Online Appendices: Introducing the Modes and Agents of Election-Related Violence in Côte d'Ivoire and Kenya (MAVERICK) dataset

Sebastian van Baalen[†] Kristine Höglund[‡]

October 15, 2025

Abstract

These Online Appendices accompany the article "Introducing the Modes and Agents of Election-Related Violence in Côte d'Ivoire and Kenya (MAVERICK) dataset," published in *Journal of Peace Research*. Appendix A situates the cases Côte d'Ivoire and Kenya in the larger universe of countries with violent elections. Appendix B compares MAVERICK with other datasets that focus on or capture variables associated with electoral violence, and outlines MAVERICK's relative strengths and weaknesses. Appendix C provides additional descriptive statistics referred to in the text. Appendix D presents the MAVERICK-specific robustness tests discussed in the main article. Finally, Appendix E presents several additional robustness tests.

[†]Sebastian van Baalen is Associate Professor at the Department of Peace and Conflict Research at Uppsala University. Email: sebastian.van-baalen@pcr.uu.se.

[‡]Kristine Höglund is Professor at the Department of Peace and Conflict Research at Uppsala University. Email: kristine.hoglund@pcr.uu.se.

Appendix A Situating the cases

Table AI: 20 most electoral violence-affected countries 1989–2017

Country	Number of deaths
India	2,404
Nigeria	2,034
Kenya	1,924
Côte d'Ivoire	1,779
Afghanistan	1,485
South Africa	1,277
Ethiopia	1,247
Sudan	1,171
DR Congo (Zaire)	948
Indonesia	883
Peru	873
Iraq	699
Colombia	607
Pakistan	584
Somalia	579
Philippines	571
Russia (Soviet Union)	506
Burundi	468
Algeria	413
Bangladesh	387

Note:
Based on data from DECO (Fjelde & Höglund, 2022).

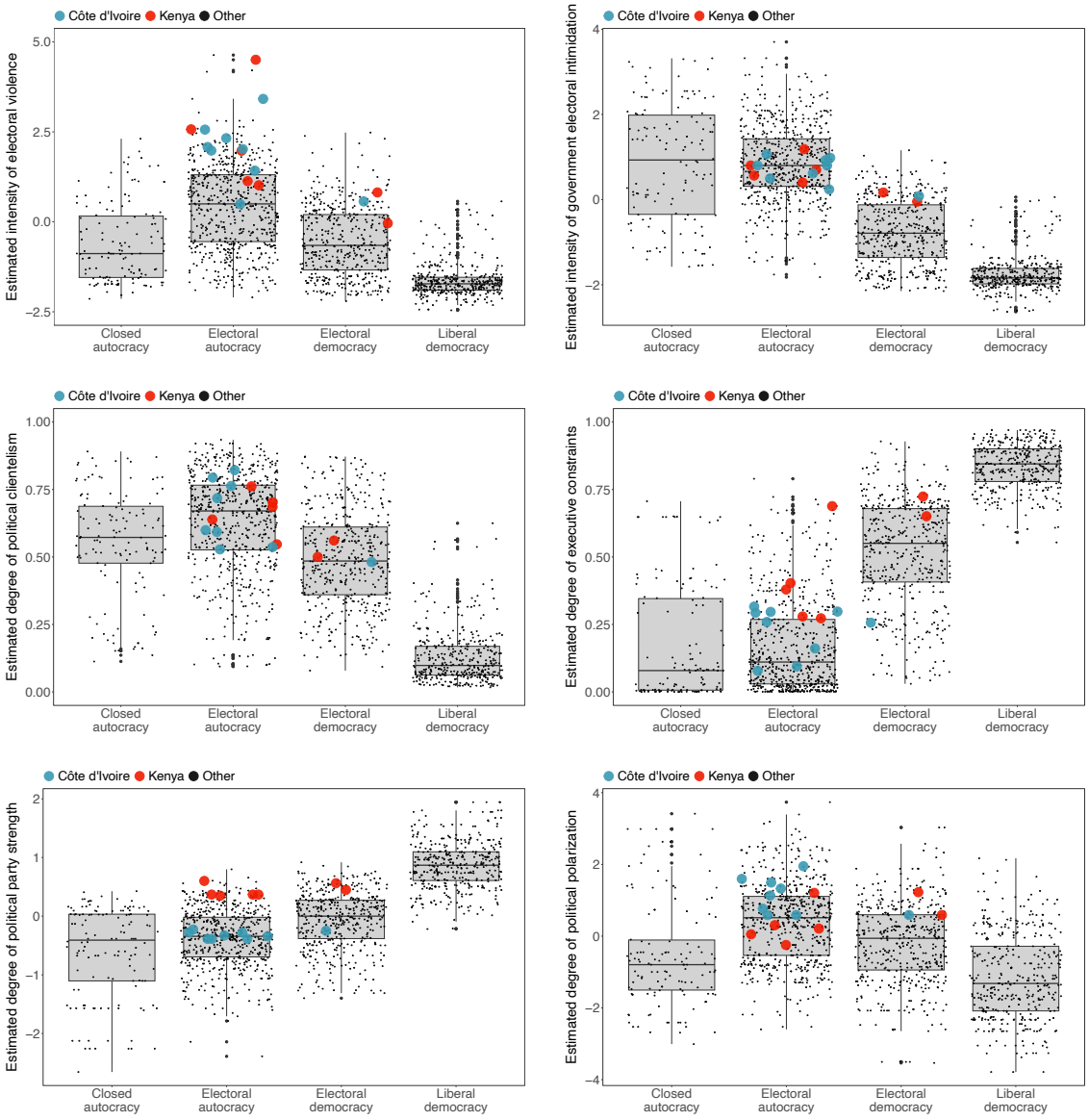


Figure A1: Boxplots of the intensity of civilian-on-civilian electoral violence (upper-left panel), government electoral intimidation (upper-right panel), estimated degree of political clientelism (middle-left panel), executive constraints (middle-right panel), political party strength (lower-left panel), and political polarization (lower-right panel) by election year and regime type, 1992–2023. The horizontal line represents the median, and the lower and upper hinges correspond to the 25th and 75th percentiles. Data points represent country-years. Based on data from V-Dem (Coppedge et al., 2025).

Appendix B Comparison with other datasets

There are several existing datasets that either focus explicitly on electoral violence or that capture variables associated with electoral violence, including the Deadly Electoral Conflict (DECO) dataset (Fjelde and Höglund, 2022), the Electoral Contention and Violence (ECAV) dataset (Daxecker, Amicarelli, and Jung, 2019), the Countries at Risk of Electoral Violence (CREV) dataset (Birch and Muchlinski, 2017), the National Elections across Democracy and Autocracy (NELDA) dataset (Hyde and Marinov, 2012), and the Varieties of Democracy (V-Dem) dataset (Coppedge et al., 2025). While we draw inspiration from all these datasets, the different datasets have relative strengths and weaknesses.

Table BI compares the main features of MAVERICK to some of the other most frequently used datasets on electoral violence and lists their relative strengths. As the table indicates, comprehensive datasets such as NELDA and V-Dem are superior when researchers aim to analyze the occurrence of electoral violence over long time periods and globally, but provide relatively limited detail given their crude coding of electoral violence. Researchers that require global or semi-global coverage but also need georeferenced event data are better off using DECO or ECAV, two georeferenced event datasets that enable scholars to study the occurrence, location, timing, and (to some extent) intensity of electoral violence across many countries. While these datasets enable far more nuanced assessments of electoral violence than NELDA or V-Dem, they both contain some limitations in assessing more fine-grained patterns of electoral violence. DECO, for example, is constrained by the fact that it only records fatal electoral violence events (Fjelde and Höglund, 2022: 172), hence omitting less violent but often impactful forms of electoral violence (see Wahman, 2024). ECAV also comes with several limitations, including that it does not include a count of electoral violence fatalities and is based exclusively on a few Western news agencies.

Given its exclusive focus on Côte d'Ivoire and Kenya, MAVERICK is unsuitable for testing general proposition on a global sample and cannot compete with datasets such as DECO and ECAV in that regard. Instead, our data is more useful for fine-grained analysis than DECO and ECAV. While all three datasets provide temporally and spatially disaggregated data at the event level, MAVERICK contains more electoral violence events and recorded deaths than its competitors. Figure B1 plots the number of electoral violence events by month, country, and dataset. The graph illustrates that MAVERICK captures a longer time frame and records more events than either DECO or ECAV. This observation is especially stark for the 2017 general election in Kenya. This election falls outside ECAV's temporal scope,

and compared to DECO, MAVERICK records many more violent events, likely because DECO only records fatal events and therefore omits non-fatal events (Fjelde and Höglund, 2022: 172). Figure B2 plots the number of electoral violence deaths by month, country, and dataset, excluding ECAV, which does not record death counts. Again, the graph shows that MAVERICK documents violence over a longer time frame and records more deaths than DECO, with the exception of the 2010–2011 post-electoral crisis in Côte d'Ivoire, an artifact of DECO considering all violent events during the second Ivorian civil war election-related. Finally, Figure B3 plots the number of deaths by location, dataset, and country, and shows that MAVERICK maps violence across larger geographical areas than DECO. Thus, for scholars aiming to test theories about electoral violence frequency, severity, and location using these two paradigmatic cases, MAVERICK offers unrivaled benefits.

Table BI: Comparison between MAVERICK and other electoral violence datasets

	MAVERICK	Deadly Electoral Conflict (DECO) dataset (Fjelde and Höglund, 2022)	Electoral Contention and Violence (ECAV) dataset (Daxecker, Amicarelli, and Jung, 2019)	Countries at Risk of Electoral Violence (CREV) dataset (Birch and Muchlinski, 2017)	National Elections across Democracy and Autocracy (NELDA) dataset (Hyde and Marinov, 2012)	Varieties of Democracy (V-Dem) dataset (Coppedge et al., 2025)
Relative strengths	Temporally and spatially disaggregated analysis of electoral violence occurrence, intensity, actor repertoires, and non-lethal violence intensity in Côte d'Ivoire and Kenya	Temporally and spatially disaggregated analysis of electoral violence occurrence and intensity	Temporally and spatially disaggregated analysis of electoral violence occurrence and intensity	Compatible with machine-learning data collection techniques	Cross-sectional analysis with long time series of electoral violence occurrence	Cross-sectional analysis with long time series of electoral violence occurrence
Geographical scope	Côte d'Ivoire and Kenya	Global	Global, except OECD countries	101 countries deemed at risk of electoral violence	Global	Global
Temporal scope	1995–2022 (Côte d'Ivoire); 1992–2022 (Kenya)	1989–2017	1990–2012, but only includes events occurring 6 months prior and 3 months after day of national election	1995–2018	1945–2020	1900–2024
Unit of analysis	Event report, event-location-day	Event-location-day	Event-location-day	Election, Electoral cycle-month	National election round	Country-year

continued on next page

Table BI *Comparison between MAVERICK and other electoral violence datasets – continued from previous page*

	MAVERICK	DECO	ECAV	CREV	NELDA	V-Dem
Death estimates	Best, high, low	Best, high, low	Ordinal (none, less than 10, 10–100, 100 or more)	–	Several binary variables (Government harassment, Government violence against demonstrators, Significant civilian deaths)	Two ordinal variables (Election government intimidation, Election other electoral violence)
Injury estimates	Best, high, low	–	–	–	–	–
Actor structure	Multi-party actor structure with up to 6 actors per event	Binary actor structure (Side A vs. Side B)	Binary actor structure (Side A vs. Side B)	Binary actor structure (5 dyads)	–	–
Actor categories	17 (including "Other" and "Unknown")	10 (including "Other" and "Unclear")	6 (including "Other" and "Unknown")	–	–	
Actor-related variables	72 (12 per actor: Actor X name, Actor X id, Actor X type, Actor X subtype, Actor X party affiliation, Actor X initiator, Actor X perpetrator, Actor X intervener, Actor X passive bystander, Actor X victim, Actor X violence repertoire, Actor X precision)	11 (Side A name, Side A id, Side A type, Side A type 2, Side A incumbent ties, Side B name, Side B id, Side B type, Side B type 2, Side B incumbent ties, Perpetrator)	7 (Actor 1 type, Actor 1 side, Actor 1 name, Actor 2 type, Actor 2 side, Actor 2 name, Initiator)	5 (Dyad 1, Dyad 2, Dyad 3, Dyad 4, Dyad 5)	–	–

continued on next page

Table BI *Comparison between MAVERICK and other electoral violence datasets – continued from previous page*

	MAVERICK	DECO	ECAV	CREV	NELDA	V-Dem
Sources	International and domestic news reports and other reports in English and French	Event data from the Uppsala Conflict Data Program (UCDP) back-end	News reports from Associated Press, Agence France-Presse, and BBC Monitoring	Automatically coded event data from the Integrated Crisis Early Warning System (ICEWS)	Multiple sources, including news reports, handbooks, scholarly studies, and the Inter-Parliamentary Union's data	Country expert coding

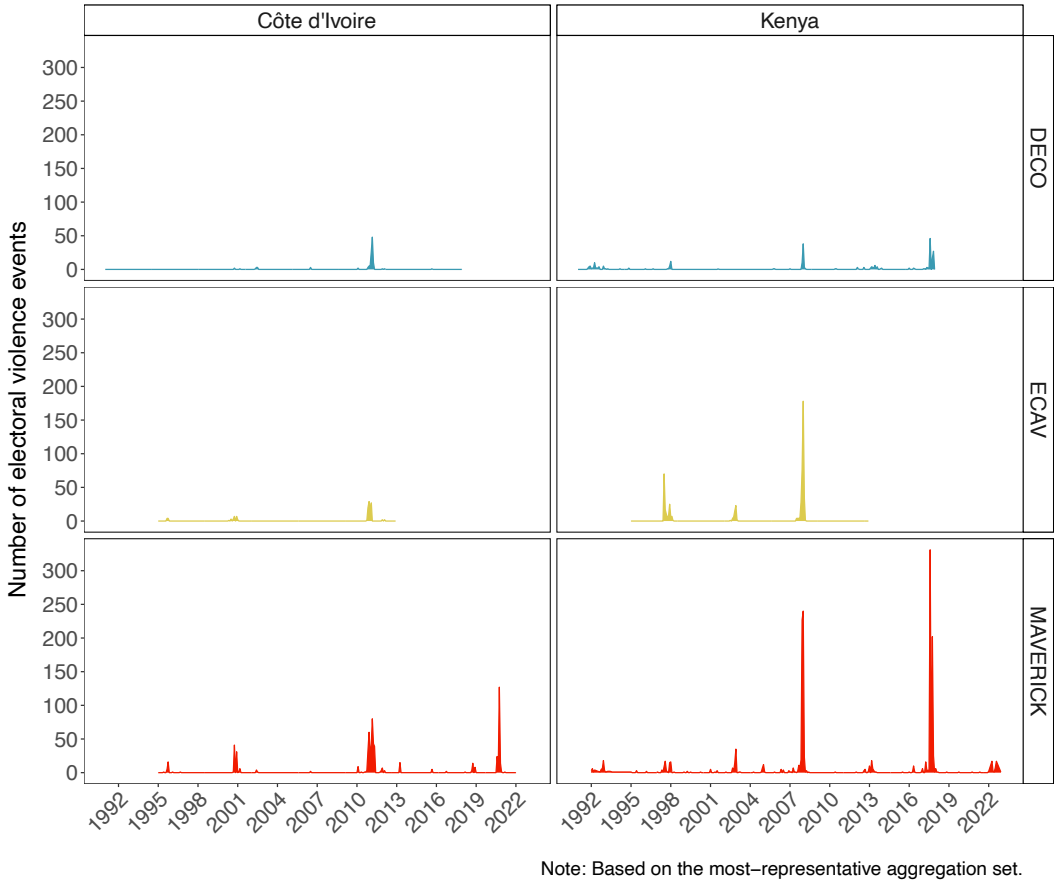


Figure B1: Number of events by month, country, and dataset

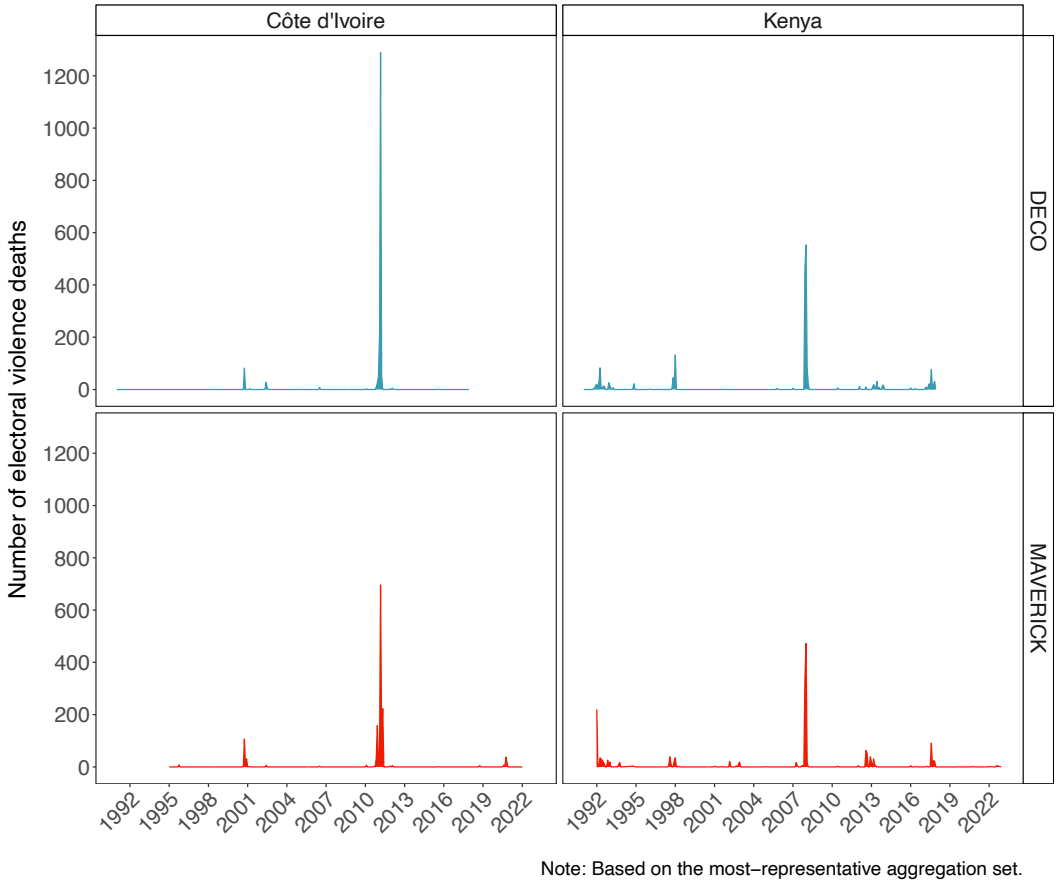
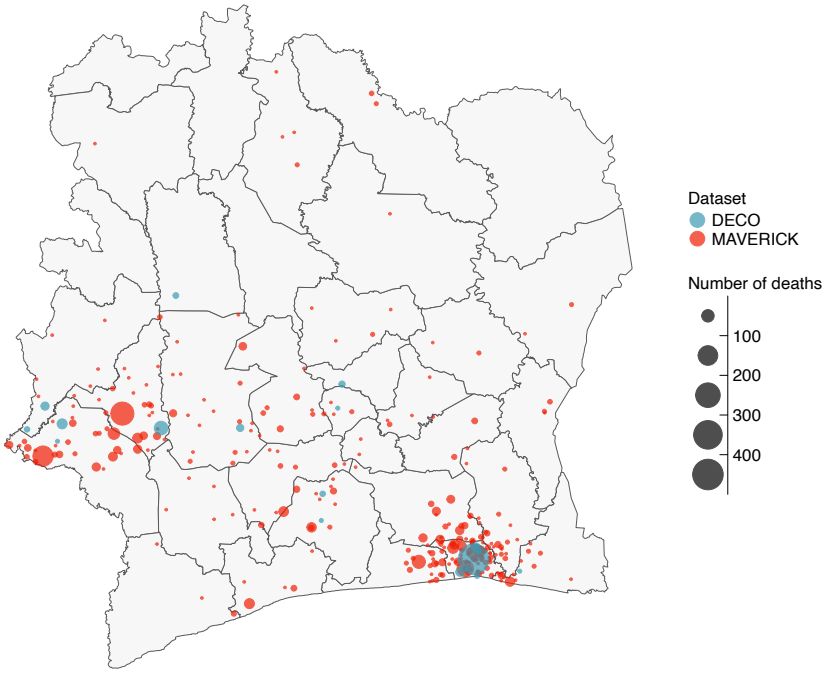
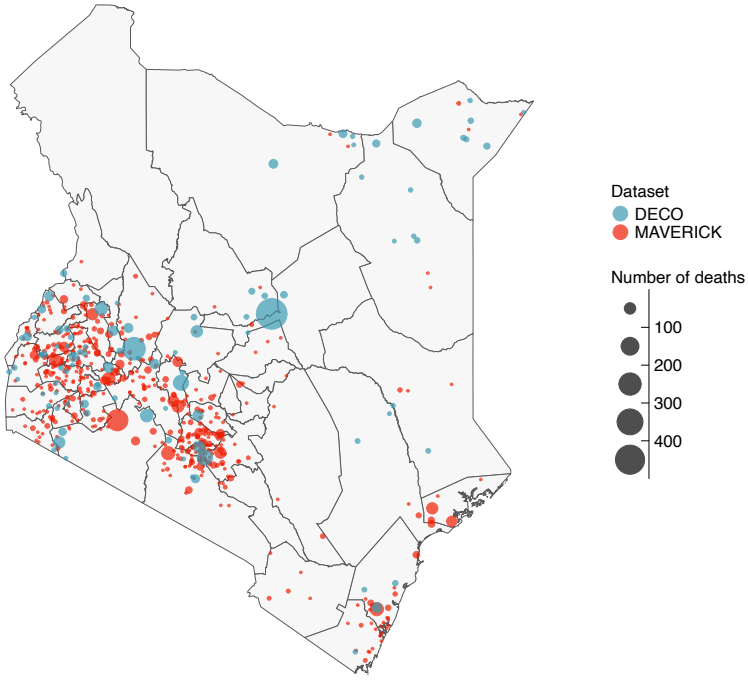


Figure B2: Number of deaths by month, country, and dataset



Note: Based on the most−representative aggregation set.



Note: Based on the most−representative aggregation set.

Figure B3: Number of deaths by location and dataset in Côte d'Ivoire (upper panel) and Kenya (bottom panel)

Appendix C Additional descriptive statistics

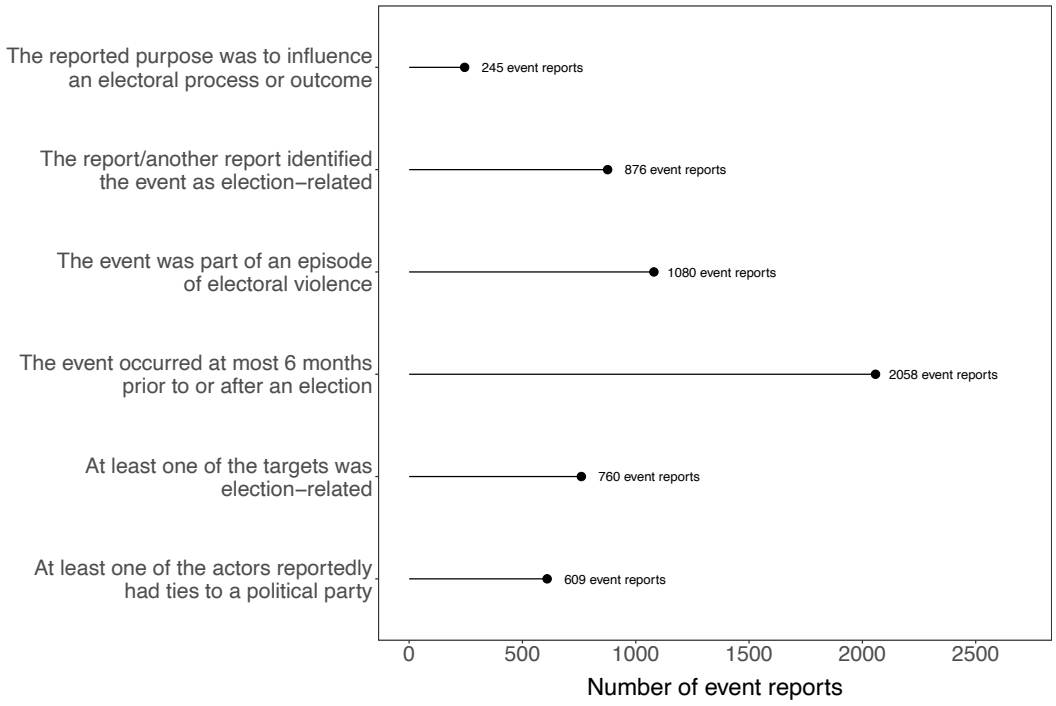


Figure C1: Number of event reports by inferential clue

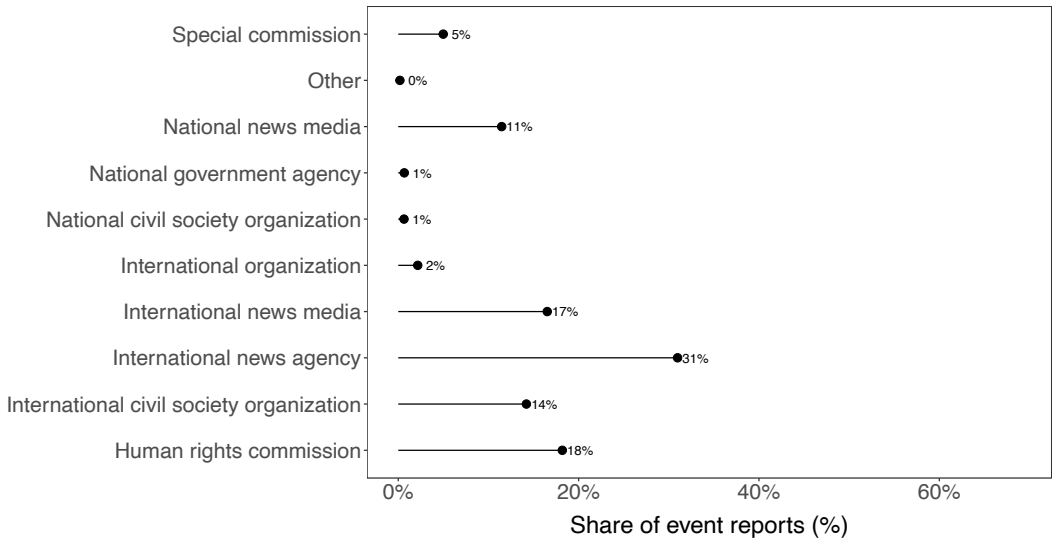
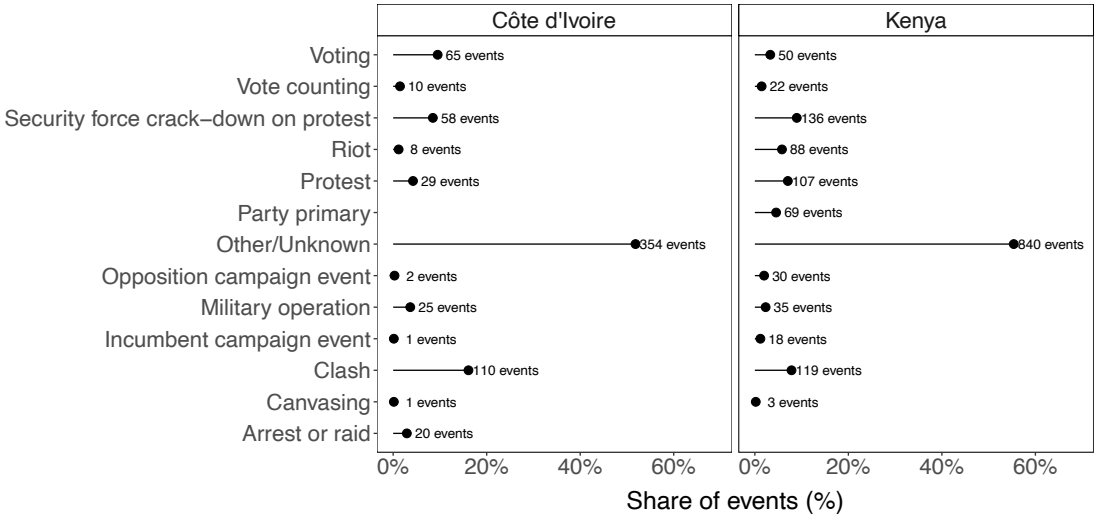


Figure C2: Number of event reports by source type



Note: Based on the most–representative aggregation set.

Figure C3: Number and share of electoral violence events by event context and country

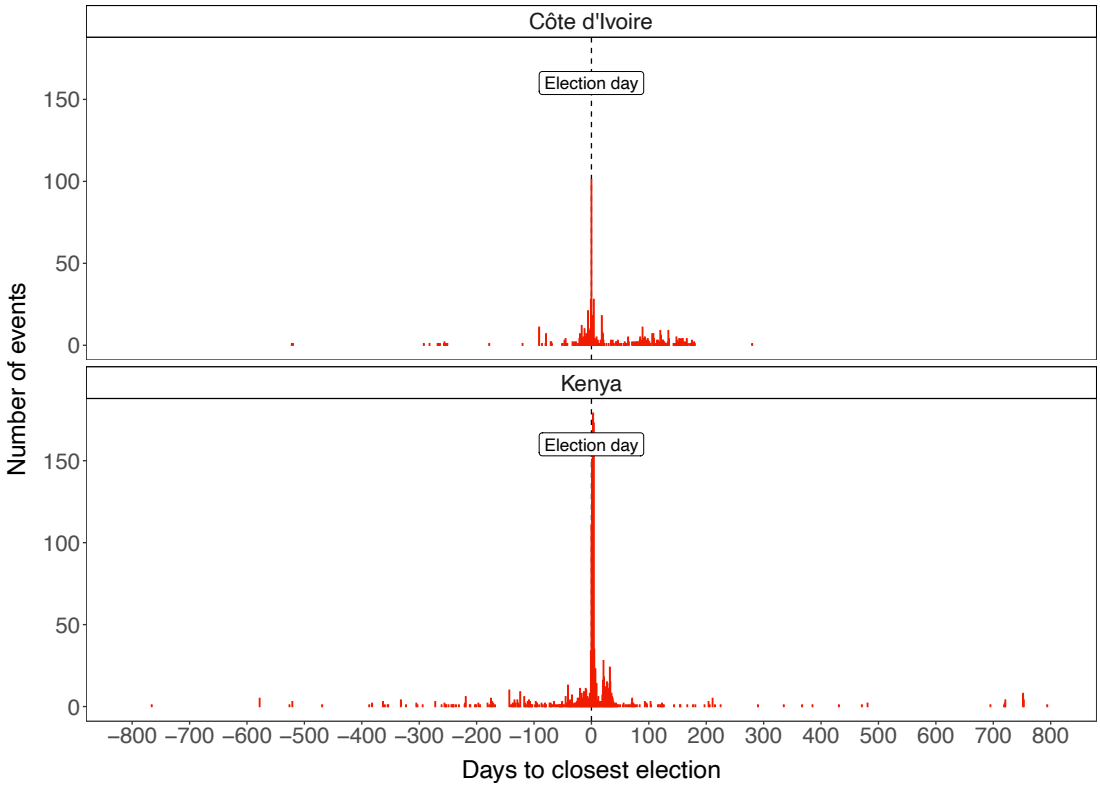


Figure C4: Number of electoral violence events by days to closest election

Table CI: Ten most frequently involved actors by country (most-representative aggregation set)

Actor name (Côte d'Ivoire)	Number of events	Actor name (Kenya)	Number of events
Unknown actor (Côte d'Ivoire)	200	Police (Kenya)	622
Forces Nouvelles (FN)	77	Unknown actor (Kenya)	275
Unknown security force (Côte d'Ivoire)	64	Unknown citizens (Kenya)	150
Police (Côte d'Ivoire)	62	Youth (Kenya)	97
Unknown political actor (Côte d'Ivoire)	54	Unknown political actor (Kenya)	90
Gendarmerie (Côte d'Ivoire)	45	Supporters of the Orange Democratic Movement (ODM)	85
Supporters of the Rassemblement des Républicains (RDR)	42	Members of the Kalenjin ethnic group	72
Supporters of the Front Populaire Ivoirien (FPI)	36	Supporters of the Kenya African National Union (KANU)	66
Opposition supporters (Côte d'Ivoire)	31	Administration Police (AP)	56
Indeterminate	30	Unknown non-state armed actor (Kenya)	52

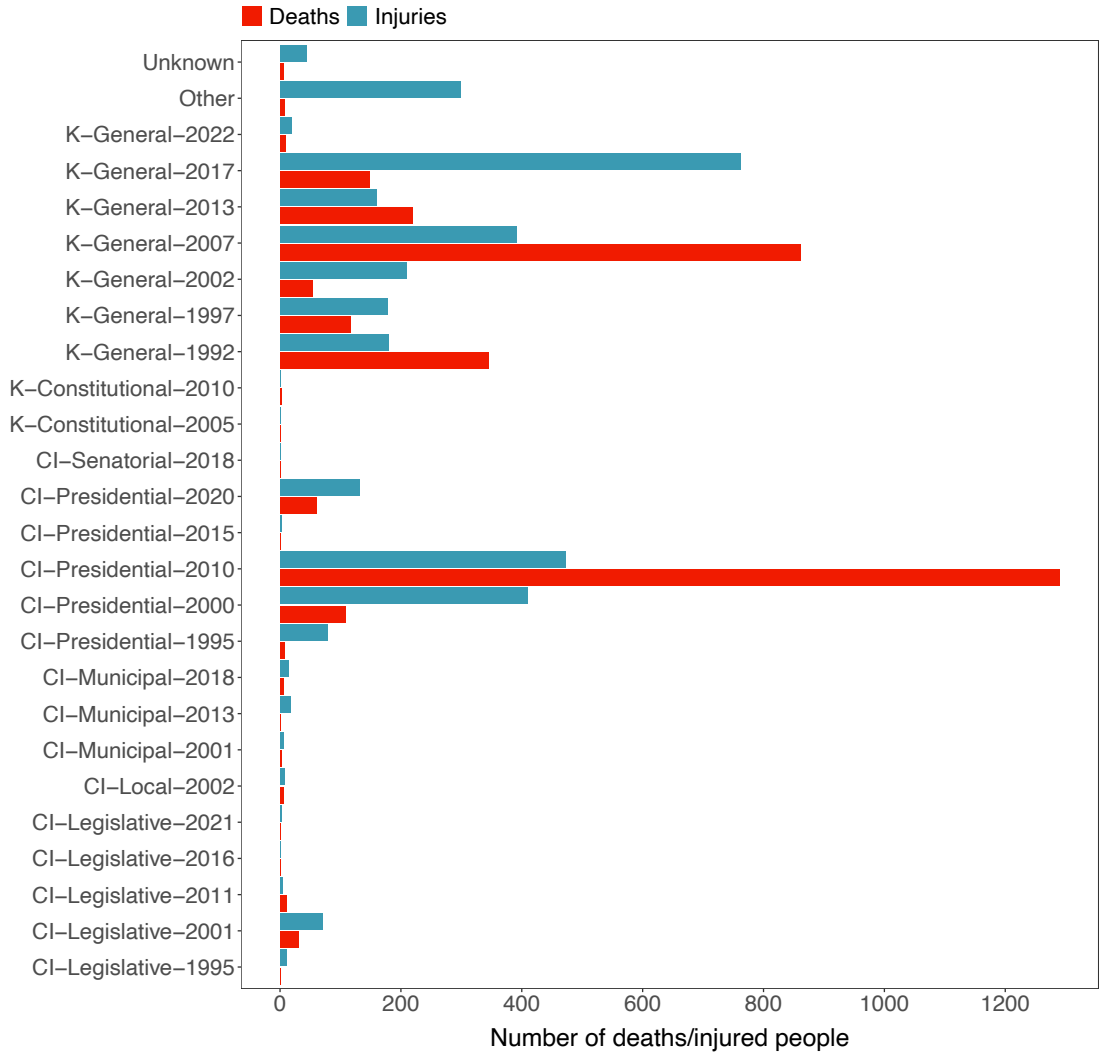


Figure C5: Number of electoral violence deaths and injured people by election

Table CII: Descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Median	Max
Best estimated number of deaths	1,780	1.53	8.56	0	0	220
Best estimate number of injured people	1,780	1.70	8.91	0	0	200
Material destruction	1,780	0.26	0.44	0	0	1
Security force involvement	1,780	0.46	0.50	0	0	1
Police involvement	1,780	0.36	0.48	0	0	1
Paramilitary police involvement	1,780	0.06	0.23	0	0	1
Army involvement	1,780	0.01	0.09	0	0	1
Agricultural land cover	1,780	52.14	27.09	0.33	40.16	99.92
Travel time to the nearest urban center	1,780	162.54	110.74	89.15	114.32	1,082.55

Appendix D MAVERICK-specific robustness checks

MAVERICK comes with specific features that enable unique robustness checks. Below, we demonstrate these features by checking the robustness of the main finding—that police involvement is associated with less lethal electoral violence, while paramilitary police involvement is associated with more lethal electoral violence—in four novel ways: by accounting for aggregation sensitivity, accounting for more and less election-related events, omitting events with missing actor information, and by including only events sourced from news sources.

Accounting for aggregation sensitivity

A key strength of the MAVERICK event report data structure is that it makes it possible to assess whether the findings are robust to aggregation sensitivity (Cook and Weidmann, 2019), that is, that the findings depend on how we combine different reports about an event into a single event observation (van Baalen and Höglund, 2025). The main results are based on the most-representative aggregation set, an event level dataset that takes the most commonly reported values for each variable. To assess whether our main findings are robust to other aggregation approaches, we replicate the models in Table I using the most-informative and most-conservative aggregation sets. The results are in Table DI.

The table reveals several insights. First, because different aggregation sets yield different geo-referencing precisions and because we only include events coded at the city level, the three datasets yield a differential number of observations. Second, Models 7 and 8 show that our main finding that police involvement is associated with less deaths remains statistically significant ($p < 0.01$) and negative when aggregating the event report level data using the most-conservative aggregation set, but not when using the most-detailed aggregation set (although the coefficient remains negative). Models 11 and 12 further indicate that the negative association between police involvement and the risk of material destruction remains statistically significant ($p < 0.01$) regardless of aggregation approach. These findings imply that two of our main results are robust to accounting for aggregation sensitivity. Third, Models 7 and 8 indicate that the positive association between paramilitary police involvement and lethal violence remains statistically significant ($p < 0.01$) across both aggregation sets. Thus, with the exception of the significance of one coefficient, our main result is robust to aggregating the event report level data using different aggregation rules.

Table DI: Accounting for aggregation sensitivity

	Dependent variable:					
	Lethal violence		Injurious violence		Material destruction	
	(7)	(8)	(9)	(10)	(11)	(12)
Police involvement	−0.12	−0.59**	−0.03	0.08	−0.49**	−0.64**
	(0.11)	(0.11)	(0.31)	(0.31)	(0.12)	(0.15)
Paramilitary police involvement	0.72**	0.90**	0.56+	0.47	−0.25	−0.09
	(0.14)	(0.17)	(0.32)	(0.33)	(0.27)	(0.38)
Army involvement	0.56*	0.06	0.22	0.04	−0.52+	−1.57
	(0.27)	(0.83)	(0.36)	(0.62)	(0.30)	(1.30)
Agricultural land cover	−0.001	−0.002	−0.01	−0.01**	0.004	0.01*
	(0.01)	(0.004)	(0.01)	(0.004)	(0.003)	(0.003)
Travel time to nearest urban center	0.002**	0.002**	0.002	0.003*	0.001*	0.002**
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Election fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Aggregation	Most-inf.	Most-con.	Most-inf.	Most-con.	Most-inf.	Most-con.
Observations	1,916	1,845	1,916	1,845	1,916	1,845
Log Likelihood	−2,615.66	−1,994.45	−3,267.86	−2,506.85	−950.41	−831.49
θ	0.22** (0.01)	0.23** (0.02)	0.21** (0.01)	0.24** (0.01)		
Akaike Inf. Crit.	5,295.32	4,052.89	6,599.72	5,077.71	1,964.81	1,726.98

Note: Robust standard errors in parentheses, clustered by district/county:
+ p < 0.1, * p < 0.05, ** p < 0.01

Using inclusion certainty weights

Another concern is that the results are driven by the inclusion of violent events that were not election-related. Political violence datasets often need to make difficult decisions regarding which events to include or exclude from the data. In our case, we are forced to infer an intention to influence an electoral process or outcome from observable clues. Although we can never know the true number of false positives or negatives, MAVERICK's documentation of what unique inferential clue each event fulfilled allow us to test how sensitive our results are to our inclusion criteria. For example, it is possible that the police is less involved in more communal violence than in more political violence. Because communal violence in electoral contexts often has a less clear link to elections but can yield high death counts, our results could be an artifact of including communal violence events that are not election-related. To assess this possibility, we use a normalized measure of the cumulative count of the number of unique inferential clues that each event fulfilled as weights in the regression analysis. These weights allow us to put more statistical emphasis on events that fulfilled many unique inferential clues than on events that fulfilled only one or a few. The results are in Table DII and show that police involvement remains associated with less lethal violence ($p < 0.01$) and a lower risk of material destruction ($p < 0.01$), whereas paramilitary police involvement remains associated with more lethal violence ($p < 0.01$). Thus, our main result is robust to accounting for events with a less unequivocal link to elections.

Table DII: Using inclusion certainty weights

	Dependent variable:		
	Lethal violence	Injurious violence	Material destruction
	(13)	(14)	(15)
Police involvement	−0.32**	−0.25	−0.84**
	(0.10)	(0.29)	(0.14)
Paramilitary police involvement	0.74**	0.28	−0.15
	(0.15)	(0.38)	(0.36)
Army involvement	−0.29	−0.24	−1.99
	(0.78)	(0.48)	(1.27)
Agricultural land cover	−0.002	−0.01**	0.01*
	(0.004)	(0.004)	(0.003)
Travel time to nearest urban center	0.003**	0.002	0.001*
	(0.001)	(0.001)	(0.001)
Election fixed effects	Yes	Yes	Yes
Inclusion certainty weights	Yes	Yes	Yes
Aggregation	Most-rep.	Most-rep.	Most-rep.
Observations	1,780	1,780	1,780
Log Likelihood	−2,207.69	−2,367.08	−923.03
θ	0.24** (0.01)	0.20** (0.01)	
Akaike Inf. Crit.	4,455.38	4,774.16	1,886.07

Note: All models include inclusion certainty weights.
Robust standard errors in parentheses, clustered by district/county:
+ p < 0.1, * p < 0.05, ** p < 0.01

Omitting events with missing actor information

As indicted in Figure 2, a large share of the events have missing information about at least one of the involved actors. The large share of unknown actors, however, raises the concern that events with unknown actors bias the analysis, as such events may have involved the security forces (or one of its branches) even if they were not reportedly involved. To account for this possibility, we replicate the main analysis in Table I on a sample of events that omits all events that included an unknown actor as an initiator, perpetrator, or intervener. Models 16, 18, and 20 omit all events that include at least one actor for which the actor type was recorded as "Unknown," thereby removing all events that could have included the security forces. Models 17, 19, and 21 omit all events that include at least one actor for which the actor subtype was recorded as "Unknown," thereby removing all events that could have included the police, paramilitary police, or army. The results are in Table DIII and show that police involvement remains associated with less lethal violence ($p < 0.05$) and a lower risk of material destruction ($p < 0.01$), whereas paramilitary police involvement remains associated with more lethal violence ($p < 0.01$). Hence, our main result is robust to removing events for which at least one of the actors is unknown.

Table DIII: Omitting events with missing actor information

	Dependent variable:					
	Lethal violence		Injurious violence		Material destruction	
	(16)	(17)	(18)	(19)	(20)	(21)
Security force involvement	−0.15		0.02		−0.81**	
	(0.12)		(0.18)		(0.16)	
Police involvement		−0.22*		−0.004		−0.97**
		(0.11)		(0.20)		(0.16)
Paramilitary police involvement		0.71**		0.19		−0.23
		(0.12)		(0.38)		(0.34)
Army involvement		−1.07**		−0.09		−15.32**
		(0.24)		(0.52)		(0.74)
Agricultural land cover	−0.005	−0.01**	−0.01*	−0.01*	0.004	0.01+
	(0.005)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)
Travel time to nearest urban center	0.004**	0.003**	−0.0005	−0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Election fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Aggregation	Most-rep.	Most-rep.	Most-rep.	Most-rep.	Most-rep.	Most-rep.
Observations	1,419	1,275	1,419	1,275	1,419	1,275
Log Likelihood	−1,633.53	−1,341.69	−2,027.03	−1,863.63	−615.37	−545.18
θ	0.24** (0.02)	0.26** (0.02)	0.29** (0.02)	0.32** (0.02)		
Akaike Inf. Crit.	3,303.06	2,723.37	4,090.05	3,767.25	1,266.74	1,130.36

Note: Models 16, 18, and 20 omit events that include at least one unknown actor type.
Models 17, 19, and 21 omit events that include at least one unknown actor subtype.
Robust standard errors in parentheses, clustered by district/county:
+ p < 0.1, * p < 0.05, ** p < 0.01

Including only events sourced from news sources

A strength of collecting event data at the event report level is that the user can apply different filters on the event report level data and then re-aggregate the data to the event level, allowing researchers to assess if their results are sensitive to the inclusion of particular sources (van Baalen and Höglund, 2025). Aggregating the event report level data using only event reports sourced from international and domestic news agencies and media, on the one hand, and only event reports sourced from governmental and non-governmental reports and special commissions, on the other hand, matters for what is included in the final data. As we show in Figure D1, some subtypes of actors are more frequently recorded in some types of sources. Moreover, reports are more likely to identify the subtype of the involved actors than media sources. To illustrate how MAVERICK's event report structure can be used for analysis, we filter the event report level data so that it only includes event reports sourced from international and domestic news agencies and media. This reduces the number of event reports from 3,287 reports to 1,938 reports. The results are in Table DIV and show that police involvement remains associated with less lethal violence ($p < 0.05$) and a lower risk of material destruction ($p < 0.01$). Paramilitary police involvement remains associated with more lethal violence, but this association is not statistically significant at conventional levels ($p \approx 0.2$).

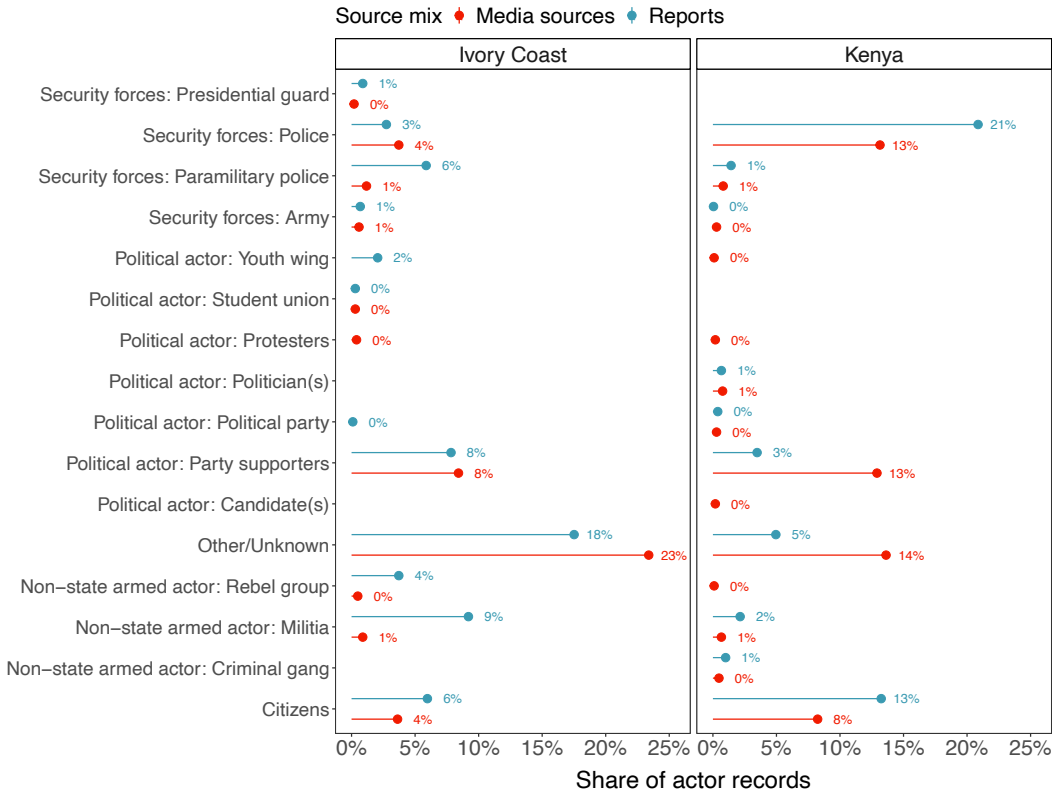


Figure D1: Share of actor records by actor subtype, country, and source mix

Table DIV: Using only events sourced media reports

	Dependent variable:					
	Lethal violence		Injurious violence		Material destruction	
	(22)	(23)	(24)	(25)	(26)	(27)
Security force involvement	−0.78**		−0.19		−0.77**	
	(0.20)		(0.34)		(0.21)	
Police involvement		−0.92**		−0.16		−0.70**
		(0.20)		(0.41)		(0.23)
Paramilitary police involvement		0.43		0.97		0.27
		(0.33)		(0.66)		(0.42)
Army involvement		1.19		−0.23		−1.38
		(0.99)		(0.98)		(1.43)
Agricultural land cover	0.004	0.01	−0.02**	−0.02**	0.004	0.004
	(0.01)	(0.01)	(0.01)	(0.01)	(0.004)	(0.004)
Travel time to nearest urban center	0.002+	0.002+	0.005**	0.005**	0.001	0.001
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)
Election fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Aggregation	Most-rep.	Most-rep.	Most-rep.	Most-rep.	Most-rep.	Most-rep.
Observations	788	788	788	788	788	788
Log Likelihood	−1,090.86	−1,088.69	−854.14	−853.47	−387.69	−389.29
θ	0.22** (0.02)	0.22** (0.02)	0.09** (0.01)	0.09** (0.01)		
Akaike Inf. Crit.	2,213.73	2,213.37	1,740.28	1,742.94	807.38	814.58

Note: The models include only events aggregated based on media event reports.
Robust standard errors in parentheses, clustered by district/county:
+ p < 0.1, * p < 0.05, ** p < 0.01

Appendix E Robustness checks

Accounting for event context

One concern is that our main result that police involvement is associated with less lethal violence, whereas paramilitary police involvement is associated with more lethal violence, is driven by the differential involvement of certain security force branches in certain types of events. It is, for instance, possible that paramilitary police are more often deployed to deal with riots and clashes than the police, hence biasing the main correlation we observe. To account for this possibility, we replicate our main models but include event context fixed effects. The results are in Table EI and show that police involvement remains associated with less lethal violence ($p < 0.01$) and a lower risk of material destruction ($p < 0.01$), whereas paramilitary police involvement remains associated with more lethal violence ($p < 0.01$). Thus, our main result is robust to accounting for the context in which different security force branches intervene.

Table EI: Accounting for event context

	Dependent variable:					
	Lethal violence		Injurious violence		Material destruction	
	(28)	(29)	(30)	(31)	(32)	(33)
Security force involvement	−0.21		0.17		−0.92**	
	(0.16)		(0.31)		(0.14)	
Police involvement		−0.38**		0.09		−0.88**
		(0.10)		(0.32)		(0.14)
Paramilitary police involvement		0.83**		0.42		−0.12
		(0.17)		(0.48)		(0.29)
Army involvement		0.08		−0.26		−1.81
		(0.84)		(0.52)		(1.34)
Agricultural land cover	−0.002	−0.001	−0.01**	−0.01**	0.01+	0.01*
	(0.004)	(0.004)	(0.005)	(0.005)	(0.003)	(0.003)
Travel time to nearest urban center	0.002**	0.002**	0.002	0.002	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Election fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Event context fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Aggregation	Most-rep.	Most-rep.	Most-rep.	Most-rep.	Most-rep.	Most-rep.
Observations	1,762	1,762	1,762	1,762	1,762	1,762
Log Likelihood	−2,090.34	−2,082.96	−2,315.35	−2,314.97	−815.59	−817.35
θ	0.26** (0.02)	0.27** (0.02)	0.25** (0.01)	0.25** (0.01)		
Akaike Inf. Crit.	4,230.68	4,219.91	4,680.71	4,683.93	1,681.18	1,688.69

Note: Robust standard errors in parentheses, clustered by district/county:
+ p < 0.1, * p < 0.05, ** p < 0.01

Using Conley-adjusted standard errors

Another concern with our main analysis is that the models suffer from spatial autocorrelation in the error terms, as electoral violence events are often clustered in space. To account for this spatial autocorrelation, we use Conley standard errors that adjust the variance-covariance matrix to allow for spatial correlation up to 50 km from the event location. This approach provides more reliable inference when spatial dependence is likely but the exact form of that dependence is unknown. The results are in Table EII and show that police involvement remains associated with less lethal violence ($p < 0.01$) and a lower risk of material destruction ($p < 0.01$), whereas paramilitary police involvement remains associated with more lethal violence ($p < 0.01$). Thus, our main result is robust to accounting for spatial autocorrelation in the error terms.

Table EII: Using Conley standard errors

Model	(34)	(35)	(36)	(37)	(38)	(39)
Dependent Variables:	Lethal violence		Injurious violence		Material destruction	
Variables						
Security force involvement	-0.22		-0.07		-0.82**	
	(0.19)		(0.17)		(0.13)	
Police involvement		-0.41**		-0.13		-0.78**
		(0.10)		(0.17)		(0.09)
Paramililitary police involvement		0.81**		0.27		-0.19
		(0.15)		(0.36)		(0.38)
Army involvement		-0.02		-0.21		-1.80
		(0.78)		(0.51)		(1.31)
Agricultural land cover	0.00	0.00	-0.01**	-0.01**	0.01*	0.01*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Travel time to nearest urban center	0.00**	0.00**	0.00	0.00	0.00**	0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Fixed-effects						
Election	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	1,765	1,765	1,780	1,780	1,780	1,780
AIC	4,289.74	4,279.65	4,879.43	4,881.60	1,706.71	1,714.53

Custom standard-errors in parentheses

*Signif. Codes: **: 0.01, *: 0.05, +: 0.1*

Using Poisson regression

We also estimate Poisson regression models to assess the robustness of our findings. While our primary specification uses negative binomial regression to account for overdispersion in the death and injury counts, Poisson models provide a useful benchmark as they are consistent under weaker assumptions and often perform well when combined with cluster-robust standard errors. This check addresses concerns that the results may be driven by the additional dispersion parameter in the negative binomial model or by potential instability in fixed effects estimation. The results are in Table EIII and show that police involvement remains associated with less lethal violence ($p < 0.01$), whereas paramilitary police involvement remains associated with more lethal violence ($p < 0.05$).

Table EIII: Using Poisson regression

	Dependent variable:			
	Lethal violence		Injurious violence	
	(40)	(41)	(42)	(43)
Security force involvement	−0.21		0.08	
	(0.25)		(0.48)	
Police involvement		−0.49**		−0.01
		(0.14)		(0.52)
Paramilitary police involvement		0.39*		0.35
		(0.17)		(0.47)
Army involvement		0.11		−0.69
		(0.79)		(0.57)
Agricultural land cover	0.003	0.003	−0.01+	−0.01*
	(0.004)	(0.004)	(0.01)	(0.01)
Travel time to nearest urban center	0.004**	0.004**	0.002	0.002
	(0.001)	(0.001)	(0.002)	(0.002)
Election fixed effects	Yes	Yes	Yes	Yes
Aggregation	Most-rep.	Most-rep.	Most-rep.	Most-rep.
Observations	1,780	1,780	1,780	1,780
Log Likelihood	−5,190.65	−5,157.84	−6,838.39	−6,819.48
Akaike Inf. Crit.	10,417.30	10,355.68	13,712.78	13,678.96

Note: Robust standard errors in parentheses, clustered by district/county:
+ p < 0.1, * p < 0.05, ** p < 0.01

Using a Linear Probability Model and binary outcome variables

A final concern is that our main results are sensitive to the functional form assumptions of the negative binomial model. To alleviate this concern, and to ensure that the results are not driven solely by variation in the scale of violence, we replicate our main models using a Linear Probability Model (LPM) using dichotomized outcome variables that capture the occurrence of electoral violence deaths and injuries, respectively. Although LPMs have known limitations, they offer a straightforward way to incorporate fixed effects and cluster-robust standard errors. The results are in Table EIV and show that police involvement is associated with a lesser risk of both lethal violence ($p < 0.05$) and material destruction ($p < 0.01$), whereas paramilitary police involvement is associated with a greater risk of both lethal violence and injurious violence ($p < 0.01$).

Table EIV: Using a Linear Probability Model and binary outcome variables

	Dependent variable:					
	Lethal violence		Injurious violence		Material destruction	
	(44)	(45)	(46)	(47)	(48)	(49)
Security force involvement	−0.01		0.04		−0.13**	
	(0.02)		(0.04)		(0.03)	
Police involvement		−0.04*		0.03		−0.12**
		(0.02)		(0.04)		(0.02)
Paramilitary police involvement		0.15**		0.13**		−0.03
		(0.05)		(0.04)		(0.05)
Army involvement		0.002		0.18		−0.19+
		(0.11)		(0.12)		(0.11)
Agricultural land cover	−0.001	−0.001	0.001*	0.001*	0.001+	0.001+
	(0.001)	(0.001)	(0.001)	(0.001)	(0.0005)	(0.0005)
Travel time to nearest urban center	−0.0001	−0.0001	−0.0001	−0.0001	0.0003**	0.0003**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Election fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Aggregation	Most-rep.	Most-rep.	Most-rep.	Most-rep.	Most-rep.	Most-rep.
Observations	1,780	1,780	1,780	1,780	1,780	1,780
R^2	0.12	0.13	0.20	0.20	0.20	0.20
Adjusted R^2	0.12	0.12	0.19	0.19	0.19	0.19
Residual Std. Error	0.44 (df = 1762)	0.43 (df = 1760)	0.44 (df = 1762)	0.43 (df = 1760)	0.39 (df = 1762)	0.39 (df = 1760)
F Statistic	14.61** (df = 17; 1762)	13.84** (df = 19; 1760)	25.70** (df = 17; 1762)	23.50** (df = 19; 1760)	25.60** (df = 17; 1762)	22.59** (df = 19; 1760)

Note: Robust standard errors in parentheses, clustered by district/county:
+ p < 0.1, * p < 0.05, ** p < 0.01

References

Birch, Sarah and David Muchlinski (2017). "The Dataset of Countries at Risk of Electoral Violence". *Terrorism and Political Violence* 32 (2), 217–236.

Cook, Scott J. and Nils B. Weidmann (2019). "Lost in Aggregation: Improving Event Analysis with Report-Level Data". *American Journal of Political Science* 63 (1), 250–264.

Coppedge, Michael et al. (2025). "V-Dem Codebook v15". Codebook. Gothenburg: Varieties of Democracy (V-Dem) Project.

Daxecker, Ursula, Elio Amicarelli, and Alexander Jung (2019). "Electoral Contention and Violence (ECAV): A New Dataset". *Journal of Peace Research* 56 (5), 714–723.

Fjelde, Hanne and Kristine Höglund (2022). "Introducing the Deadly Electoral Conflict Dataset (DECO)". *Journal of Conflict Resolution* 66 (1), 162–185.

Hyde, Susan D. and Nikolay Marinov (2012). "Which Elections Can be Lost?" *Political Analysis* 20 (2), 191–210.

van Baalen, Sebastian and Kristine Höglund (2025). "Trials and Triangulations: Analyzing Aggregation Sensitivity in Event Data on Political Violence". Working paper. Uppsala: Uppsala University.

Wahman, Michael (2024). "How Strategic Violence Distorts African Elections". *Journal of Democracy* 35 (2), 108–121.