Sowing Hatred: Local Ethno-Political Competition and Pre-Election Violence in Majoritarian Elections

Carl Müller-Crepon*

International Conflict Research Group, ETH Zurich

December 2018

Abstract

How do local ethnic demographies affect the conduct of majoritarian elections? Because legislative elections in majoritarian systems are contested locally, local ethno-political polarization increases the risk of pre-election violence. Violence in such cases can be targeted with comparative ease at opposing voters, and can, if perpetrated collectively, mobilize the perpetrators' co-ethnics. No such dynamic is expected in PR systems where political competition plays out at higher geographical levels. Empirically, I combine new data on the ethnic composition of local populations in 22 African countries with monthly data on riots as well as survey data on citizens' fear of campaign violence. Ethno-politically polarized districts in majoritarian and mixed electoral systems see substantively higher (1) increases in the number of riots prior to elections and (2) levels of fear of pre-election violence than non-polarized districts in the same country and at the same time. Pure PR systems do not exhibit this pattern.

*email: carl.mueller-crepon@icr.gess.ethz.ch. This paper has profited from very constructive comments made by Robert Bates, Nils-Christian Bormann, Lars-Erik Cederman, Michael Gilligan, Philipp Hunziker, Yannick Pengl, Espen Geelmuyden Rød, and participants of the 2016 ISA Conference in Atlanta, as well as of workshops at Harvard University and in Konstanz and Witten (Germany).

Introduction

Choices over the design of electoral systems in ethnically divided societies are most influential in determining the fate of democracy and peace in a polity. Addressing a vital threat to young democracies around the globe, this paper analyzes the impact of local ethnic demographies on violence preceding majoritarian elections in Africa. I argue that the risk of pre-election violence increases in the degree of local polarization between politically mobilized ethnic groups before majoritarian but not proportional legislative elections.

The literature on the vices and virtues of majoritarian and PR systems in ethnically divided societies is, beginning with the seminal contributions of Horowitz (1985, 1990, 1991, 1994) and Lijphart (1985) and Lijphart and Aitkin (1994), extensive. It mostly focuses on the effects of electoral systems on political parties, post-conflict stability, and the risk of civil war in general.[1] Importantly however, arguments about the effects of majoritarian systems in ethnically diverse societies for the most part disregard the intricate nexus between ethno-political geographies and legislative elections which are conducted in local constituencies (Barkan et al., 2006; Wagner and Dreef, 2013). In addition, only relatively coarse evidence exists on the effect of electoral systems on electoral violence. While Birch (2007) and Fjelde and Höglund (2016) present country-level evidence that majoritarian elections come with more misconduct and campaign violence than PR systems, we lack evidence on the strategic incentives of candidate MPs to campaign violently. Addressing these two gaps in the literature, I analyze high resolution spatio-temporal data on violence and local ethnic demographies to test the effect of local ethno-political competition on campaign violence in majoritarian systems. The analysis does not only contribute to the relevant literature but also presents actors that try to prevent electoral violence with evidence on the conditions under which it is likely to break out.

The literature on campaign violence describes the phenomenon as the violent counterpart to monetary campaign expenditures. Pre-election violence can increase the odds of victory of its instigator through the polarization of the electorate (Dercon and Gutiérrez-Romero, 2012; Horowitz, 2001; Wilkinson, 2004) and the demobilization of his opponent's voters by means of intim-

[1] On the effects of the electoral system in ethnically diverse polities see for example Neto and Cox (1997),Ordeshook and Shvetsova (1994), and Mozaffar et al. (2003); on post-conflict stability: Bogaards (2013) with an encompassing review of the debate; on civil wars: Reynal-Querol (2005), Reynal-Querol (2002), Saideman et al. (2002), and Schneider and Wiesehomeier (2008).

idation, displacement, and death (Bratton, 2008; Collier and Vicente, 2014; Klopp, 2001; Steele, 2011). It not only affects country-wide official elections, but also intra-party nomination processes (Bech Seeberg et al., 2018). In parallel to the incentives underlying peaceful campaigning, contested elections are at highest risk of being preceded by violence (Hafner-Burton et al., 2013; Salehyan and Linebarger, 2015).

Building on this literature, as well as drawing on the structure and geographic locus of legislative electoral contests in majoritarian systems, I argue that local political competition between ethnic groups increases the risk of violent campaigning in majoritarian systems. In ethno-politically polarized constituencies, violence can be effectively targeted at opposed voters and, especially when it comes in the form of a riot, serves the purpose of polarizing the electorate. In contrast, no effect of local ethno-political competition is expected in PR systems, where the locus of electoral competition is either regional or national and the ethnic composition of localities does not carry much information about where to best incite violence.

This argument is tested using new spatial data on the ethnic composition of local populations in 22 African countries in many of which commonly experience electoral violence (e.g. Goldsmith, 2015). The analysis of the monthly number of riots at the level of districts between 1990 and 2013 shows that districts in majoritarian systems that are polarized between politically mobilized ethnic groups experience steeper increase in the number of riots prior to elections than non-polarized districts. This effect is not found in PR systems. Rigid fixed effects on the country-month and district-year levels, as well as controls of spatio-temporal autocorrelation of the data severely restrict the potential for spuriousness of the results. In addition, an analysis of geocoded survey data on citizens' fear to experience pre-election violence or intimidation across 19 African countries exposes the same pattern: individuals who live in polarized districts under majoritarian voting fear violent campaigns more than their co-nationals in non-polarized districts. Again, responses from PR systems do not exhibit such a pattern.

The geography of ethno-political competition and violence before legislative elections

In line with the potential of pre-election violence to increase the electoral chances of perpetrators, a broad range of studies (Collier and Vicente, 2014,

2011; Horowitz, 2001; Wilkinson, 2004) argues that violence during electoral campaigns is oftentimes orchestrated or "produced" (Brass, 2011) by political elites and their henchmen trying to maximize their chances at the ballot box. In particular in ethnically divided constituencies, candidates might choose to deliberately incite ethno-nationalist discourses and plan inter-ethnic violence. Such patterns have been observed, for example, in India (Brass, 2011; Wilkinson, 2004) and by Throup and Hornsby (1998) in the 1992 Kenyan legislative election. Here, local MPs of the *Kenya African National Union* (KANU) actively recruited those who later attacked 200'000 ethnic Kikuyu, Kisii, Luo, and Luhya which where associated with their electoral opponents (see also Klopp, 2001; Klopp and Zuern, 2007).

In parallel to monetary expenditures (Cox and Munger, 1989; Erikson and Palfrey, 2000; Pattie et al., 1995), campaigns are more likely to come with substantial bloodshed where races are expected to be close (Hafner-Burton et al., 2013; Klopp and Zuern, 2007; Salehyan and Linebarger, 2015; Wilkinson, 2004). Only then do the expected benefits of violence outweigh its costs which consists in material payments for those who perpetrate the violence, the risks of alienating part of one's support-base, and the potential of being persecuted by the judiciary.

However, for pre-election violence to be effective, it must be targeted at the voters of the perpetrator's opponent(s). While ideologically motivated electoral preferences are difficult to observe, prospective vote choices that are based upon ethnic identities are much more easily discerned. Since voters in many multi-ethnic societies base their vote to a significant degree on ethnic attributes of candidates such as language, religion, or name (Adida, 2012; Basedau et al., 2011; Bratton et al., 2012; Bratton and Kimenyi, 2008; Chandra, 2004), perpetrators of campaign violence can use the same characteristics to increase the precision of their targeting.[2] But not only does the politicization of ethnicity ease the violent demobilization of opponents, it also increases the potential of violence to foster the mobilization of the perpetrator's ethnic support base by highlighting ethnic differences and inciting ethno-nationalist sentiments (Dercon and Gutiérrez-Romero, 2012; Horowitz, 2001; Wilkinson, 2004). This argument coincides with the general tendency of contested electoral campaigns in Africa to increase the salience of ethnic identities (Eifert et al., 2010).

[2]See Fearon (1999) for a similar argument on the incentives for politicians in multi-ethnic societies to deliver 'pork' to their co-ethnics.

Not all forms of violence are equally suited to achieve the aims of pre-election violence in ethnicized polities. To achieve the first goal of demobilizing opposing voters, violence has to be ethnically targeted to such an extent as to induce sufficient fear among them and their co-ethnics. As to pursue the second goal, the rise of the salience of ethnic identities among the voters of the violence-inducing candidate himself, the demographic basis of those who perpetrate the violence has to be equally broad. Only if a sufficient number of people participate in the violence can a public arousal of sentiment be achieved (Brass, 2011). With these two goals of pre-election violence in ethnicized polities in mind, the ethnic riot fits the incentives of politicians much better than other forms of collective violence, because it combines ethnicized popular mobilization with selective targeting of ordinary members of the ethnic 'other' (Horowitz, 2001; Wilkinson, 2004). It is, in addition, a form of medium-scale violence with a low risk of punishment. Given its broad demographic basis, even independent prosecutors may find it difficult to expose the planners behind riotous masses after the fact. In contrast, violence executed by organized structures such as the police, political parties, and militias is much less anonymous, leaving more traces for prosecution and punishment.

From the previous reasoning it emerges that political competition between ethnic groups increases the risk of pre-election violence, in particular in the form of riots. Since the degree to which the competition for political power is ethnicized is strongly related to the electoral design of a multi-ethnic society,[3] the electoral system likely also affects the extent to which one should expect violent legislative campaigns. This literature on electoral systems can be roughly divided into two camps. The first holds that PR leads to equal representation of all ethnic groups, facilitating power-sharing, preventing the political domination of single groups, and thus fostering peace (Lijphart and Aitkin, 1994).[4] Critics of this view hold that PR encourages ethnic mobilization and perpetuates divisions along main cleavages (Horowitz, 1991, 167-172). Instead, they argue that ethnically divided societies should conduct elections following plurality rules, in particular using the 'single transferable' or 'alternative' vote that encourages cross-ethnic alliances and intra-ethnic divisions (Horowitz, 1991, 1985). Empirically however, PR systems are found to exhibit lower degrees of ethnicization of political preferences than majoritarian systems (Huber, 2012). This coincides with Fjelde and Höglund's (2016) finding

[3]See Bogaards (2013) for a review.

[4]Correspondingly, Schneider and Wiesehomeier (2008) suggest that PR has pacifying effects on ethnically fractionalized societies.

that majoritarian countries in Africa exhibit more electoral violence than proportional ones, especially where large ethnic groups are excluded from political power.

Notwithstanding its merits, the grand debate between proponents of proportional vs. majoritarian voting does not sufficiently consider the importance of the geography of ethnic cleavages (cf. Barkan et al., 2006; Wagner and Dreef, 2013) when analyzing the consequences of the two systems.[5] Since majoritarian elections are contested locally, the geography of political preferences is a key determinant of the degree of competition in the electoral districts (Sartori, 1997).[6] Consequently, the risk of pre-election violence under ethnicized voting in a majoritarian system is co-determined by the extent to which local constituencies are divided between ethnic groups. The risk of violent campaigns will be highest where two politically mobilized ethnic groups of equal size populate a district, making it ethno-politically polarized. The risk of electoral violence decreases as the number of groups and/or their heterogeneity in size increases.

Hypothesis: *In plurality systems, the risk of pre-election violence increases in the local degree of ethno-political polarization.*

The link between local political competition between ethnic groups is thus contingent on the nature of majoritarian systems and the locus of their electoral contests. Fundamentally different geographical patterns of pre-election violence should therefore be observed in pure PR systems which are here analyzed as a control group only. Under PR voting, competition takes place at a higher geographical level, mostly at the region- or country-level. Thus, regional or national characteristics will shape incentives for violent campaigning before elections. In contrast, the degree of *local* ethno-political polarization will not influence considerations about where to best incite violence before an election since it does not strongly determine the share of votes won by parties. In a proportional contest it might be much less 'effective' to target ethnically mixed areas than those homogenously inhabited by one's opponents – a strat-

[5]The same disregard of geographical factors interacting with the electoral system can be seen in studies of the impact of ethnic heterogeneity on the number of parties in a polity (Neto and Cox, 1997; Ordeshook and Shvetsova, 1994). An exception is found in Mozaffar et al. (2003).

[6]In extension, the geographic distribution of partisan preferences influences the complex translation of votes to assembly seats in majoritarian polities (Barkan et al., 2006; Calvo and Rodden, 2015; Gudgin and Taylor, 2012; Rodden, 2010).

egy that South Africa's ANC pursued in the first post-apartheid election in 1994 (Klopp and Zuern, 2007). In addition and as Birch (2007) points out, parties in PR systems pool the risks and benefits of electoral campaigning. They are thus less vulnerable to the collective action problems faced in majoritarian systems (Carey and Shugart, 1995) and have greater powers to maintain their credibility and avoid violence during electoral campaigns.

In order to comprehensively test this argument, I analyze two sets of data on different levels of observation and with different measures of the extent of local pre-election violence. The first combines unique data on the ethnic composition of African districts with monthly riot data to model local increases of the number of riots prior to legislative elections. The second analysis combines the district-level data on ethnic compositions with Afrobarometer data on citizens' reported fear of pre-election violence. I proceed by introducing and analyzing the first dataset to then come to the second one.

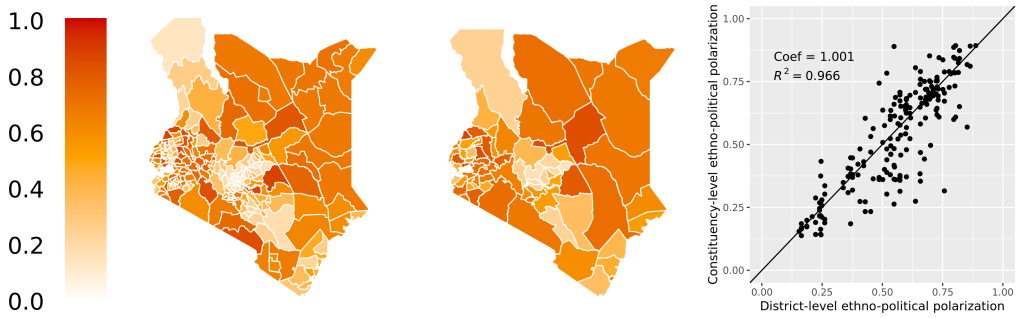
Local ethno-political competition and pre-election riots

Using data on riots in Africa, this section analyzes the effect of local ethno-political polarization on the extent to which districts experience an increase in rioting prior to majoritarian elections. The empirical analysis finds that ethno-politically polarized districts see an escalation of rioting before legislative, majoritarian elections that is much more severe than the escalation observed in non-polarized districts or before legislative elections under proportional voting.

Data

The main empirical strategy models differences in the pre-election increase of the monthly number of riots as legislative elections approach in ethno-politically polarized and non-polarized districts. The focus on the pre-election *increase* in violence comes much closer to the hypothesized causal mechanism than other models on the link between electoral competition and local-level violence. These either compare average levels of violence across units of analysis[7] or restrict the sample to election periods only (e.g. Hafner-Burton et al.,

[7]Most prominently, Wilkinson (2004) argues that party-competition in Indian states increases the odds of Hindu-Muslim riots. In his empirical analysis however, riots are modeled only as a function of the additive effects of electoral proximity and party-competition. This leaves the conjoint impact of the two variables, which is at the heart of the theoretical



(a) Constituency-level (b) District-level (c) Comparison

Figure 1: Ethno-political polarization in Kenya

2013; Daxecker et al., 2018). The latter strategy lacks the crucial comparison of violence during months directly before an election with violence occurring at other points in time.[8]

Therefore, the district-month in 22 African countries between 1990 and 2013 constitutes the fundamental unit of analysis. Yearly varying data on the spatial extent of districts, defined as the second administrative level in states, comes from FAO's (2014) GAUL data. Administrative units as units of analysis might seem inferior to using electoral districts where electoral competition takes place. However, there is no comprehensive cross-national data on electoral districts available to date. More importantly, taking electoral districts as units of analysis would make a comparison between majoritarian and PR systems all but impossible, because electoral districts created for majoritarian elections do not exist in PR systems. Because electoral districts in majoritarian systems are typically nested within administrative units, measures for administrative districts are expected to be a reliable proxy for those on the level of majoritarian constituencies – real ones in majoritarian systems and 'counterfactual' constituencies in PR systems. To vindicate this notion, Figure 1 plots the degree of ethno-political polarization, which is introduced below, in Kenyan districts and 2007 constituencies. A statistical comparison of the two measures indicates that district-level ethno-political polarization explains 97% of the variation found at the constituency-level.[9]

The main dependent variable, the monthly count of riots, is retrieved from the geocoded Social Conflict in Africa Data (SCAD 1990-2013; Salehyan et al.,

argument, unexplored.

[8]In particular, it might be that polarized districts *always* experience more violence. Such constant violence might be unrelated to electoral dynamics.

[9]See also the results in Table A8 that compares district- with constituency-level results on electoral violence before Kenyan elections.

2012). Riot-events are spatio-temporally matched to district-polygons and aggregated to the monthly level.[10] To compare the robustness of the results with different conflict data (Hegre and Sambanis, 2006), I complement the analysis with counts of riots and riot-fatalities from the ACLED data (1997-2015; Raleigh et al., 2010). Furthermore, I will also draw on the new ECAV data on local electoral violence (Daxecker et al., 2018) for a robustness check.[11] Throughout, I take the natural logarithm of the count of riots and riot fatalities +1 as the dependent variable to alleviate the right-skew in the dependent variable and to take into account the intuition that the step from 0 to 1 riot is much larger than moving from 3 to 4 riots.

To model the increase of rioting prior to legislative elections, each district-month is associated with its temporal distance to the next legislative election. Data on the date of elections are taken from the National Elections across Democracies and Autocracies data (NELDA v4; 1989-2012; Hyde and Marinov, 2011).[12] Because the 'effectiveness' of violence likely increases exponentially as elections come closer, the variable time to election is calculated as the inverse of the distance to the next legislative election (after adding 1 so as not to divide by 0 in election months). The value of the variable thus increases as an election comes closer. The exponential setup is more realistic (see Figure 3 in the next section), more flexible, and does fit the data better than a simpler pre-election dummy. Lastly, most analyses are conducted on the sample of countries under majoritarian election rules only. To differentiate these from proportional systems, I rely on the World Bank Data on Political Institutions (Beck et al., 2001). The data encode whether legislators are elected using a first-past-the-post or winner-takes-all rule. This coding includes 5 mixed majoritarian and PR systems[13] for which, according to the argument presented above, incentives for pre-election violence should be higher in ethno-politically polarized single-member-districts as well.[14]

[10] Events without coordinates are dropped from the sample. The few riots that span more than one month are attributed to the month of their onset.

[11] As highlighted above, this data does not allow for observing *increases* in violence as elections approach as it lacks data on the intensity of local violence in non-campaign periods.

[12] The data have been extended up to December 2014 to make use of the full set of SCAD events. All months further away from the next election than 5 years are dropped from the sample.

[13] Cameroon, DRC Congo, Guinea, Niger, and Senegal; cf. Figure A1.

[14] The data cover all years up to 2012. The electoral system under which elections in 2013 have taken place has been hand-coded for the given countries. Since the data are coded on the basis of a January 1st rule, I additional make sure that years of elections following new electoral rules, as for example in Togo 2007, are recoded so as to reflect the systemic incentives for electoral violence.

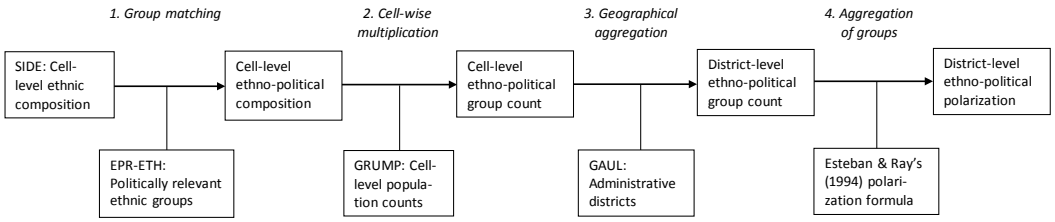


Figure 2: Flowchart of the construction of the measure of district-level ethno-political polarization.

I measure the degree of local ethno-political polarization by computing a polarization index with data on local ethnic demographies and the political relevance of ethnic groups. The first input consists in newly available maps of the ethnic composition of local populations in Africa (Spatially Imputed Data on Ethnicity SIDE; Müller-Crepon and Hunziker, 2018). These data address the need for detailed cross-national data on ethnic geographies, which arises from a lack of micro-level census data in many developing countries as well as the inadequacy of polygon-based data (such as GREG and GeoEPR; Weidmann et al., 2010; Wucherpfennig et al., 2011) on the matter. The data are constructed by spatially imputing the ethnic composition of geocoded survey-clusters enumerated in USAID's Demographic and Health Surveys (DHS, 2015). Using non-parametric modeling techniques, Müller-Crepon and Hunziker (2018) impute this point-level data over a grid with a resolution of 8.3×10^{-3} degrees ($\sim$1km). As an indication of its reliability, the SIDE data exhibits substantial overlap with local level census data from Uganda and Senegal. Since the SIDE maps are available for different years, I take the most recent map available for every district-month. Where no past maps are available, the most proximate map from the following years is used (Figure A1 in the Appendix).

Based on the SIDE data, the measure for local ethno-political polarization is constructed in four steps visualized in Figure 2. To move from ethnic compositions of local populations to their composition in terms of politically mobilized ethnic groups, I first match the SIDE data with the Ethnic Power Relations dataset (EPR-ETH, 1945-2013; Cederman et al., 2010; Vogt et al., 2015). This dataset provides a time-varying list of ethnic groups which are politically mobilized by at least one actor at the national level or which are politically discriminated against by the state. The coding of political mobilization is most often based on the existence of ethnically mobilizing parties or

politicians. Because MPs are often part of larger ethnic coalitions[15], the data fits the proposed theoretical argument well. Each ethnic group in the SIDE data is matched with its counterpart in the EPR-ETH for every year between 1990 and 2013.[16] Groups from SIDE which cannot be matched to EPR-ETH are coded as being politically irrelevant.

In a second step, each grid-cell is weighted with its respective population in a given year.[17] The resulting grid of head-counts of politically relevant ethnic groups is then aggregated to the district-year polygons introduced above (step 3). In the fourth step, the yearly ethno-political composition of districts is used to derive the measure for local ethno-political polarization (see Figure 1), applying the standard formula for polarization introduced by Esteban and Ray (1994).[18]

Empirical strategy

Using the resulting dataset, the effect of ethno-political polarization on the increase in rioting prior to elections is modeled as follows:

$$\begin{aligned}\text{riot}_{dcym} = \boldsymbol{E}_{dcym} + \beta_1 \text{time to election}_{mc} + \beta_2 \text{ethno-political polarization}_{dy} + \\ \beta_3 \text{time to election}_{mc} \times \text{ethno-political polarization}_{dy} + \delta \boldsymbol{X}_{dm} + \epsilon_d,\end{aligned}$$

where $riot_{dcym}$ is the logged count of riots in district d and month m of year y which is associated with a time to the next election in its country c as well as with a level of ethno-political polarization. Since ethnic heterogeneity and thus potentially also polarization is expected to be higher in populated districts which might also experience most pre-election riots, the natural logarithm of districts' population and its interaction with time to election are added as

[15] For the case of Kenya, see e.g. Throup and Hornsby (1998).

[16] The matching procedure is base on either (1) string matching, (2) a search on the Joshua Project's and Ethnologue's websites, (3) or, lastly, a Wikipedia search. Burkina Faso is dropped from the sample since it does not have politically relevant ethnic groups. Maps for which one or more SIDE groups can only be attributed to multiple EPR-ETH groups are disregarded since this keeping them would introduce bias in the measure of local ethno-political polarization. This leads to omitting maps from Ghana after 2002, from Cameroon after 2010 and from the Côte d'Ivoire after 1993.

[17] The respective population data are taken from CIESIN et al. (2011), providing estimated population counts for the years 1990, 1995, and 2000 at a spatial resolution of 8.3×10^{-3} degrees.

[18] $\text{Ethno-political polarization}_{dy} = 4 * \sum_{i \in I_{dy}} (size_i^2 * (1 - size_i))$, with $size_i$ being the size of ethnic group i relative to all politically relevant groups I populating a district d in a particular year y.

control variables in $\boldsymbol{X}_{dm}$.

The matrix $\boldsymbol{E}_{dcym}$ includes variations of spatio-temporally defined fixed effects. In particular, the main model includes fixed effects for district-years and the country-months. They serve four purposes. First, as the Arab Spring and common adjournments of electoral contests evidence, elections might be reversely caused or inhibited by violence preceding them. The country-month fixed effects effectively block this link by netting the data of all variation that is constant on the country-month level. Second, they account for omitted variables which are constant at this level and influence both, the timing of elections and the occurrence of riots. These covariates include all national-level socio-economic factors. Third, the use of time-varying data on the spatial extent of districts and the related danger of boundary changes that are endogenous to elections or riots presents the Modifiable Areal Unit Problem in its time-varying form. By using district-year fixed effects, the problem is alleviated insofar, as for each district-year only one stable areal unit is observed and local causes of past changes are controlled for. Fourth, the district-year fixed effects reduce the impact of locally varying spatial- and temporal auto-correlation. They account for the intermediate past of districts-years and their yearly environment composed of other districts observed in the data and thereby limit the bias any spatio-temporal auto-correlation can introduce.

Further spatio-temporal auto-correlation is taken into account by a set of control variables. To account for temporal auto-correlation, I follow Carter and Signorino (2010) and approximate the decay of riot-risk after a riot as a cubic polynomial, capturing the time to the last event in a district and the potential non-linearity of its influence. To model spatial auto-correlation, I add the number of riots in neighboring districts at time $t-1$, $t-2$, and $t-3$ as additional controls to all models.[19]

Standard-errors are clustered on the district-level. Using different levels of clustering such as the region or the country-year, as well as non-paramteric spatial and temporal error-clustering à la Bester et al. (2011) and Conley (1999) does not change the statistical interpretation of the results (Table A4).

Results

The hypothesis that ethno-politically polarized districts in majoritarian systems experience steeper increases in the number of riots prior to elections

[19]Since district areas are varying across years, all such controls are calculated on the basis of past riot events in a district's current area and that of its neighbors.

than their non-polarized counterparts is substantiated descriptively in Figure 3. The Figure shows that districts that are above-average polarized do not only see slightly higher numbers of riots during non-election times. They also experience a starker escalation of riots during electoral campaigns if compared to districts with a below-average degree of polarization. This pattern of pre-election violence is statistically tested in Table 1. The table summarizes the association of local ethno-political polarization in majoritarian and mixed electoral systems with pre-election increases in the number of riots at the district level. I iteratively introduce fixed effects on the country and district levels in Models 2 and 3 and finally combine country-month and district-year fixed effects for reasons outlined above. With the full set of fixed effects, the constitutive terms of the main interaction term time to election $\times$ ethno-political polarization cannot be identified because only the product of both variables varies within the clusters of the fixed effects.[20]

The results do not only indicate that ethno-politically polarized districts see more riots over the entire period[21] but crucially, that they see a markedly higher increase of rioting preceding elections than non-polarized districts. As evidenced by Figure 4, the substantive effect of ethno-political polarization on pre-election increases in the number of riots is large and precisely estimated.[22] While Model 1 indicates that non-polarized districts see an increase of the average number of riots by a factor of 2.7 over the year preceding a legislative election, the number of riots in polarized districts increases by a factor of 4.9.[23]

As a first indication of the robustness of the result, the difference in the local escalation of the number of riots prior to elections seen between polarized and non-polarized districts remains very stable once the country-month and district-year fixed effects are added to the model (Model 4, Table 1). They control for unobserved heterogeneity that might influence the timing of elections, spatio-temporal auto-correlation not captured by the respective controls, as well as endogenous changes of district borders. Since this model does only capture relative increases in the number of riots prior to elections as a function

[20] I.e. districts have a constant value of ethno-political polarization within a year, and all districts within the same country and month have the same distance to the next election. However, within district-years, there is variation in the product of the time to the next election and its degree of ethno-political polarization.

[21] See also Montalvo and Reynal-Querol (2005b) on ethnic polarization and civil conflict.

[22] Unless otherwise noted, all results reported below are associated with p-values below .05.

[23] All covariates other than time to election and ethno-political polarization are set to their sample mean.

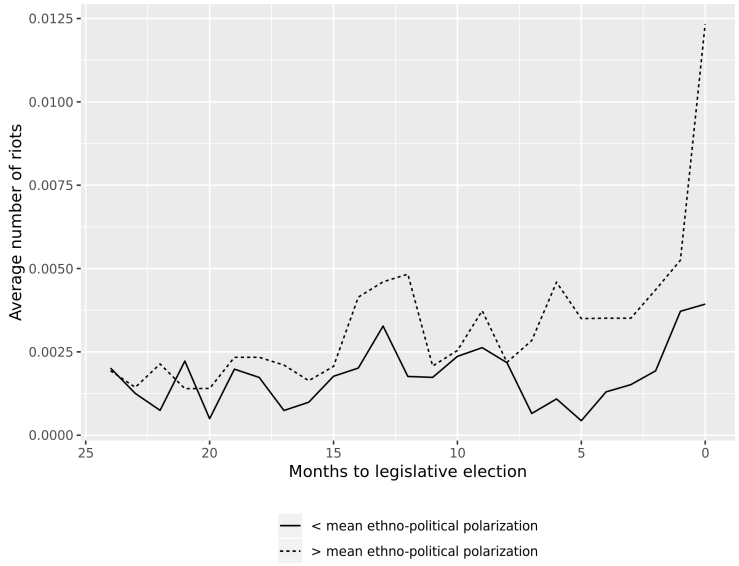


Figure 3: Mean number of riots in polarized and non-polarized districts over the 24 months prior to elections in majoritarian and mixed systems.

Table 1: Local ethnic polarization & pre-election violence in majoritarian and mixed systems

	Dependent variable:			
	Riots (SCAD)	Riots (SCAD)	Riots (SCAD)	Riots (SCAD)
	(1)	(2)	(3)	(4)
Constant	−0.0003			
	(0.0012)			
Time to election	−0.0386***	−0.0394***	−0.0375***	
	(0.0109)	(0.0111)	(0.0117)	
Ethno-pol. polarization	0.0008**	0.0011***	−0.0011	
	(0.0003)	(0.0004)	(0.0010)	
Time to elec. × Ethno-pol. polar.	0.0079***	0.0078***	0.0072***	0.0076***
	(0.0022)	(0.0022)	(0.0023)	(0.0029)
Population (log)	0.0007***	0.0010***	−0.0008**	
	(0.0001)	(0.0002)	(0.0003)	
Time to elec. × Population	0.0034***	0.0035***	0.0033***	0.0045***
	(0.0009)	(0.0009)	(0.0010)	(0.0015)
Sample:	Maj. & Mix.	Maj. & Mix.	Maj. & Mix.	Maj. & Mix.
Fixed effects:	–	country	district	district-year & country-month
Spatial $\text{lag}_{t-1,t-2,t-3}$:	yes	yes	yes	yes
Polynomial $\text{DV}^{1,2,3}$:	yes	yes	yes	yes
Mean DV:	0.0014	0.0014	0.0014	0.0014
Observations	434,303	434,303	434,303	434,303
R^2	0.0054	0.0066	0.0418	0.2311

Notes: OLS linear models. Standard errors clustered on the district-level in parentheses. Significance codes: *p<0.1; **p<0.05; ***p<0.01

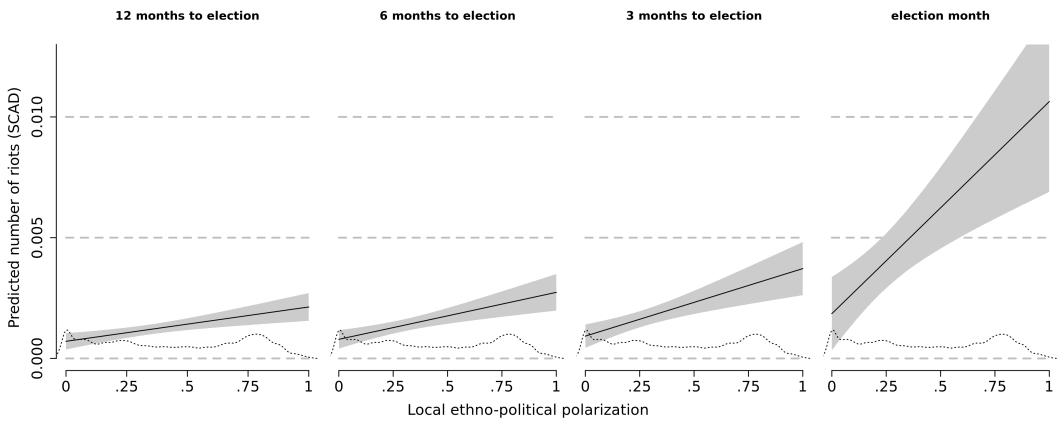


Figure 4: Prediction of the number of riots over the pre-election period in majoritarian polities, varying the degree of local ethno-political polarization. Based on Model 1 in Table 1. All covariates are held at their sample mean. The dotted line indicates distribution of ethno-political polarization in the sample.

of ethno-political polarization, it supplies strong evidence for Hypothesis 1.

So far, the results indicate that, in plurality systems, local ethno-political polarization heightens the risk of pre-election increases in the number of riots. However, as argued above, this finding can only be related to a theory of electoral competition in majoritarian system if no such effect is found in pure PR elections. The estimates for the electoral violence-inducing effect of local ethno-political polarization in PR systems coefficients reported in Table 2 are small, negative and statistically insignificant. No evidence is found for that local ethno-political polarization drives pre-election increases in rioting as documented in the SCAD and ACLED data. Using a dummy variable to distinguish between pure PR elections and those following majoritarian or a mix of majoritarian and PR rules, the results also show that there is a marked and statistically significant difference in the effects of local ethno-political polarization between the two ideal types of electoral systems (Figure 5). While the effects found in majoritarian/mixed systems are almost identical to those reported above, local ethno-political polarization has no substantive effect on pre-election violence in pure PR systems.

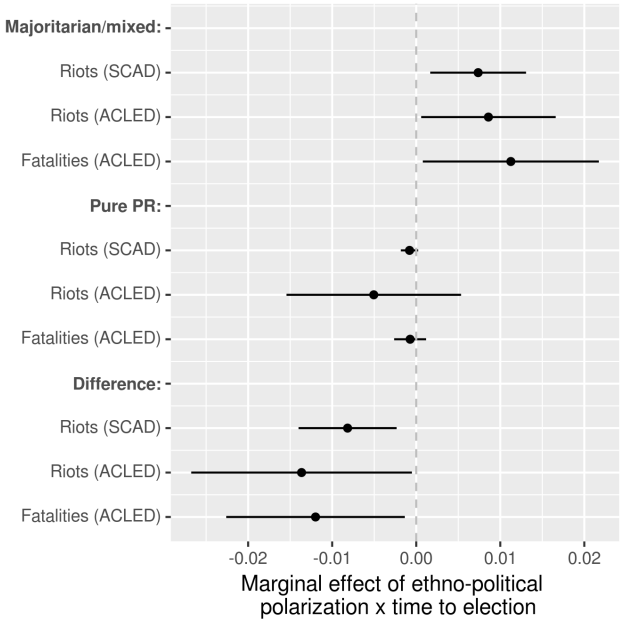


Figure 5: Comparison of the effect of ethno-political polarization $\times$ time to election on the monthly number of riots under majoritarian and PR voting, derived from Models 2,4, and 6 in Table 2. Bars indicate 95% CIs.

Table 2: Local ethnic polarization & various forms of pre-election violence

	Dependent variable:					
	Riots (SCAD)	Riots (SCAD)	Riots (ACLED)	Riots (ACLED)	Fatalities (ACLED)	Fatalities (ACLED)
	(1)	(2)	(3)	(4)	(5)	(6)
Time to elec. × Ethno-pol. polar.	−0.0008	0.0076***	−0.0050	0.0086**	−0.0008	0.0112**
	(0.0005)	(0.0029)	(0.0052)	(0.0041)	(0.0010)	(0.0053)
Time to elec. × Population	0.0004	0.0045***	0.0069	0.0081***	0.0004	0.0083***
	(0.0004)	(0.0015)	(0.0042)	(0.0026)	(0.0005)	(0.0030)
Time to elec. × Ethno-pol. polar. × PR		−0.0084***		−0.0137**		−0.0120**
		(0.0030)		(0.0067)		(0.0054)
Time to elec. × Population × PR		−0.0040**		−0.0011		−0.0080***
		(0.0016)		(0.0050)		(0.0030)
Sample:	PR	all	PR	all	PR	all
Spatial $\text{lag}_{t-1,t-2,t-3}$:	yes	yes	yes	yes	yes	yes
Polynomial $\text{DV}^{1,2,3}$:	yes	yes	yes	yes	yes	yes
District-year FE:	yes	yes	yes	yes	yes	yes
Country-month FE:	yes	yes	yes	yes	yes	yes
Mean DV:	3e-04	0.0012	9e-04	0.0024	2e-04	0.001
Observations	108,381	542,684	81,351	394,360	81,351	394,360
R^2	0.2567	0.2323	0.1582	0.2521	0.1185	0.1670

Notes: OLS linear models. Standard errors clustered on the district-level in parentheses. Significance codes:
*p<0.1; **p<0.05; ***p<0.01

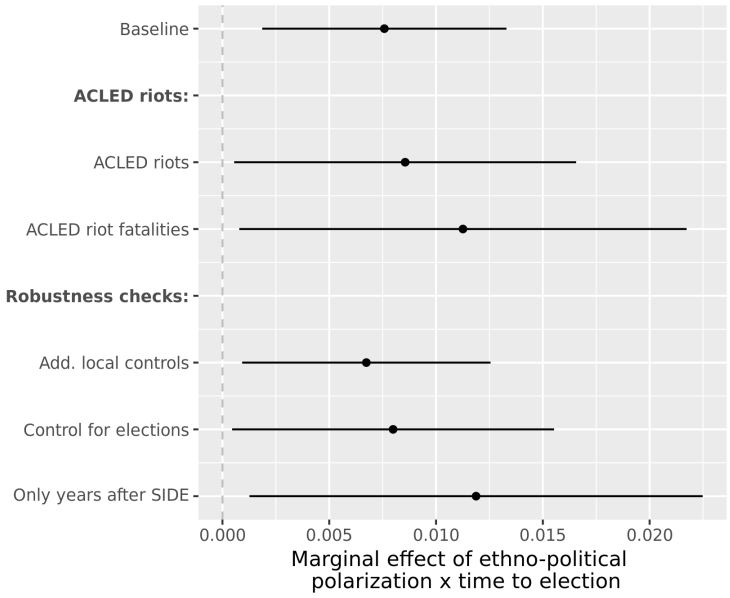


Figure 6: Robustness checks on Model 4 in Table 1 with 95% CIs.

Robustness checks:

In the following, I summarize the results of a number of robustness checks on the models reported above using the sample from majoritarian and mixed systems only. The same tests are conducted on the models that compare majoritarian with PR systems. These lead to very similar results and are reported in Table A3.

Given that the SCAD data are based on media reports in international media which might not report on every riot in the hinterlands of every African country, I first replicate the above presented results using data on riots and riot fatalities retrieved from ACLED (1997–2013; Raleigh et al., 2010) which also draws on national media reports. The marginal effect of ethno-political polarization on the pre-election increase of the number of local riots remains significant and is slightly greater in size than at the baseline. Since ACLED codes riot-days rather than riot episodes as SCAD, this suggests that riots occur not only more often as elections come closer, but also that they might last longer. Such increased intensity is also indicated by the increase of riot fatalities before elections in ethno-politically polarized districts, which is observed less strongly in non-polarized districts.

Second, I add additional control variables to (1) reduce the risk of omitted variable bias and (2) increase the confidence that the results reported above stem from ethno-political competition rather than its most proximate corre-

lates. All variables introduced in this paragraph are interacted with the proximity of the next legislative election time to election. With regard to the first concern, the size of districts' population might be an imperfect control for the ethnic diversity which might be driven by urban populations. These might be at the same time at greatest risk of experiencing pre-election violence. I therefore add a measure of the logged size of the urban population in a district to the model.[24] With regard to the concern that it is not ethno-political competition, but rather another correlate of ethno-political polarization that is driving the reported results, the first candidate for such a potential omitted variable is pure *ethnic* polarization. In response, I add the degree of ethnic polarization as calculated directly from the SIDE maps to the model. Additionally, ethnic polarization might increase local levels of poverty, which in turn heightens the odds of campaign violence.[25] To block this potential path, I add the logged per-capita night-lights emissions (National Geophysical Data Center, 2014) of a district-year as a proxy for economic activity (Chen and Nordhaus, 2011). Of these three additional variables only the proportion of the urban population has a substantive (and positive) effect on the degree to which the number of riots changes as elections approach. Bolstering the confidence in the results previously reported, the estimate of the impact of ethno-political polarization remains almost unchanged.

Third, the results reported above might be driven by observations of election months. For these months I cannot – in this aggregated setup – distinguish pre- from post-election riots. I therefore add an interaction of ethno-political polarization with a dummy for election months to the model, leading to an almost unchanged point estimate, which is indistinguishable from the baseline. It is thus pre-election violence that drives this finding.

Fourth, the data used up to now include many district-months for which the data used for computing the SIDE data has been collected after they have been observed. District years from the sample in which no past or contemporaneous map is available make up about $^2/_3$ of the sample. Dropping them to avoid potential reverse causality slightly increases the estimated effect of ethno-political polarization on pre-election rioting. However, if reverse causal-

[24] Geo-coded urban population counts are available in the GRUMP data (CIESIN et al., 2011)

[25] The literature arguing that poverty increases the odds of violent conflict is large, cf. Collier and Hoeffler (2004); Fearon and Laitin (2003). However, although the link between economic development and country-level ethnic diversity is well established (Alesina et al., 1999; Montalvo and Reynal-Querol, 2005a), local level evidence points in the opposite direction (Gerring et al., 2015).

ity would strongly affect the results reported above one would expect the point estimate to drop towards zero.

I conduct three additional robustness checks. First, Table A5 addresses the challenge of disentangling the effects of upcoming presidential and legislative elections held at the same time. The results suggests slightly greater but statistically indistinguishable effects of ethno-political polarization before general legislative and presidential than before pure legislative elections. Second, I address the question whether the results above capture *electoral* violence. Using the ECAV data on local electoral violence as an outcome in a cross-sectional setting reveals similar effects than reported above (Table A6). In comparison with the SCAD or ACLED data, these results come however with the caveat that we cannot identify increases in the level of violence before elections. Furthermore, the data do not include violence that is driven by upcoming elections but not designated as electoral violence in the media. Third and lastly, reestimating the main specification using SCAD's count of demonstrations, strikes, and violent attacks by militias as dependent variables supports the argument that riots can have unique and tangible electoral benefits for its instigators. The other forms of violence show no pre-election increases associated with local ethno-political polarization (Table A7). Assuming that media bias would affect reporting on all event types, these non-results also suggest that biased reporting does not drive the results.

The fear of pre-election victimization and local ethno-political polarization

As shown above, local ethno-political competition is strongly associated with district-level escalations of rioting prior to legislative elections in majoritarian and mixed but not PR systems. Moving beyond event counts, the following section sheds light on the individual-level incidence of pre-election violence in ethno-politically polarized districts.

Data and empirical strategy

The 4^{th}, 5^{th}, and 6^{th} round of the Afrobarometer (Afrobarometer, 2018) inquired individuals about their fear of falling victim to political intimidation

or violence during electoral campaigns.[26] Additionally, the Afrobarometer has conducted extracurricular surveys in Nigeria (2007) and Uganda (2010/2011). Here, citizens have been asked in pre-election surveys whether they themselves or members of their community have been subject to recent campaign threats relating, inter alia, to their physical well-being. Following the data, the level of analysis thus moves to the individual respondent included in Afrobarometer surveys taken in the 19 countries with data on local ethno-political polarization (see Figure A1 in the Appendix).

I match the above used district-level measure of ethno-political polarization to Afrobarometer respondents, using the recently published geocodes of Afrobarometer survey clusters for rounds 4 to 6 (Ben Yishay, Ariel Rotberg et al., 2017).[27] Each respondent is matched to the polarization measure of its districts from the year prior to the survey. Furthermore, respondents that were interviewed after 2013 are matched to the polarization measure from 2013.[28] To ease the interpretation of the results, the main analyses are conducted using simple OLS regressions:[29]

$$y_{idcr} = \delta_{cr} + \beta_1 \text{ethno-political polarization}_d + \delta \boldsymbol{X}_{id} + \epsilon_{id},$$

where outcomes y of an individual i who lives in district d of country c and was interviewed in round r is regressed on the district's level of ethno-political polarization and a vector of individual and district level covariates.[30] Through the use of country-round fixed effects δ_{cr}, the estimations only compare respondents interviewed in the same survey round and country with each other. In addition, I include a vector of control variables $\boldsymbol{X}_{id}$ which includes the size of district's population, as well as respondents' sex, their age and its square, their level of education as well as a dummy for urban respondents. Standard

[26] The respective question reads 'During election campaigns in this country, how much do you personally fear becoming a victim of political intimidation or violence?' Ordinal answers range from 0 (not at all) to 3 (a lot).

[27] All observations which could be matched to only to the regional level or above are dropped from the sample. For the additional rounds from Uganda and Nigeria, I use the surveys' identifiers of districts and automatically geocode the approximate survey location. See Technical Appendix.

[28] Dropping these observations does not change the results. See Table A13 in the Appendix.

[29] All results hold when estimating (ordered) logistic regressions with country-round dummies. See Tables A14 to A15 in the Appendix.

[30] On the individual-level, the models include a respondents sex, her age and age squared, a dummy for urban residents, and her level of education. In parallel to the previous analysis, I include the logged population count on the district level.

errors are clustered on the district-level.[31]

Results

The results of the main analysis of the fear of Afrobarometer respondents to be victimized in electoral campaigns are presented in Table 3 and corroborate the insights gained in the previous section. In majoritarian systems, the fear of respondents is positively and significantly associated with the ethno-political polarization of their home district (Model 1). The effect is of substantive size, amounting to a change in the reported level of fear by 0.2 on a scale from 0 (no fear at all) to 3 (a lot; mean=1.1) as one moves the polarization measure from 0 to 1. As above, the results indicate that in pure proportional systems this effect of local ethno-political polarization does not exist (Model 2). The difference between pure PR and majoritarian systems is however not statistically significant, as evidenced by Model 3. This stems from larger standard errors estimated in the comparatively small PR sample.

Table 3: Local ethnic polarization & fear of pre-election victimization

	Dependent variable:		
	Fear	Fear	Fear
	(1)	(2)	(3)
Ethno-pol. polarization	0.199***	0.107	0.199***
	(0.056)	(0.078)	(0.056)
Ethno-pol. polar. × PR			−0.092
			(0.096)
Sample:	Maj. & Mix.	PR	all
Covariates:	yes	yes	yes
Country-round FE:	yes	yes	yes
Mean DV:	1.1478	0.7189	1.0545
Observations	40,945	11,382	52,327
R^2	0.103	0.054	0.116

Notes: OLS linear models. Control variables include the district population (logged), and urban and female dummy, age and its square, as well as the respondent's education. Model 3 also includes interactions of all covariates with the PR dummy. Standard errors clustered on the district-level in parentheses. Significance codes: *p<0.1; **p<0.05; ***p<0.01

Robustness Checks

For the robustness checks presented in the following, I use Model 1 in Table 3 as the baseline, limiting the analysis to respondents from majoritarian polities

[31] Clustering standard errors on the region- or country-round-level, as well as employing Conley's clustering of errors (Bester et al., 2011; Conley, 1999) does not change the statistical interpretation of the main results reported below. See Table A12 in the Appendix.

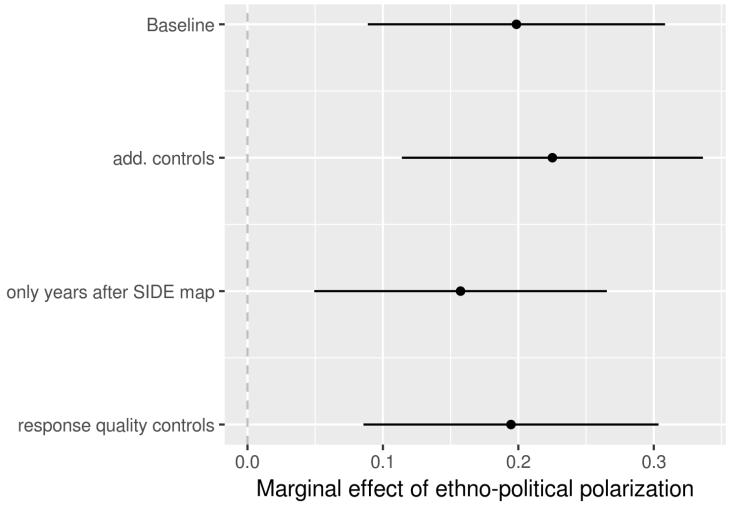


Figure 7: Robustness checks on Model 1 in Table 3.

only. All robustness checks are summarized by Figure 7 and reported in Table A11 in the Appendix. As a first sensitivity analysis, I follow the approach used above and also control for the degree of pure ethnic polarization, the size of the logged urban population, as well as the level of local development proxied by a districts' logged nightlight emissions per capita in a given survey-year. The inclusion of these measures does not alter the results by much. Only the level of pure ethnic polarization has a weakly significant and negative association with the reported fear of pre-election violence ($p < .1$).

In the next step, I again drop observations from the survey data that have been made before the DHS data underlying the SIDE maps have been collected.[32] Bolstering the confidence that the reported results are not caused by reverse causality, the baseline results hold. While the point estimate slightly drops in size to 0.16, the effect remains significant and statistically indistinguishable from the baseline.

Finally, the results might be affected by biased answers of respondents, effected by the presence of bystanders, co-ethnicity with the interviewer (Adida et al., 2016), or mistrust in the government. I therefore include related survey-items in the model.[33] The results indicate that these factors do not bias the baseline results.

[32]This leads to dropping four surveys from Zambia, Nigeria and Uganda from the sample.

[33]I specifically include five variables, a dummy for linguistic co-ethnicity of respondents with the interviewer, three dummy for whether other persons were present, and checking or influencing respondents' answers, as well as a categorical coding of answers to the question whom respondents' think the interviewer was working for.

The information on the fear of respondents to become victimized in electoral campaigns is useful to uncover differences in the dynamics of pre-election violence in a larger number of majoritarian polities, as well as between majoritarian and proportional systems. However, the measure is based on subjective *perceptions* of fear, and might therefore only be an imperfect proxy of the *experience* of pre-election intimidation and violence. I therefore turn towards data that are available for two countries only, Nigeria and Uganda. The Afrobarometer surveys conducted in 2007 Nigeria and 2010/2011 Uganda happened closely before the countries' general elections held under majoritarian rules. Relating to the ongoing electoral campaigns, respondents were asked how often they had been targeted by threats and how often they had heard of any threats being issued at members of their communities. Furthermore, respondents who indicated that they were targeted by or had heard of intimidation, were inquired about the nature of those threats, in particular if they were related to their personal or their families' safety. While the first two variables are ordered on a scale from 0 (never threatened) to 3 (often threatened), the second two are dummies that indicate whether respondents and their communities have been violently intimidated or not. Arguably, the experience of threats, especially those related to physical integrity, is a much more viable proxy for electoral violence than the mere fear of it. I therefore estimate the effect of local ethno-political polarization on these four outcomes.

The results are reported in Table 4. They indicate that respondents who live in ethno-politically polarized districts in Nigeria 2007 and Uganda 2010/2011 report that they and members of their community have been targeted significantly more often by campaign threats (0.11 and 0.10, on a 0 (never) to 3 (often) scale; means=0.13 and 0.17) than those living in non-polarized areas. The effect is even stronger if the experience of security-related threats is analyzed. Here, respondents living in fully polarized areas are on average 5.1 percentage points more likely to have received such a threat. Unsurprisingly, a very similar effect is found with regard to reports about security-related threats being issued at members of their communities. The size of these effects is large, given that only 4.9% (6.1%) of all respondents report that they (members of their community) have been targeted by such threats.

Table 4: Local ethnic polarization & pre-election threats: Nigeria 2007 and Uganda 2010/2011

	Dependent variable:			
	Personal threat: frequency	Personal threat: security	Community threat: frequency	Community threat security
	(1)	(2)	(3)	(4)
Ethno-pol. polarization	0.105**	0.051***	0.103**	0.054**
	(0.047)	(0.018)	(0.051)	(0.021)
Sample:	NIG & UGA	NIG & UGA	NIG & UGA	NIG & UGA
Controls	yes	yes	yes	yes
Country-round FE:	yes	yes	yes	yes
Mean DV:	0.1329	0.0488	0.1697	0.0606
Observations	6,071	6,071	5,745	5,745
R^2	0.011	0.017	0.012	0.020

Notes: OLS linear models. Control variables include the district population (logged), and urban and female dummy, age and its square, as well as the respondent's education. Standard errors clustered on the district-level in parentheses. Significance codes: *p<0.1; **p<0.05; ***p<0.01

Conclusion

Local political competition between ethnic groups can increase the odds of pre-election violence in majoritarian elections in Africa. By focusing on the nexus between local ethno-political cleavages, the electoral system, and campaign violence, the preceding analysis highlights the importance of socio-political geographies for gauging the merits and effects of majoritarian systems. Echoing arguments made by Barkan et al. (2006) and Wagner and Dreef (2013), I show that – at least pertaining to the question of how peacefully legislative campaigns are conducted – there is likely no 'one-size-fits-all' electoral system to be preferred. Rather, constitutional engineers aiming for peaceful elections are well advised to take the impact of ethnic geographies into account when drafting electoral institutions.

While majoritarian contests are more peaceful in ethnically homogeneous constituencies, areas settled by politically mobilized ethnic groups in a polarized manner are at risk of suffering from markedly increased rioting as elections approach. Similarly, citizens who live in polarized districts in mixed and majoritarian polities systematically report substantially higher levels of fear of pre-election violence than their co-nationals living in non-polarized districts. These patterns of more severe pre-election violence in ethno-politically polarized districts under mixed or majoritarian voting do not perturbate elections in pure PR systems.

This evidence suggests a promising research agenda in comparative politics. Since more and more geographically disaggregated data become available, the analysis of local factors such as ethnic geographies and their interaction with

macro-level electoral systems may increase our understanding about the effects of democratic institutions under varying conditions. Ultimately, such knowledge can inform future constitutional designs and those who want to prevent divisive and violence elections.

References

Adida, C. L. (2012). Do African Voters Favor Co-Ethnics? A new identification strategy. *Working Paper University of California San Diego. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2054853.*

Adida, C. L., Ferree, K. E., Posner, D. N., and Robinson, A. L. (2016). Who's Asking? Interviewer Coethnicity Effects in African Survey Data. *Comparative Political Studies*, 49(12):1630–1660.

Afrobarometer (2018). Afrobarometer Data. *Available at http://www.afrobarometer.org.*

Alesina, A., Baqir, R., and Easterly, W. (1999). Public goods and ethnic divisions. *Quarterly Journal of Economics*, 1999(November):1243–1284.

Barkan, J. D., Densham, P. J., and Rushton, G. (2006). Space Matters: Designing Better Electoral Systems for Emerging Democracies. *American Journal of Political Science*, 50(4):926–939.

Basedau, M., Erdmann, G., Lay, J., and Stroh, A. (2011). Ethnicity and party preference in sub-Saharan Africa. *Democratization*, 18(2):462–489.

Bech Seeberg, M., Wahman, M., and Skaaning, S. E. (2018). Candidate nomination, intra-party democracy, and election violence in Africa. *Democratization*, 25(6):959–977.

Beck, T., Clarke, G., Groff, A., Keefer, P., and Walsh, P. (2001). New tools in comparative political economy: The Database of Political Institutions. *The World Bank Economic Review*, 15(1):165–176.

Ben Yishay, Ariel Rotberg, R., Wells, J., Lv, Z., Goodman, S., Kovacevic, L., and Runfola, D. (2017). *Geocoding Afrobarometer Rounds 1-6: Methodology and Data Quality.* AidData, Williamsburg, VA.

Bester, C. A., Conley, T. G., and Hansen, C. B. (2011). Inference with dependent data using cluster covariance estimators. *Journal of Econometrics*, 165(2):137–151.

Birch, S. (2007). Electoral Systems and Electoral Misconduct. *Comparative Political Studies*, 40(12):1533–1556.

Bogaards, M. (2013). The Choice for Proportional Representation: Electoral System Design in Peace Agreements. *Civil Wars*, 15(sup1):71–87.

Brass, P. R. (2011). *The production of Hindu-Muslim violence in contemporary India*. University of Washington Press, Washington.

Bratton, M. (2008). Vote buying and violence in Nigerian election campaigns. *Electoral Studies*, 27(4):621–632.

Bratton, M., Bhavnani, R., and Chen, T.-H. (2012). Voting intentions in Africa: ethnic, economic or partisan? *Commonwealth & Comparative Politics*, 50(1):27–52.

Bratton, M. and Kimenyi, M. S. (2008). Voting in Kenya: Putting Ethnicity in Perspective. *Journal of Eastern African Studies*, 2(2):272–289.

Calvo, E. and Rodden, J. (2015). The Achilles Heel of Plurality Systems: Geography and Representation in Multiparty Democracies. *American Journal of Political Science*, 59(4):789–805.

Carey, J. M. and Shugart, M. S. (1995). Incentives to cultivate a personal vote: A rank ordering of electoral formulas. *Electoral Studies*, 14(4):417–439.

Carter, D. B. and Signorino, C. S. (2010). Back to the Future: Modeling Time Dependence in Binary Data. *Political Analysis*, 18(3):271–292.

Cederman, L.-E., Wimmer, A., and Min, B. (2010). Why do ethnic groups rebel? New data and analysis. *World Politics*, 62(1):87–119.

Chandra, K. (2004). *Why Ethnic Parties Succeed*. Cambridge University Press, Cambridge.

Chen, X. and Nordhaus, W. D. (2011). Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences*, 108(21):8589–8594.

CIESIN, IFPRI, World Bank, and CIAT (2011). Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Population Density Grid. *Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). Accessed from: http://sedac.ciesin.columbia.edu/data/collection/grump-v1.*

Collier, P. and Hoeffler, A. (2004). Greed and grievance in civil war. *Oxford Economic Papers*, 56(4):563–595.

Collier, P. and Vicente, P. C. (2011). Violence, bribery, and fraud: the political economy of elections in Sub-Saharan Africa. *Public Choice*, 153(1-2):117–147.

Collier, P. and Vicente, P. C. (2014). Votes and Violence: Evidence from a Field Experiment in Nigeria. *The Economic Journal*, 124(574):327–355.

Conley, T. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, 92(1):1–45.

Cox, G. W. and Munger, M. C. (1989). Closeness, expenditures, and turnout in the 1982 US House elections. *The American Political Science Review*, 83(1):217–231.

Daxecker, U., Amicarelli, E., and Jung, A. (2018). Electoral Contention and Violence (ECAV): A New Dataset. *Journal of Peace Research, conditionally accepted*.

Dercon, S. and Gutiérrez-Romero, R. (2012). Triggers and Characteristics of the 2007 Kenyan Electoral Violence. *World Development*, 40(4):731–744.

DHS (2015). Demographic and Health Surveys. *Integrated Demographic and Health Series (IDHS), version 2.0, Minnesota Population Center and ICF International. Accessed from http://idhsdata.org on 2015/07/28*.

Eifert, B., Miguel, E., and Posner, D. N. (2010). Political competition and ethnic identification in Africa. *American Journal of Political Science*, 54(89):494–510.

Erikson, R. S. and Palfrey, T. R. (2000). Equilibria in campaign spending games: Theory and data. *American Political Science Review*, 94(3):595–609.

Esteban, J.-M. and Ray, D. (1994). On the Measurement of Polarization. *Econometrica*, 62(4):819–851.

FAO (2014). Global Administrative Unit Layers. *Available from: http://data.fao.org/map?entryId=f7e7adb0-88fd-11da-a88f-000d939bc5d8&tab=metadata*.

Fearon, J. D. (1999). Why ethnic politics and 'pork' tend to go together. *Unpublished Working Paper. Available from:*

https://web.stanford.edu/group/fearon-research/cgi-bin/wordpress/wp-content/uploads/2013/10/Pork.pdf.

Fearon, J. D. and Laitin, D. D. (2003). Ethnicity, insurgency, and civil war. *American Political Science Review*, 97(1):75–90.

Fjelde, H. and Höglund, K. (2016). Electoral Institutions and Electoral Violence in Sub-Saharan Africa. *British Journal of Political Science*, 46(2):297–320.

Gerring, J., Thacker, S. C., Lu, Y., and Huang, W. (2015). Does Diversity Impair Human Development? A Multi-Level Test of the Diversity Debit Hypothesis. *World Development*, 66:166–188.

Goldsmith, A. A. (2015). Elections and civil violence in new multiparty regimes: Evidence from Africa. *Journal of Peace Research*, 52(5):607–621.

Gudgin, G. and Taylor, P. (2012). *Seats, votes, and the spatial organisation of elections*. ECPR Press, Colchester.

Hafner-Burton, E. M., Hyde, S. D., and Jablonski, R. S. (2013). When Do Governments Resort to Election Violence? *British Journal of Political Science*, 44(01):149–179.

Hegre, H. and Sambanis, N. (2006). Sensitivity analysis of empirical results on civil war onset. *Journal of conflict resolution*, 50(4):508–535.

Horowitz, D. L. (1985). *Ethnic groups in conflict*. University of California Press, Berkeley.

Horowitz, D. L. (1990). Comparing Democratic Systems. *Journal of Democracy*, 1(4):73–79.

Horowitz, D. L. (1991). *A Democratic South Africa? Constitutional Engineering in a Divided Society*. University of California Press, Berkeley.

Horowitz, D. L. (1994). Democracy in Divided Societies. In Diamond, L. and Plattner, M. F., editors, *Nationalism, Ethnic Conflict, and Democracy*, pages 35–55. John Hopkins University Press, London.

Horowitz, D. L. (2001). *The deadly ethnic riot*. University of California Press, Berkeley.

Huber, J. D. (2012). Measuring Ethnic Voting: Do Proportional Electoral Laws Politicize Ethnicity? *American Journal of Political Science*, 56(4):986–1001.

Hyde, S. D. and Marinov, N. (2011). Which Elections Can Be Lost? *Political Analysis*, 20(2):191–210.

Klopp, J. M. (2001). "Ethnic Clashes" and Winning Elections: The Case of Kenya's Electoral Despotism. *Canadian Journal of African Studies*, 35(3):473–517.

Klopp, J. M. and Zuern, E. (2007). The politics of violence in democratization: Lessons from Kenya and South Africa. *Comparative Politics*, 39(2):127–146.

Lijphart, A. (1985). *Power-Sharing in South Africa*. University of California, Institute of International Studies, Berkeley.

Lijphart, A. and Aitkin, D. (1994). *Electoral systems and party systems: A study of twenty-seven democracies*. Oxford University Press, Oxford.

Montalvo, J. G. and Reynal-Querol, M. (2005a). Ethnic diversity and economic development. *Journal of Development Economics*, 76(2):293–323.

Montalvo, J. G. and Reynal-Querol, M. (2005b). Ethnic polarization, potential conflict, and civil wars. *American Economic Review*, 95(3):796–816.

Mozaffar, S., Scarritt, J. R., and Galaich, G. (2003). Electoral Institutions, Ethnopolitical Cleavages, and Party Systems in Africa's Emerging Democracies. *American Political Science Review*, 97(03):379–390.

Müller-Crepon, C. and Hunziker, P. (2018). New Spatial Data on Ethnicity: Introducing SIDE. *Journal of Peace Research, forthcoming*.

National Geophysical Data Center (2014). DMSP-OLS Nighttime Lights Time Series, Version 4. *Electronic resource. Available at http://ngdc.noaa.gov/eog/dmsp/ downloadV4composites.html.*

Neto, O. and Cox, G. (1997). Electoral institutions, cleavage structures, and the number of parties. *American Journal of Political Science*, 41(1):149–174.

Ordeshook, P. and Shvetsova, O. (1994). Ethnic heterogeneity, district magnitude, and the number of parties. *American Journal of Political Science*, 38(1):100–123.

Pattie, C. J., Johnston, R. J., and Fieldhouse, E. (1995). Winning the Local Vote: The Effectiveness of Constituency Campaign Spending in Great Britain, 1983-1992. *The American Political Science Review*, 89(4):969–983.

Raleigh, C., Linke, A., Hegre, H., and Karlsen, J. (2010). Introducing ACLED: An Armed Conflict Location and Event Dataset. *Journal of Peace Research*, 47(5):651–660.

Reynal-Querol, M. (2002). Ethnicity, Political Systems, and Civil Wars. *Journal of Conflict Resolution*, 46(1):29–54.

Reynal-Querol, M. (2005). Does democracy preempt civil wars? *European Journal of Political Economy*, 21(2):445–465.

Rodden, J. (2010). The Geographic Distribution of Political Preferences. *Annual Review of Political Science*, 13(1):321–340.

Saideman, S. M., Lanoue, D. J., Campenni, M., and Stanton, S. (2002). Democratization, Political Institutions, and Ethnic Conflict. A Pooled Time-Series Analysis, 1985-1998. *Comparative Political Studies*, 35(1):103–129.

Salehyan, I., Hendrix, C. S., Hamner, J., Case, C., Linebarger, C., Stull, E., and Williams, J. (2012). Social Conflict in Africa: A New Database. *International Interactions*, 38(4):503–511.

Salehyan, I. and Linebarger, C. (2015). Elections and Social Conflict in Africa, 1990-2009. *Studies in Comparative International Development*, 50(1):23–49.

Sartori, G. (1997). *Comparative Constitutional Engineering. An Inquiry into Structures, Incenntives and Outcomes.* New York University Press, New York, 2 edition.

Schneider, G. and Wiesehomeier, N. (2008). Rules That Matter: Political Institutions and the Diversity–Conflict Nexus. *Journal of Peace Research*, 45(2):183–203.

Steele, A. (2011). Electing Displacement: Political Cleansing in Apartado, Colombia. *Journal of Conflict Resolution*, 55(3):423–445.

Throup, D. and Hornsby, C. (1998). *Multi-party politics in Kenya: The Kenyatta & Moi States & the triumph of the system in the 1992 election.* James Currey, Oxford.

Vogt, M., Bormann, N.-C., Rüegger, S., Cederman, L.-E., Hunziker, P. M., and Girardin, L. (2015). Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Dataset Family. *Journal of Conflict Resolution, forthcoming*, 59(7):1327–1342.

Wagner, W. and Dreef, S. (2013). Ethnic Composition and Electoral System Design: Demographic Context Conditions for Post-conflict Elections. *Ethnopolitics*, 13(3):288–307.

Weidmann, N. B., Rød, J. K., and Cederman, L.-E. (2010). Representing ethnic groups in space: A new dataset. *Journal of Peace Research*, 47(4):491–499.

Wilkinson, S. I. (2004). *Votes and violence*. Cambridge University Press, Cambridge.

Wucherpfennig, J., Weidmann, N. B., Girardin, L., Cederman, L.-E., and Wimmer, A. (2011). Politically relevant ethnic groups across space and time: Introducing the GeoEPR dataset. *Conflict Management and Peace Science*, 28(5):423–437.

Sowing Hatred: Local Ethno-Political Competition and Pre-Election Violence in Majoritarian Elections

ONLINE APPENDIX

A1 Data overview

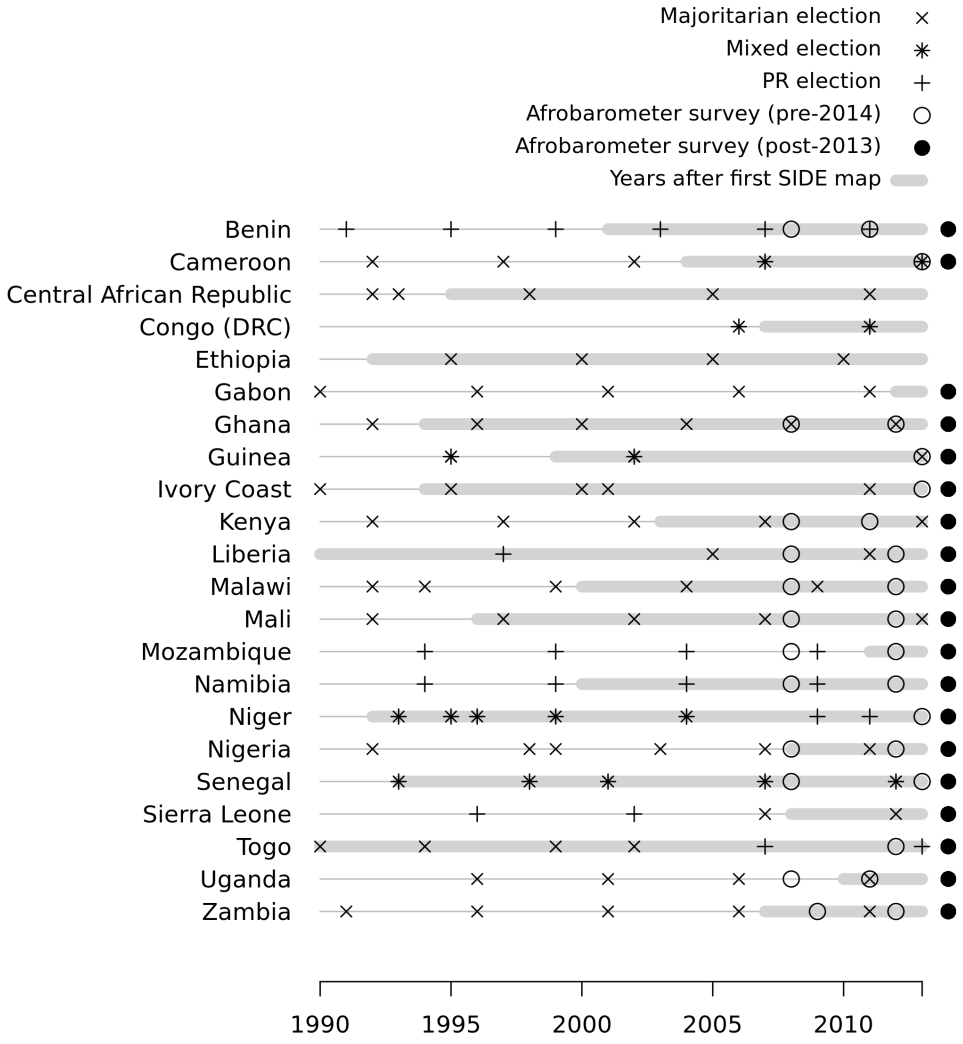


Figure A1: Summary of the samples used in the empirical analyses.

A2 District-month analysis

A2.1 Summary statistics

Table A1: Summary statistics: Districts 1990–2013

Statistic	N	Mean	St. Dev.	Min	Max
Time to election	547,577	0.092	0.163	0.016	1.000
Ethno-pol. polarization	547,577	0.462	0.295	0.000	1.000
Population (log)	547,577	11.468	1.270	3.086	15.452
Ethnic polarization	547,577	0.567	0.184	0.000	0.954
Urban population (log)	547,577	5.610	5.187	0.000	15.440
Nightlights per capita	503,532	0.013	0.146	0.000	8.542
Riots (SCAD; log)	547,577	0.001	0.030	0.000	1.946
Riots (ACLED; log)	394,360	0.002	0.047	0.000	2.639
Fatalities (ACLED; log)	394,360	0.001	0.046	0.000	6.909
Viol. events (ECAV; log)	68,551	0.004	0.067	0.000	2.996
Fatalities (ECAV; log)	68,551	0.002	0.040	0.000	2.079
Majoritarian & mixed	547,577	0.801	0.399	0	1
PR (pure)	547,577	0.199	0.399	0	1

A2.2 Robustness checks

This section presents all robustness checks of the riot analysis which are discussed and/or briefly mentioned in the main paper. Each table is accompanied by a short discussion of the sensitivity analyses and results it contains.

Main robustness checks: Table A2 presents the results of the main robustness checks conducted for the sample of majoritarian/mixed electoral systems only. It presents the full results behind Figure 6 in the main paper. Models 1 and 2 show how ethno-politically polarized and populous districts experience a comparatively steep increase in rioting – measured with event data from ACLED in Model 1 and casualties from these events in Model 2. Model 3 returns to using the SCAD data on riots for constructing the dependent variable and adds interactions of the time to the next election with (1) pure ethnic polarization, (2) nightlights per capita, and (3) the size of the local urban population. This is to control for additional omitted variables and mechanisms of how ethno-political polarization might lead to higher levels of pre-election violence that are not connected to electoral competition. While urban districts indeed see higher levels of pre-election violence, the addition of these interactions does not affect the main coefficient of interest. Model 4 adds an interaction of an election-month dummy with the level of ethno–political polarization to gauge whether the effects of the time to election are driven by

election months for which one cannot distinguish pre- from post-election violence. This is not the case. Lastly, Model 5 drops all observations for which the measure of ethno-political polarization is based on SIDE maps which are constructed on the basis of survey data from 'the future', that is on the basis of information collected after the observation of a district month (see Figure A1). Dropping these observations (about 60% of the entire sample) does if at all *increase* the estimated effect of ethno-political polarization on pre-election violence. This suggests that the findings are not driven by reverse causality.

Table A2: Robustness Checks: Local ethnic polarization & pre-election violence in maj. & mix. systems

	Dependent variable:				
	Riots (ACLED)	Fatalities (ACLED)	Riots (SCAD)	Riots (SCAD)	Riots (SCAD)
	(1)	(2)	(3)	(4)	(5)
Time to elec. × Ethno-pol. polar.	0.0086**	0.0113**	0.0067**	0.0080**	0.0119**
	(0.0041)	(0.0053)	(0.0030)	(0.0038)	(0.0054)
Time to elec. × Population	0.0081***	0.0084***	0.0046***	0.0045***	0.0075***
	(0.0026)	(0.0029)	(0.0016)	(0.0015)	(0.0026)
Time to elec. × Eth. polarization			−0.0009		
			(0.0050)		
Time to elec. × Nightlights p.c. (log)			0.0018**		
			(0.0008)		
Time to elec. × Urban pop. (log)			0.0002		
			(0.0001)		
Election dummy × Ethno-pol. polar.				−0.0004	
				(0.0048)	
Check:	ACLED riots	ACLED deaths	geo controls	elec months	$t \geq t_{SIDE}$
Sample:	Maj. & Mix.	Maj. & Mix.	Maj. & Mix.	Maj. & Mix.	Maj. & Mix.
Spatial $\text{lag}_{t-1,t-2,t-3}$:	yes	yes	yes	yes	yes
Polynomial $\text{DV}^{1,2,3}$:	yes	yes	yes	yes	yes
District-year FE:	yes	yes	yes	yes	yes
Country-month FE:	yes	yes	yes	yes	yes
Mean DV:	0.0027	0.0012	0.0015	0.0014	0.0015
Observations	313,009	313,009	401,537	434,303	175,298
R^2	0.2592	0.1685	0.2329	0.2311	0.2677

Notes: OLS linear models. Standard errors clustered on the district-level in parentheses. Significance codes: *p<0.1; **p<0.05; ***p<0.01

Table A3 conducts the same robustness checks as presented in Table A2 for the direct comparison between the effect of ethno-political polarization on pre-election violence in majoritarian and PR elections. Note that the comparisons based on ACLED riot events and fatalities are presented in Table 2 in the main paper. All models prove to be robust to the changes made. The only exception consists in the rather imprecisely estimated difference between majoritarian and PR systems ($p < .1$) once I drop all observations with 'future' SIDE map

information only in Model 3. This is likely due to the drastically reduced sample size. However, the absolute size of the difference between the effect of ethno-political polarization in majoritarian and PR systems remains stable.

Table A3: Robustness Checks: Local ethnic polarization & pre-election violence

	Dependent variable:		
	Riots (SCAD)	Riots (SCAD)	Riots (SCAD)
	(1)	(2)	(3)
Time to elec. × Ethno-pol. polar.	0.0083***	0.0080**	0.0118**
	(0.0031)	(0.0038)	(0.0054)
Time to elec. × Population	0.0044***	0.0045***	0.0075***
	(0.0016)	(0.0015)	(0.0026)
Time to elec. × Ethno-pol. polar. × PR	−0.0090***	−0.0093**	−0.0107*
	(0.0032)	(0.0041)	(0.0055)
Time to elec. × Population × PR	−0.0038**	−0.0040**	−0.0064**
	(0.0016)	(0.0016)	(0.0028)
Check:	geo controls	elec. months	$t \geq t_{SIDE}$
Sample:	all	all	all
Spatial $\text{lag}_{t-1,t-2,t-3}$:	yes	yes	yes
Polynomial $\text{DV}^{1,2,3}$:	yes	yes	yes
District-year FE:	yes	yes	yes
Country-month FE:	yes	yes	yes
Mean DV:	0.0012	0.0012	0.0012
Observations	503,532	542,684	230,213
R^2	0.2341	0.2323	0.2655

Notes: OLS linear models. Model 1 includes interactions of urban population, nightlights per capital and pure ethnic polarization with the time to the next election, seperate for PR and majoritarian/mixed systems. Model 3 includes a three way interaction of a dummy for an election month × ethno-political polarization × a dummy for pure PR systems. Standard errors clustered on the district-level in parentheses. Significance codes: *p<0.1; **p<0.05; ***p<0.01

Standard error specifications: To gauge whether the results are sensitive to the manner of clustering standard errors, Table A4 presents four variations which increase the level on which errors are clustered. First, I present the baseline model with errors clustered on the district level. Model 2 uses Conley's standard errors which account for spatial and temporal clustering in a non-paramteric manner (Bester et al., 2011; Conley, 1999). Model 3 clusters on the regional (first-level administrative unit) level, and Model 4 on the country-year level.[1] The analyses show that the results are insensitive to the kind of standard error clustering applied, with the Conley clustering even reducing the uncertainty attributed to the estimates.

[1]Clustering on the country-level would lead to an insufficiently low number of clusters.

Table A4: Local ethnic polarization & pre-election violence in majoritarian and mixed systems: Standard error specifications

	Dependent variable:			
	Riots (SCAD)	Riots (SCAD)	Riots (SCAD)	Riots (SCAD)
	(1)	(2)	(3)	(4)
Time to elec. × Ethno-pol. polar.	0.0076***	0.0076***	0.0076**	0.0076**
	(0.0029)	(0.0023)	(0.0031)	(0.0032)
Time to elec. × Population	0.0045***	0.0045***	0.0045***	0.0045***
	(0.0015)	(0.0011)	(0.0016)	(0.0017)
SE clustering:	District	Conley	Region	Country-year
Sample:	Maj. & Mix.	Maj. & Mix.	Maj. & Mix.	Maj. & Mix.
District-year FE:	yes	yes	yes	yes
Country-month FE:	yes	yes	yes	yes
Spatial $\text{lag}_{t-1,t-2,t-3}$:	yes	yes	yes	yes
Polynomial $\text{DV}^{1,2,3}$:	yes	yes	yes	yes
Mean DV:	0.0014	0.0014	0.0014	0.0014
Observations	434,303	434,303	434,303	434,303
R^2	0.2311	0.2311	0.2311	0.2311

Notes: OLS linear models. Significance codes: *p<0.1; **p<0.05; ***p<0.01

Alignment with presidential elections: Legislative elections that are aligned and non-aligned with presidential elections might come with differnt forms of pre-election violence. In particular, the main results might be drivne by presidential, rathern than legislative contests held at the same time. Distinguishing between aligned and non-aligned elections,[2] Table A5 reestimates the baseline analyses. The results show that higher levels of ethno-political polarization are associated with steeper increases in the number of riots before no matter whether of legislative elections in majoritarian systems are aligned or not. Although the coefficient is slightly smaller in non-aligned elections, the difference between the two coefficients is not statistically significant. On a theoretical level, this slight difference is not too surprising, given that the (ethno-political) tensions arising in presidential elections might add to those of close local legislative contests.

[2]Elections are coded as being aligned if a legislative election is accompanied by a presidential one in the same month.

Table A5: Local ethnic polarization & pre-election violence in majoritarian and mixed systems: alligend vs. non.alligned elections

	Dependent variable:			
	Riots (SCAD)	Riots (SCAD)	Riots (SCAD)	Riots (SCAD)
	(1)	(2)	(3)	(4)
Constant	−0.0049**			
	(0.0021)			
Aligned: Time to elec. × Ethno-pol. polar.	0.0090***	0.0093***	0.0090**	0.0081**
	(0.0033)	(0.0034)	(0.0036)	(0.0038)
Not aligned: Time to elec. × Ethno-pol. polar.	0.0073**	0.0072**	0.0064**	0.0069*
	(0.0029)	(0.0029)	(0.0030)	(0.0040)
Sample:	Maj. & Mix.	Maj. & Mix.	Maj. & Mix.	Maj. & Mix.
Fixed effects:	–	country	district	district-year & country-month
Covariates:	yes	yes	yes	yes
Spatial $\text{lag}_{t-1,t-2,t-3}$:	yes	yes	yes	yes
Polynomial $\text{DV}^{1,2,3}$:	yes	yes	yes	yes
Mean DV:	0.0015	0.0015	0.0015	0.0015
Observations	399,406	399,406	399,406	399,406
R^2	0.0086	0.0100	0.0433	0.2312

Notes: OLS linear models. Standard errors clustered on the district-level in parentheses. Significance codes: *p<0.1; **p<0.05; ***p<0.01

ECAV data: Table A6 presents the results of the analyses of violent events of electoral violence as coded by Daxecker et al. (2018). These events are coded only for the six months preceding elections. It is therefore impossible to model the *increase* of violence as districts move from between-election to campaign periods. For this reason, the models draw on cross-sectional variation between polarized and non-polarized districts in the six months before legislative elections and do not model the increase of violence. The events coded by the data include only events that are can be attributed to the election. This might introduce under-counting and/or bias if violent events which have been planned with electoral motives in mind are not described as such in the news articles the data relies on.

The results show that ethno-politically polarized districts see more violent and election-related events in the six months before legislative elections in majoritarian countries (Model 1, p< .1). As evidenced by the negative and statistically significant interaction term of polarization × PR in Model 2, this relationship is absent in pure PR systems. While the coefficient of the effect of ethno-political polarization on the count of all violent events is associated with somewhat higher levels of uncertainty, standard-errors are much smaller once

we move to the count of victims in Models 3 and 4.[3] Substantively more people die from pre-election violence in polarized districts under majoritarian voting than in non-polarized ones. Again, this pattern is absent in PR systems.

Table A6: Local ethnic polarization & pre-election violence from ECAV

	Dependent variable:			
	Violent Events (ECAV)	Violent Events (ECAV)	Fatalities (ECAV)	Fatalities (ECAV)
	(1)	(2)	(3)	(4)
Ethno-pol. polarization	0.0026*	0.0029*	0.0027***	0.0028***
	(0.0015)	(0.0015)	(0.0010)	(0.0010)
Population (log)	0.0026***	0.0028***	0.0019***	0.0020***
	(0.0007)	(0.0007)	(0.0004)	(0.0005)
Ethno-pol. polarization × PR		−0.0049**		−0.0033***
		(0.0020)		(0.0011)
Population (log) × PR		−0.0027***		−0.0020***
		(0.0011)		(0.0006)
Sample:	Maj. & Mix.	all	Maj. & Mix.	all
Spatial $\text{lag}_{t-1,t-2,t-3}$:	yes	yes	yes	yes
Polynomial $\text{DV}^{1,2,3}$:	yes	yes	yes	yes
District-year FE:	no	no	no	no
Country-month FE:	yes	yes	yes	yes
Mean DV:	0.0049	0.0042	0.0022	0.0019
Observations	56,236	68,488	56,236	68,488
R^2	0.1134	0.1101	0.0775	0.0802

Notes: OLS linear models. Standard errors clustered on the district-level in parentheses. Significance codes: *p<0.1; **p<0.05; ***p<0.01

Testing for biased reporting: Because elections involve a generally increased focus of international media on a country – and within it on the particularly troublesome areas – the results might be driven by a higher propensity to report on violence in polarized districts prior to an election. Such media bias should however affect reporting of all types of social conflict events. Pre-election violence driven by ethno-political polarization should instead pertain mainly to rioting which involves popular mobilization, rather than more organized forms of violence conducted by militias, or non-violent events such as demonstrations and strikes. Using the information from SCAD on these latter event types, Table A7 thus conducts a 'placebo' test to discern potential bias caused by increased media attention. The results from Models 2–4 indicate that ethno-political polarization before elections is not associated with increases in the number of reported demonstrations, strikes, or militia-related

[3]ECAV only offers ordinal estimates for the number of victims. To derive a sum of fatalities in a district-month, I take the sum of the lowest number of victims in a given fatality-bracket coded by ECAV (e.g. 10 for the bracket that ranges from 10 to 100 victims).

events. This suggests that the results are not driven by biased media reports.

Table A7: Local ethnic polarization & various forms of pre-election violence

	Dependent variable:			
	Riots (SCAD)	Demonstrations (SCAD)	Strikes (SCAD)	Militia (SCAD)
	(1)	(2)	(3)	(4)
Time to elec. × Ethno-pol. polar.	0.0076***	0.0004	−0.0005	−0.0011
	(0.0029)	(0.0018)	(0.0007)	(0.0032)
Time to elec. × Population	0.0045***	−0.0001	−0.0003	0.0013
	(0.0015)	(0.0010)	(0.0005)	(0.0013)
Sample:	Maj. & Mix.	Maj. & Mix.	Maj. & Mix.	Maj. & Mix.
District-year FE:	yes	yes	yes	yes
Country-month FE:	yes	yes	yes	yes
Spatial $\text{lag}_{t-1,t-2,t-3}$:	yes	yes	yes	yes
Polynomial $\text{DV}^{1,2,3}$:	yes	yes	yes	yes
Mean DV:	0.0014	0.0014	4e-04	0.0019
Observations	434,303	434,303	434,303	324,214
R^2	0.2311	0.2639	0.2035	0.2890

Notes: OLS linear models. Standard errors clustered on the district-level in parentheses. Significance codes: *p<0.1; **p<0.05; ***p<0.01

Constituency-level evidence from Kenya: As discussed in the main paper, the choice of administrative *districts*, although facilitating the comparison of majoritarian and PR systems, might be inadequate for capturing the dynamics of electoral competition in majoritarian systems. Using maps of electoral constituencies from Kenya, Table A8 shows that the results remain comparable once the geographical unit of analysis is changed to constituenceis where the actual competition for votes takes place. Using the baseline model that includes country-month and district-year fixed effects, all results point towards a positive effect of ethno-political polarization on pre-election violence in Kenya. The respective constituency-level analysis is about three times smaller than that from the district-level analysis. This difference is directly related to the three times lower average number of riots observed in constituencies, which in turn mirrors the fact that each district contains about three constituencies. Given these results and the high correlation between the district- and constituency-level measures of ethno-political polarization (see Figure 1 in the main paper), it seems very unlikely that the results are a mere artifact of the choice of unit of analysis.

Table A8: Local ethnic polarization & various forms of pre-election violence: Kenya constituencies

	Dependent variable:					
	Riots (SCAD)	Riots (SCAD)	Riots (ACLED)	Riots (ACLED)	Fatalities (ACLED)	Fatalities (ACLED)
	(1)	(2)	(3)	(4)	(5)	(6)
Time to elec. × Ethno-pol. polar.	0.0321*	0.1141*	0.1690***	0.3842***	0.1747**	0.6102***
	(0.0185)	(0.0625)	(0.0527)	(0.1369)	(0.0751)	(0.2144)
Time to elec. × Population	0.0069	0.0273	0.0275*	0.0716**	0.0274	0.1042**
	(0.0046)	(0.0177)	(0.0150)	(0.0324)	(0.0186)	(0.0499)
Sample:	Kenya constituencies	Kenya districts	Kenya constituencies	Kenya districts	Kenya constituencies	Kenya districts
District-year FE:	yes	yes	yes	yes	yes	yes
Country-month FE:	yes	yes	yes	yes	yes	yes
Spatial $\text{lag}_{t-1,t-2,t-3}$:	yes	yes	yes	yes	yes	yes
Polynomial $\text{DV}^{1,2,3}$:	yes	yes	yes	yes	yes	yes
Mean DV:	0.0017	0.0053	0.0064	0.0173	0.002	0.0055
Observations	57,814	19,220	40,723	14,090	40,723	14,090
R^2	0.2427	0.2819	0.2441	0.3407	0.1417	0.2362

Notes: OLS linear models. Standard errors clustered on the district-level in parentheses. Significance codes:
*p<0.1; **p<0.05; ***p<0.01

A3 Evidence from the Afrobarometer

A3.1 Geocoding Afrobarometer for Nigeria (3.5) and Uganda (4.5)

Since its 3rd round, Afrobarometer has coded the 'district' of respondents. However, depending on the country and the respective round, the Afrobarometer districts refer to administrative units on different levels. Most of the time, they can be matched to level-2 units, but sometimes only to lower or higher levels. In order to geographically locate the respondents of the additional rounds of the Afrobarometer conducted in Uganda (round 4.5, 2010/2011) and Nigeria (round 3.5, 2007), I implement a geographical matching procedure consisting of the following steps, each using cleaned ASCII strings as an input. Each step is implemented on those Afrobarometer districts that have not been matched in the previous steps. Districts are only matched within their countries.

1. Match districts to 2nd-level administrative unit names as indicated in the GAUL data of the respective year (FAO, 2014). Fuzzy string-matching using the a maximum Jaro-Winker distance (Winkler, 1990) of 0.1.

2. Match Afrobarometer regions to 1st-level administrative unit names as indicated in the GAUL data of the respective year (FAO, 2014). Fuzzy string-matching using a maximum Jaro-Winker distance of 0.1.

3. Search the `geonames.org` API to access the coordinates of an Afrobarometer district using a maximum Jaro-Winker distance of 0.1. If multiple coordinates are returned, the one with the 1st-level administrative unit name closest to the one indicated by Afrobarometer is chosen.

4. Search the Google Maps API for the Afrobarometer district nested within its region as indicated by the survey. This second parameter has to be specified since no string-distance parameter can be passed to the database. Results are only kept if they indicate that the engine has found a place at a level below the respective 1st-level administrative unit.

The coordinates returned in step 3 and 4 are then mapped to level-2 administrative units from a given survey year, again using the GAUL data (FAO, 2014).

A3.2 Summary statistics

Table A9: Summary statistics: Afrobarometer, rounds 4-6

Statistic	N	Mean	St. Dev.	Min	Max
Fear of elec. violence	53,212	1.052	1.164	0	3
Ethno-pol. polarization	54,119	0.356	0.330	0.000	1.000
Majoritarian & mixed	54,119	0.781	0.414	0	1
PR (pure)	54,119	0.219	0.414	0	1
Population (log)	54,119	11.985	1.284	6.526	14.937
Urban	53,864	0.622	0.485	0	1
Female	54,119	0.500	0.500	0	1
Age	53,554	35.524	13.828	18	105
Education	54,002	2.386	0.958	1	4
Same language as interviewer	54,119	0.413	0.492	0	1
Others checked during interview	54,080	0.045	0.206	0	1
Others influenced	54,047	0.038	0.191	0	1
Others present	54,027	0.322	0.467	0	1

Table A10: Summary statistics: Afrobarometer, Nigeria (3.5) & Uganda (4.5)

Statistic	N	Mean	St. Dev.	Min	Max
Personal threat: frequency	6,138	0.133	0.502	0	3
Personal threat: security	6,138	0.048	0.215	0	1
Community threat: frequency	5,810	0.170	0.563	0	3
Community threat security	5,810	0.060	0.238	0	1
Ethno-pol. polarization	6,305	0.471	0.261	0.0004	0.918
Majoritarian & mixed	6,305	1.000	0.000	1	1
PR (pure)	6,305	0.000	0.000	0	0
Population (log)	6,305	12.310	0.819	8.030	14.029
Urban	6,305	0.270	0.444	0	1
Female	6,305	1.500	0.500	1	2
Age	6,254	33.834	12.755	18	93
Education	6,279	2.571	0.927	1	4
Same language as interviewer	6,305	0.531	0.499	0	1
Others checked during interview	6,295	0.018	0.133	0	1
Others influenced	6,295	0.013	0.112	0	1
Others present	6,287	0.321	0.467	0	1

A3.3 Survey data analysis: Robustness checks

This section presents the robustness checks conducted to gauge the sensitivity of the analysis of the fear of Afrobarometer respondents to fall victim of pre-election violence. Each paragraph provides a short summary of the table that ensues.

Main robustness checks: Table A11 summarizes the main robustness checks discussed in the article. Model 1 summarizes the model in which I add three additional control variables to assess whether the results are driven by pure ethnic polarization, local economic activities captured through nightlight

emissions, or the size of the local urban population. While the level of pure ethnic polarization is negatively associated with the fear of pre-election violence and intimidation ($p < .1$),[4] the effect of ethno-political polarization remains stable and significant. Model 2 drops all observations of Afrobarometer respondents that were interviewed before the first SIDE map for their country is available. This does lower the effect of ethno-political polarization but does not change the substantive interpretation of the results. Lastly, Model 3 includes five variables to control for potential bias of the Afrobarometer responses, in particular the co-ethnicity of respondents to their interviewers, dummies for whether others were present, checked with, or influencing the respondent, and lastly a factor of the institutions respondents believe to conduct the survey. The inclusion of these items into the regression model does not change the results.

Table A11: Local ethnic polarization & fear of pre-election victimization

	Dependent variable:		
	Fear	Fear	Fear
	(1)	(2)	(3)
Ethno-pol. polarization	0.224***	0.157***	0.195***
	(0.058)	(0.055)	(0.056)
Ethnic polarization	−0.142		
	(0.088)		
Nightlights/capita (log)	0.009		
	(0.013)		
Urban population (log)	0.0004		
	(0.003)		
Sample:	Maj. & Mix.	Maj. & Mix. & $t \geq t_{SIDE}$	Maj. & Mix.
Covariates	yes	yes	yes
Add. controls			Quality items
Controls	yes	yes	yes
Country-round FE:	yes	yes	yes
Mean DV:	1.1858	1.121	1.1477
Observations	27,590	38,591	40,834
R^2	0.098	0.101	0.106

Notes: OLS linear models. Control variables include the district population (logged), and urban and female dummy, age and its square, as well as the respondent's education. Standard errors clustered on the district-level in parentheses. Significance codes: *p<0.1; **p<0.05; ***p<0.01

Standard error specifications: In parallel to the riot-analysis above, Table A12 presents the results from the survey data analysis with standard erorrs

[4]Note that this relation turns insignificant when the an ordinal logistical regression is estimated. See Table A15 below.)

clustered on the district level, Conley's clustered standard errors, and clustering on the regional, and country-survey-round level. While standard errors slightly increase in the level of clustering, all results remain highly statistically significant above the 1% level.

Table A12: Local ethnic polarization & fear of pre-election victimization

	Dependent variable:			
	Fear	Fear	Fear	Fear
	(1)	(2)	(3)	(4)
Ethno-pol. polarization	0.199***	0.199***	0.199***	0.199***
	(0.056)	(0.058)	(0.063)	(0.071)
Cluster-level	District	Conley	Region	Country-round
Sample:	Maj. & Mix.	Maj. & Mix.	Maj. & Mix.	Maj. & Mix.
Controls	yes	yes	yes	yes
Country-round FE:	yes	yes	yes	yes
Mean DV:	1.1478	1.1478	1.1478	1.1478
Observations	40,945	40,945	40,945	40,945
R^2	0.103	0.103	0.103	0.103

Notes: OLS linear models. Control variables include the district population (logged), and urban and female dummy, age and its square, as well as the respondent's education.

Pre-2014 sample: As discussed in the data section of the main paper, the Afrobarometer round 6 was conducted after 2013, the year for which the last data on the political mobilization of ethnic groups is available. Because this relevance barely changes over time, the 2013 data on ethno-political polarization is matched to all respondents interviewed thereafter for the main analyses. Table A13 analyzes whether this coding decision is driving the results. It appears that dropping all observations from after 2013 does not change the results. In majoritarian systems, local ethno-political polarization is significantly associated with the fear of pre-election violence but not so in PR systems. However, the difference between both is not statistically significant.

Table A13: Local ethnic polarization & fear of pre-election victimization: Pre-2014

	Dependent variable:		
	Fear	Fear	Fear
	(1)	(2)	(3)
Ethno-pol. polarization	0.199***	0.111	0.199***
	(0.056)	(0.077)	(0.056)
Ethno-pol. polar. × PR			−0.088
			(0.095)
Sample:	Maj. & Mix.	PR	all
Covariates:	yes	yes	yes
Country-round FE:	yes	yes	yes
Mean DV:	1.1858	0.7321	1.0838
Observations	27,590	8,000	35,590
R^2	0.098	0.046	0.112

Notes: OLS linear models. Control variables include the district population (logged), and urban and female dummy, age and its square, as well as the respondent's education. Model 3 also includes interactions of all covariates with the PR dummy. Standard errors clustered on the district-level in parentheses. Significance codes: *p<0.1; **p<0.05; ***p<0.01

(Ordered) logistic regressions: Lastly, Tables A14 to A16 summarize the results of conducting the main analyses and robustness checks in a logistic rather than the linear regression setup used above. While not as intuitively to interpret, the ordered logit models are a better fit to the outcome indicators. These are ordinal in the case of the fear of pre-election violence and reports of experienced intimidation, and binary for the case of reports of threats of physical safety (see Table A16). However, moving from OLS to ordered logits does not change the substantive conclusions drawn from the Afrobarometer data that fear and reports of pre-election violence is more common in ethno-politically polarized districts in majoritarian, but not PR systems.

Table A14: Local ethno-political polarization and fear: ordered logit

	(1)	(2)	(3)
	Fear	Fear	Fear
Ethno-pol. polarization	0.299	0.169	0.301
	(0.0895)	(0.145)	(0.0900)
Ethno-pol. polarization × PR			−0.135
			(0.169)
Sample	Maj. & Mix.	PR	all
Country-round FE	yes	yes	yes
Controls	yes	yes	yes
Observations	40945	11382	52327
χ^2	1993.3	405.4	2981.7

Notes: OLS linear models. Control variables include the district population (logged), and urban and female dummy, age and its square, as well as the respondent's education. Model 3 also includes interactions of all covariates with the PR dummy. Standard errors clustered on the district-level in parentheses. Significance codes: $^{*}p < 0.1;^{**} p < 0.05;^{***} p < 0.01$

Table A15: Local ethno-political polarization and fear, robustness checks: ordered logit

	(1) Fear	(2) Fear	(3) Fear
Ethno-pol. polarization	0.344	0.235	0.294
	(0.0920)	(0.0892)	(0.0893)
Nightlights p.c.	0.470		
	(0.647)		
Ethnic polarization	−0.213		
	(0.138)		
Urban population (log)	0.00120		
	(0.00495)		
Sample	Maj. & Mix.	Maj. & Mix. $t \geq t_{SIDE}$	Maj. & Mix.
Add. controls			Quality items
Country-round FE	yes	yes	yes
Controls	yes	yes	yes
Observations	27590	38591	40834
χ^2	1259.5	1925.3	2284.1

Notes: OLS linear models. Control variables include the district population (logged), and urban and female dummy, age and its square, as well as the respondent's education. Standard errors clustered on the district-level in parentheses.

Table A16: Local ethno-political polarization and pre-election intimidation: (ordered) logit

	(1) Personal threat: frequency	(2) Personal threat: safety	(3) Community threat frequency	(4) Community threat: safety
Ethno-pol. polarization	0.792	1.163	0.752	1.019
	(0.371)	(0.394)	(0.291)	(0.341)
Sample	NIG & UGA	NIG & UGA	NIG & UGA	NIG & UGA
Model	ologit	logit	ologit	logit
Country-round FE	yes	yes	yes	yes
Controls	yes	yes	yes	yes
Observations	6071	6071	5745	5745
χ^2	35.59	50.92	48.34	71.97

Notes: OLS linear models. Control variables include the district population (logged), and urban and female dummy, age and its square, as well as the respondent's education. Standard errors clustered on the district-level in parentheses.

References

Bester, C. A., Conley, T. G., and Hansen, C. B. (2011). Inference with dependent data using cluster covariance estimators. *Journal of Econometrics*, 165(2):137–151.

Conley, T. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, 92(1):1–45.

Daxecker, U., Amicarelli, E., and Jung, A. (2018). Electoral Contention and Violence (ECAV): A New Dataset. *Journal of Peace Research, conditionally accepted.*

FAO (2014). Global Administrative Unit Layers. *Available from: http://data.fao.org/map?entryId=f7e7adb0-88fd-11da-a88f-000d939bc5d8&tab=metadata.*

Winkler, W. E. (1990). String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage. *Proceedings of the Section on Survey Research*, 1990:354–359.