log]					1.57 (0.89)	0.467
Country Pop. [log]			0.88 (0.09)	0.226	0.86 (0.09)	0.185
Country Size [log]			1.07 (0.11)	0.539	1.06 (0.11)	0.605
Democracy Level			0.98 (0.02)	0.463	0.99 (0.02)	0.518
Ongoing War			0.01 (0.01)	<0.001	0.01 (0.01)	<0.001
Post Cold War			1.67 (0.60)	0.146	1.64 (0.59)	0.164
New Country			15.19 (6.74)	<0.001	15.32 (6.78)	<0.001
Instability			2.78 (0.75)	<0.001	2.79 (0.75)	<0.001

Group Pop. [log]			1.08 (0.08)	0.282	1.10 (0.08)	0.228
Group GDP/capita [log]			0.91 (0.21)	0.706	1.02 (0.28)	0.935
Group Political Power			0.60 (0.06)	<0.001	0.60 (0.06)	<0.001
Past Wars			1.87 (0.25)	<0.001	1.88 (0.25)	<0.001
Country GDP/capita [log]			0.61 (0.15)	0.067	0.53 (0.15)	0.076
Country GDP/capita outlier			7.11 (15.89)	0.414	4.48 (10.26)	0.536
Grievance Score Outlier					0.95 (2.26)	0.983
(Intercept)	0.01 (0.00)	<0.001	1.92 (2.97)	0.688	2.45 (3.85)	0.585
Observations	13644			12883		12883
R ²	-0.000			0.037		0.037

Appendix F, Table 4: Regression results corresponding to Models 20, 21, and 22 in the main paper but using PMLE instead of GLM.

Appendix G: Aggregate Group Regression Results

Appendix G, Table 1: Aggregate Group Regression Results										
Dependent Variable: Onset										
<i>Predictors</i>	(25) GLM		(26) GLM		(27) GLM		(28) GLM		(29) PMLE	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Aggregate Group	1.06 (1.07)	0.953	0.74 (0.77)	0.770	0.70 (0.73)	0.735	0.00 (0.00)	0.428	0.29 (0.51)	0.444
Grievance Score [log]					1.44 (0.89)	0.556	1.18 (0.77)	0.803	1.31 (0.79)	0.683
Grievance Score [log] x Aggregate Group							501218.26 (7076362.62)	0.353	24.01 (59.05)	0.170
Country Pop. [log]			0.88 (0.10)	0.229	0.87 (0.10)	0.197	0.88 (0.10)	0.247	0.88 (0.09)	0.241
Country Size [log]			1.07 (0.12)	0.549	1.06 (0.12)	0.624	1.06 (0.12)	0.605	1.06 (0.11)	0.603
Democracy Level			0.98 (0.02)	0.482	0.99 (0.02)	0.541	0.99 (0.02)	0.569	0.99 (0.02)	0.537
Ongoing War			0.00 (0.00)	0.971	0.00 (0.00)	0.971	0.00 (0.00)	0.971	0.01 (0.01)	<0.001
Post Cold War			1.74 (0.66)	0.145	1.71 (0.65)	0.156	1.71 (0.65)	0.161	1.66 (0.59)	0.154

New Country	14.59 (6.87)	<0.001	14.67 (6.91)	<0.001	14.90 (7.03)	<0.001	15.47 (6.85)	<0.001		
Instability	2.81 (0.79)	<0.001	2.79 (0.79)	<0.001	2.81 (0.79)	<0.001	2.79 (0.75)	<0.001		
Group Pop. [log]	1.09 (0.08)	0.268	1.10 (0.09)	0.218	1.10 (0.09)	0.219	1.09 (0.08)	0.244		
Group GDP/capita [log]	0.88 (0.21)	0.590	0.94 (0.29)	0.854	0.90 (0.30)	0.749	0.96 (0.28)	0.888		
Group Political Power	0.59 (0.06)	<0.001	0.59 (0.06)	<0.001	0.59 (0.06)	<0.001	0.60 (0.06)	<0.001		
Past Wars	1.86 (0.26)	<0.001	1.87 (0.26)	<0.001	1.81 (0.26)	<0.001	1.84 (0.25)	<0.001		
Country GDP/capita [log]	0.62 (0.16)	0.065	0.57 (0.18)	0.081	0.59 (0.20)	0.122	0.56 (0.17)	0.132		
Country GDP/capita outlier	0.02 (206.80)	1.000	0.01 (150.24)	1.000	0.02 (179.87)	1.000	5.37 (12.43)	0.495		
Grievance Score Outlier			0.00 (0.00)	0.999	0.00 (0.00)	0.999	1.44 (3.54)	0.888		
(Intercept)	0.01 (0.00)	<0.001	1.87 (3.05)	0.700	2.38 (3.93)	0.601	2.05 (3.42)	0.668	2.25 (3.56)	0.626
Observations	13644		12883		12883		12883		12883	
R ² Tjur	0.000		0.035		0.035		0.036		0.038	

Appendix G, Table 1: Model 25 is a bivariate regression of onset on aggregate group and Model 26 is the multivariate version of the same regression. Model 27 regresses onset on both aggregate group and grievance score, with controls. Models 28 and 29 show the

regression of onset on the interaction between grievance score and aggregate group, with controls. Models 25, 26, 27, and 28 use the GLM method with standard errors clustered by country and Model 29 uses PMLE with non-clustered standard errors. Corresponds to Table 7 in the main paper.

Appendix G, Table 2: Aggregate Group Extra Controls Regressions

<i>Predictors</i>	Dependent Variable: Onset					
	(1) GLM		(2) GLM		(3) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Aggregate Group	0.67 (0.73)	0.716	0.62 (0.69)	0.665	0.00 (0.00)	0.985
Grievance Score [log]			1.17 (0.79)	0.819	0.82 (0.61)	0.794
Grievance Score [log] x Aggregate Group					~Infinity (~Infinity)	0.985
Country Pop. [log]	0.90 (0.15)	0.536	0.89 (0.15)	0.505	0.91 (0.16)	0.590
Country Size [log]	1.15 (0.17)	0.369	1.13 (0.17)	0.423	1.12 (0.17)	0.449
Democracy Level	0.98 (0.03)	0.582	0.99 (0.03)	0.676	0.99 (0.03)	0.859
Ongoing War	0.00 (0.00)	0.976	0.00 (0.00)	0.976	0.00 (0.00)	0.985
Post Cold War	2.40 (1.02)	0.040	2.35 (1.01)	0.046	2.31 (0.99)	0.051
New Country	13.78 (7.62)	<0.001	13.66 (7.57)	<0.001	13.75 (7.63)	<0.001
Instability	2.69 (0.87)	0.002	2.63 (0.86)	0.003	2.42 (0.80)	0.008
Group Pop. [log]	0.98 (0.10)	0.835	0.99 (0.10)	0.915	0.98 (0.10)	0.864
Group GDP/capita [log]	0.75 (0.18)	0.229	0.73 (0.23)	0.309	0.64 (0.22)	0.204
Group Political Power	0.64 (0.08)	<0.001	0.64 (0.08)	<0.001	0.66 (0.08)	0.001

Past Wars	1.72 (0.36)	0.009	1.72 (0.36)	0.009	1.70 (0.35)	0.011
Country GDP/capita [log]	0.64 (0.18)	0.109	0.64 (0.21)	0.171	0.67 (0.24)	0.262
Military Expenditures [log]	1.04 (0.06)	0.438	1.05 (0.06)	0.385	1.06 (0.06)	0.272
Total Ethnic Groups	0.98 (0.02)	0.400	0.98 (0.03)	0.387	0.98 (0.03)	0.363
Mountainous Territory	1.19 (0.56)	0.717	1.17 (0.55)	0.734	1.29 (0.61)	0.594
Oil Presence	1.18 (0.38)	0.616	1.20 (0.39)	0.574	1.26 (0.42)	0.481
Alluvial Diamonds	1.31 (0.50)	0.477	1.33 (0.51)	0.460	1.48 (0.57)	0.305
Distance to Border [log]	0.88 (0.06)	0.070	0.89 (0.06)	0.090	0.88 (0.06)	0.071
Grievance Score Outlier			0.00 (0.00)	0.999	0.00 (0.00)	0.999
(Intercept)	2.44 (6.66)	0.743	3.32 (9.18)	0.665	3.89 (10.89)	0.627
Observations	8321		8321		8321	
R ² Tjur	0.033		0.034		0.051	

Appendix G, Table 2: Model 1 in this table corresponds to Model 26 in the main paper. Model 2 corresponds to Model 27 in the main paper. Model 3 corresponds to Model 28 in the main paper. In each instance, the models in this table add a series of extra controls that reduce N.

Appendix G, Table 3: Aggregate Group 10-Year Lag Regressions								
Dependent Variable: Onset								
<i>Predictors</i>	(1) GLM		(2) GLM		(3) GLM		(4) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Aggregate Group	1.20 (1.22)	0.854	0.72 (0.76)	0.758	0.70 (0.73)	0.732	0.00 (0.00)	0.984

Grievance Score [log]			1.66 (1.02)	0.409	1.34 (0.88)	0.653
Grievance Score [log] x Aggregate Group					~Infinity (~Infinity)	0.984
Country Pop. [log]	0.84 (0.09)	0.115	0.82 (0.10)	0.090	0.84 (0.10)	0.124
Country Size [log]	1.10 (0.13)	0.386	1.09 (0.13)	0.469	1.09 (0.13)	0.472
Democracy Level	0.99 (0.02)	0.597	0.99 (0.02)	0.680	0.99 (0.02)	0.814
Ongoing War	0.00 (0.00)	0.973	0.00 (0.00)	0.973	0.00 (0.00)	0.982
Post Cold War	1.68 (0.64)	0.176	1.65 (0.63)	0.190	1.61 (0.62)	0.214
New Country	15.10 (7.18)	<0.001	15.23 (7.25)	<0.001	15.44 (7.37)	<0.001
Instability	2.55 (0.76)	0.002	2.52 (0.76)	0.002	2.34 (0.72)	0.005
Group Pop. [log]	1.12 (0.09)	0.174	1.13 (0.09)	0.127	1.14 (0.09)	0.127
Group GDP/capita [log]	0.82 (0.20)	0.423	0.91 (0.29)	0.766	0.86 (0.29)	0.657
Group Political Power	0.59 (0.06)	<0.001	0.59 (0.06)	<0.001	0.59 (0.06)	<0.001
Past Wars	1.58 (0.40)	0.074	1.58 (0.40)	0.070	1.57 (0.40)	0.074
Country GDP/capita [log]	0.64 (0.17)	0.098	0.56 (0.19)	0.087	0.58 (0.21)	0.126
Country GDP/capita outlier	0.02 (305.18)	1.000	0.02 (193.08)	1.000	0.02 (335.80)	1.000

Grievance Score Outlier					0.00 (0.00)	0.999	0.00 (0.00)	0.999
(Intercept)	0.01 (0.00)	<0.001	2.74 (4.65)	0.554	3.85 (6.65)	0.436	3.83 (6.70)	0.443
Observations	13331		12597		12597		12597	
R ² Tjur	0.000		0.019		0.020		0.034	

Appendix G, Table 3: Regression results corresponding to Models 25, 26, 27, and 28 in the main paper but with a 10-year lag for groups to re-enter the dataset instead of a 5-year lag.

Appendix G, Table 4: Aggregate Group PMLE Regression Results						
Dependent Variable: Onset						
<i>Predictors</i>	(1) PMLE		(2) PMLE		(3) PMLE	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Aggregate Group	1.58 (1.31)	0.608	1.11 (0.93)	0.906	1.07 (0.89)	0.939
Grievance Score [log]					1.62 (0.91)	0.441
Country Pop. [log]			0.88 (0.09)	0.230	0.86 (0.09)	0.185
Country Size [log]			1.07 (0.11)	0.536	1.06 (0.11)	0.602
Democracy Level			0.98 (0.02)	0.483	0.99 (0.02)	0.538
Ongoing War			0.01 (0.01)	<0.001	0.01 (0.01)	<0.001
Post Cold War			1.67 (0.60)	0.149	1.63 (0.59)	0.167
New Country			15.13 (6.71)	<0.001	15.26 (6.75)	<0.001
Instability			2.79 (0.75)	<0.001	2.80 (0.75)	<0.001

Group Pop. [log]			1.08 (0.08)	0.289	1.10 (0.08)	0.227
Group GDP/capita [log]			0.89 (0.21)	0.657	1.02 (0.28)	0.960
Group Political Power			0.60 (0.06)	<0.001	0.60 (0.06)	<0.001
Past Wars			1.87 (0.25)	<0.001	1.87 (0.25)	<0.001
Country GDP/capita [log]			0.61 (0.15)	0.078	0.53 (0.15)	0.080
Country GDP/capita outlier			7.26 (16.28)	0.410	4.44 (10.20)	0.539
Grievance Score Outlier					0.83 (1.97)	0.938
(Intercept)	0.01 (0.00)	<0.001	1.99 (3.10)	0.670	2.60 (4.09)	0.562
Observations	13644			12883		12883
R ²	-0.000			0.037		0.037

Appendix G, Table 4: Regression results corresponding to Models 25, 26, and 27 in the main paper but using PMLE instead of GLM.

Appendix H: Group Level Tables Shown in the Paper

Appendix H, Table 1: Group Level Descriptive Statistics						
Variable	N	Mean	S. Dev.	Min.	Median	Max.
Onset	686	0.109	0.312	0	0	1
Avg. Grievance Score	682	1.35	1.95	1	1.01	49.8
Avg. Legibility Score	483	0.361	0.499	0.0259	0.228	4.46
Avg. Administrative Capacity	573	-0.177	0.597	-1.58	-0.252	1.39
Ethnicity Collection Ever	623	0.822	0.383	0	1	1
Avg. Transnational Kin	688	3.14	5.92	0	1	32.3
Aggregate Group Ever	688	0.0247	0.155	0	0	1
Avg. Country GDP/capita	662	6475	7756	146	3042	44376
Avg. Country Population (mil.)	667	139.3	310.1	0.168	18.3	1244.2
Avg. Military Expenditures [log]	642	14.3	4.66	3.54	14	25.7
Avg. Country Size (km ²)	688	2640941	5328449	629	464748	22065955
Avg. Total Ethnic Groups	688	10.8	12.6	2	5.5	53
Avg. Democracy Level	679	1.08	5.93	-10	1.38	10
New Country	679	0.187	0.39	0	0	1
Avg. Group Population (mil.)	686	8.089	49.3	0.000621	0.96	1157.9
Avg. Group GDP/capita	682	6261	10338	13.3	2811	186239
Avg. Group Political Power	688	3.24	1.63	1	2.55	7
Previous Wars	688	0.119	0.474	0	0	5
Avg. Mountainous Terrain	631	0.372	0.339	0	0.276	1
Oil Ever	605	0.494	0.5	0	0	1
Diamonds Ever	606	0.177	0.382	0	0	1
Avg. Distance to Border	632	39.6	109	0.00338	3.01	855

Appendix H, Table 1: Descriptive statistics for all group level variables. Corresponds to Table 8.

Appendix H, Table 2: Grievance Score Group Level Regression Results						
Predictors	Dependent Variable: Onset					
	(30) GLM		(31) GLM		(32) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Avg. Grievance Score [log]	1.02 (0.36)	0.958	1.24 (0.70)	0.705	2.11 (1.32)	0.233
Avg. Country Pop. [log]			0.89 (0.11)	0.347	1.06 (0.19)	0.763
Avg. Country Size [log]			0.98 (0.12)	0.861	1.20 (0.19)	0.244
Avg. Democracy Level			0.95 (0.03)	0.068	0.96 (0.03)	0.135
New Country			1.62 (0.59)	0.183	1.67 (0.68)	0.214
Avg. Group Pop. [log]			1.06 (0.09)	0.466	0.90 (0.09)	0.298
Avg. Group GDP/capita [log]			1.38 (0.65)	0.493	1.50 (0.75)	0.414
Avg. Group Political Power			0.80 (0.09)	0.038	0.87 (0.11)	0.251
Past Wars			1.53 (0.31)	0.033	1.40 (0.33)	0.148
Avg. Country GDP/capita [log]			0.42 (0.21)	0.080	0.44 (0.23)	0.111
Avg. Military Expenditures [log]					0.95 (0.07)	0.510
Avg. Total Ethnic Groups					0.97 (0.02)	0.191
Avg. Mountainous Terrain					1.18 (0.56)	0.724

Oil Presence Ever					1.60 (0.53)	0.152
Alluvial Diamonds Ever					1.45 (0.57)	0.343
Distance to Border [log]					0.85 (0.06)	0.022
(Intercept)	0.12 (0.02)	<0.001	60.45 (103.91)	0.017	1.36 (4.31)	0.922
Observations	680		649		582	
R ² Tjur	0.000		0.088		0.086	

Appendix H, Table 2: Model 1 shows the bivariate GLM regression of onset on average grievance score at the group level. Model 2 adds standard controls to make this multivariate and Model 3 adds extra controls on top of that. All standard errors are clustered by country. Corresponds to Table 9.

Appendix H, Table 3: Group Level Legibility Score Regression Results								
Dependent Variable: Onset								
<i>Predictors</i>	(33) GLM		(34) GLM		(35) GLM		(36) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Avg. Legibility Score [log]	0.50 (0.08)	<0.001	0.67 (0.13)	0.042	0.69 (0.14)	0.065	0.70 (0.18)	0.161
Avg. Grievance Score [log]					0.30 (0.19)	0.061	0.24 (0.43)	0.424
Avg. Grievance Score [log] x Avg. Legibility Score [log]							0.92 (0.57)	0.888
Avg. Country Pop. [log]			0.91 (0.14)	0.540	0.88 (0.14)	0.411	0.88 (0.15)	0.465
Avg. Country Size [log]			0.98 (0.18)	0.914	1.13 (0.22)	0.535	1.13 (0.22)	0.540
Avg. Democracy Level			0.97 (0.04)	0.405	0.96 (0.04)	0.305	0.96 (0.04)	0.303

New Country			0.35 (0.38)	0.331	0.29 (0.32)	0.265	0.29 (0.32)	0.266
Avg. Group Pop. [log]			1.06 (0.12)	0.564	1.04 (0.11)	0.752	1.03 (0.11)	0.765
Avg. Group GDP/capita [log]			0.55 (0.14)	0.020	0.50 (0.13)	0.010	0.50 (0.14)	0.010
Avg. Group Political Power			0.70 (0.11)	0.024	0.67 (0.11)	0.012	0.67 (0.11)	0.013
Past Wars			1.72 (0.36)	0.009	1.78 (0.38)	0.007	1.76 (0.39)	0.010
(Intercept)	0.03 (0.01)	<0.001	30.81 (86.53)		0.222	36.42 (103)	0.204	36.52 (103.3)
Observations	483		475		475		475	
R ² Tjur	0.053		0.138		0.157		0.158	

Appendix H, Table 3: Model 1 shows the bivariate GLM regression of onset on average legibility score at the group level. Model 2 adds standard controls to make this multivariate and Model 3 adds average grievance score as well. In Model 4, onset is regressed on the interaction of average grievance score and average legibility score. All standard errors are clustered by country. Corresponds to Table 10 in the main paper.

Appendix H, Table 4: Group Level Administrative Capacity Regression Results								
Dependent Variable: Onset								
<i>Predictors</i>	(37) GLM		(38) GLM		(39) GLM		(40) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Avg. Admin. Capacity	0.19 (0.06)	<0.001	0.22 (0.09)	<0.001	0.22 (0.09)	<0.001	0.25 (0.11)	0.002
Avg. Grievance Score [log]					0.82 (0.42)	0.691	0.59 (0.48)	0.515
Avg. Grievance Score [log] x Avg. Admin Capacity							0.54 (0.59)	0.571

Avg. Country Pop. [log]			0.98 (0.17)	0.903	0.98 (0.16)	0.897	1.00 (0.17)	0.991
Avg. Country Size [log]			0.99 (0.15)	0.931	1.01 (0.16)	0.969	1.00 (0.16)	0.978
Avg. Democracy Level			0.96 (0.03)	0.209	0.96 (0.03)	0.199	0.96 (0.03)	0.194
New Country			1.19 (0.60)	0.729	1.18 (0.60)	0.736	1.17 (0.59)	0.760
Avg. Group Pop. [log]			1.08 (0.10)	0.423	1.07 (0.11)	0.473	1.07 (0.11)	0.512
Avg. Group GDP/capita [log]			0.92 (0.16)	0.619	0.90 (0.16)	0.566	0.90 (0.16)	0.542
Avg. Group Political Power			0.67 (0.09)	0.004	0.67 (0.09)	0.004	0.67 (0.09)	0.004
Past Wars			1.59 (0.36)	0.038	1.59 (0.36)	0.039	1.57 (0.35)	0.045
(Intercept)	0.06 (0.01)	<0.001	0.20 (0.51)	0.523	0.21 (0.53)	0.537	0.19 (0.48)	0.510
Observations	572		557		557		557	
R ² Tjur	0.081		0.150		0.150		0.151	

Appendix H, Table 4: Model 1 shows the bivariate GLM regression of onset on average administrative capacity at the group level. Model 2 adds standard controls to make this multivariate and Model 3 adds average grievance score as well. In Model 4, onset is regressed on the interaction of average grievance score and average administrative capacity. All standard errors are clustered by country. Corresponds to Table 11 in the main paper.

Appendix H, Table 5: Group Level Ethnicity Collection Regression Results								
Dependent Variable: Onset								
Predictors	(41) GLM		(42) GLM		(43) GLM		(44) GLM	
	Odds Ratios	<i>p</i>	Odds Ratios	<i>p</i>	Odds Ratios	<i>p</i>	Odds Ratios	<i>p</i>
Ethnicity Collection Ever	2.18 (0.90)	0.058	2.70 (1.26)	0.033	2.73 (1.27)	0.031	2.53 (1.33)	0.078

Avg. Grievance Score [log]			1.24 (0.72)	0.707	0.82 (1.28)	0.902
Avg. Grievance Score [log] x Ethnicity Collection Ever					1.56 (2.40)	0.774
Avg. Country Pop. [log]	0.95 (0.13)	0.693	0.95 (0.13)	0.686	0.94 (0.13)	0.668
Avg. Country Size [log]	0.88 (0.11)	0.290	0.87 (0.11)	0.268	0.87 (0.11)	0.283
Avg. Democracy Level	0.92 (0.03)	0.006	0.92 (0.03)	0.006	0.92 (0.03)	0.007
New Country	2.79 (1.14)	0.012	2.80 (1.15)	0.012	2.77 (1.14)	0.013
Avg. Group Pop. [log]	1.11 (0.10)	0.257	1.11 (0.10)	0.244	1.11 (0.10)	0.245
Avg. Group GDP/capita [log]	1.27 (0.43)	0.493	1.43 (0.67)	0.447	1.43 (0.67)	0.445
Avg. Group Political Power	0.77 (0.09)	0.018	0.77 (0.09)	0.022	0.77 (0.09)	0.022
Past Wars	1.52 (0.32)	0.047	1.51 (0.32)	0.050	1.52 (0.32)	0.047
Avg. Country GDP/capita [log]	0.49 (0.18)	0.050	0.43 (0.21)	0.089	0.43 (0.21)	0.088
(Intercept)	0.07 (0.03)	<0.001	15.01 (26.53)	0.126	15.79 (27.99)	0.113
Observations	621	610	610	610	610	
R ² Tjur	0.006	0.110	0.111	0.111	0.111	

Appendix H, Table 5: Model 1 shows the bivariate GLM regression of onset on the binary variable capturing whether ethnicity data was ever collected on a recent census. Model 2 adds standard controls to make this multivariate and Model 3 adds average grievance score as well. In Model 4, onset is regressed on the interaction of average grievance score and ethnicity collection ever. All standard errors are clustered by country. Corresponds to Table 12.

Appendix I: Group Level Extra Regressions and Robustness Checks

Appendix I, Table 1: Group Level Transnational Kin Regression Results								
Dependent Variable: Onset								
<i>Predictors</i>	(1) GLM		(2) GLM		(3) GLM		(4) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Avg. Transnational Kin	1.00 (0.02)	0.943	1.00 (0.02)	0.895	1.00 (0.02)	0.910	0.99 (0.03)	0.840
Avg. Grievance Score [log]					1.24 (0.70)	0.710	1.21 (0.70)	0.739
Avg. Grievance Score [log] x Avg. Transnational Kin							1.05 (0.21)	0.800
Avg. Country Pop. [log]			0.89 (0.11)	0.345	0.89 (0.11)	0.344	0.89 (0.11)	0.349
Avg. Country Size [log]			0.99 (0.12)	0.908	0.98 (0.12)	0.861	0.98 (0.12)	0.863
Avg. Democracy Level			0.95 (0.03)	0.062	0.95 (0.03)	0.069	0.95 (0.03)	0.070
New Country			1.62 (0.59)	0.187	1.63 (0.59)	0.182	1.64 (0.60)	0.179
Avg. Group Pop. [log]			1.06 (0.09)	0.480	1.07 (0.09)	0.461	1.06 (0.09)	0.478
Avg. Group GDP/capita [log]			1.23 (0.43)	0.555	1.38 (0.65)	0.489	1.40 (0.66)	0.478
Avg. Group Political Power			0.79 (0.09)	0.033	0.80 (0.09)	0.038	0.80 (0.09)	0.046
Past Wars			1.55 (0.31)	0.030	1.54 (0.31)	0.033	1.54 (0.31)	0.032
Avg. Country GDP/capita [log]			0.48 (0.17)	0.043	0.42 (0.21)	0.080	0.42 (0.21)	0.077

(Intercept)	0.12 (0.02)	<0.001	56.70 (97.39)	0.019	59.81 (102.98)	0.018	60.23 (103.84)	0.017
Observations	686		649		649		649	
R ² Tjur	0.000		0.087		0.088		0.087	

Appendix I, Table 1: Model 1 shows the bivariate GLM regression of onset on average transnational kin at the group level. Model 2 adds standard controls to make this multivariate and Model 3 adds average grievance score as well. In Model 4, onset is regressed on the interaction of average grievance score and average transnational kin. All standard errors are clustered by country.

Appendix I, Table 2: Group Level Aggregate Group Regression Results								
Dependent Variable: Onset								
<i>Predictors</i>	(1) GLM		(2) GLM		(3) GLM		(4) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Aggregate Group Ever	0.54 (0.56)	0.550	0.37 (0.39)	0.350	0.37 (0.40)	0.353	0.00 (0.00)	0.360
Avg. Grievance Score [log]					1.23 (0.70)	0.717	1.09 (0.64)	0.888
Avg. Grievance Score [log] x Aggregate Group Ever							2 trillion (61 trillion)	0.322
Avg. Country Pop. [log]			0.87 (0.11)	0.290	0.87 (0.11)	0.288	0.89 (0.11)	0.344
Avg. Country Size [log]			0.99 (0.12)	0.963	0.99 (0.12)	0.918	0.98 (0.12)	0.898
Avg. Democracy Level			0.95 (0.03)	0.058	0.95 (0.03)	0.064	0.95 (0.03)	0.046
New Country			1.61 (0.59)	0.192	1.62 (0.59)	0.186	1.65 (0.60)	0.174
Avg. Group Pop. [log]			1.07 (0.09)	0.423	1.08 (0.09)	0.405	1.06 (0.09)	0.489

Avg. Group GDP/capita [log]			1.18 (0.41)	0.636	1.33 (0.64)	0.550	1.26 (0.62)	0.640
Avg. Group Political Power			0.79 (0.09)	0.037	0.80 (0.09)	0.042	0.81 (0.09)	0.064
Past Wars			1.53 (0.31)	0.034	1.52 (0.31)	0.037	1.51 (0.31)	0.043
Avg. Country GDP/capita [log]			0.49 (0.18)	0.051	0.43 (0.22)	0.094	0.45 (0.23)	0.118
(Intercept)	0.12 (0.02)	<0.001	66.61 (114.5)	0.015	69.68 (120)	0.014	75.94 (131.6)	0.012
Observations	686		649		649		649	
R ² Tjur	0.001		0.090		0.090		0.097	

Appendix I, Table 2: Model 1 shows the bivariate GLM regression of onset on aggregate group ever. Model 2 adds standard controls to make this multivariate and Model 3 adds average grievance score as well. In Model 4, onset is regressed on the interaction of average grievance score and aggregate group ever. All standard errors are clustered by country.

Appendix I, Table 3: Group Level Grievance Score Functional Form Robustness						
Dependent Variable: Onset						
<i>Predictors</i>	(1) GLM		(2) GLM		(3) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Avg. Grievance Score	0.97 (0.11)	0.774	0.97 (0.11)	0.799	1.01 (0.09)	0.909
Avg. Country Pop. [log]			0.89 (0.11)	0.346	1.02 (0.18)	0.903
Avg. Country Size [log]			0.99 (0.12)	0.909	1.21 (0.19)	0.216
Avg. Democracy Level			0.95 (0.03)	0.066	0.95 (0.03)	0.130
New Country			1.61 (0.59)	0.188	1.58 (0.64)	0.260

Avg. Group Pop. [log]			1.06 (0.09)	0.480	0.90 (0.09)	0.292
Avg. Group GDP/capita [log]			1.15 (0.47)	0.736	1.05 (0.47)	0.921
Avg. Group Political Power			0.79 (0.09)	0.031	0.85 (0.10)	0.182
Past Wars			1.55 (0.31)	0.029	1.41 (0.33)	0.145
Avg. Country GDP/capita [log]			0.51 (0.21)	0.109	0.60 (0.29)	0.286
Avg. Military Expenditures [log]					0.97 (0.07)	0.656
Avg. Total Ethnic Groups					0.97 (0.02)	0.262
Avg. Mountainous Terrain					1.16 (0.55)	0.755
Oil Presence Ever					1.60 (0.53)	0.155
Alluvial Diamonds Ever					1.36 (0.52)	0.426
Distance to Border [log]					0.85 (0.06)	0.025
(Intercept)	0.13 (0.02)	<0.001	60.61 (104.61)	0.017	2.53 (8.14)	0.772
Observations	680			649		582
R ² Tjur	0.000			0.087		0.084

Appendix I, Table 3: This table displays results that correspond to Table 9, regressing onset on average grievance score with different sets of control variables. The one alteration in these models is that average grievance score is *not* logged.

Appendix I, Table 4: Group Level Legibility Score Extra Controls Robustness						
Dependent Variable: Onset						
<i>Predictors</i>	(1) GLM		(2) GLM		(3) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Avg. Legibility Score [log]	0.51 (0.14)	0.013	0.51 (0.14)	0.013	0.41 (0.14)	0.009
Avg. Grievance Score [log]			0.90 (0.82)	0.911	5.76 (10.94)	0.357
Avg. Grievance Score [log] x Avg. Legibility Score [log]					2.04 (1.33)	0.272
Avg. Country Pop. [log]	0.98 (0.28)	0.945	0.97 (0.29)	0.922	0.92 (0.27)	0.775
Avg. Country Size [log]	1.57 (0.37)	0.055	1.58 (0.37)	0.056	1.63 (0.39)	0.042
Avg. Democracy Level	0.98 (0.04)	0.669	0.98 (0.05)	0.664	0.98 (0.04)	0.666
New Country	0.30 (0.36)	0.310	0.29 (0.35)	0.308	0.28 (0.34)	0.290
Avg. Group Pop. [log]	0.81 (0.11)	0.131	0.81 (0.11)	0.135	0.81 (0.11)	0.128
Avg. Group GDP/capita [log]	1.36 (0.90)	0.642	1.29 (1.06)	0.760	1.47 (1.22)	0.645
Avg. Group Political Power	0.71 (0.14)	0.077	0.71 (0.14)	0.081	0.70 (0.14)	0.071
Past Wars	1.64 (0.44)	0.066	1.65 (0.45)	0.066	1.70 (0.46)	0.053
Avg. Country GDP/capita [log]	0.38 (0.24)	0.133	0.39 (0.31)	0.230	0.35 (0.27)	0.181
Avg. Military Expenditures [log]	1.21 (0.14)	0.107	1.21 (0.15)	0.119	1.23 (0.15)	0.090

Avg. Total Ethnic Groups	0.92 (0.04)	0.058	0.92 (0.04)	0.065	0.92 (0.04)	0.053
Avg. Mountainous Terrain	0.61 (0.46)	0.509	0.60 (0.46)	0.503	0.57 (0.44)	0.471
Oil Presence Ever	0.59 (0.28)	0.265	0.59 (0.28)	0.268	0.56 (0.27)	0.225
Alluvial Diamonds Ever	1.56 (0.78)	0.371	1.55 (0.78)	0.383	1.58 (0.80)	0.370
Distance to Border [log]	0.85 (0.08)	0.086	0.85 (0.08)	0.088	0.85 (0.08)	0.083
(Intercept)	0.18 (0.82)	0.707	0.20 (0.94)	0.732	0.15 (0.70)	0.683
Observations	426		426		426	
R ² Tjur	0.201		0.201		0.204	

Appendix H, Table 4: This table displays results that correspond to Table 10, regressing onset on average legibility score. Model 1 in this table corresponds to Model 34 in the main paper, Model 2 goes with Model 35, and Model 3 is with Model 36. This table adds extra control variables that reduce N as a robustness check.

Appendix I, Table 5: Group Level Legibility Score Functional Form Robustness								
Dependent Variable: Onset								
<i>Predictors</i>	(1) GLM		(2) GLM		(3) GLM		(4) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Avg. Legibility Score	0.05 (0.04)	0.001	0.39 (0.42)	0.382	0.44 (0.48)	0.450	3.82 (12.08)	0.671
Avg. Grievance Score					0.57 (0.20)	0.102	0.67 (0.28)	0.336
Avg. Grievance Score x Legibility Score							0.18 (0.44)	0.483
Avg. Country Pop. [log]			0.87 (0.14)	0.394	0.84 (0.13)	0.286	0.88 (0.15)	0.441

Avg. Country Size [log]			1.02 (0.19)	0.902	1.16 (0.23)	0.458	1.15 (0.23)	0.468
Avg. Democracy Level			0.99 (0.04)	0.738	0.98 (0.04)	0.633	0.98 (0.04)	0.571
New Country			0.32 (0.34)	0.284	0.28 (0.31)	0.243	0.28 (0.31)	0.244
Avg. Group Pop. [log]			1.09 (0.12)	0.419	1.07 (0.12)	0.534	1.05 (0.12)	0.636
Avg. Group GDP/capita [log]			0.47 (0.13)	0.005	0.43 (0.12)	0.002	0.42 (0.12)	0.002
Avg. Group Political Power			0.69 (0.11)	0.018	0.67 (0.11)	0.011	0.67 (0.11)	0.013
Past Wars			1.71 (0.36)	0.011	1.76 (0.38)	0.008	1.70 (0.37)	0.015
(Intercept)	0.22 (0.05)	<0.001	227.54 (595.51)	0.038	469.43 (1249.23)	0.021	284.95 (782.75)	0.040
Observations	483		475		475		475	
R ² Tjur	0.036		0.133		0.147		0.149	

Appendix I, Table 5: This table displays results that correspond to Table 10, regressing onset on average legibility score. Model 1 in this table corresponds to Model 33 in the main paper, Model 2 goes with Model 34 in the main paper, Model 3 is with Model 35, and Model 4 with Model 36. In this table, average legibility score and average grievance score are *not* logged.

Appendix I, Table 6: Group Level Administrative Capacity Extra Controls Robustness						
Dependent Variable: Onset						
<i>Predictors</i>	(1) GLM		(2) GLM		(3) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Avg. Admin. Capacity	0.22 (0.10)	0.001	0.22 (0.10)	0.001	0.22 (0.11)	0.002
Avg. Grievance Score [log]			1.33 (1.00)	0.704	1.41 (1.56)	0.754
Avg. Grievance Score [log] x Avg. Admin Capacity					1.09 (1.24)	0.941

Avg. Country Pop. [log]	1.17 (0.24)	0.455	1.18 (0.25)	0.426	1.18 (0.25)	0.430
Avg. Country Size [log]	1.20 (0.23)	0.337	1.19 (0.23)	0.367	1.19 (0.23)	0.365
Avg. Democracy Level	0.96 (0.04)	0.241	0.96 (0.04)	0.244	0.96 (0.04)	0.244
New Country	1.48 (0.80)	0.471	1.51 (0.82)	0.449	1.51 (0.83)	0.447
Avg. Group Pop. [log]	0.92 (0.11)	0.503	0.93 (0.11)	0.513	0.93 (0.11)	0.514
Avg. Group GDP/capita [log]	1.17 (0.55)	0.741	1.35 (0.81)	0.616	1.37 (0.86)	0.617
Avg. Group Political Power	0.73 (0.11)	0.033	0.73 (0.11)	0.038	0.73 (0.11)	0.038
Past Wars	1.36 (0.34)	0.219	1.35 (0.34)	0.228	1.36 (0.34)	0.228
Avg. Country GDP/capita [log]	0.74 (0.37)	0.544	0.65 (0.39)	0.470	0.64 (0.41)	0.478
Avg. Military Expenditures [log]	0.99 (0.09)	0.956	0.99 (0.09)	0.892	0.99 (0.09)	0.896
Avg. Total Ethnic Groups	0.97 (0.02)	0.197	0.97 (0.02)	0.185	0.97 (0.03)	0.192
Avg. Mountainous Terrain	0.79 (0.45)	0.682	0.81 (0.46)	0.707	0.80 (0.47)	0.701
Oil Presence Ever	1.12 (0.46)	0.782	1.13 (0.47)	0.760	1.13 (0.47)	0.763
Alluvial Diamonds Ever	0.92 (0.42)	0.861	0.96 (0.44)	0.920	0.95 (0.44)	0.915
Distance to Border [log]	0.81 (0.07)	0.009	0.81 (0.07)	0.009	0.81 (0.07)	0.009
(Intercept)	0.02 (0.06)	0.284	0.01 (0.05)	0.266	0.01 (0.05)	0.265
Observations	512		512		512	

R ² Tjur	0.154	0.154	0.154
---------------------	-------	-------	-------

Appendix I, Table 6: This table displays results that correspond to Table 11, regressing onset on average administrative capacity. Model 1 in this table corresponds to Model 38 in the main paper, Model 2 goes with Model 39, and Model 3 is with Model 40. This table adds extra control variables that reduce N as a robustness check.

Appendix I, Table 7: Group Level Ethnicity Collection Extra Controls Regressions						
<i>Predictors</i>	Dependent Variable: Onset					
	(1) GLM		(2) GLM		(3) GLM	
	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>	<i>Odds Ratios</i>	<i>p</i>
Ethnicity Collection Ever	2.63 (1.38)	0.065	2.66 (1.40)	0.064	3.12 (1.92)	0.065
Avg. Grievance Score [log]			1.98 (1.26)	0.285	8.04 (20.54)	0.415
Avg. Grievance Score [log] x Ethnicity Collection Ever					0.24 (0.61)	0.571
Avg. Country Pop. [log]	1.17 (0.22)	0.403	1.21 (0.23)	0.322	1.20 (0.23)	0.339
Avg. Country Size [log]	1.09 (0.18)	0.599	1.08 (0.18)	0.647	1.08 (0.18)	0.635
Avg. Democracy Level	0.92 (0.03)	0.019	0.93 (0.03)	0.019	0.93 (0.03)	0.019
New Country	3.01 (1.43)	0.020	3.18 (1.52)	0.016	3.16 (1.52)	0.017
Avg. Group Pop. [log]	0.91 (0.10)	0.400	0.91 (0.10)	0.405	0.92 (0.10)	0.438
Avg. Group GDP/capita [log]	1.08 (0.43)	0.844	1.53 (0.77)	0.400	1.56 (0.79)	0.379
Avg. Group Political Power	0.83 (0.10)	0.137	0.85 (0.11)	0.188	0.85 (0.11)	0.195
Past Wars	1.30 (0.31)	0.280	1.29 (0.31)	0.289	1.29 (0.31)	0.286

Avg. Country GDP/capita [log]	0.60 (0.27)	0.258	0.44 (0.23)	0.121	0.43 (0.23)	0.112
Avg. Military Expenditures [log]	0.97 (0.07)	0.694	0.96 (0.07)	0.557	0.96 (0.07)	0.568
Avg. Total Ethnic Groups	0.96 (0.02)	0.092	0.95 (0.02)	0.067	0.95 (0.02)	0.068
Avg. Mountainous Terrain	1.09 (0.53)	0.866	1.11 (0.54)	0.833	1.10 (0.54)	0.840
Oil Presence Ever	1.57 (0.55)	0.193	1.57 (0.54)	0.193	1.55 (0.54)	0.206
Alluvial Diamonds Ever	1.16 (0.46)	0.703	1.24 (0.49)	0.596	1.22 (0.49)	0.625
Distance to Border [log]	0.83 (0.06)	0.010	0.83 (0.06)	0.008	0.83 (0.06)	0.008
(Intercept)	0.33 (1.07)	0.732	0.18 (0.58)	0.598	0.15 (0.49)	0.564
Observations	546		546		546	
R ² Tjur	0.100		0.102		0.102	

Appendix I, Table 7: This table displays results that correspond to Table 11, regressing onset on ethnicity collection ever. Model 1 in this table corresponds to Model 42 in the main paper, Model 2 goes with Model 43, and Model 3 is with Model 44. This table adds extra control variables that reduce N as a robustness check.