Repression and Individual-Level Ethnic Identification

David Bowden

June 20, 2017

# Repression and Individual-Level Ethnic Identification

In this chapter I test the theory articulated in Chapter 2 at the individual level, using data from the Afrobarometer survey. This analysis allows me to directly test whether the hypothesized mechanisms linking repression to rebel movement structure are in fact at work. Findings consistent with my expectations in this chapter would allow me to rule out many alternative explanations for the findings in the national-level analyses in subsequent chapters.

I argue that the size and structure of the rebel movement is shaped by two parameters - the number of dissidents willing to engage in violence, and the extent to which dissidents are oriented toward broadly inclusive ideologies or identities, rather than particularistic identities or causes. One process that should influence these attributes is the repression of dissident civilians. As the risk of violence to civilians increases, the relative cost of fighting decreases, increasing the number of individuals willing to participate in rebellion.

*Hypothesis 1: Individuals who experience repression should be more willing to participate in political violence themselves*

As repression is often targeted on the basis of sub-national identities such as ethnicity or religion, and as these groups often offer a basis for collective defense, repression should induce individuals to identify more strongly with sub-national groups.

*Hypothesis 2: Repression should increase the extent to which an individual identifies with their ethnic group*

I proceed with a discussion of the general attributes of the Afrobarometer survey, followed by descriptions of the variables of interest, and finally analyze multilevel models of individual attitudes toward political mobilization and ethnic identity.

## Research Design

### The Afrobarometer Survey

To examine the relationship between repression and individual attitudes toward political participation and ethnic identities, I use waves 3-6[[1]](#footnote-1) of the Afrobarometer survey. The Afrobarometer is administered by researchers at Michigan State University, the Institute for Democracy in South Africa, and the Center for Democratic Development in Ghana. In each wave the survey attempts to obtain a nationally-representative sample of 20-25 African countries. This is accomplished by randomly sampling geographic areas (villages, neighborhoods, etc), with selection probabilities weighted by population. Within each geographic area a starting point is chosen at random, from which interviews begin randomly selecting households. Individuals are then randomly selected within households, alternating between men and women to ensure gender balance. The sample in each country usually numbers either 1,200 or 2,400, depending on the size and diversity of the country. Respondents are asked over 300 questions on their demographics and background, and their opinions on a wide range of political and cultural questions. One advantage of using such a general survey is the relatively low likelihood that individuals will be primed to answer questions about ethnicity in a way that is not representative of their normal opinions (Eifert, Miguel, and Posner 2010).

Attributes for each survey wave are summarized in Table 1. The four waves span the period 2005--2016, cover 38 countries, and collect a total of 158,362 individual responses. Response rates are generally quite high, averaging 76.5% in wave 6, and 77.7% in wave 5. A detailed summary of included countries is provided in Table ?? of the Appendix.

Table 1: The Afrobarometer Survey by Wave

|  |  |  |  |
| --- | --- | --- | --- |
| Wave | Years | Total Responses | Countries |
| 3 | 2005 | 25,397 | 18 |
| 4 | 2008 | 27,713 | 20 |
| 5 | 2011--2013 | 51,587 | 34 |
| 6 | 2016 | 53,935 | 36 |
| Total | 2005--2016 | 158,362 | 36 |

One concern relevant to the present application is that in rare cases the Afrobarometer excludes geographic areas experiencing significant violence or other factors that would pose a danger to interviewers. Additionally, questions about ethnic identity are not asked in some countries where doing so is deemed to be potentially harmful to the sampled communities. Each of these attributes is suboptimal, but I argue that each introduces bias *against* my hypotheses, rather than for them. I expect that repression will induce individuals to be more willing to engage in violent mobilization, and more likely to identify with their ethnic group. By excluding areas experiencing high levels of violence, the sample is likely to exclude many of the areas experiencing the highest levels of repression. Thus, my hypotheses face a hard test --- I must find an effect for repression in a sample where repression levels are mostly low or moderate.

### Dependent Variables

#### Attitude Toward Violence

The first dependent variable explored in this chapter is willingness to use violence. To distinguish factors that influence an individual's willingness to engage in violence from those that make them more active generally, I include several other forms of political behavior, including voting, attending community meetings, and protesting. I collapse each variable into binary categories with individuals who engaged in the activity at least once coded as one, and individuals who did not participate in the activity for any reason, including those who are willing but have not actually done the activity, coded as zero. The violence, meeting, and protest questions share a common stub:

Question text: "Here is a list of actions that people sometimes take as citizens when they are dissatisfied with government performance. For each of these, please tell me whether you, personally, have done any of these things during the past year. If not, would you do this if you had the chance: Participated in a demonstration or protest march / Attended a community meeting / Used force or violence for a political cause."

Responses: "No, would never do this," "No, but would do if had the chance," "Yes, once or twice," "Yes, several times," and "Yes, often."

The voting question differs as individuals generally vote only once, and it solicits explanations for why individuals did not vote:

Question text: "Understanding that some people were unable to vote in the most recent national election in [20xx], which of the following statements is true for you?"

Responses "You were not registered to vote," "You voted in the elections," "You decided not to vote," "You could not find the polling station," "You were prevented from voting," "You did not have time to vote," "You did not vote because you could not find your name in the voters' register," "Did not vote for some other reason," "You were too young to vote," "Don't Know/ Can't Remember."

Each of these measures is self-reported, and not subject to any independent verification. It is well-established in the US context that self-reported surveys overestimate the prevalence of voting, introducing bias to models of political participation (Bernstein, Chadha, and Montjoy 2001). It is unclear whether such effects are prevalent in Africa. 70.4% of respondents in the full sample reported voting, though the number varies wildly across countries in a manner that generally matches variation in actual voter turnout (see Kuenzi and Lambright 2007). While social desirability bias might lead some individuals to claim they have voted while in fact they have not, the effect is likely to be in the opposite direction for less desirable behaviors such as protesting and violence. These effects are most likely to introduce bias into my analysis if individuals misreport behavior on both the independent in dependent variables. For example, if individual falsely claimed both to have been attacked and engaged in violence themselves, they would contribute to the relationship between those variables being overstated.

Table 2 Summary of Participation (waves 3-6)

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | Yes | No | Percentage Yes |
| Violence | 1473 | 50077 | 2.9% |
| Protest | 16492 | 142056 | 10.4% |
| Meeting | 92849 | 65719 | 58.6% |
| Vote | 111632 | 46952 | 70.4% |

The participation measures are summarized in Table 2. Participation in violence is rare, with only 2.9% of respondents reporting to have done it at least once in the past year. Participation rates increase as the degree of commitment required decreases - 10.4% participated in at least one protest, 58.6% participated in a community meeting, and 70.4% voted.

#### Ethnic Identity

The second dependent variable is ethnic identification. The Afrobarometer asks individuals about the extent to which they identify with their ethnic group, relative to their nation. The question text is as follows:

"Let us suppose that you had to choose between being a [ENTER NATIONALITY] and being a \_\_\_\_\_\_\_\_ [Respondent’s Ethnic Group]. Which of the following best expresses your feelings?"

Respondents place themselves on a five point scale with the possible responses: "I feel only [ethnicity]", "I feel more [ethnicity] than [nationality]", "I feel equally [ethnicity] and [nationality]", "I feel more [nationality] than [ethnicity]", and "I feel only [nationality]". I collapse the measure into a binary variable, with respondents in the first two categories coded as ethnic identifiers, and all others as non-ethnic identifiers.

Individuals self-report their ethnicity earlier in the survey. The question is open-ended, allowing for the possibility that respondents may conceive of ethnicity in ways that do not comport with scholarly definitions. Indeed, around 1.5% of respondents provide answers such as "African" or the name of a sub-national region. The vast majority, however, choose ethnicities that appear in externally-imposed classifications, such as the Ethnic Power Relations data (Vogt et al. 2015).

Relatively few respondents identify with their ethnic group, with 4.3% answering that they feel only an ethnic identity, and 7.4% saying that their ethnic identity was more prevalent than their national identity (see Table ??). A plurality of respondents (35.8%) said that they felt equally attached to their national and ethnic identities, and a large percentage (34.1%) said that they feel only a national identity.

### Independent Variables

*H1* predicts that individuals who experience violent repression should be more likely to participate in violence than others, and *H2* predicts that repression should increase the extent that individuals identify with their ethnic group. I test these propositions using both individual-level and national-level measures of repression. At the individual level, I use an Afrobarometer question that asks respondents whether they or a family member has been attacked in the past year:

"During the past year, have you or anyone in your family: Been physically attacked?"

The possible responses are: "no," "once," "twice," "three or more times," and "don't know." I recode the variable into a binary measure with individuals who experienced any attacks coded as 1, and individuals who experienced no attacks coded as 0. This question has two noteworthy limitations. First, it does not differentiate between individuals who were personally attacked from family members of people who were attacked. However, I expect that this feature is more likely to introduce bias against my hypotheses, than in their favor. The effect of violence on family members of people who are attacked should be less than or equal to that on people who personally experience violence. If this assumption holds, including family members should either have no effect or understate the effect of being attacked. Second, the question does not identify the source of the attack. While government repression may account for some attacks, the measure likely also includes violence from non-state actors including rebel groups, as well as common criminal activity. Again, however, I expect this to create bias against my hypotheses. Attacks that clearly should not be characterized as repression, such as domestic violence, should be less likely to influence willingness to engage and violence or identify with an ethnic group. Thus, including these types of attacks in the measure is more likely to understate the effect of repression than overstate it. With these coding decisions, 10.4% of respondents report having experienced an attack.

I also include a measure of threat perception, as I expect that the *belief* that the risk of non-violence is approaching that of violence should be sufficient to alter an individual's attitudes. While the question is somewhat limited in scope, only asking about election-related violence,[[2]](#footnote-2) this does bring the advantage of shedding light on the reason why an individual might be targeted. 28.6% of respondents reported at least some fear of being attacked during an election.

At the country level, I employ I use the Latent Human Protection Scores, version 2 (Fariss 2014; Schnakenberg and Fariss 2014). The project uses a Bayesian measurement model to estimate latent human rights scores using several data sources including US State Department and Amnesty International country reports, and several scholarly datasets on repression and mass killing. This data improves on previous approaches to measuring human rights by accounting for the fact that the standards by which government and NGO reports have judged countries have generally improved both over time and cross-nationally. The result is an aggregate measure that ranges from roughly -3 (most repressive) to 3 (most respectful of human rights). The score is calculated yearly, 1946--2015 for each country. I match the Latent Protection Human Protection Scores to each Afrobarometer respondent by the respondent's country and the year in which the survey was conducted. Within the sample, the measure ranges from -2.18 (Sudan in 2013) to 1.81 (Botswana in 2012), with mean of 0.26.[[3]](#footnote-3) The sample thus lacks any cases with the exceptionally levels of respect for human rights, as would be seen in many European democracies. The average, however, is quite close to the full sample mean of 0.29.

### Control Variables

I draw on previous studies of participation in rebellion (e.g. Humphreys and Weinstein 2008) and ethnic identity (e.g. Eifert, Miguel, and Posner 2010; Gibler, Hutchison, and Miller 2012; Masella 2013; Robinson 2014) to identify a set of relevant control variables. Each of these measures comes from the Afrobarometer, though some are not included in all waves. These include the respondent's gender, logged age, and an ordinal measure of educational attainment,[[4]](#footnote-4) and binary indicators of whether they are employed at least part-time, whether they reside in an urban area, and whether they support the ruling party. I have examined several other controls, but exclude them from the models reported here as they are neither statistically significant nor do they alter the performance of my variables of interest. These include a binary indicator for individuals who work in agriculture (farming and fishing), an index of the level of economic development in the respondent's community, and the size of the respondent's ethnic group.

At the country level I control for a curvilinear effect for ethnolinguistic fractionalization, using data from Fearon and Laitin (2003). The intuition behind this choice is that at very low levels of fractionalization, meaning most individuals belong to the same ethnic group, ethnicity is not likely to be an important social cleavage. The same is likely to hold at the opposite extreme, where individuals might fragmented into a sufficiently large number of groups that ethnicity is unlikely to be a salient. A curvilinear effect should thus identify the cases in the middle of the spectrum where ethnicity is likely to matter. Additionally, I include the country's Polity IV regime score (Marshall, Gurr, and Jaggers 2016), as Eifert, Miguel, and Posner (2010) find that elections can induce greater levels of ethnic identification. Finally, as forms of violence besides repression might influence I include indicators of whether the country had a separatist war or civil war over the central government during the year the respondent was interviewed, constructed from the Uppsala Conflict Data (Melander, Pettersson, and Themnér 2016).

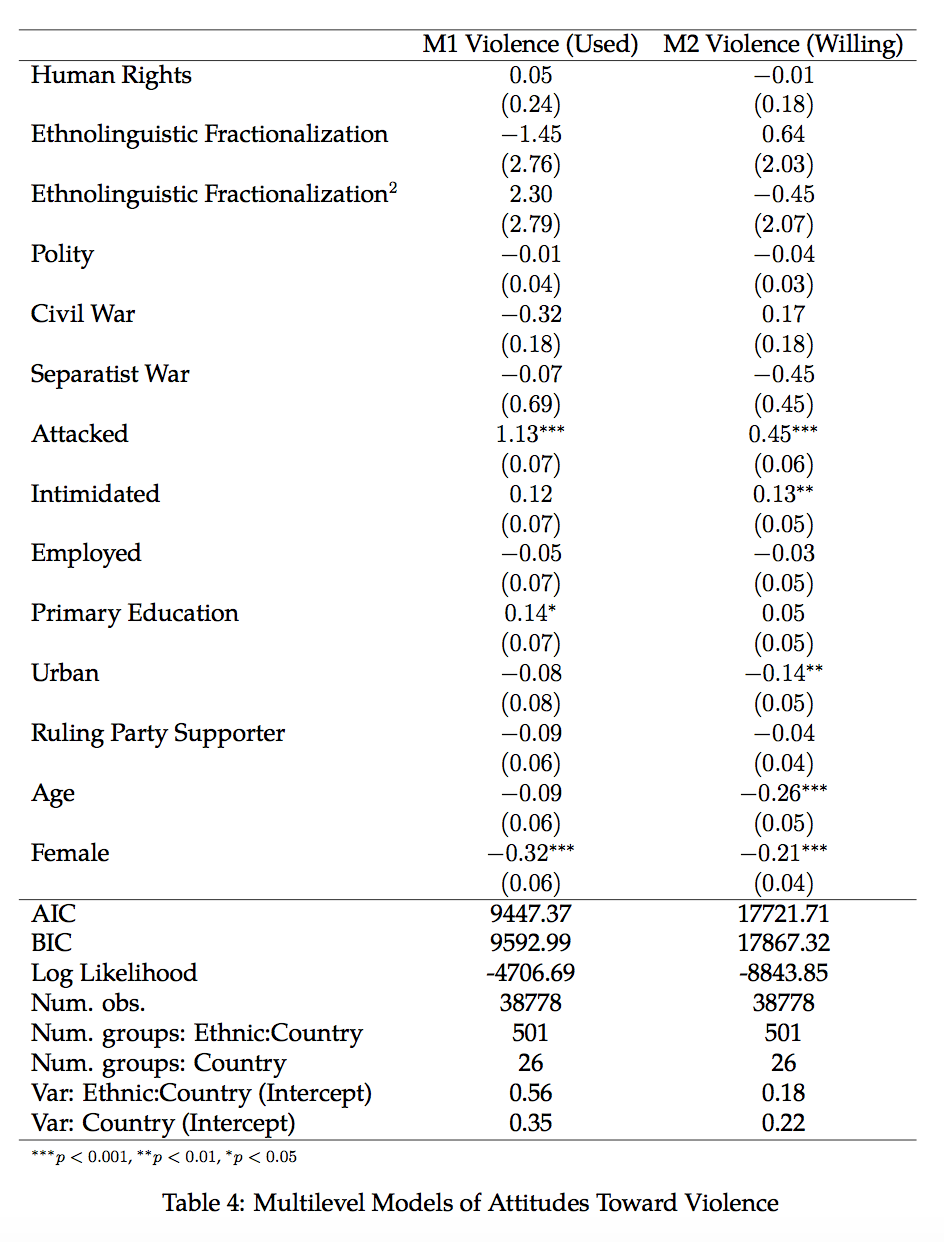
### The Model

As I am interested in the effects of variables measured at both the individual and country levels, and my dependent variables are all binary, a multilevel logistic regression model is the appropriate method of analysis. I begin with a relatively simple model with random intercepts for each country. The intuition behind this model is that the baseline values for each dependent variable vary by country, while the independent variables have a consistent effect in each country. For example, the baseline probability of ethnic identification might vary from country to country, but the model assumes that the effect of repression will be the same across all countries. For robustness, I estimate more complex models with random intercepts for each ethnic group nested within each country, and with random intercepts for each survey wave. Additionally, I utilize the survey weights provided by Afrobarometer, meaning that individuals from under-sampled groups are weighted more heavily in the regressions.

Individual-level variables are interpreted normally, with coefficients representing the increase in the logged odds ratio of the dependent variable associated with a one-unit increase in the independent variable. The country-level variables are used in a separate model, estimated simultaneously, in which the dependent variable is the group-level intercept. Thus, country-level coefficients predict the change in baseline probability in a country associated with a one-unit increase in the independent variable.

## Use of Violence Results

The results for the use of violence are reported in Table 2, Model 1. Consistent with *H1*, repression at the individual level is associated with an increased probability that a respondent has engaged in violence, or is willing to do so. The effect is substantively large, with the probability that an individual engaged in violence increasing from around 0.02 for individuals who have not been attacked, to 0.09 for individuals who have (see Figure 1). While this increase is small in absolute terms, it represents a large percentage change given how rare violence is in general. Furthermore, it is statistically significant at the 99.9% level. Neither the individual-level threat of violence, nor the country-level degree of respect for human rights significantly influences violent



mobilization. Collectively these results suggest that the presence of violence does not generally make individuals more willing to engage in violence themselves. Experiencing violence personally, however, produces such a drastic change in one's outlook that they are likely to increase their own willingness to engage in violence.

It is possible, however, that this result is endogenous. The Afrobarometer is not a panel survey, meaning that I am unable to track individuals over time. I therefore cannot determine whether attitudes toward the use of violence change in response to repression, or whether such attitudes might predate being attacked. It could be that individuals experience violence *because* they have engaged in violence themselves. Such individuals might be especially likely to be targeted with repression by the government. Furthermore, if individuals have engaged in violence, perhaps as members of a rebel group or in a riot, there is a strong possibility that their opponent will have fought back, leading the individual to report being attacked in the survey. I argue, however, that the potential for endogeneity should be substantially lower among individuals who are willing to, but have not yet engaged in violence. Identifying insurgents who intermix with the civilian population is immensely challenging for governments (Kalyvas 2006). Thus if the effect of repression on attitudes towards violence is endogenous, we might expect a weak or non-existent relationship between repression and the willingness to use violence, as it would be difficult for the government to target such individuals. As Model 2 shows, this is somewhat true. The effect of repression on willingness to use violence is weaker than the effect on the actual use of violence. Yet, the effect is still relatively large and statistically significant, suggesting that the relationship matches my causal story in at least a portion of cases.

Perhaps due in part to its rarity, only a few control variables are significantly related to violence. Consistent with previous findings (e.g. Humphreys and Weinstein 2008), women are less likely than men to engage in violence, and are less likely to report a willingness to use violence. Individuals with at least a primary school education are slightly more likely than others to participate in violence, perhaps reflecting the fact that student organizations generally account for a substantial portion of political violence. This pattern does not hold for willingness to use violence, however. Urban individuals are less likely to be willing to use violence, while the measure is unrelated to the use of violence. Finally, willingness to use violence declines with age, while participation in violence is unrelated to age.

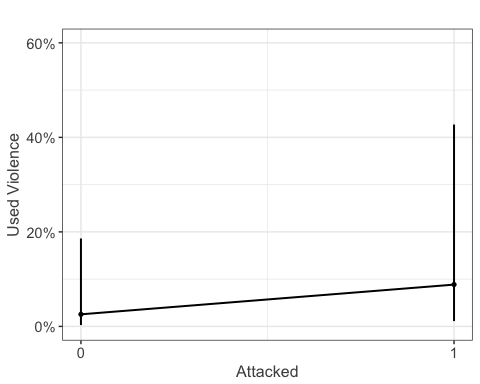
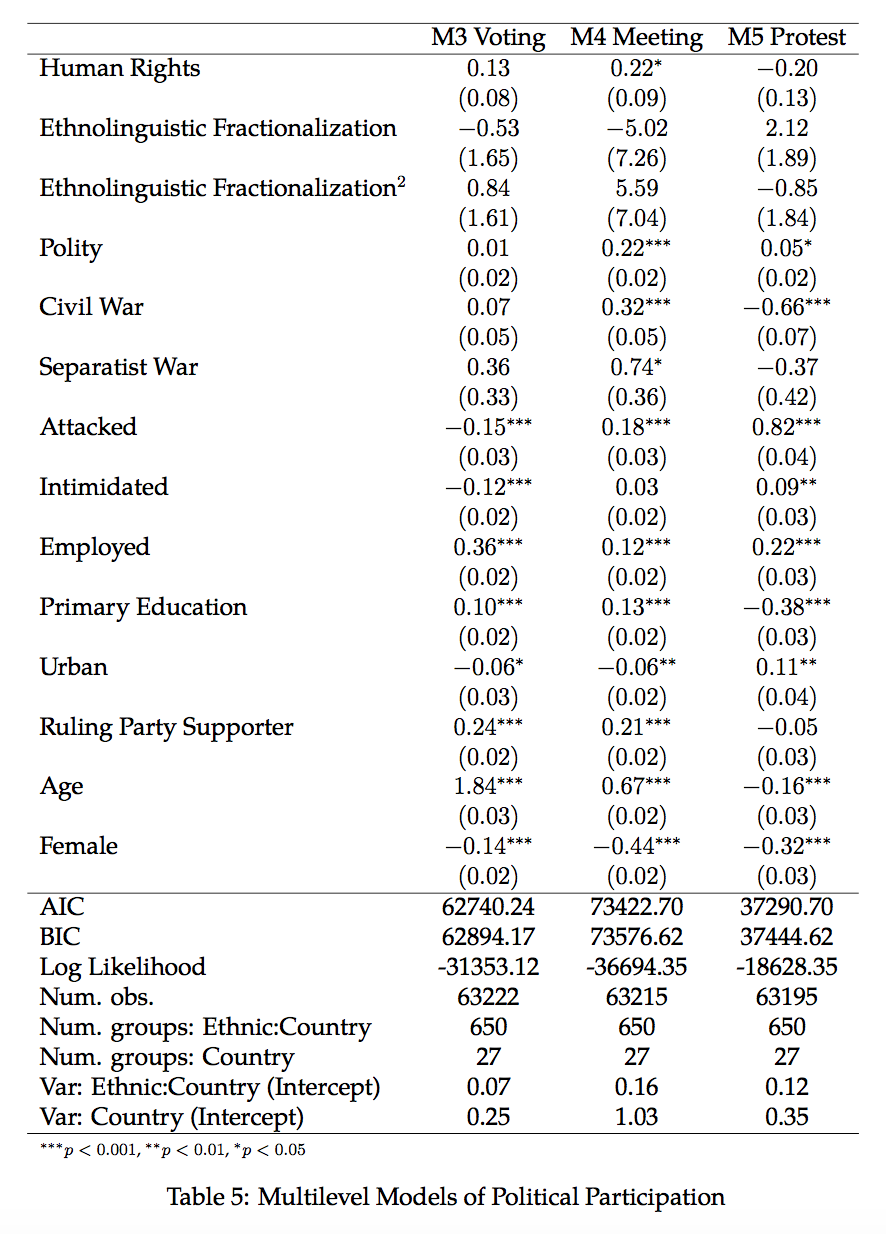


Figure 1 Predicted Probability of the Use of Violence (Model 1)

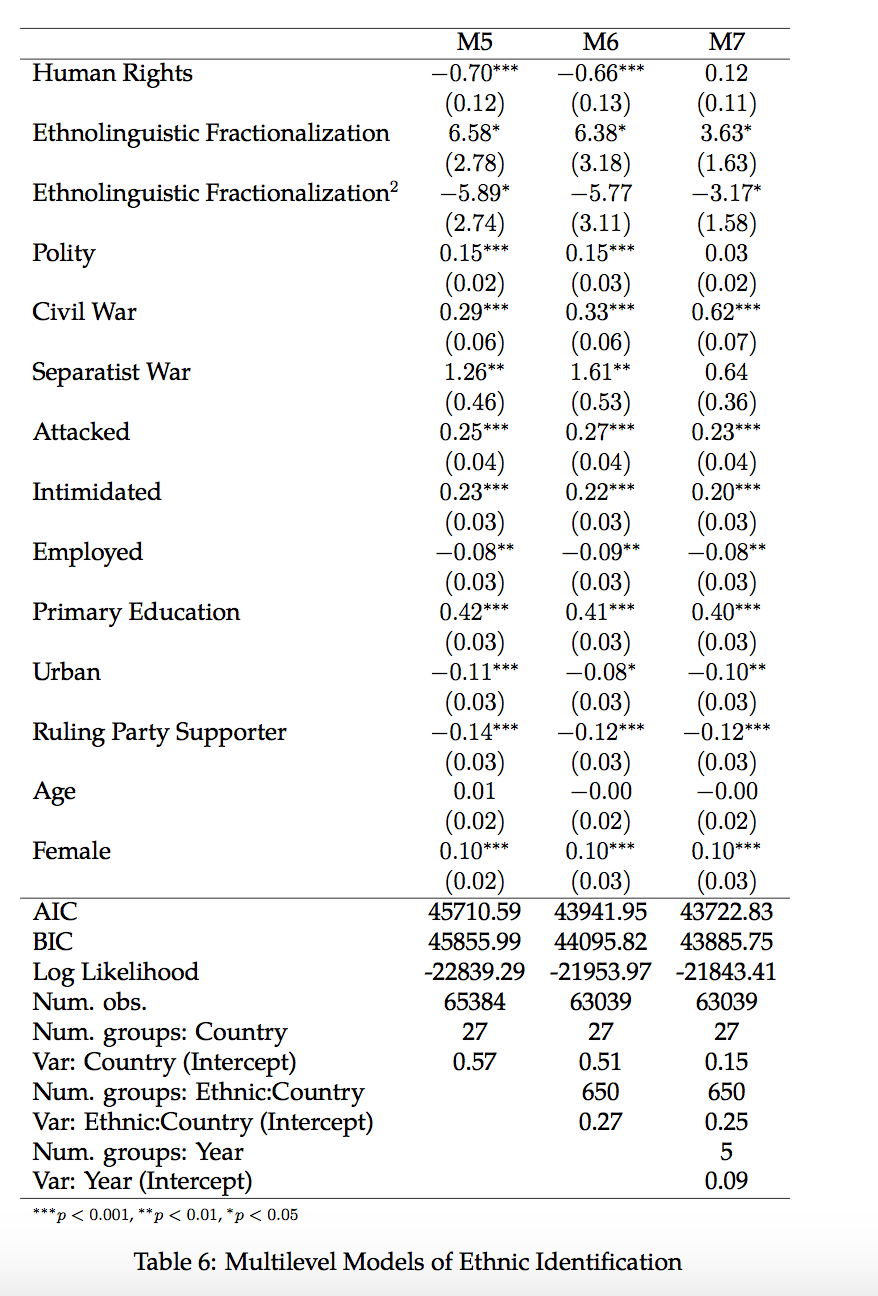


I also examine the effects of repression on three other forms of political participation to shed light on alternative explanations. For example, it may be the case that violent individuals are simply very active in general, and thus have more opportunities than others to be repressed. Comparative homebodies might see lower rates of repression simply because they spend less time in public locations where repression tends to occur, and any apparent association with lower levels of political participation would likely be coincidental. In Model 3 I examine voting. Repression measured at the individual level has a negative relationship with voting. Individuals who reported an attack on themselves or a family member were 20.4% less likely than others to have voted. The threat of election-related violence has a smaller, but still statistically significant effect. At the national level better human rights practices are associated with higher baseline rates of voting, but the effect just misses the 90% level of statistical significance. These results suggest a potential explanation to the puzzle of why governments use repression despite the negative consequences I predict - it appears that repression is effective at deterring individuals from voting. For two more involved forms of political action, however, attacks are associated with increased levels of political participation. Experiencing an attack is associated with a statistically significant, though substantively modest increase in the probability than an individual has participated in community meetings, as is the country-level human rights situation. The country-level measure is not significantly related to protest activity, but being attacked once again is, with individuals who have experienced an attack being more than twice as likely to have participated in protest. It should be noted, however, that the same endogeneity concerns that exist for violence apply to these forms of participation as well, as these results are consistent

Collectively, these results allow me to reject the null hypothesis of no relationship between repression and willingness to engage in violence associated with *H1*. Individuals who have been attacked are more than three times more likely than others to engage in violence, and roughly 20% more likely to express a willingness to use violence. The country-level human rights measure is not significantly related to either outcome, however, suggesting that the effect of repression is specific to the individuals who are targeted, and does not produce a widespread spillover effect leading large swaths of society to change their behavior. Several caveats must be noted, however. First, the repression measure is imprecise, as individuals who experienced violence personally are grouped with individuals with family members who experienced violence, and the actor who perpetrated the attack is not specified. Second, these results could be endogenous, with individuals being attacked because they used or were known to be willing to use violence. I address this possibility in Section 1.4.

## Ethnic Identification Results

The ethnic identification results are reported in Table ??. Model 5 includes a random intercept for each country, while Model 6 adds an intercept for each ethnic group nested within each country, and to that Model 7 adds a random intercept for year.[[5]](#footnote-5) In all three models, individuals who have experienced an attack are more likely to identify with their ethnic group than their nation, relative to individuals who have not experienced an attack. This effect is substantively modest, increasing the probability from rough 0.12 to 0.17 (see Figure 2), but is statistically significant at the 99.9% level. Political intimidation has a similar effect. The country-level



human rights measure is statistically significant in Models 5 and 6, with a similar substantive effect. Among the most repressive cases in the sample, individuals have a roughly 0.15 probability of identifying with their ethnic group, with the probability decreasing to 0.10 among the cases with the greatest respect for human rights (see Figure 3). The human rights variable is not significant in Model 7, likely because as a relatively constant measure, it has little ability to predict intercepts that vary by year.

The control variables provide a number of interesting results. As I expected, ethnolinguistic fractionalization has a substantively strong and statistically significant curvilinear relationship with ethnic identification in Models 5 and 7, and a linear one in Model 6 (indicated by the squared term not being statistically significant). The curvilinear pattern is consistent with my expectation that ethnic identification will be unlikely at extreme levels of diversity. Consistent with the findings of Eifert, Miguel, and Posner (2010), Models 5 and 6 show that the probability of ethnic identification increases as a country becomes more democratic. Countries experiencing civil war have a somewhat higher baseline level of ethnic identification, and the effect is considerable for separatist wars. At the individual level, urban-dwellers, ruling party supporters, and employed individuals are less likely to emphasize an ethnic identity. Having at least a primary education increases the probability that an individual with identify ethnically, and women are slightly more likely than men to adopt such an identity.

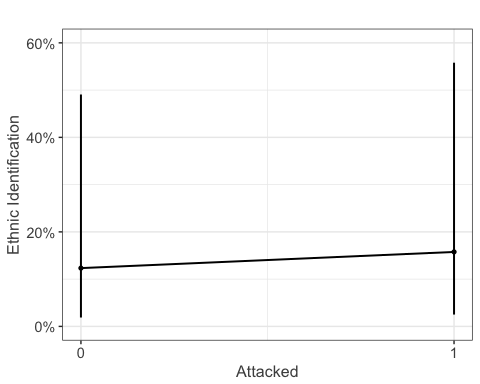


Figure 2 Predicted Probability of Ethnic Identification (Model 5)

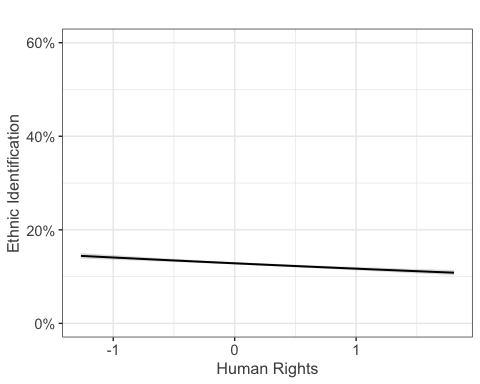


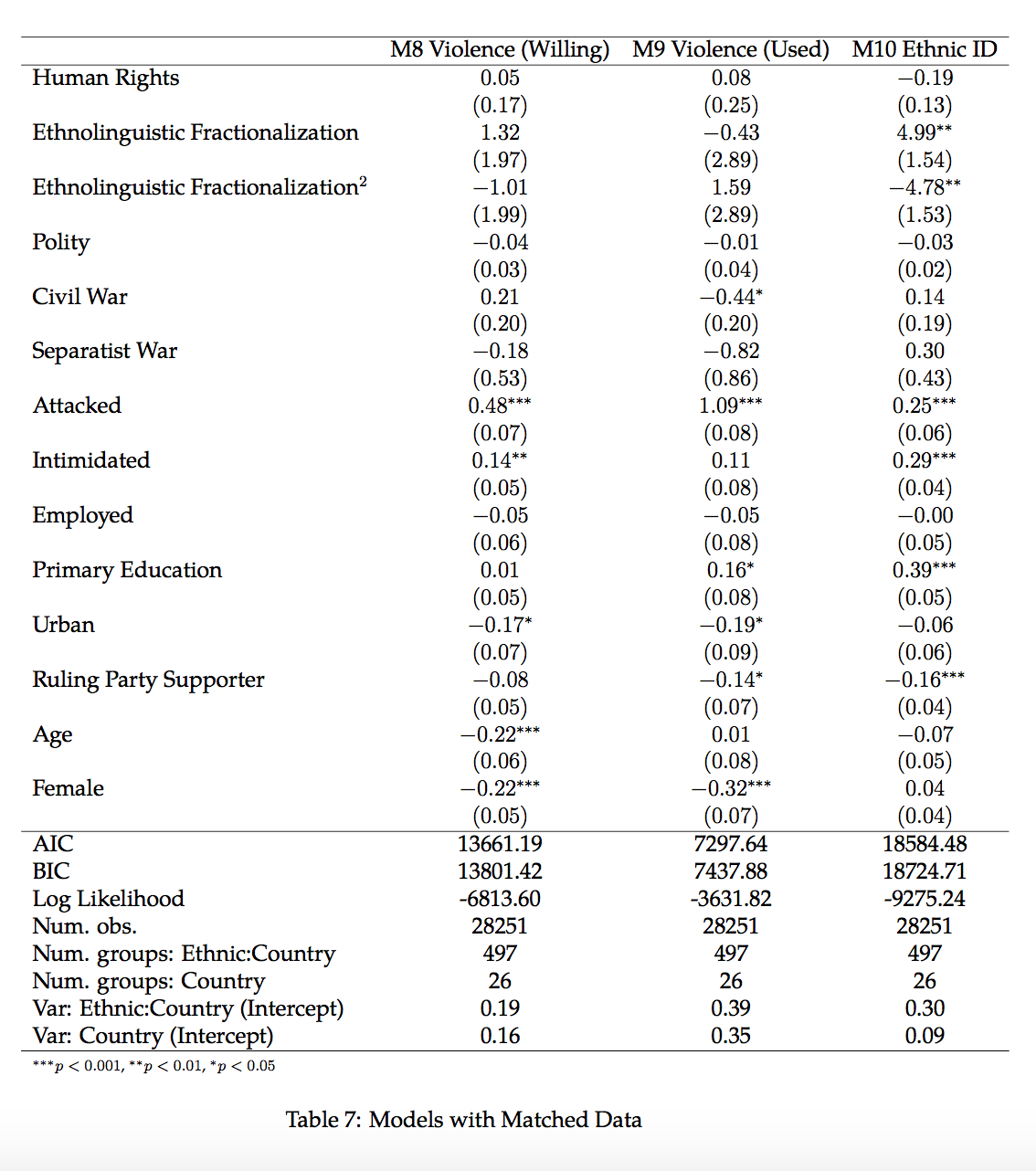
Figure 3 Predicted Probability of Ethnic Identification (Model 5)

While further analysis is need to establish the direction of the causal relationship, these results do allow me to tentatively reject the null hypothesis of no relationship between repression and ethnic identification associated with *H2*. Individuals who have been attacked are more than 40% more likely than others to identify with their ethnic group. The country-level human rights measure tells a similar story, with the probability of ethnic identification being lower in countries with greater respect for human rights. As is the case with the violence results, I cannot here rule out the possibility of endogeneity. It is quite plausible that individuals who identify strongly with an ethnic group are most likely to be targeted with repression. Many governments repress ethnic minorities to prevent them from undermining national unity. For example, the Turkish government has denied the claim that the Kurds are a distinct ethnicity from Turks, and have repressed them to prevent a secessionist movement. I address the concern in the following section.

## Causal Identification

As discussed above, the preceding results do not account for the possibility of endogeneity. The ideal solution would be an instrumental variable. Unfortunately, few if any measures included in the Afrobarometer meet the requirements of a valid instrument. For instance, previous work has often used distance from the capital to instrument for an individual or location's probability of experiencing violence (e.g. Voors et al. 2012). While this measure may meet the exclusion restriction for some outcomes, there is reason to believe that it does not for ethnic identification. Robinson (2014) finds that orientation toward national identities is driven in part by modernization. Thus living in a remote location may affect ethnic identification directly, rather than only through the variable it is intended to instrument.

As an alternative I use coarsened exact matching (Iacus, King, and Porro 2012). Matching seeks to create a subset of the data with a "treatment" (in this case the individual-level attack variable) and "control" group with similar values on a set of observable covariates. In this case I seek to balance the sample on individual-level measures of education, age, urban residence, support for the ruling party, and employment status, and country-level measures of ethnolinguistic fractionalization, Polity IV score, the Latent Human Protection Scores, and indicators for the presence of civil and separatist wars. Coarsened exact matching achieves balance by collapsing each continuous and categorical variable into a smaller number of strata, and identifying pairs of



treated and control units that fall into the same strata on each variable. While there was a statistically significant difference of means between the treated and control groups on each of the covariates prior to matching, there are no significant differences on any variable after matching, and the mean difference between the groups reduces to zero for each variable except age and the categorical education measure, which each differ by less than 0.1. The trade-off for pursuing such exact matches is a loss of observations, as cases with no close are match are discarded. The problem is not especially dire in this case, however, as the number of cases reduces from 38,681 (the number of cases with no missing values on any covariate) to 28,251. The limitation of matching is its inability to address unobservable sources of bias. Thus, if certain individuals are disproportionately likely to be attacked for reasons that are not entirely captured by the included covariates, this bias is likely to remain in the post-matching sample.

The results using the matched data are reported in Table ??, and the estimates for the attack variable are very similar to those seen in the raw data. Individuals who have experienced an attack are substantially more likely to express willingness to engage in violence, with the effect being statistically significant at the 99.9% level (Model 8). The substantive effect is modest, however, increasing the probability from 0.08 to 0.11 (see Figure 4). These individuals also report higher probabilities of having engaged in violence, and the effect is significant at the 99.9% level (Model 9). Individuals who have been attacked are three times more likely than others to have used violence themselves (0.09 vs. 0.03, see Figure 5). Additionally, being attacked is associated with a modest increase (0.14 vs. 0.11, see Figure 6) in the probability of ethnic identification, which is again significant at the 99.9% level. Many of the covariates are no longer significant after matching, as attack and non-attack subsets have identical means on these variables.

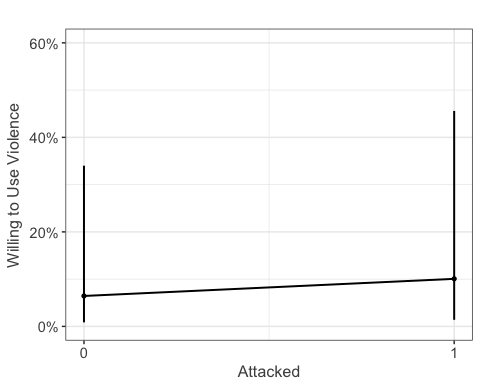


Figure 4 Predicted Probability of the Use of Violence (Model 8)

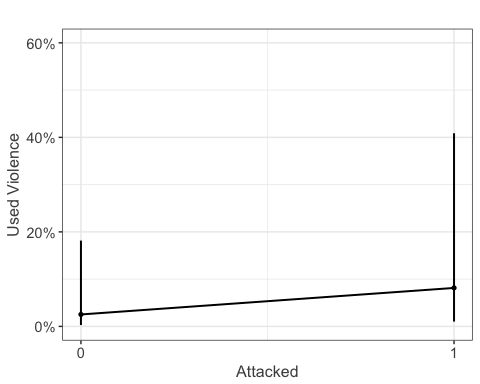


Figure 5 Predicted Probability of the Use of Violence (Model 8)

As noted above, matching cannot guard against all potential threats to causal inference. If some individuals are disproportionately likely both to be attacked and to engage in violence or identify ethnically for reasons that are not captured by the covariates, this bias will remain. One might imagine, however, that governments (and other actors) often make decisions of who to repress based on the sort of observable characteristics such as age and sex that are included in the matching. The matching analysis ensures that these observable measures do not bias the results. Thus while there is still some possibility of endogeneity, these results should increase our confidence that individuals who are attacked are not systematically different from others.

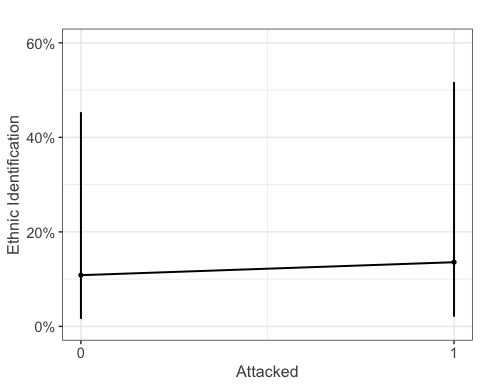


Figure 6 Predicted Probability of Ethnic Identification (Model 9)

## Conclusion

The results in this chapter provide strong support for the microfoundations of my theory. I expected that repression would make individuals more willing to engage in violence. Consistent with this hypothesis, I find that while such sentiments are generally rare, individuals who have experienced a violent attacked are roughly 30% more likely than others to report a willingness to use violence, and are nearly three times more likely to report having used violence. I also predicted that repression should induce greater levels of ethnic identification among its targets. Indeed, I find that individuals who have experienced an attack are 62% more likely than others to identify more with their ethnic group than with their nation. The results hold after conducting coarsened exact matching, meaning that the results are not driven by any observable differences between the individuals who have been attacked and those who have not.

This analysis has several important practical and theoretical implications. First, it suggests that repression is often counterproductive. Presumably, governments use repression to mitigate and deter threats to their rule. Yet, my findings suggest that repression could *increase* the number of individuals using violence, and entrench identities that could form the basis of an opposition to the government. As discussed in Chapter @ref(#theory), this makes the government's use of repression puzzling. My other findings on political participation hint at an answer, however. Repression does seem to reduce the probability than individual will vote, suggesting that governments may be accepting increased numbers of violent individuals in exchange for the opportunity to shape the electorate. Second, the results suggest that repression can trigger a vicious cycle, in which the government responds to an initial threat in a way that further entrenches the opposition, leading to ever-greater levels of violence.

This analysis could be refined on several dimensions in future work. One limitation of the existing results is their inability to identify the source of repression. It would be possible to make inferences about the likely perpetrator by matching the survey results, which include the respondent's city, to a geocoded dataset of battles, such as ACLED (Raleigh 2012). If most of the violent events in a particular locale are perpetrated by the government, it might be reasonable to assume that it is the source of most attacks on individuals in that area. By contrast, this would not be a safe assumption in territory that is clearly controlled by a rebel group. The use of an external conflict data source could also address the issue of temporal ordering. The Afrobarometer data does not specify whether individuals were attacked before they engaged in violence. With geocoded conflict data one could examine whether the average probability of participation in violence or of ethnic identification in a geographic area changes after violent events there. Finally, a more robust method of causal inference that can account for unobservable sources of bias would enhance the validity of the results. While finding a valid instrument at the individual level may be difficult, it should be possible to instrument for the country-level human rights measure.

In the remaining chapters, I build on the foundations established here to explain variations at the level of the rebel movement. I find evidence that the dynamics discussed here shape the formation of new rebel groups, the fragmentation of existing ones, and the formation of alliances between previously independent groups.

# References

Bernstein, Robert, Anita Chadha, and Robert Montjoy. 2001. “Overreporting Voting.” *Public Opinion Quarterly* 65 (1): 22–44. doi:[10.1086/320036](https://doi.org/10.1086/320036).

Eifert, Benn, Edward Miguel, and Daniel N. Posner. 2010. “Political competition and ethnic identification in Africa.” *American Journal of Political Science* 54 (2): 494–510.

Fariss, Christopher J. 2014. “Respect for Human Rights has Improved Over Time: Modeling the Changing Standard of Accountability.” *American Political Science Review* 108 (2): 297–318.

Fearon, James D., and David D. Laitin. 2003. “Ethnicity, Insurgency, and Civil War.” *American Political Science Review* 97 (1): 75–90.

Gibler, Douglas M., Marc L. Hutchison, and Steven V. Miller. 2012. “Individual Identity Attachments and International Conflict: The Importance of Territorial Threat.” *Comparative Political Studies* 45 (12): 1655–83. doi:[10.1177/0010414012463899](https://doi.org/10.1177/0010414012463899).

Humphreys, Macartan, and Jeremy M Weinstein. 2008. “Who Fights ? in Civil War The Determinants of Participation.” *American Journal of Political Science* 52 (2): 436–55. doi:[10.1111/j.1540-5907.2008.00322.x](https://doi.org/10.1111/j.1540-5907.2008.00322.x).

Iacus, Stefano M., Gary King, and Giuseppe Porro. 2012. “Causal inference without balance checking: Coarsened exact matching.” *Political Analysis* 20 (1): 1–24.

Kalyvas, Stathis N. 2006. *The Logic of Violence in Civil War*. Cambridge: Cambridge University Press.

Kuenzi, Michelle, and Gina M. S. Lambright. 2007. “Voter Turnout in Africa’s Multiparty Regimes.” *Comparative Political Studies* 40 (6): 665–90.

Marshall, Monty G., Ted Robert Gurr, and Keith Jaggers. 2016. “Polity IV Project Dataset Users’ Manual, v.2015.” *Polity IV Project*, 1–86. doi:[10.1177/0738894213499673](https://doi.org/10.1177/0738894213499673).

Masella, Paolo. 2013. “National identity and ethnic diversity.” *Journal of Population Economics* 26 (2): 437–54. doi:[10.1007/s00148-011-0398-0](https://doi.org/10.1007/s00148-011-0398-0).

Melander, Erik, Therése Pettersson, and Lotta Themnér. 2016. “Organized violence, 1989–2015.” *Journal of Peace Research* 53 (5). SAGE PublicationsSage UK: London, England: 727–42.

Raleigh, Clionadh. 2012. “Violence Against Civilians: A Disaggregated Analysis.” *International Interactions* 38 (4): 462–81. doi:[10.1080/03050629.2012.697049](https://doi.org/10.1080/03050629.2012.697049).

Robinson, Amanda Lea. 2014. “National Versus Ethnic Identification in Africa: Modernization, Colonial Legacy, and the Origins of territorial Nationalism.” *World Politics* 66 (4): 709–46. doi:[10.1017/S0043887114000239](https://doi.org/10.1017/S0043887114000239).

Schnakenberg, Keith E., and Christopher J. Fariss. 2014. “Dynamic Patterns of Human Rights Practices.” *Political Science Research and Methods* 2 (1): 1–31.

Vogt, Manuel, Nils-Christian Bormann, Seraina Ruegger, 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. doi:[10.1177/0022002715591215](https://doi.org/10.1177/0022002715591215).

Voors, Maarten J., Eleonora E M Nillesen, Philip Verwimp, Erwin H. Bulte, Robert Lensink, and Daan P. Van Soest. 2012. “Violent conflict and behavior: A field experiment in Burundi.” doi:[10.1257/aer.102.2.941](https://doi.org/10.1257/aer.102.2.941).

1. The ethnic vs. national identity question was not asked in waves 1 and 2. [↑](#footnote-ref-1)
2. Question text: "During election campaigns in this country, how much do you personally fear becoming a victim of political intimidation or violence?" [↑](#footnote-ref-2)
3. Calculated with all country-years included in the data weighted equally. [↑](#footnote-ref-3)
4. Possible responses: No formal schooling, Informal schooling only (including Koranic schooling), Some primary schooling, Primary school completed, Intermediate school or Some secondary school / high school, Secondary school / high school completed , Post-secondary qualifications, other than university e.g. a diploma or degree from a polytechnic or college, Some university, University completed, Post-graduate. [↑](#footnote-ref-4)
5. I use year instead of survey wave as the country-level variables are measured in yearly intervals. [↑](#footnote-ref-5)