Online Appendix

The Reach of Radio: Ending Civil Conflict through Rebel Demobilization

Alex Armand, Paul Atwell, Joseph F. Gomes

This appendix is divided into eight sections. Section A presents a summary of the data used in the paper. Section B shows additional results that were omitted from the main paper due to space constraints. Sections C and D present additional information about the surveys of radio stations and of LRA returnees. Section E provides additional details about the content of defection messages. Section F presents the processes and institutions involved in the return and reintegration of LRA rebels into civil society. Section G discusses the generalizability of the results presented in the paper. Section H discusses differences between the recruitment and defection processes.

A Summary of data and variable definition

The following table presents a description (including sources and variable transformations) for the variables and data used in the paper. The variables from the LRACT database used as main outcome variables are described under *Conflict intensity*. Several variables are obtained from the PRIO-GRID version 2.0 database (Tollefsen et al., 2012), an open access gridded dataset compiling data from multiple third-party sources, detailed in the table.

Variable (Source)	Description
Basemaps (Esri)	Basemaps throughout the paper were created using ArcGIS® software by Esri®. Basemaps are used in line with the Esri Master License Agreement, specifically for the inclusion of screen captures in academic publications. We make use of the *World Topographic Map* (sources: Esri, HERE, Garmin, Intermap, INCREMENT P, GEBCO, USGS, FAO, NPS, NRCAN, GeoBase, IGN, Kadaster NL, Ordnance Survey, Esri Japan, METI, Esri China (Hong Kong), ®OpenStreetMap contributors, GIS User Community) and the *World Light Gray Base* (sources: Esri, HERE, Garmin, ®OpenStreetMap contributors, and the GIS User Community).
Commodity prices in international markets (GEM and USGS)	Data obtained from the Global Economic Monitor (The World Bank, 2019) and the U.S. Geological Survey's Historical Statistics for Mineral and Material Commodities (USGS, 2016a).
Conflict intensity (LRACT, ACLED, UCDP)	Number of violent events (and fatalities) in each cell for a specific year. Data are obtained from three distinct event-based databases: the LRA Crisis Tracker (LRACT) (The Resolve, 2015), the Uppsala Conflict Data Program (UCDP) (Sundberg and Melander, 2013) and the Armed Conflict Location & Event Data Project (ACLED) (Raleigh et al., 2010). For the LRACT, the following variables are used: - *Fatalities*: total number of civilians killed in an incident. - *Returnees*: sum of adult and child returnees. These are defined as adult and child escapees, who return from armed group captivity or enrollment willingly. This category excludes armed group members captured and civilians released. - *Abductees*: number of people "taken captive against their will by the LRA for any period of time, including short-term abductions".

(continued on next page)

Variable (Source)	Description
	- *Violence against civilians*: number of events characterized by "any physical violence committed against civilians which resulted in death or injury, including sexual or gender based violence". - *Clash*: number of events in which "an armed group violently engage with one or more armed groups or security forces (any organized, armed, non-rebel or terrorist group)". - *Looting*: number of events in which "LRA members commit robbery, extortion, or destruction of property".
Crop coverage (Monfreda et al., 2008)	Share of the cell covered by a crop. M3-Crops Data offers a raster dataset at the $5' \times 5'$ latitude/longitude grid for 175 crops in the period 1997–2003.
Distance from border (PRIO-GRID)	Spherical distance (in kilometer) from the PRIO-GRID cell centroid to the territorial outline of the country the cell belongs to. Data is downloaded from the PRIO-GRID version 2.0 database (Tollefsen et al., 2012).
Forest and urban cover (PRIO-GRID)	Share of the cell covered by forested area or urban areas in 2009. Data extracted from the Globcover 2009 dataset v.2.3 (Bontemps et al., 2009) following FAO land cover classification system. Data is downloaded from the PRIO-GRID version 2.0 database (Tollefsen et al., 2012).
Ethnolinguistic diversity (Desmet et al., 2018)	Cell-level ethnolinguistic distance, fractionalization (Easterly and Levine, 1997; Alesina et al., 2003) and polarization (Esteban and Ray, 1994; Montalvo and Reynal-Querol, 2005) computed using the distribution of language groups at the resolution of 5 km $\times$ 5 km from Desmet et al. (2018).
General FM radio coverage (UKW/TV-Arbeitskreis, 2018)	Percentage of cell covered by topography-corrected FM radio signals from any FM radio station (excluding defection messaging broadcasting stations) in the region. When technical parameters are missing, we impute them using an iterative imputation using the median value of the missing parameter. We start at the smallest level (region), and we increase the level if the number of observations is smaller than 10. The largest level is the full study area.
Infant mortality (PRIO-GRID)	Number of children per 10,000 live births that die before reaching their first birthday. Measured as average rate within the grid cell at a resolution of 0.5×0.5 decimal degrees and only for the year 2000. Data available from the SEDAC Global Poverty Mapping project (Storeygard et al., 2008). Information available only for a sub-set of the study area. Data is downloaded from the PRIO-GRID version 2.0 database (Tollefsen et al., 2012).
Military presence (ACLED, UCDP)	Indicator variable equal to 1 if at time t in a cell at least one event is recorded in either ACLED or UCDP in which a security force is involved and 0 otherwise.
Mineral presence (USGS, 2016b)	Dummy variable indicating whether a given mineral is present in the cell. Data is obtained from the Mineral Resource Data System (MRDS) database (USGS, 2016b). It provides geo-located extraction sites by type of mineral and the magnitude of production.
Mobile phone coverage (GSMA, 2012)	Dummy variable equal to 1 if at time t the cell is covered by the 2G (GSM) network. Data come from the Collins Mobile Coverage Explorer, supplied by GSMA and Collins Bartholomew. The dataset provides geo-located information on yearly mobile phone coverage for 2G (GSM), 3G and 4G (LTE) networks on a global basis. It is built using submissions from Mobile Network Operators and is then aggregated. The resolution of coverage varies from 1 km^2 to 15–23 km^2.
Nightlight (PRIO-GRID)	Average nighttime light emission from the DMSP-OLS Nighttime Lights Time Series v.4 (National Oceanic and Atmospheric Administration, 2014). Image and data processing by NOAA's National Geophysical Data Center. Data is downloaded from the PRIO-GRID version 2.0 database (Tollefsen et al., 2012).

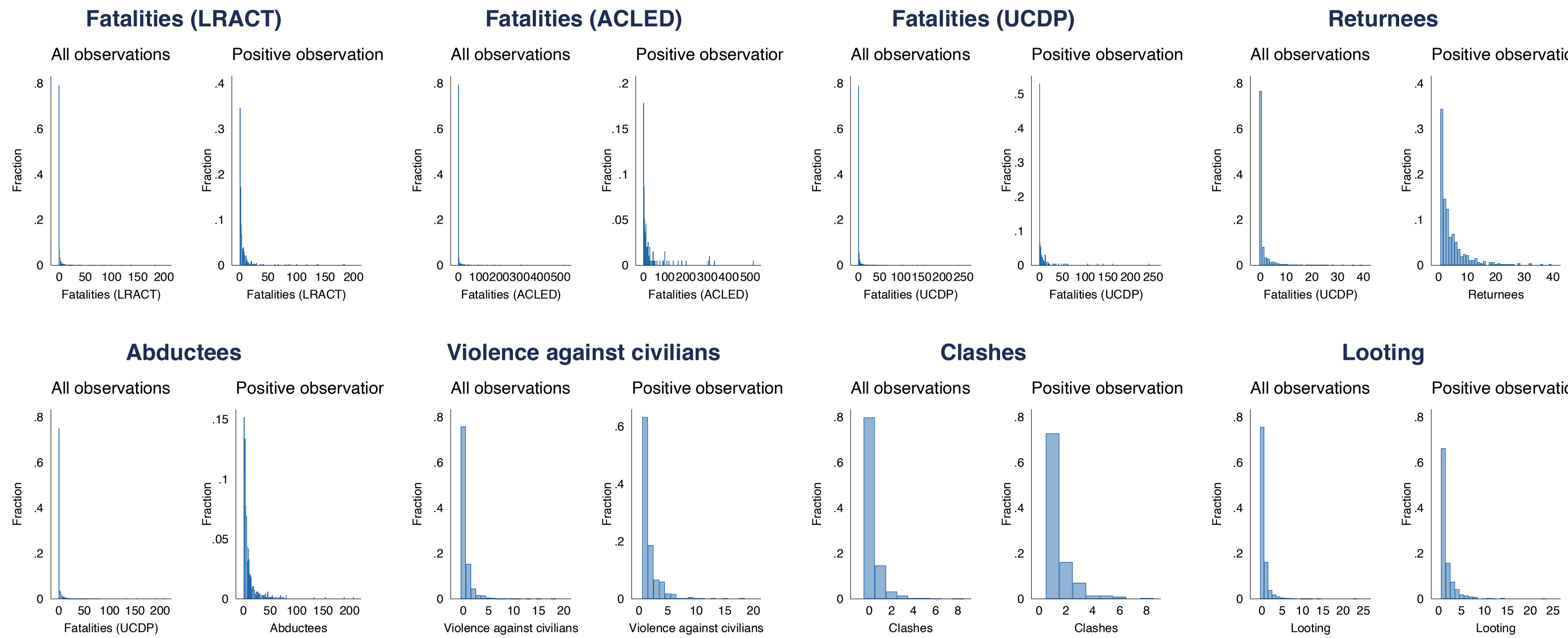
(continued on next page)

Variable (Source)	Description
Population (PRIO-GRID)	Population in each cell over time. Data available from the Gridded Population of the World version 3 (CIESIN-CIAT, 2005). Available for the years 1990, 1995, 2000, and 2005. Data is downloaded from the PRIO-GRID version 2.0 database (Tollefsen et al., 2012).
Precipitation (CHIRPS)	Average amount of daily precipitations (in mm) in the cell, based on daily precipitations data provided by the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) database (Funk et al., 2015). CHIRPS provides $0.05° \times 0.05°$ resolution satellite imagery supplemented with in-situ monitoring station data. To ensure comparability of the measure across cells, we use double-standardized rainfall deviations (Hidalgo et al., 2010). We first account for seasonal patterns by standardizing monthly rain totals by cell and month for the period 2000–2015. For each cell, these indicators are then summed up by year and standardized over the same period. We then use the absolute value of standardized rainfall as the main measure to capture the non-monotonic relationship between rainfall and income changes. We also use alternative functions of rainfall deviations (such as its square), linearly and non-linearly de-trended rainfall (Fujiwara et al., 2016), measures of growing-season-specific rainfall, or current and lagged year-on-year precipitation growth (Miguel et al., 2004; Ciccone, 2011).
Share of year experiencing drought (PRIO-GRID)	Proportion of months out of 12 that are part of the longest streak of consecutive months ending in the given year with SPI1 values below -1.5. Data are built using the Standardized Precipitation and Evapotranspiration Index SPEI1 from the SPEI Global Drought Monitor (Beguería et al., 2014). Data is downloaded from the PRIO-GRID version 2.0 database (Tollefsen et al., 2012).
Temperature (PRIO-GRID)	Yearly mean temperature (in degrees Celsius) in the cell, based on monthly meteorological statistics from GHCN/CAMS (Fan and van den Dool, 2008). Data are available for the period 1948–2014. To ensure comparability of the measure across cells, we use standardized temperature deviations, by restricting the standardization to the year level. Data is downloaded from the PRIO-GRID version 2.0 database (Tollefsen et al., 2012).
Terrain Ruggedness (Nunn and Puga, 2012)	Terrain ruggedness calculated at the level of 30 arc-second cells on a regular geographic grid covering the Earth.

Notes. For time-varying variables, missing values are linearly interpolated.

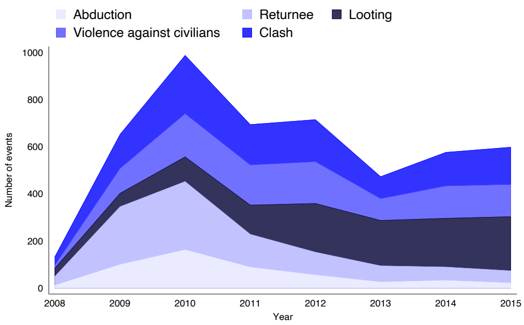
Figure A1 presents the distributions of the different types of LRA-related events including fatalities, returnees, and abductees, among others. For each category, the figure presents the unconditional distribution and the distribution conditional on the concerned variable being positive (i.e. excluding observations with value 0). Figure A2 presents the series of total events associated with the LRA by type of incident (left panel), and the series of the number of fatalities, returnees, and abductees over the period of analysis (right panel).

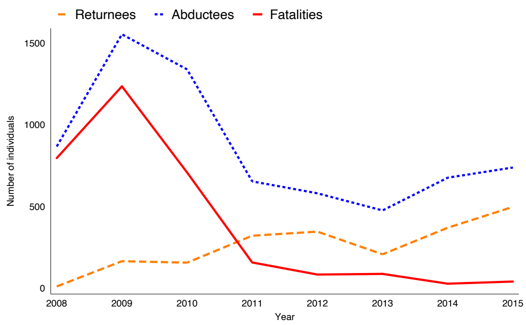
Figure A1: Distribution of the number of fatalities and events



Notes. The figure plots the distributions of the number of fatalities, returnees and abductees, and of different types of events. For each category, the unconditional distribution and the distribution conditional on being positive (excluding cells with values equal to zero) are presented. Observations that are equal to zero throughout the period of analysis are excluded from the distribution. The time period is restricted to 2008–2015 and the cell resolution is equal to $0.125° \times 0.125°$. Data source: LRACT, ACLED and UCDP for fatalities, LRACT for the other variables. See Appendix A for further information on the variables.

Figure A2: Composition of LRA-related events and the number of individuals involved





Notes. The figure plots the time series of different conflict events. The left panel presents the composition of total events per year, while the right panel focuses on the number of returnees, abductees and fatalities. Source: LRACT database. See Appendix A for further information on the variables.

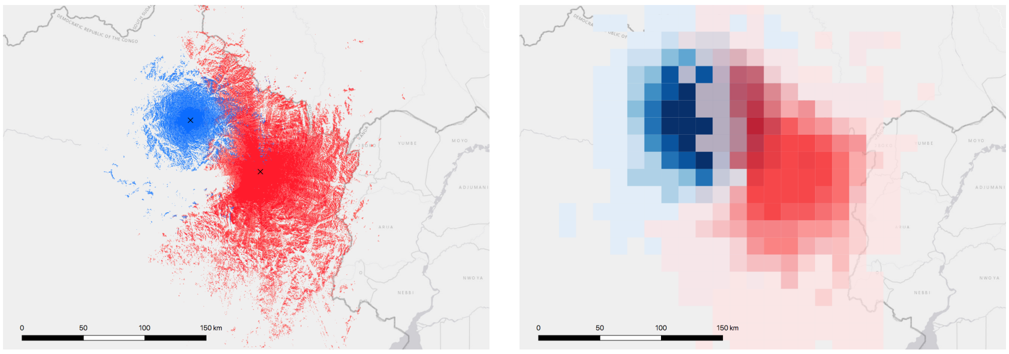
B Additional analysis

B.1 Radio coverage

To compute radio coverage for each radio station we apply the field standard radio propagation model, i.e., the Longley–Rice or Irregular Terrain Model. The model uses the geographic coordinates and technical parameters as inputs to generate coverage of each radio station over space. The technical parameters primarily include the mast/tower height and the transmitter power which we collected from the radio survey. We partnered with a radio engineer to finalize all other parameters needed to adapt the model to the specific setting. Coverage is calculated using CloudRF (`cloudrf.com`), a commercial radio planning tool. Figure B1 presents an example output from the model for two different radio stations. Radio coverage is first computed at the finest resolution of 90 meters (left panel), and then merged to the grid in the study area in order to build percent coverage of each radio at the cell level (right panel). Conditional on distance from each antenna, variation in topography generates not only plausibly random coverage of the signal, but also plausibly random overlap of different signals. These features, together with frequencies of messaging, are exploited when building intensity of messaging (equation 1). Figure B2 plots the evolution of exposure or coverage (left figure) and the intensity (right figure) of defection messaging content over time.

Figure B1: Radio coverage: an illustration

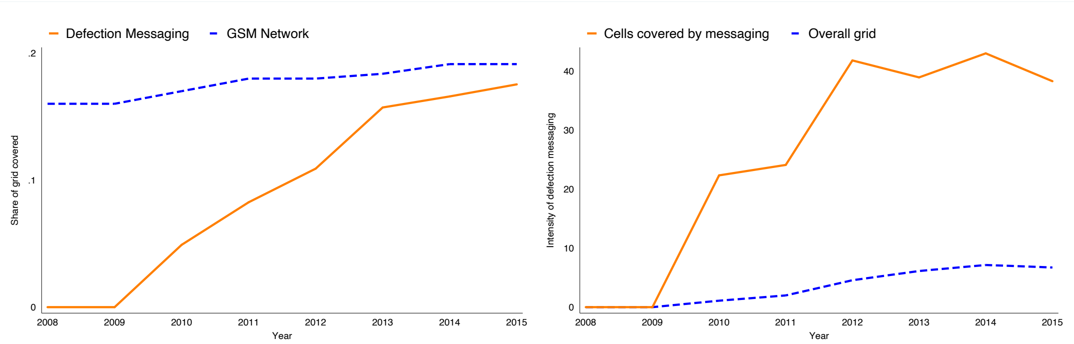
90-meter resolution **0.125° × 0.125° resolution**



Notes. Example of topography-corrected radio coverage (computed with the Longley–Rice or Irregular Terrain Model) for two antennas ("×" indicates the location of an antenna). The left figure shows radio coverage at a 90-meter resolution, while the right figure shows radio coverage using grid cells of 0.125° × 0.125° resolution. Each cell is assigned with the value of the share of the cell covered by the signal, with darker colors indicating larger shares. See Appendix A for further information on the variables. Basemap source: Esri (see Appendix A for details and attributions).

Figure B2: Annual coverage and intensity of defection messaging

Coverage | Intensity of defection messaging



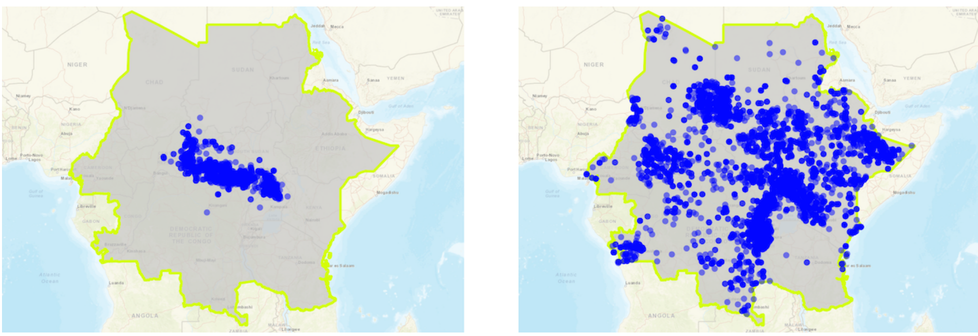
Notes. The left figure shows the share of cells that are covered by radio signals from defection messaging stations and by the GSM mobile-phone network. The right figure presents the intensity (averaged at the grid-cell level) of defection messaging, as defined by equation (1). Source: own elaboration. See Appendix A for further information on the variables.

B.2 Geographical extent of study

Figure B3 presents the geographical extent of violence in the region of the study. As no clear precedent exists in the literature, we select the geographic extent of the analysis using a rule based on the geographic distribution of LRA-related events during the 1997–2015 period. While the study focuses on the 2008–2015 period, we use a longer period to take into account the entire area in which the LRA has historically operated. We select a geographical area that is defined by the 1st percentile minus 0.5° and 99th percentile plus 0.5° of both latitude and longitude of the events. The parameter 0.5 is chosen to allow a buffer around the events that fall on the edge of the grid. Estimates are robust to variation in this parameter (Figure B4).

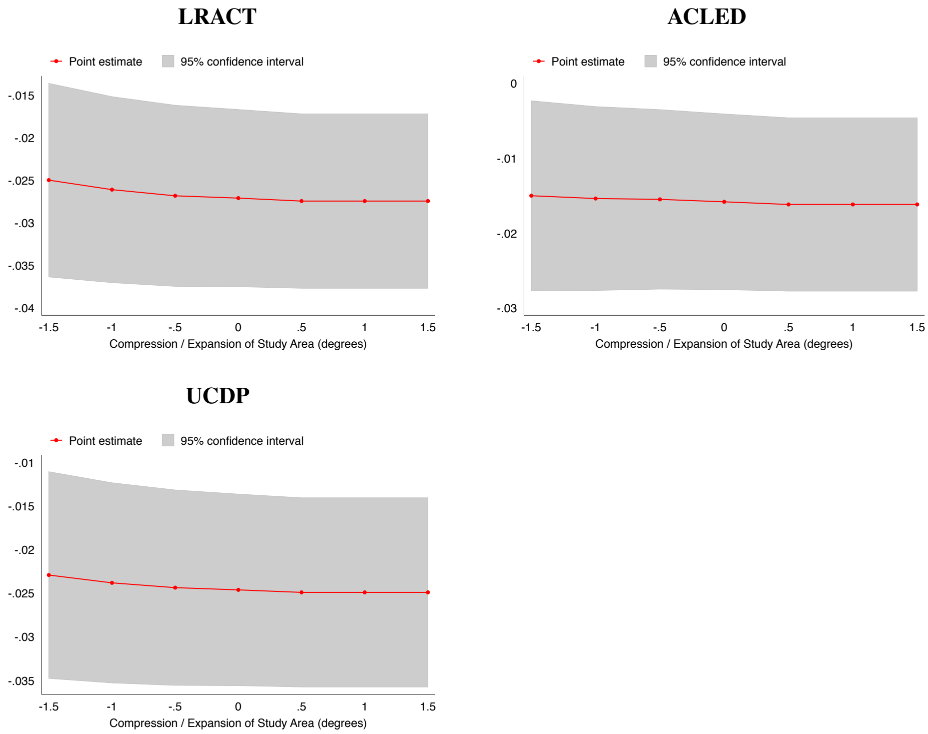
Figure B3: Extent of violence in the region (1989–2015)

LRA violence | **Non-LRA violence**



Notes. The figures present the geographical distribution of violent events throughout central Africa. Each dot represents an event as defined in the UCDP dataset. In the left figure, dots are LRA-related violent events, while in the right figure, dots are non-LRA events. Basemap source: Esri (see Appendix A for details and attributions).

Figure B4: Sensitivity of estimates to the size of the study area



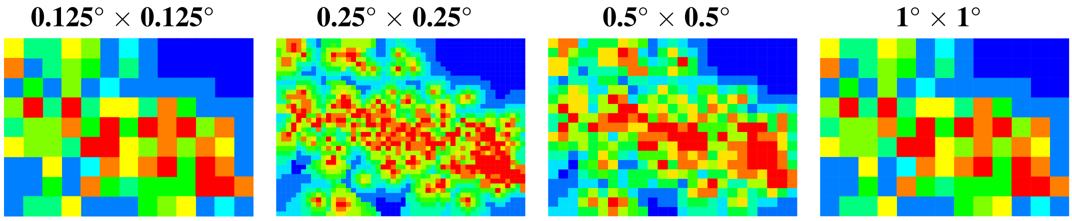
Notes. The figures plot estimates and 95% confidence intervals of the effect of intensity of messaging on the number of fatalities when the size of the study area changes. The dependent variables are the log-transformed number of fatalities using the LRACT, ACLED and UCDP datasets (adding one unit before taking logarithms to accommodate 0 values). The expansion and compression of the area is reported on the horizontal axis in degrees. Estimates are based on equation (2). See Appendix A for further information on the variables.

B.3 Cell size analysis and the Modifiable Areal Unit Problem

The objective of this section is to determine the correct grid to be used for the analysis. We are analyzing a two-dimensional spatial point pattern S, defined as a set of points s_i $(i = 1, ..., n)$ and located in a two-dimensional region R. Each point represents the location in R of a violent event where the LRA is an actor and has coordinates (s_{i1}, s_{i2}). In this setting, a grid is a regular tessellation of the study region R that divides it into a set of contiguous cells. We discuss issues related to the selection of the region R in Section B.2. Figure B5 presents the geographic distribution of the probability of observing an LRA event in a specific cell in the study area for four alternative resolutions: $0.125° \times 0.125°$ (high resolution), $0.25° \times 0.25°$, $0.5° \times 0.5°$, and $1° \times 1°$ (low resolution). Given the clustering and the full extent of the events observed, a finer resolution allows capturing a much larger variation compared to lower resolutions. Aggregating cells affects the overall information contained by the grid.

While no ideal resolution exists, grid resolution can be related to the geometry of point patterns (Hengl, 2006). According to Boots and Getis (1988), the grid resolution should be at most half the average of the mean/median shortest distance, i.e. the mean spacing between the closest point

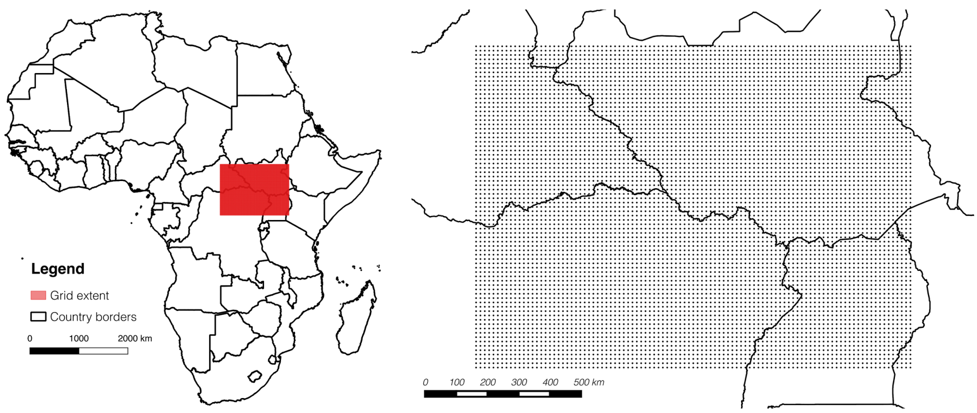
Figure B5: Probability of observing a violent event per cell, by cell resolution



Notes. The distribution of the probability of observing an LRA-related event in a specific cell for four alternative resolutions. The probability density function is estimated using a Kernel estimator assuming a quartic distribution where the bandwidth is determined using the minimum number of data points method (k = 1). Cell resolution is expressed in degrees per side. Results are produced using Stata command *spkde*. We use the ACLED dataset for the period 1996–2015 to compute these statistics, since it allows observing events in the period before 2008 and it provides higher geographical dispersion of events with respect to UCDP (see Section B.2). Results are similar using the LRACT for the post-2008 period only.

pairs. When we include (exclude) events taking place in the same location, the corresponding median is approximately 22 km (25 km). We therefore select a cell resolution of 0.125° × 0.125° (approximately 14 km × 14 km at the equator). Figure B6 plots the area covered by the grid and illustrates its resolution.

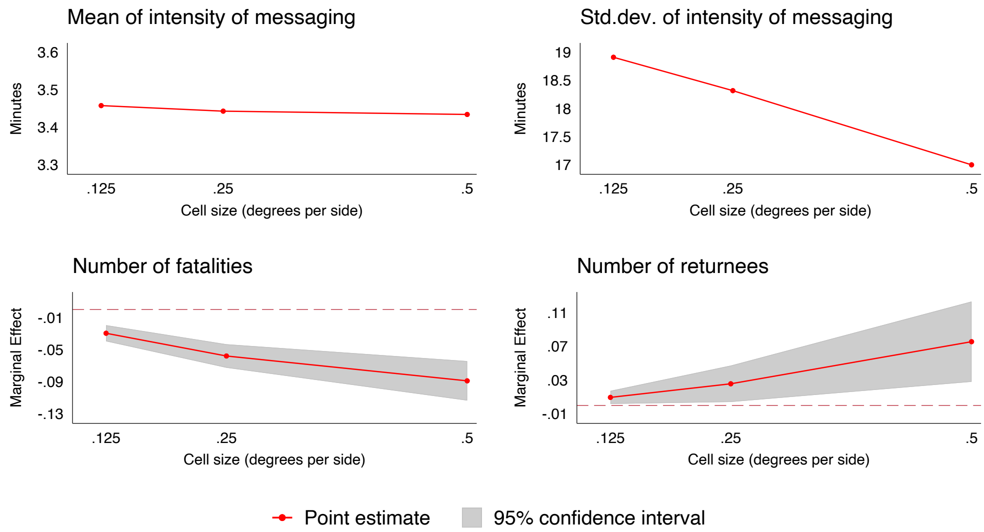
Figure B6: Geographical coverage and grid resolution



Notes. The left panel shows the area covered by the grid. The right panel shows its resolution. Each dot represents the centroid of a 0.125° × 0.125° cell.

The Modifiable Areal Unit Problem (MAUP) occurs when cell sizes are chosen in order to provide a pre-selected type of result. We construct cells of 0.125° × 0.125°, 0.25° × 0.25°, and 0.5° × 0.5°, and we estimate the main specification for each of these grids. Figure B7 shows how intensity of messaging and estimates for the effect of intensity of messaging on fatalities and returnees change with different cell sizes. While the mean intensity of defection messaging is relatively unchanged across resolutions, the standard deviation decreases from around 19 when the cell size is 0.125° × 0.125° to 17 when the cell size is 0.5° × 0.5° (upper panel). Estimates of the marginal effect of intensity of messaging (lower panels) increase with the cell size. In line with Fotheringham and Wong (1991), the increase in the coefficient following aggregation is explained by the reduction in variation due to averaging across cells. In this case, the correlation between two variables is expected to increase when the variance is reduced and the covariance is stable.

Figure B7: Cell size, intensity of messaging and estimates of the effect



Notes. The upper panels show how the mean and standard deviation of intensity of messaging, defined by equation (1), varies with the cell size (reported on the horizontal axis). The lower panels show how estimates of equation (2) vary when the cell size changes. We consider as outcome variables the number of fatalities and the number of returnees from the LRACT database. Both variables are log-transformed (adding one unit before taking logarithms to accommodate 0 values). See Appendix A for further information on the variables.

B.4 Variable-specific trends

To control for differential trends associated with determinants of conflict, we add interaction terms between the year of observation and cell-level terrain ruggedness, ex-ante income proxied by nightlight, ex-ante log-population, urban and forest cover, and country indicator dummies. Table B1 presents the results. Similar results are obtained with interactions with year dummies to allow for non-linear time effects. In both cases, results are unaffected.

Table B1: Robustness to adding variable-specific trends

Dependent variable:	**Number of fatalities linked to LRA activity**				
	(1)	(2)	(3)	(4)	(5)
Intensity of messaging	-0.028	-0.028	-0.028	-0.028	-0.028
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Ruggedness x Year	0.001				
	(0.001)				
Nightlight (2007) x Year		0.000			
		(0.000)			
Population (2005) x Year			0.000		
			(0.000)		
Urban cover (2009) x Year				0.001	
				(0.001)	
Forest cover (2009) x Year				0.000	
				(0.000)	
CAR x Year					0.003
					(0.001)
DRC x Year					0.000
					(0.001)
Uganda x Year					0.002
					(0.001)
South Sudan x Year					0.001
					(0.001)
Observations	60600	60600	60600	60600	60600
Number of cells	7575	7575	7575	7575	7575

Notes. Standard errors in parentheses are allowed to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). The intensity of messaging variable is defined by equation (1), and is standardized. The dependent variable is the log number of total fatalities (adding 1 to accommodate zero values) linked to the LRA. All specifications include cell and year FE, propagation controls, and additional controls. Additional controls include commodity price and weather shocks (see Section III). The time period is restricted to 2008–2015. See Appendix A for further information on the variables.

B.5 Robustness to alternative normalizations of fatalities and alternative samples

Table B2 presents results of the effect of intensity of messaging on LRA-related fatalities, when the dependent variable is normalized using alternative measures. We focus on population, available land, density, and non-urban land. Results are robust to using these alternative definitions.

Table B2: Robustness to alternative normalizations of fatalities

Dependent variable:	Number of fatalities linked to LRA activity				
Normalization by:	Population	Available land	Density	Density (corrected)	Non-urban land
	(1)	(2)	(3)	(4)	(5)
Intensity of messaging	-0.026	-0.002	-0.013	-0.011	-0.021
	(0.005)	(0.000)	(0.002)	(0.002)	(0.004)
Observations	60600	60600	60600	60600	60600
Number of cells	7575	7575	7575	7575	7575

Notes. Standard errors in parentheses are allowed to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). The intensity of messaging variable is defined by equation (1), and is standardized. The dependent variable is the number of fatalities linked to LRA, computed from the LRACT dataset, and normalized according to different criteria. Normalization by "Population" refers to number of fatalities per thousand inhabitants. Normalization by "Available land' refers to number of fatalities per square kilometer of land not covered by forest. Normalization by "Density" refers to number of fatalities divided by inhabitants per square kilometer. Normalization by "Density (corrected)" refers to number of fatalities divided by inhabitants per square kilometer of land not covered by forest. Normalization by "Non-urban land" refers to number of fatalities per square kilometer not covered by urban area. All specifications include cell and year FE, propagation controls, additional controls and macro-region-specific time fixed effects. Additional controls include commodity price and weather shocks (see Section III). The time period is restricted to 2008–2015. See Appendix A for further information on the variables.

Table B3 presents results of the effect of intensity of messaging on LRA-related fatalities by restricting the sample to cells away from the location of the antennas (at least 50km away) and to cells within 250km from the antennas. Results are robust to these alternative samples.

Table B3: Robustness to alternative samples

Dependent variable:	Number of fatalities linked to LRA activity							
Event Dataset:	LRACT		ACLED		UCDP		Combined	
Distance from antenna (km):	>50	0-250	>50	0-250	>50	0-250	>50	0-250
Specification:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intensity of messaging	-0.028	-0.025	-0.017	-0.016	-0.024	-0.022	-0.033	-0.024
	(0.007)	(0.006)	(0.004)	(0.003)	(0.006)	(0.005)	(0.008)	(0.005)
Observations	58072	21357	58072	21357	58072	21357	58072	21357
Number of cells	7492	3911	7492	3911	7492	3911	7492	3911

Notes. The table reports marginal effects estimated using a fixed effects model. Standard errors in parentheses are allowed to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). The independent variable is intensity of defection messaging, defined by equation (1), and is standardized. The dependent variable is the number of fatalities linked to LRA. Columns 1 to 6 use fatalities data from the LRACT, ACLED and UCDP databases. Columns 7-8 use the mean number of fatalities from across the three datasets. The number of fatalities is reported in logarithms (adding one unit before taking logarithms to accommodate 0 values). All specifications include cell and year FE, propagation controls, additional controls and interactions between year dummies and macro-region dummies. Additional controls include commodity price and weather shocks (see Section III). The time period is restricted to 2008–2015. See Appendix A for further information on the variables.

B.6 Contemporaneous versus lagged intensity of messaging

In the main text we focused on the contemporaneous effect of defection messaging on different indicators of violence and of LRA strategic behavior. Table B4 presents estimates of equation (2) using lagged values of intensity of messaging, and using both contemporaneous and lagged values of intensity of messaging. Table B5 presents serial correlation of intensity of messaging in the period 2008-2015.

Table B4: Defection messaging and timing of intensity of messaging

Dependent variable:	**Fatalities**	**Number of individuals...**		**Number of events involving...**		
		Returning	Being abducted	Violence against civilians	Clashes	Looting
	(1)	(2)	(3)	(4)	(5)	(6)
A. Lagged						
Intensity of messaging (t - 1)	-0.035	0.005	-0.011	-0.020	-0.011	0.012
	(0.006)	(0.004)	(0.005)	(0.005)	(0.003)	(0.004)
B. Contemporaneous and lagged						
Intensity of messaging	-0.006	0.011	0.005	0.007	0.003	0.015
	(0.007)	(0.005)	(0.007)	(0.005)	(0.004)	(0.004)
Intensity of messaging (t - 1)	-0.031	-0.003	-0.014	-0.025	-0.013	0.002
	(0.009)	(0.005)	(0.007)	(0.007)	(0.005)	(0.004)
Observations	60600	60600	60600	60600	60600	60600
Number of cells	7575	7575	7575	7575	7575	7575

Notes. Standard errors in parentheses are allowed to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). The intensity of messaging variable is defined by equation (1), and is standardized. The dependent variables are the number of fatalities linked to LRA (column 1), the number of individuals returning or being abducted (columns 2–3) and the number of violent events involving different LRA activities (columns 4–6). The dependent variables are measured in logarithms (adding one unit before taking logarithms to accommodate 0 values). All specifications include cell and year FE, propagation controls, additional controls and interactions between year dummies and macro-region dummies. Additional controls include commodity price and weather shocks (see Section III). The time period is restricted to 2008–2015. Panel A estimates equation (2) using only the lagged value of intensity of messaging, while panel B includes both contemporaneous and lagged intensities. See Appendix A for further information on the variables.

Table B5: Serial correlation in intensity of messaging

Period	-	1 lag	2 lags	3 lags	4 lags
-	1.000				
1 lag	0.863	1.000			
2 lag	0.766	0.833	1.000		
3 lag	0.636	0.714	0.756	1.000	
4 lag	0.572	0.600	0.653	0.721	1.000

Notes. Serial correlation is computed using the study area. The time period is restricted to 2008–2015.

B.7 Spillovers across fighters

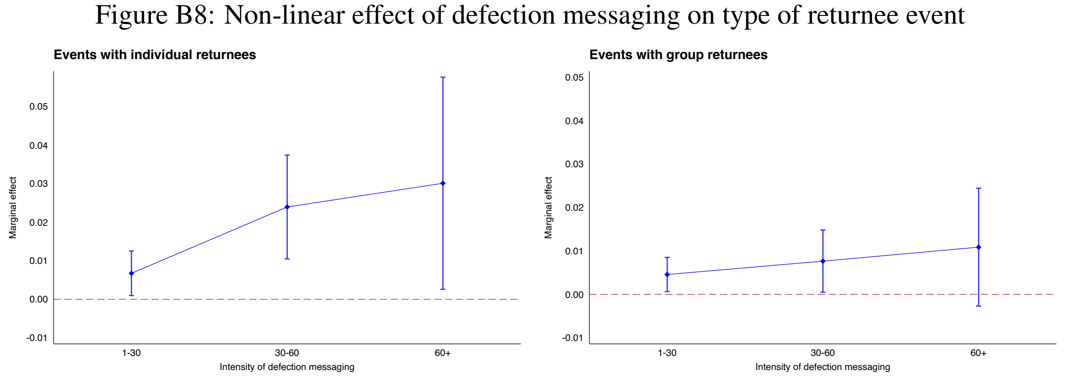
Table B6 presents estimates of the effect of intensity of messaging on the number of events characterized by returnees, distinguishing by the number of returnees in each event. We consider the events in which one or two individuals return as being motivated by individual behavior (individual returnees). We consider events characterized by a larger number of returnees as being motivated by social interaction (group returnees). Table B6 shows that defection messaging primarily increases the number of events characterized by individual returnees rather than group defections. Figure B8 plots the non-linear effects.

Table B6: Effect of defection messaging on the type of returnee event

Dependent variable:	Number of events characterized by...					
	Any number of returnees		Individual returnees		Group returnees	
	(1)	(2)	(3)	(4)	(5)	(6)
Intensity of messaging	0.006	0.006	0.005	0.005	0.001	0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)
Observations	60600	60600	60600	60600	60600	60600
Number of cells	7575	7575	7575	7575	7575	7575

Notes. Standard errors in parentheses are allowed to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). The intensity of messaging variable is defined by equation (1), and is standardized. The dependent variable is the number of events characterized by returnees, distinguishing by the number of returnees per event. The number of events is reported in logarithms (adding one unit before taking logarithms to accommodate 0 values). Individual returnees are events characterized by 1 or 2 returnees. Group returnees are events characterized by 3 or more returnees. All specifications include cell and year FE, propagation controls, additional controls and interactions between year dummies and macro-region dummies. Additional controls include commodity price and weather shocks (see Section III). The time period is restricted to 2008–2015. See Appendix A for further information on the variables.

Figure B8: Non-linear effect of defection messaging on type of returnee event



Notes. The figure plots the coefficients of equation (2) against the intensity of defection messaging. The intensity of defection messaging is decomposed into four dummy variables denoting different intervals of radio messaging intensity: 0 minutes, 0–30 minutes, 30–60 minutes and more than 60 minutes of daily messaging. The dummy for "0 minutes" serves as the excluded category. The dependent variable is the number of events characterized by returnees, distinguishing by the number of returnees per event. Individual returnees are events characterized by 1 or 2 returnees. Group returnees are events characterized by 3 or more returnees. The dependent variables are measured in logarithms (adding one unit before taking logarithms to accommodate 0 values). For comparison, we allow the vertical axis to vary at the same scale for all outcome variables. Confidence intervals are computed at 95% of confidence, standard errors are allowed to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). All specifications include cell and year fixed effects, propagation controls, additional controls and interaction terms between year and macro-region indicators. Additional controls include commodity price and weather shocks (see Section III). The time period is restricted to 2008–2015. See Appendix A for further information on the variables.

This result stands in contrast to findings from the survey of ex-combatants we conducted in DRC (Appendix D), which suggests high exposure to radio messages through peers (see discussion in Section IV.A). While radio messages reach the individual fighters either directly because they themselves listen to the radio or indirectly from fellow combatants who share the information they have heard on the radio, the final decision to defect is primarily an individual decision.

Any planned or failed defection bears the highest penalty than any other infraction for an LRA member. It is common knowledge that if superiors were to discover such a plan or catch a defector, the offender would be tortured and possibly killed. Gates and Nordås (2015) highlight the harsh nature of punishments meted out by the LRA as a deterrent to defection. Thus, the risks of attempting to coordinate any defection, holding aside a group defection, are notably high. LRA members would have known that if a plan to defect reaches one individual who for any reason is more likely to inform superiors, the would-be defectors would pay the ultimate price. By maintaining a private and individual escape plan, this risk is minimized. In contrast, consuming and sharing defection information, while also proscribed by the LRA, is associated with lower penalties. For example, leadership has responded in the past by confiscating radios from rank and file troops. We find that only 16% of respondents said they knew of someone being punished for listening to the radio (Appendix D). This offense clearly belongs to a different category than defection itself. Discussing defection itself can be re-construed in purpose by the information sharer. Whereas the motives for an actual escape are unmistakable, sharing information could even conceivably be part of establishing the LRA's counter-narrative.

B.8 Spillovers on other violent events

Table B7 presents estimates of the effect of intensity of messaging on the number of violent events separated by whether the LRA was involved or not, in columns 1–4, and by whether the LRA is the perpetrator or in the receiving end of the violence, in columns 5–8.

Table B7: Effect of defection messaging on LRA versus non-LRA activity

Dependent variable:	**Number of violent events by ...**							
	Actor involved in the violent event				**LRA role in the violent event**			
	LRA was involved		*LRA was not involved*		*LRA attacked*		*LRA was attacked*	
Event dataset:	ACLED	UCDP	ACLED	UCDP	ACLED	UCDP	ACLED	UCDP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intensity of messaging	-0.007	-0.011	-0.001	-0.001	-0.007	-0.010	-0.003	-0.003
	(0.002)	(0.002)	(0.001)	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)
Observations	60600	60600	60600	60600	60600	60600	60600	60600
Number of cells	7575	7575	7575	7575	7575	7575	7575	7575

Notes. Standard errors in parentheses are allowed to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). The intensity of messaging variable is defined by equation (1), and is standardized. The dependent variable is the number of violent events in which at least one actor is the LRA, in columns 1–2, in which none of the actors is the LRA, in columns 3–4, in which LRA is attacking, in columns 5–6 and in which LRA is being attacked, in columns 7–8. The number of events is reported in logarithms (adding one unit before taking logarithms to accommodate 0 values). All specifications include cell and year FE, propagation controls, additional controls and interactions between year dummies and macro-region dummies. Additional controls include commodity price and weather shocks (see Section III). The time period is restricted to 2008–2015. See Appendix A for further information on the variables.

Table B8 investigates whether intensity of defection messaging is capturing military presence. Since military operations are not directly observable to us, we rely on event-based data from ACLED and UCDP identifying events in which the security forces are involved (including nonviolent events) to construct a proxy for military presence. Columns 1 and 2 provide estimates of equation (2) where the dependent variable is the number of events in which a security force is the perpetrator of the action. In columns 3–6, we estimate the effect of intensity of messaging on the number of fatalities linked to LRA activity using equation (2) and controlling for (potentially endogenous) military presence. We define army presence using a dummy variable equal to 1 if at time t in a cell at least one event is recorded in either ACLED or UCDP in which a security force is involved and 0 otherwise. Estimates are unaffected.

Table B8: Effect of defection messaging and army presence

Dependent variable:	**N. of violent events in which the army is the perpetrator**		**Number of fatalities linked to LRA activity**			
Event dataset:	UCDP	ACLED	LRACT	LRACT	LRACT	LRACT
	(1)	(2)	(3)	(4)	(5)	(6)
Intensity of messaging	-0.003	-0.003	-0.027	-0.027	-0.027	-0.026
	(0.001)	(0.001)	(0.005)	(0.005)	(0.005)	(0.005)
Army presence (ACLED)			0.061	0.197		
			(0.014)	(0.050)		
* Minimum distance				-0.000		
				(0.000)		
Army presence (UCDP)					0.178	0.453
					(0.042)	(0.110)
* Minimum distance						-0.001
						(0.000)
Observations	60600	60600	60600	60600	60600	60600
Number of cells	7575	7575	7575	7575	7575	7575

Notes. Standard errors in parentheses are allowed to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). The dependent variables are the number of violent events where the perpetrator is the army (columns 1 and 2) and the number of LRA-associated fatalities reported in logarithm (columns 3-6). The number of events is reported in logarithms (adding one unit before taking logarithms to accommodate 0 values). All specifications include cell and year FE, propagation controls, additional controls and interactions between year dummies and macro-region dummies. Additional controls include commodity price and weather shocks (see Section III). The time period is restricted to 2008–2015. See Appendix A for further information on the variables.

B.9 Analysis of looting

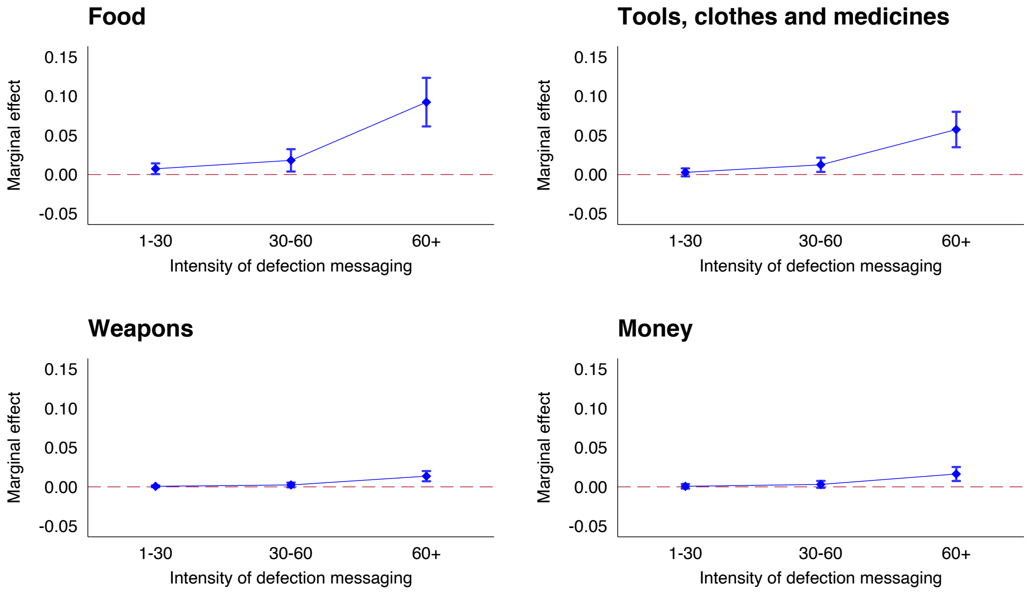
Table B9 shows the effect of intensity of messaging on the number of events that are characterized by either zero killings or at least one killing (columns 1–2), and on the number of looting events associated to other violent events (columns 3–6). We then look at the effect of intensity of messaging on type of goods looted by estimating equation (2) using the number of events characterized by looting of a specific good as dependent variable and allowing the effect of intensity of messaging to be non-linear. We distinguish between food, tools, clothes and medicines, weapons, and money. Figure B9 plots the coefficients.

Table B9: Effect of defection messaging on violent looting

Dependent variable:	Number of LRA events...		Number of events characterized by looting and...			
	without death	with death	no death	at least one death	at least one injury	at least one abduction
	(1)	(2)	(3)	(4)	(5)	(6)
Intensity of messaging	0.009	-0.013	0.017	-0.001	0.002	0.008
	(0.005)	(0.004)	(0.004)	(0.001)	(0.001)	(0.002)
Observations	60600	60600	60600	60600	60600	60600
Number of cells	7575	7575	7575	7575	7575	7575

Notes. Standard errors in parentheses are allowed to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). The dependent variables are the number of events where LRA is the perpetrator, depending on the number of fatalities associated with the event, in columns 1–2, and the number of events where looting occurs concurrently (or not) with other violent events, in columns 3–6. The number of events is reported in logarithms (adding one unit before taking logarithms to accommodate 0 values). All specifications include cell and year FE, propagation controls, additional controls and interactions between year dummies and macro-region dummies. Additional controls include commodity price and weather shocks (see Section III). The time period is restricted to 2008–2015.

Figure B9: Non-linear effect of defection messaging on goods looted



Notes. The figure plots the coefficients of equation (2) against the intensity of defection messaging. The intensity of defection messaging is decomposed into four dummy variables denoting different intervals of radio messaging intensity: 0 minutes, 0–30 minutes, 30–60 minutes and more than 60 minutes of daily messaging. The dummy for "0 minutes" serves as the excluded category. The dependent variables are the number of events characterized by looting, by looted good. The dependent variables are measured in logarithms (adding one unit before taking logarithms to accommodate 0 values). For comparison, we allow the vertical axis to vary at the same scale for all outcome variables. Confidence intervals are computed at 95% of confidence, standard errors are allowed to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). All specifications include cell and year fixed effects, propagation controls, additional controls and interaction terms between year and macro-region indicators. Additional controls include commodity price and weather shocks (see Section III). The time period is restricted to 2008–2015. See Appendix A for further information on the variables. See Appendix A for further information on the variables.

B.10 Commodities and price shocks

Table B10 presents the list of the main cash crops and natural resources in the four countries affected by LRA violence based on the CIA World Factbook (CIA, 2019). We combine information about the geographical distribution of commodities with their prices on the international market.

Geographical distribution of agricultural crops is based on the M3-Crops dataset (Monfreda et al., 2008), which offers a raster dataset at the $5' \times 5'$ latitude/longitude grid and information about harvested area for 175 crops in the 1997–2003 period. Spatially disaggregated information combining national-, state-, and county-level census statistics with satellite imagery for land cover provides an improvement from just using survey data. This dataset allows significantly increasing the variation observed in each cell by providing crop-level information for all major crops in the area. Price-series data are based on the Global Economic Monitor Commodities dataset (The World Bank, 2019). We also consider extractive resources using information about the geographical distribution of minerals (Mineral Resource Data System-MRDS, USGS, 2016b, and the PRIO-GRID datasets) and their prices (Historical Statistics for Mineral and Material Commodities in the United States, USGS, 2016a). Minerals are either not widespread in the study area or the Lasso procedure rejects shocks to mineral prices as relevant (Section III).

Table B10: Main exported crops and natural resources present in LRA-affected countries

Type	**Commodity**	**Price (Source)**	**Geo-location**
Cash Crops	Coffee	Coffee, Robusta, \$/kg, real 2010\$ (GEM)	M3-Crops
	Cotton	Cotton, A Index, \$/kg, real 2010\$ (GEM)	M3-Crops
	Oil palm	Palm oil, \$/mt, real 2010\$ (GEM)	M3-Crops
	Groundnut	Groundnut oil, \$/mt, real 2010\$ (GEM)	M3-Crops
	Rubber	Rubber, Singapore, \$/kg, real 2010\$ (GEM)	M3-Crops
	Sesame	Grains, 2010=100, real 2010\$ (GEM)	M3-Crops
	Sugar	Sugar, world, \$/kg, real 2010\$ (GEM)	M3-Crops
	Tea	Tea average, \$/kg, real 2010\$ (GEM)	M3-Crops
	Tobacco	Tobacco, \$/mt, real 2010\$ (GEM)	M3-Crops
Extractive resources	Cobalt	Cobalt, \$/mt, real 1998\$ (USGS)	MRDS
	Copper	Copper, \$/mt, real 2010\$ (GEM)	MRDS
	Gold	Gold, \$/toz, real 2010\$ (GEM)	PRIO Goldata

Notes. Commodities are listed in order of relative importance. South Sudan includes the information for Sudan. Price for groundnut is proxied by the price of groundnut oil. Source: CIA World Factbook (CIA, 2019). We exclude diamonds and crude oil from the analysis, since they are not present in the area of analysis, and wood since no information is available on the type of forest cover that could be exploited for the international market.

Commodity price shocks are computed as the product of the share of a cell cropped with a commodity and the log-price difference between time t and time $t-1$. Prices are normalized using the year 2010 with the base value equal to 100. The dashed line is a moving average of the time series using a plus/minus 5 year window. Table B11 presents the descriptive statistics of the extension of commodities in the study area and on commodity price shocks (conditional on cells where the commodity is either produced or extracted).

From the list of many potential candidates presented in Table B11, we are interested in selecting the most relevant shocks for violence associated with the LRA. To avoid subjectively *cherry picking* shocks, we rely on Lasso (least absolute shrinkage and selection operator) regressions for objective criteria for covariate selection. The Lasso "estimator imposes a linear model for outcomes as a function of covariates and attempts to minimize an objective that includes the sum of square residuals as in ordinary least squares, but also adds on an additional term penalizing the magnitude of regression parameters" (Athey and Imbens, 2017). By allowing some of coefficients to be exactly zero, Lasso selects an optimal subset of covariates. For further reading on Lasso regressions, please refer to Tibshirani, 1996; Friedman et al., 2001; Tibshirani et al., 2015; Athey and Imbens, 2017; Varian, 2014.

Table B12 compares estimates of equation (2) using the number of fatalities as the dependent variable and different selection criteria for control variables. In columns 1 and 2, the specification includes all commodity prices shocks and all climate shocks (share of the year with drought, rainfall and temperature absolute deviations). In columns 3-6, the specifications includes controls selected with Lasso regressions. We use two distinct approaches to determine the degree of penalization in the Lasso regression. In columns 3 and 4, the degree of penalization is the corrected Akaike Information Criteria (Sugiura, 1978; Hurvich and Tsai, 1989). In columns 5 and 6, the degree of penalization is selected with a 10-fold cross-validation procedure, which optimizes the out-of-sample prediction performance. In the main text, we keep the controls selected by both approaches (highlighted in bold in Table B12). Figure B10 shows the geographic distribution of areas farmed with cotton and groundnut, the main crops selected for analysis in the main text, and the times series of their prices on the international market.

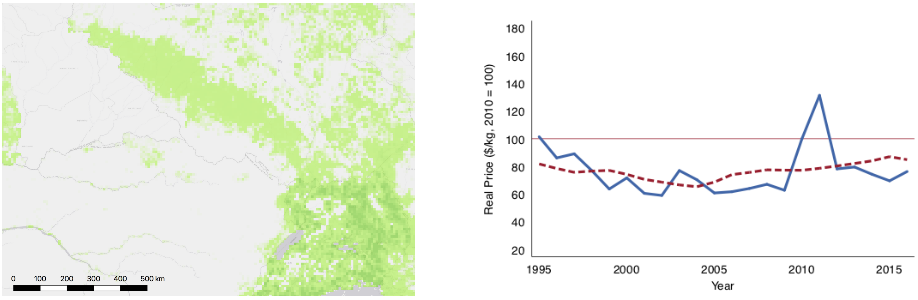
Table B11: Descriptive statistics on cash crops price shocks

	Extension			**Commodity price shock**			
Commodity	Share of study area covered	Cells covered	Share within covered area	Mean	Standard deviation	Min	Max
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Cobalt	0.000	1	1.000	-0.107	0.274	-0.689	0.189
Coffee	0.037	478	0.583	-0.001	0.087	-0.281	0.223
Copper	0.000	2	1.000	-0.035	0.167	-0.237	0.345
Cotton	0.033	409	0.620	0.006	0.172	-0.519	0.466
Gold	0.001	5	1.000	0.061	0.125	-0.164	0.194
Groundnut	0.031	1154	0.204	-0.001	0.070	-0.524	0.380
Palm oil	0.001	5	0.851	-0.027	0.143	-0.265	0.241
Sesame	0.016	1022	0.119	-0.000	0.026	-0.214	0.310
Sugar	0.016	152	0.808	0.027	0.158	-0.193	0.413
Tea	0.001	13	0.648	0.021	0.056	-0.092	0.182

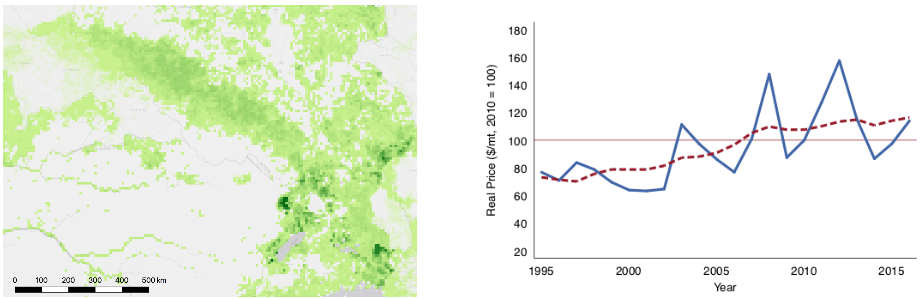
Notes. The table presents descriptive statistics on the extent of commodities in the study area and on commodity price shocks. For minerals, the share of a cell covered is set to 1 if a mine is present, and 0 otherwise. Statistics relative to the price shocks are conditional on cells where the crop is produced or extracted. A shock is defined as the product between the share of cell farmed with the commodity (ranging 0–100) and the log-price difference between t and $t-1$.

Figure B10: Geographic distribution and price series of main cash crops

A. Cotton



B. Groundnut



Notes. The geographic distribution (left panels) and the price series (right panels) for cotton and groundnut. Price for groundnut refers to groundnut oil. The geographic extent of the figure is restricted to the study area. Prices are reported in real values using US$ per corresponding unit. Prices are normalized using the year 2010 as base. The horizontal line shows the base value of 100. The dashed line is a moving average of the time series using a plus/minus 5 year window. Source: M3-Crops Data (Monfreda et al., 2008), GEM Commodities dataset (The World Bank, 2019). Basemap source: Esri (see Appendix A for details and attributions).

Table B12: Alternative procedures for control variables selection

Dependent variable:	Number of fatalities linked to LRA activity					
Criteria for selection of controls:	**Full list of controls**		**AIC Lasso**		**Cross-validation Lasso**	
	Coefficient	Std. error	Coefficient	Std. error	Coefficient	Std. error
	(1)	(2)	(3)	(4)	(5)	(6)
Intensity of messaging	-0.028	0.005	-0.028	0.005	-0.028	0.005
Cobalt price shock	0.018	0.009	-	-	0.018	0.008
Coffee price shock	0.013	0.009	-	-	0.013	0.009
Cotton price shock	-0.018	0.009	-0.016	0.008	-0.018	0.009
Gold price shock	-0.126	0.047	-	-	-0.126	0.047
Groundnut price shock	0.044	0.020	0.054	0.019	0.044	0.020
Copper price shock	0.000	0.011	-	-	-	-
Palm oil price shock	-0.095	0.076	-	-	-0.095	0.076
Sugar price shock	-0.004	0.006	-	-	-0.004	0.006
Sesame price shock	0.047	0.029	-	-	0.047	0.029
Tea price shock	0.035	0.030	-	-	0.035	0.030
Share of the year with drought	-0.015	0.010	-0.012	0.009	-0.015	0.010
Absolute rainfall deviation	-0.003	0.002	-0.003	0.002	-0.003	0.002
Absolute temperature deviation	-0.002	0.003	-0.002	0.003	-0.002	0.003

Notes. The table reports estimates using equation (2) estimated using a fixed effects model. "-" indicates that the Lasso procedure sets the coefficient to zero. Standard errors in parentheses are allowed to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). Intensity of defection messaging, defined by equation (1), is standardized. The dependent variable is the number of fatalities linked to LRA activity (adding 1 unit before taking logarithms to accommodate zero values). The specification in column 1 and 2 includes cell and year fixed effects, all commodity price shocks, weather shocks and macro-region-specific time fixed effects (see Section III). In columns 3 and 4, controls are selected using Lasso with a degree of penalization selected according to the corrected Akaike Information Criteria (Sugiura, 1978; Hurvich and Tsai, 1989). In columns 5 and 6, controls are selected using Lasso with a degree of penalization selected according to 10-fold cross-validation. The time period is restricted to 2008–2015. See Appendix A for further information on the variables. The variables selected in all criteria are reported in bold.

Table B13 presents estimates of the effect of intensity of messaging using equation (2) and focusing in areas not affected by the production of either cotton or groundnut. Panel A restricts the sample to areas not covered by the production of cotton, panel B restricts the sample to areas not covered by the production of groundnut, and panel C restricts the sample to areas not covered by the production of both commodities. Estimates presented for the whole study area are not driven by areas affected by price shocks for cotton or groundnut.

Table B13: The effect of defection messaging in areas not covered by shocks

Dependent variable:	**Fatalities**	**Number of individuals...**		**Number of events involving...**		
		Returning	Being abducted	Violence against civilians	Clashes	Looting
	(1)	(2)	(3)	(4)	(5)	(6)
A. Sample restricted to areas not covered by cotton						
Intensity of messaging	-0.027	0.009	-0.005	-0.010	-0.006	0.016
	(0.005)	(0.004)	(0.005)	(0.004)	(0.003)	(0.004)
Observations	57328	57328	57328	57328	57328	57328
Number of cells	7166	7166	7166	7166	7166	7166
B. Sample restricted to areas not covered by groundnut						
Intensity of messaging	-0.027	0.009	-0.005	-0.010	-0.006	0.016
	(0.005)	(0.004)	(0.005)	(0.004)	(0.003)	(0.004)
Observations	51368	51368	51368	51368	51368	51368
Number of cells	6421	6421	6421	6421	6421	6421
C. Sample restricted to areas not covered by cotton and groundnut						
Intensity of messaging	-0.027	0.009	-0.005	-0.010	-0.006	0.016
	(0.005)	(0.004)	(0.005)	(0.004)	(0.003)	(0.004)
Observations	50120	50120	50120	50120	50120	50120
Number of cells	6265	6265	6265	6265	6265	6265

Notes. The table reports marginal effects estimated using a fixed effects model. Standard errors in parentheses are allowed to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). The dependent variables are the number of fatalities linked to LRA (column 1), the number of individuals returning (columns 2), the number of abductees (columns 3) and the number of other LRA-related violent events (columns 4–6). The dependent variables are measured in logarithms (adding one unit before taking logarithms to accommodate 0 values). In panel A, coefficients are estimated using equation (2) and controlling for free-space circular coverage. Cell covered (circular coverage) is a dummy variable equal to 1 if the cell is covered by free-space circular coverage of radio signal, and 0 otherwise. In panel B, coefficients are estimated using equation (2) and restricting the sample to cells within free-space circular coverage at a specific point in time (contemporaneous to the measurement of violence). In panel C, coefficients are estimated using equation (2) and restricting the sample to cells within free-space circular coverage at any point in time in the period 2008–2015. All specifications include cell and year fixed effects, propagation controls, additional controls and interactions between year dummies and macro-region dummies. Additional controls include commodity price and weather shocks (see Section III). The time period is restricted to 2008–2015. See Appendix A for further information on the variables.

To understand the role of price shocks in areas not covered by the production of the selected commodities, we generate spatial allocations of production of commodities at random by reshuffling the areas farmed with each selected commodity within the grid. Similar to the main setting of the paper, spatial distribution of production is assumed to be constant over time. The effect of intensity of messaging and of commodity price shocks is then estimated using equation (2) and the number of total fatalities as dependent variable. This procedure, which we iterate 1000 times, allows studying the effect of price shocks when the spatial distribution of production is random. Table B14 presents descriptive statistics for estimated coefficients. In panel A, equation (2) is estimated using all cells, while in panel B, the sample is restricted to cells producing cotton and groundnut in the real setting. The estimate of the effect of intensity of messaging is stable and similar to the estimate reported in the main text (table 2), supporting the exogeneity of intensity of messaging with respect to shocks. In addition, we do not observe a significant effect of commodity price shocks, when the spatial distribution of production is random, i.e., the value zero lies within the 5^{th} and 95^{th} empirical percentiles of the distribution of estimates.

Table B14: Random spatial allocation of production and the effect on fatalities

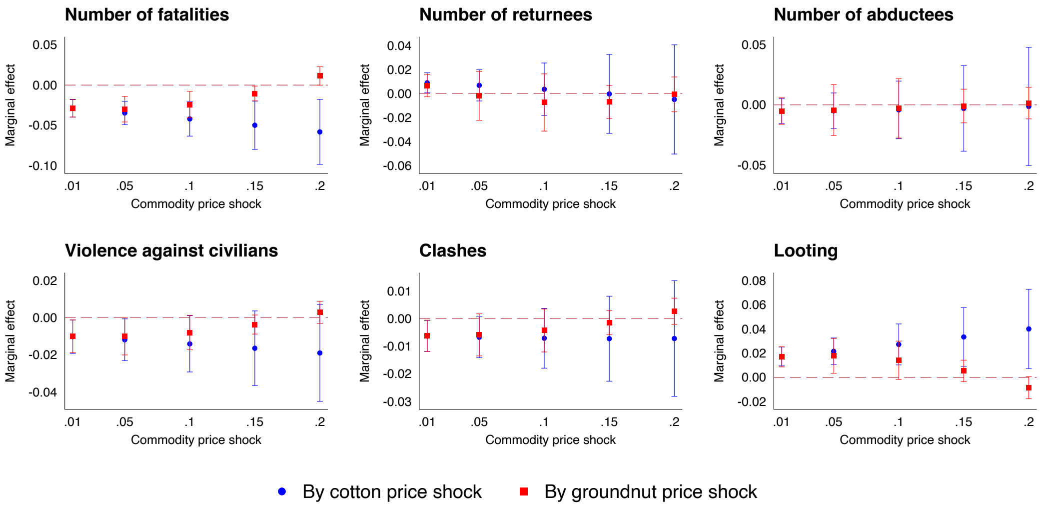
	Dep.var.: Number of fatalities linked to LRA activity				
	Mean	St.dev.	5^{th} pct.	Median	95^{th} pct.
	(1)	(2)	(3)	(4)	(5)
A. Include all cells					
Intensity of messaging	-0.032	0.000	-0.032	-0.032	-0.032
Cotton price shock	-0.000	0.014	-0.021	-0.001	0.023
Groundnut price shock	0.002	0.025	-0.042	0.003	0.042
B. Include only non-producing cells					
Intensity of messaging	-0.032	0.000	-0.032	-0.032	-0.032
Cotton price shock	-0.000	0.017	-0.026	-0.001	0.027
Groundnut price shock	0.002	0.030	-0.051	0.004	0.051

Notes. The table reports descriptive statistics of the coefficients on intensity of defection messaging and on commodity price shocks when the spatial distribution of production of cotton and groundnut is random. Each observation represents a different spatial distribution of production of commodities. These are generated by reshuffling the areas farmed with each commodity within the grid (the procedure is iterated 1000 times). The effect of intensity of messaging and of commodity price shocks is estimated using equation (2) and the number of total fatalities as dependent variable, measured in logarithms (adding one unit before taking logarithms to accommodate 0 values). Intensity of messaging is defined by equation (1) and standardized. The shock variables are measured as a product of the percentage of the cell farmed with the crop and its log-price difference with the previous year on the international market (and are not standardized). All specifications include cell and year fixed effects, propagation controls, additional controls and year $\times$ macro-region fixed effects. Additional controls include commodity price and weather shocks (see Section III). The time period is restricted to 2008–2015. See Appendix A for further information on the variables. In panel A, equation (2) is estimated using all cells, while in panel B, the sample is restricted to cells originally producing cotton and groundnut.

B.11 Defection messaging and commodity price shocks: alternative specifications

This section presents estimates of equation (2) by allowing for non-linear heterogeneity in the effects of defection messaging by the two commodity price shocks. Equation (2) is estimated interacting the intensity of messaging with cotton and groundnut price shocks and with their squared value, thus allowing for a quadratic relationship. Figure B11 depicts the marginal effects of intensity of messaging for shocks of alternative sizes of the commodity price shock.

Figure B11: Defection messaging and commodity price shocks: quadratic interactions



Notes. The figure plots marginal effects of intensity of messaging as function of commodity price shocks. Marginal effects are estimated using equation (2) interacting the intensity of messaging with cotton and groundnut price shocks and with their squared value, and assuming all other variables remain constant. The dependent variables are the number of fatalities, the number of returnees, the number of abductees, the number of events characterized by violence against civilians, clashes with security forces, and looting. The dependent variables are measured in logarithms (adding one unit before taking logarithms to accommodate 0 values). Confidence intervals are computed at 95% of confidence, standard errors are allowed to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). All specifications include cell and year fixed effects, propagation controls, additional controls and interaction terms between year and macro-region indicators. Additional controls include commodity price and weather shocks (see Section III). The time period is restricted to 2008–2015. See Appendix A for further information on the variables.

B.12 Spatial spillover

We investigate the presence of spatial spillovers using a spatial Durbin model (Anselin, 2013). In this model, the violence in each cell depends on the observable characteristics of the cell and on the same characteristics of the neighboring cells. We estimate the following model:

$$y_{it} = \gamma_i + \alpha\, dm_{it} + \alpha_2 W \mathbf{dm}_t + \mathbf{X'}_{it}\beta_1 + W\mathbf{X}_{it}\beta_2 + \alpha_t M_r + u_{it} \tag{3}$$

where the structure of spatial dependence between observations is defined through a symmetric weighting matrix W. Benchmark weighting matrix is a binary contiguity matrix in which a weight of $1/m$ is assigned to cells surrounding the cell of interest within a 0.5° distance cutoff, and a weight of 0 to other cells.[2] m corresponds to the number of cells considered in the spillover area surrounding the cell. In practice, we are controlling for the average value of the control variable in the surrounding cells. We are aware that while the model aims at measuring spatial spillovers, in this setting, averaging across multiple cells leads to reduced variation, which could invalidate the assumption of exogeneity of topography-corrected signal. Table B15 presents estimates for equation (3). We include an F-test of joint-significance of the main and spillover effect of intensity of messaging.

Table B15: Defection messaging, treatment spillover and commodity price shocks

Dependent variable:	**Fatalities**	**Number of individuals...**		**Number of events involving...**		
		Returning	Being abducted	Violence against civilians	Clashes	Looting
	(1)	(2)	(3)	(4)	(5)	(6)
Intensity of messaging	-0.027	0.000	-0.016	-0.012	-0.000	0.010
	(0.011)	(0.008)	(0.010)	(0.009)	(0.004)	(0.007)
Avg intensity of messaging (surrounding cells)	-0.001	0.012	0.012	0.003	-0.007	0.008
	(0.009)	(0.007)	(0.010)	(0.007)	(0.003)	(0.006)
Observations	60600	60600	60600	60600	60600	60600
Number of cells	7575	7575	7575	7575	7575	7575
Joint-significance F test (p-value)	0.000	0.003	0.315	0.052	0.001	0.000

Notes. Standard errors in parentheses are allowed to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). Coefficients are estimated using a spatial Durbin model (equation (3)). The dependent variables are the number of fatalities linked to LRA (column 1), the number of individuals returning (columns 2), the number of abductees (columns 3) and the number of other LRA-related violent events (columns 4–6). The dependent variables are measured in logarithms (adding one unit before taking logarithms to accommodate 0 values). All specifications include cell and year fixed effects, propagation controls, additional controls and interactions between year dummies and macro-region dummies. Additional controls include commodity price and weather shocks (see Section III). Controls are included both at the level of the cell and, as an average, at the level of the surrounding cells. The joint-significance F-test includes the main effect and the spillover effect. The time period is restricted to 2008–2015. See Appendix A for further information on the variables.

The effect of intensity of messaging on fatalities is mainly accounted by the intensity in the same cell, with no significant spatial spillover (column 1). For returnees, on the other hand, the main effect of intensity of messaging is captured as a spillover effect (column 2). This suggests that returnees might return in the main cell, but they might be operating in the surrounding cells

[2] Due to the resolution of analysis, intensity of messaging tends to present a high degree of spatial correlation. We therefore extend the area we consider to compute the spillover effect up to the point in which the correlation between intensity of messaging in the main cell and the average in the surrounding cells is lower than 0.85. This corresponds to 0.5° distance from the cell.

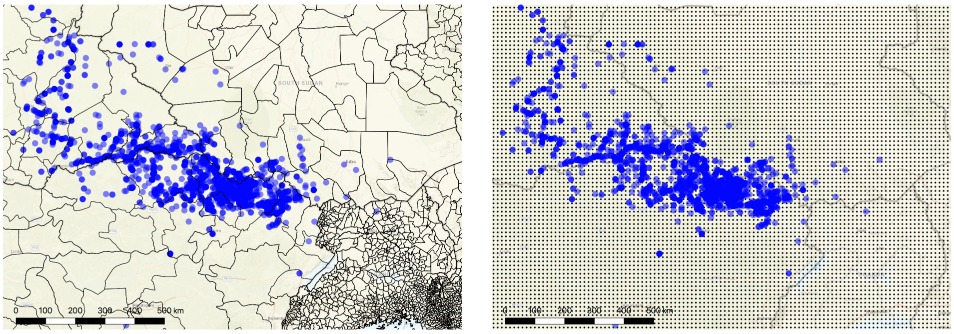
when they are exposed to the message. A similar effect is observed for clashes (columns 5). For abductees, we do not observe any effect (column 3). Finally, for looting the effect is more dispersed between main cell and surrounding cells (column 6).

B.13 Administrative-unit-level analysis

We use third-level administrative units (corresponding to districts) from the GADM dataset for the analysis. Figure B12 plots the distribution of LRA events for the 458 administrative units in the study area (left figure) and for the grid used for the main analysis (right figure).

Figure B12: Comparison between administrative boundaries and gridded dataset

Administrative boundaries **$0.125° \times 0.125°$ grid**



Notes. The geographic distribution of LRA events in the period of 2008–2015 for the 458 administrative units in the study area (left panel) and for the grid used for the main analysis (right panel). Basemap source: Esri (see Appendix A for details and attributions).

Table B16 estimates the effect of intensity of messaging on LRA fatalities using administrative boundaries. Since the use of administrative boundaries leads to units of unequal size, we use per capita fatalities as the main outcome variable, though general findings are similar using other outcome variables. At this level of analysis, one standard deviation in intensity of messaging corresponds to roughly 5.45 minutes of messaging at full district coverage.

Table B16: Effect of defection messaging using administrative divisions as units of observation

Dependent variable:	**Number of fatalities linked to LRA activity (per 1000 inhabitants)**			
	(1)	(2)	(3)	(4)
Intensity of messaging	-0.008	-0.008	-0.008	-0.008
	(0.002)	(0.002)	(0.002)	(0.002)
Observations	3664	3664	3664	3664
Number of administrative areas	458	458	458	458
Additional controls	No	Yes	Yes	Yes
Year $\times$ Country FE	No	No	Yes	No
Year $\times$ Macro-Region FE	No	No	No	Yes

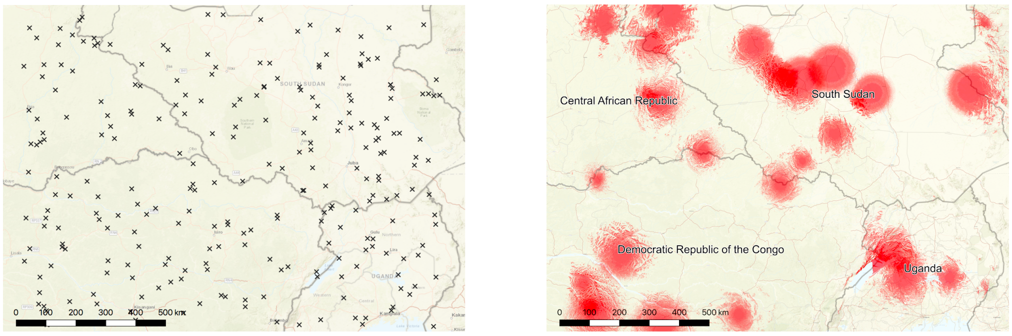
Notes. Standard errors in parentheses are allowed to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). The dependent variable is the number of LRA-associated fatalities per 1000 inhabitants. The number of fatalities is reported in logarithms (adding one unit before taking logarithms to accommodate 0 values). All specifications include cell and year FE, and propagation controls. Propagation controls are averaged at the cell level within a defined administrative area. Administrative units are computed from the GADM dataset. Additional controls include commodity price and weather shocks (see Section III). The time period is restricted to 2008–2015. See Appendix A for further information on the variables.

B.14 Placebo test

The left panel in Figure B13 shows the random position of an antenna in each of the 250 iterations used, while the right panel shows the radio coverage in one specific iteration. Table B17 presents descriptive statistics of placebo test estimates of the effect of intensity of messaging on outcome variables. Figure B14 shows the distribution of the coefficient of intensity of messaging on fatalities in the placebo samples.

Figure B13: Allocation of antennas and example of placebo coverage

Random allocation of antenna **Placebo coverage**



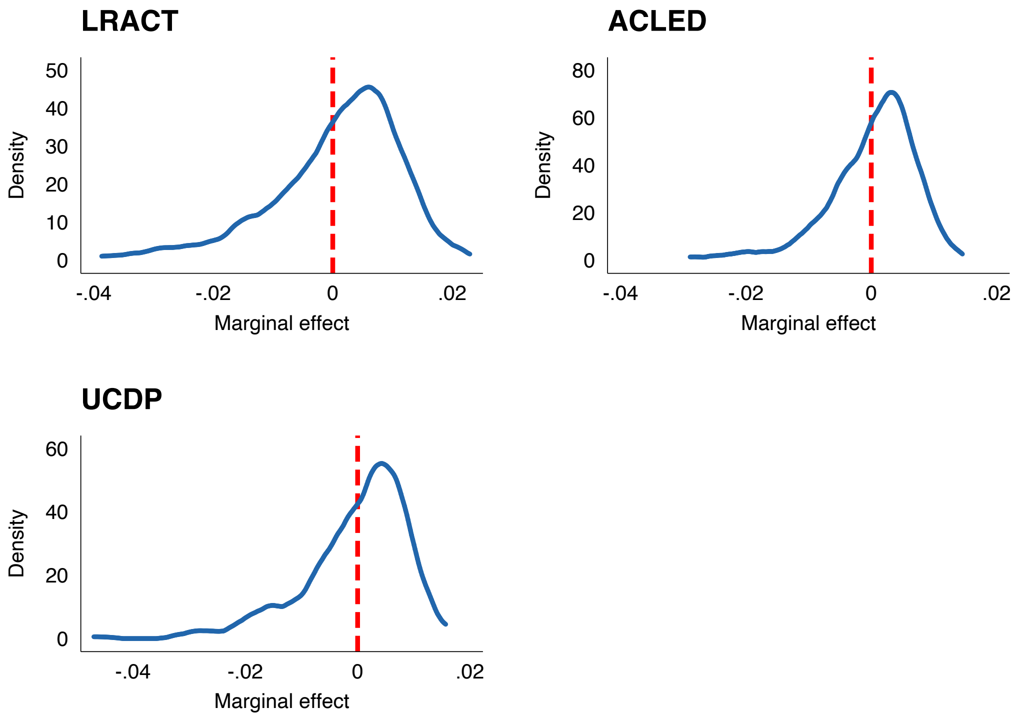
Notes. The left panel shows the random position of a single antenna in all 250 iterations. The right panel shows an example of placebo coverage of the sampled antennas for one specific iteration (for the year 2015 and the iteration number 143). Basemap source: Esri (see Appendix A for details and attributions).

Table B17: Descriptive statistics for placebo test

	Coefficient on intensity of defection messaging				
	Mean	St.dev.	5^{th} pct.	Median	95^{th} pct.
	(1)	(2)	(3)	(4)	(5)
Number of fatalities	0.000	0.003	-0.005	0.001	0.003
Number of returnees	-0.000	0.001	-0.002	-0.000	0.003
Number of abductees	-0.000	0.002	-0.003	-0.000	0.003
Number of events with violence against civilians	0.000	0.001	-0.002	0.000	0.002
Number of events with clashes	-0.000	0.001	-0.001	0.000	0.001
Number of events with looting	-0.000	0.002	-0.002	-0.000	0.003
Number of fatalities (ACLED)	0.000	0.002	-0.003	0.000	0.002
Number of fatalities (UCDP)	-0.000	0.002	-0.005	0.000	0.003

Notes. Descriptive statistics of the coefficient on intensity of defection messaging in the placebo test (250 simulations). Each observation is an estimated coefficient in equation (2) where radio coverage is generated by randomly allocating antennas in the original grid and the intensity of messaging is defined accordingly. The dependent variable is log-transformed (adding one unit before taking logarithms to accommodate 0 values). For comparability, intensity of messaging is not standardized in these regressions. See Appendix A for further information on the variables.

Figure B14: Distribution of marginal effects of intensity of messaging on violent events



Notes. The distribution of the coefficient of the intensity of defection messaging on fatalities in the placebo samples using the LRACT, ACLED and UCDP datasets. We perform 250 simulations. Each observation is coefficient that is estimated using equation (2) where radio coverage is generated by randomly allocating antennas in the original grid, and then defining the intensity of messaging accordingly. The dependent variable is log-transformed (adding one unit before taking logarithms to accommodate 0 values). For comparability, intensity of messaging is not standardized in these regressions. The dotted red line indicates zero.

B.15 Estimates using non-linear models

We focus on two families of non-linear models: count models that estimate the effect on the intensive margin of violence, and binary-outcome models that estimate the effect on the extensive margin. We start by estimating a (conditional) fixed effects negative binomial regression, which allows controlling for over-dispersion in the data (Cameron and Trivedi, 2013). We supplement this with a (conditional) fixed effects Poisson regression. The dependent variable is in levels in both cases. For both models identification relies on specific assumptions outlined in Hausman et al. (1984). In these models, the term "fixed effects" has a different connotation from that of the linear case. It applies to the distribution of the dispersion parameter c_i in the specification of the conditional mean $E[y_{it}|x_{it}, c_i] = c_i \cdot \mu(x_{it}, \beta)$. Dispersion is assumed to be constant within the same cell and to vary across cells without imposing any restriction. In the Poisson version, the dispersion parameter is assumed to cancel out by conditioning on a sufficient statistic, i.e., $\sum_{t=1}^{T} y_{it}$. In the negative binomial version, given certain assumptions, it cancels out in the derivation of the joint density for the dependent variable y_{it} (see Cameron and Trivedi (2013) for further details). In the Poisson regression, we follow the procedure by Silva and Tenreyro (2010) to ensure the existence of the maximum likelihood estimates. To estimate the extensive margin, we use binary dependent variables indicating at least one fatality or at least one event in the case of other outcome variables. We estimate the effect using a fixed effects logistic regression.

Table B18 displays the non-linear estimates of the effect of defection messaging on the number of fatalities linked to LRA activity (column 1) and other outcomes related to LRA violence (columns 2–5). Results are in line with estimates presented in the main text using linear models (Tables 2 and 3). However, the non-linear fixed effects estimators lead to loss of data as cells not experiencing any fatality (or any event) in the period considered do not contribute to the estimator. This is the case, for instance, of cells that never experience violence, but might have been covered by defection messaging. Estimates should be interpreted taking into account this restriction (Wooldridge, 2010).

Table B18: Effect of defection messaging on fatalities and additional outcomes: non-linear models

Dependent variable:	Number of fatalities	Number of individuals...		Number of events involving...		
		Returning	Being abducted	Violence against civilians	Clashes	Looting
Specification:	(1)	(2)	(3)	(4)	(5)	
Intensive margin (dep. var. in levels)						
Fixed effects negative binomial	-0.049	0.034	0.032	-0.044	-0.024	0.072
	(0.022)	(0.025)	(0.018)	(0.016)	(0.021)	(0.019)
N	1920	1928	2648	2216	1688	2576
Fixed effects Poisson	-0.082	0.057	-0.062	-0.024	0.011	0.091
	(0.025)	(0.027)	(0.036)	(0.017)	(0.020)	(0.023)
N	1920	1928	2648	2216	1688	2576
Extensive margin						
Fixed effects logistic	-0.024	0.014	0.006	-0.009	-0.009	0.017
	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)
N	1920	1912	2624	2216	1688	2560

Notes. The table reports coefficients of the effect of intensity of messaging estimated using different models (for the Logistic regression, marginal effects are reported). The intensity of messaging variable is standardized in all the specifications. The dependent variables are the total number of LRA-related fatalities (column 1), the number of individuals returning from the LRA (column 2), the number of individuals abducted (column 3), the number of events involving violence against civilians (column 4), the number of events involving clashes between the LRA and other actors (column 5), and the number of events involving looting by the LRA (column 6). All outcomes are computed from the LRACT dataset. For count models, the number of fatalities is capped at 10 to facilitate convergence. For the intensive margin, dependent variables are used in levels, while for the extensive margin indicator variables for whether at least one individual or event is recorded are used. For the fixed effects Poisson regression standard errors are cluster-robust at the level of the cell, while for the fixed effects negative binomial and the fixed effect Logistic regressions standard errors are computed using bootstrap (200 iterations). All specifications include cell fixed effects, time, minimum distance from an antenna, and additional controls. Time effects are modeled allowing for a smooth non-linear long-term change in conflict using a linear combination of sine and cosine of the year of observation (Wilks, 2011; Cox, 2006; Stolwijk et al., 1999). Additional controls include commodity price and precipitation shocks (see Section III). For the fixed effects negative binomial regressions, controls include time-invariant characteristics at cell-level, including terrain ruggedness, baseline population and baseline violence (see Hausman et al., 1984 for the role of time-invariant controls in this model). The time period is restricted to 2008–2015. See Appendix A for further information on the variables.

B.16 Ethnic diversity and ethnic distance

To analyze the variation in the effectiveness of messaging by ethnic composition of the treated areas or the host communities, we calculate how ethnically distant the populations residing in the grid cells are from the Acholi people. We calculate ethnic distances using linguistic distances (see, for instance, Desmet et al., 2012). This approach assumes linguistic identity defines ethnic identity, which is generally true in the case of Africa. For instance, Acholi is both a language and an ethnic group. We measure the linguistic distance between any two ethnic groups by the degree of difference between the languages spoken based upon linguistic tree diagrams from the Ethnologue dataset. The distance metric between two languages i and j is defined as:

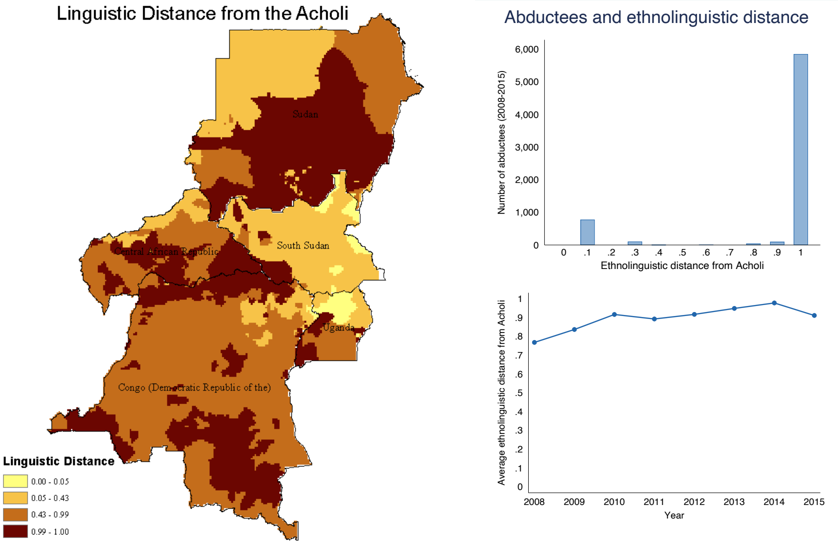
$$\tau_{ij} = 1 - \left(\frac{l}{m}\right)^{\delta} \tag{4}$$

where l is the number of shared branches between languages i and j, m is the maximum number of branches between any two languages, and δ is a decay factor determining how fast the distance declines as the number of shared branches increases. We assume $\delta = 0.05$ (Desmet et al., 2009), but results are robust to the use of alternatives, such as 0.0025 (Gomes, 2014) or 0.5 (Fearon, 2003). We then make use of recently generated cell-level data on the distribution of language groups at the resolution of 5 km $\times$ 5 km from Desmet et al. (2018) to generate measures of ethnolinguistic distance. To check for heterogeneity by ethnic diversity, we also compute indices of ethnic composition at the grid-cell level using ethnolinguistic fractionalization (Easterly and Levine, 1997; Alesina et al., 2003) and ethnolinguistic polarization (Esteban and Ray, 1994; Montalvo and Reynal-Querol, 2005).

The left panel in figure B15 shows the (grid-cell-level) geographical distribution of linguistic distance from the Acholi people across the countries in the study area. The top right panel depicts the distribution of all abductees in the period 2008–2015 according to the ethnolinguistic distance from Acholi of the population in the location of abduction. The bottom right panel plots the evolution over time of the (weighted) average ethnolinguistic distance from Acholi among abductees. Large ethnic mismatch exists between abductees and the LRA.

Combatants of Acholi ethnicity founded the LRA and, at least initially, most LRA soldiers were drawn from the Acholi population. Ethnic closeness of certain areas to the Acholi might make returning in those areas easier by facilitating forgiveness or because of lower language or cultural barriers. In the latter part of the conflict (especially during 2005–2018), the LRA relied on abductions from the bordering regions of DRC, South Sudan and CAR, which are far from the Acholi homeland. Individuals in these areas are also ethnically different from the Acholi (figure B15) and are hence less likely to be sympathetic to the LRA cause. This can lead to two opposing forces. First, people might be less sympathetic towards Acholi LRA returnees. Second, if the LRA relies on local populations for conscripts, these individuals are likely to be more sympathetic towards returnees, and the local conscripts can be made to defect more easily within these cells.

Figure B15: Ethnolinguistic distance from the Acholi people



Notes. The left panel shows the geographic distribution of ethnolinguistic distance in the study area. The top right panel presents the distribution of all abductees in the period 2008–2015 according to the ethnolinguistic distance from Acholi in the location of abduction. Values of distance are rounded to the first decimal unit. The bottom right panel shows the evolution over time of the (weighted) average ethnolinguistic distance from Acholi among abductees. Lower (higher) values indicate a lower (higher) distance. Ethnolinguistic distance is computed using equation (4) and assuming a decay factor $\delta = 0.05$ as in Desmet et al. (2009). Data on language trees come from the Ethnologue dataset.

Table B19 shows that the more ethnically distant a cell is from the Acholi, the more effective defection messaging is in the cell. This indicates that the second effect might outweigh the first. Radio messaging is also less effective in areas that are more ethnically diverse. Returnees might be more reluctant to return to areas where there are multiple ethnic groups, whose languages they might not speak, or whose cultures they might not understand. Alternatively, it could also be that they cannot predict how the different groups will treat them.

Table B19: Defection messaging and ethnolinguistic fractionalization, polarization and distance

Dependent variable:	Number of fatalities linked to LRA activity				
	(1)	(2)	(3)	(4)	(5)
Intensity of messaging	-0.018	-0.019	-0.023	-0.016	-0.018
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
* Ethnolinguistic fractionalization	0.014			0.012	
	(0.003)			(0.003)	
* Ethnolinguistic polarization		0.012			0.010
		(0.004)			(0.003)
* Ethnolinguistic distance			-0.011	-0.010	-0.009
			(0.004)	(0.004)	(0.004)
Observations	59360	59360	59360	59360	59360
Number of cells	7420	7420	7420	7420	7420

Notes. Standard errors in parentheses are allowed to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). The dependent variable is the number of fatalities linked to LRA. The number of fatalities is reported in logarithms (adding one before taking logarithms to accommodate 0 values). " Ethnolinguistic fractionalization" is defined following Alesina et al. (2003). "Ethnolinguistic polarization" is defined following Montalvo and Reynal-Querol (2005). "Ethnolinguistic distance" is computed using equation (4) and assuming a decay factor $\delta = 0.05$ as in Desmet et al. (2009). Data on language trees come from the Ethnologue dataset. All ethnolinguistic variables are standardized. All specifications include cell and year FE, propagation controls, and additional controls. Additional controls include commodity price and weather shocks (see Section III). The time period is restricted to 2008–2015. See Appendix A for further information on the variables.

B.17 Robustness to standard error assumptions

Table B20 presents estimates of the effect of intensity of messaging using equation (2) under different assumptions used to estimate standard errors. Column 1 presents the point estimate, while columns 2–8 present standard errors. Column 2 assumes clustering at cell-level, column 3 assumes clustering using district-level administrative boundaries, columns 4-8 assumes Conley (1999, 2008) and Hsiang (2010) correction allowing spatial correlation to be within 25, 50, 100, 200 and 500 kilometers, respectively. The 100-km cut-off, used in the main text, is in line with other contributions in the literature (see, for instance, Harari and La Ferrara, 2013).

Table B20: Intensity of messaging and robustness to standard error assumptions

	Coeff.	**Standard errors under different assumptions**						
		Clustering		*Conley-Hsiang correction - spatial correlation within...*				
		Cell	District	25 km.	50 km.	100 km.	200 km.	500 km.
Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of fatalities	-0.027	0.005	0.007	0.004	0.004	0.005	0.005	0.006
Number of fatalities (ACLED)	-0.016	0.006	0.002	0.003	0.002	0.002	0.002	0.002
Number of fatalities (UCDP)	-0.025	0.006	0.005	0.003	0.004	0.004	0.004	0.005
Number of returnees	0.009	0.003	0.004	0.003	0.004	0.004	0.004	0.005
Number of abductees	-0.005	0.005	0.003	0.005	0.005	0.005	0.005	0.006
Events: violence against civilians	-0.010	0.002	0.004	0.003	0.004	0.004	0.005	0.006
Events: clashes	-0.006	0.002	0.002	0.002	0.002	0.003	0.003	0.004
Events: looting	0.016	0.003	0.005	0.003	0.003	0.004	0.004	0.004

Notes. The table presents point estimates (column 1) and standard errors (columns 2–8) of the effect of intensity of messaging estimated using equation (2). Standard errors are estimated assuming clustering at the cell-level in column 2; assuming clustering at the district-level in column 3; and allowing for spatial correlation within 25, 50, 100, 200 and 500 kilometers, using the Conley (1999, 2008) and Hsiang (2010) correction, in columns 4–8. All specifications include cell and year FE, propagation controls, macro-region-specific time fixed effects, and additional controls. Additional controls include commodity price and weather shocks (see Section III). The time period is restricted to 2008–2015. The dependent variables are measured in logarithms (adding one unit before taking logarithms to accommodate 0 values). See Appendix A for further information on the variables.

B.18 Heterogeneous effects

To estimate heterogeneous effects, we split the sample into two sub-samples, one below and the other above the median value of each variable. We then estimate the effect of intensity of messaging separately for the two sub-samples using equation (2). This allows estimating the main specification by permitting the other parameters to vary according to the sub-sample. Figure B16 plots the marginal effects of intensity of messaging estimated using procedure. Table B21 reports F-statistics and p-values of joint-tests of equality of the effect of intensity of messaging across the different sub-groups, estimated using equation (2) with interaction terms between intensity of messaging and indicator variables for the corresponding groups. Appendix A provides variable definitions and data sources.

Panel A in Figure B16 focuses on variables related to geopolitics and the economy. We analyze heterogeneous effects by country, distance from border, population, nightlight, total value of cash crops, infant mortality, and the phase of conflict. Nightlight, population and infant mortality are measured ex-ante to preclude endogenous variation in these variables induced by conflict.[3] These

[3] Data on infant mortality are not available for the full study area. See Appendix A.

are measured in the years 2007, 2005 and 2000, respectively. For nightlight, 90% of the sample has a value of 0 (which also corresponds to the median), with positive numbers identifying larger urban centers. We compute the total value of cash crops by summing the products between the share of the cell farmed with a specific commodity and its price on the international market for all commodities analyzed in Section B.10 for the 2008-2015 period.

We do not observe any significant difference in effect across the three different countries, suggesting homogeneity in this dimension. The absence of heterogeneity with respect to the distance from the border, for which we also do not find significant differences, also confirms this. For nightlight we find a nil effect for cells with higher luminosity and a negative and significant effect for cells with zero values. Less populated areas and areas with lower values of cash crops also display stronger effects. We do not find evidence of heterogeneity along the other dimensions. For investigating heterogeneity by the phase of conflict, we split the sample into two periods: a first phase (2010-2012), and a second phase (2013-2015). Splitting the sample into two phases allows us to continue to use cell-level fixed effects. We observe a larger effect in the second phase.

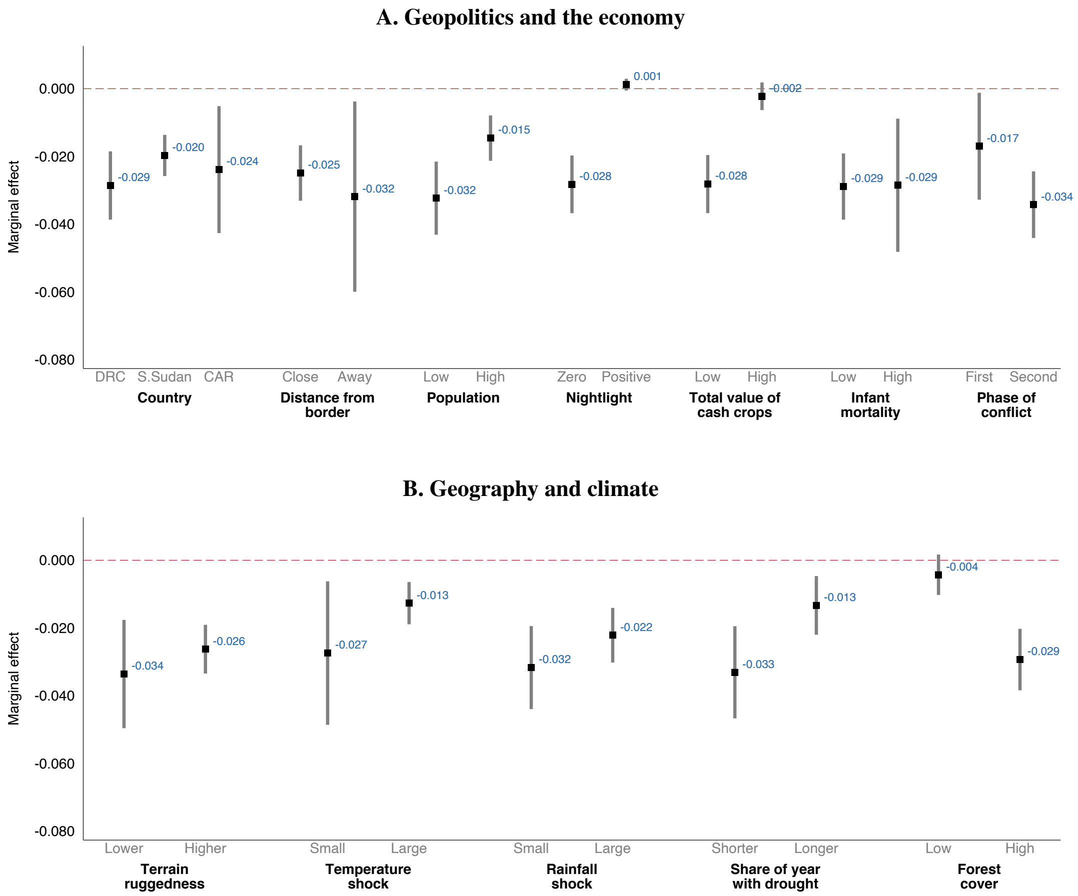
Panel B in Figure B16 focuses on variables related to geography and climate. These include terrain ruggedness, temperature and rainfall shocks, share of the year with drought, and forest cover. Apart from the forest cover variable, we do not find strong evidence of heterogeneity by any of the other variables. The effect of of messaging is stronger in cells with higher forest cover. The areas with higher forest cover are also the areas in which the LRA operates (Lancaster et al., 2011).

Table B21: Test of heterogeneous effects

Geopolitics and the economy			**Geography and climate**		
Category	F (1)	p-value (2)	Category	F (3)	p-value (4)
Country	1.56	0.21	Terrain ruggedness	0.79	0.37
Distance from border	0.00	0.95	Temperature shock	0.06	0.81
Population	10.14	0.00	Rainfall shock	1.03	0.31
Nightlight	32.43	0.00	Share of year with drought	0.13	0.72
Total value of cash crops	29.04	0.00	Forest cover	20.66	0.00
Infant mortality	0.06	0.80			
Phase of conflict	5.57	0.02			

Notes. The table presents F-statistics and p-values of Wald joint-tests of equality of the effect of intensity of messaging across the different sub-groups presented in Figure B16. For each category, the test is performed by estimating equation (2) with interaction terms between the intensity of messaging and indicator variables for each sub-group. Standard errors are allowed to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). The dependent variable is the number of fatalities linked to LRA. The number of fatalities is reported in logarithms (adding one unit before taking logarithms to accommodate 0 values). All specifications include cell and year FE, propagation controls, and additional controls. Additional controls include commodity price and weather shocks (see Section III). See Appendix A for variable definitions. The time period is restricted to 2008–2015. See Appendix A for further information on the variables.

Figure B16: Heterogeneous effects of intensity of messaging on LRA-related fatalities



Notes. The figure depicts the marginal effects of the intensity of messaging estimated using equation (2) with different sub-samples. For continuous variables, sub-groups are defined according to the sample median of the corresponding variable. Confidence intervals are reported at 90% of confidence allowing standard errors to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). The dependent variable is the number of fatalities linked to LRA. The number of fatalities is reported in logarithms (adding one unit before taking logarithms to accommodate 0 values). All specifications include cell and year FE, propagation controls, and additional controls. Additional controls include commodity price and weather shocks (see Section III). See Appendix A for variable definitions. The time period is restricted to 2008–2015. See Appendix A for further information on the variables.

B.19 Content heterogeneity

This section focuses on the effectiveness of specific messaging content. We estimate equation (2) by using LRA-related fatalities as the main dependent variable, but splitting the main independent variable – intensity of messaging – into different categories based on its content or type. We define the content-specific intensity of messaging by equation (1), using information on the share of cell covered by and the daily minutes spent broadcasting specific contents. We explain the different categories in the next paragraphs. Figure B17 presents averages and standard deviations for these variables. Figure B18 summarizes the marginal effects of the different types of content intensity on LRA-related fatalities. For ease of interpretation, when computing marginal effects all variables related to content are standardized using the mean and standard deviation of the main source of variation in the paper, that is, intensity of messaging. The figure also reports joint-tests of equality of the effects within each category.

First, in order to control for the potentially confounding effects of alternative content broadcast by radios that participated in the defection messaging effort, we collect information on the number of minutes of news, entertainment, and religious programming broadcast by these stations. We construct measures of coverage and the intensity of broadcast of alternative content (similar to equation 1), and we estimate equation (2) using both measures simultaneously. We show that the effect captured by intensity of messaging is specific to defection messages, while alternative content has no effect on fatalities. This establishes a clear link between messaging content and violence, rather than just radio broadcast.

Second, to understand further what type of defection messaging content is more effective, we classify defection messaging into two categories based on the type of messages broadcast: sensitization content and logistical content. Sensitization content primarily refers to programs that feature interviews with ex-combatants, family members, and community leaders, where the interviewees talk about their experiences returning home, and/or make emotional appeals to their friends and kin to follow suit. Logistical content, on the other hand, provides specific details on surrendering. Using information about the number of broadcast minutes in each category, we build content-specific intensity of messaging.[4] While logistical content is likely to be closer to an information/beliefs channel, sensitization content is more directly associated with preferences. Estimating the effects of intensity of both types of content using the main specification leads to comparable effects. Estimates are highly imprecise due to the high correlation between the two types of intensities (0.96). This is because, by design, defection messages are often broadcast including both types of content, rather than focusing on a specific one. It is therefore difficult to disentangle the relative effectiveness of each type of content. This is borne out by a text analysis exercise that we undertake in Appendix E on 86 digitized defection messages. The text analysis also shows that messages with both sensitization and logistical content were being broadcast.

[4] Radio survey respondents were asked to report the number of minutes in each category of defection messaging. When this information is not available, we split the total amount of defection messaging equally among the categories that have been reported as being part of the defection content for the radio station.

Third, we complement the above analysis exploiting the specialization of the radios broadcasting the messages. While content is relatively similar across all radio stations participating in defection messaging, DDR stations are managed by the UN / armed forces and are therefore more specialized in logistical content (UN-DDR, 2014). On the contrary, the other radio stations are managed by local communities, and thus specialize in sensitization content. We observe similar effects of intensity of messaging broadcast from both types of radio stations.

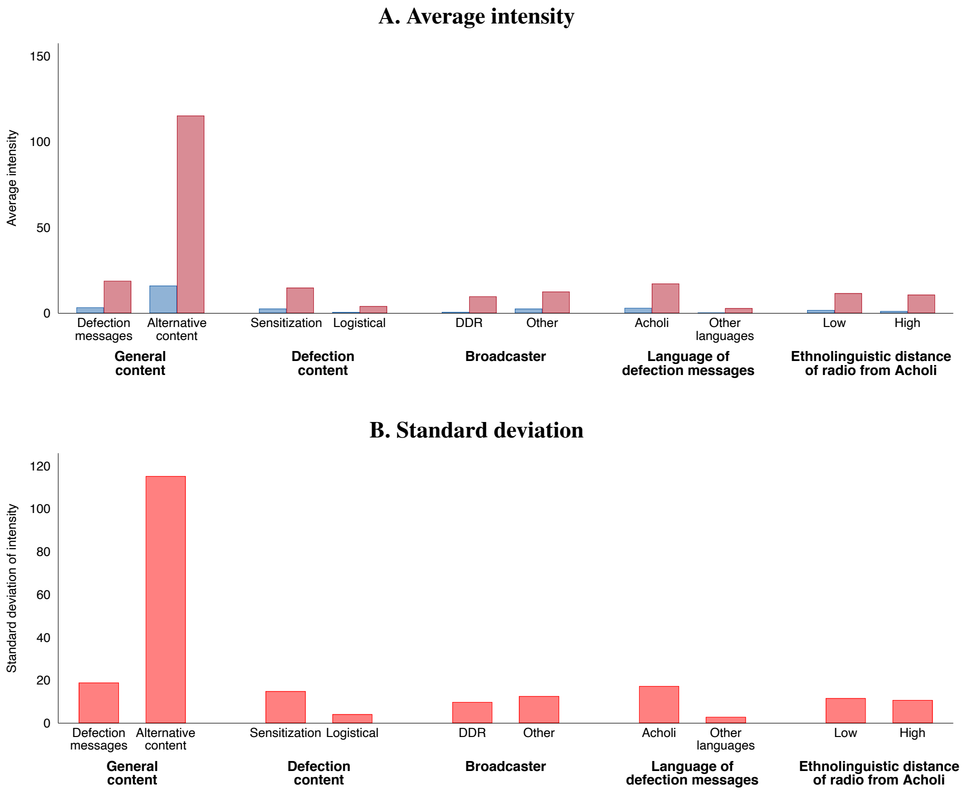
Fourth, we distinguish between defection content broadcast in Acholi (in addition to other languages), and defection content broadcast exclusively in languages other than Acholi. While most radio stations broadcast in both Acholi and a number of local languages, some broadcast exclusively in other languages. Messaging that is exclusively broadcast without using Acholi to be more effective than messaging that uses Acholi. This may indicate that most current recruits are drawn from populations who are not Acholi and hence, messaging in Acholi is less successful. It might also indicate that messaging has a lesser effect on senior members of the LRA and recruits from the earlier phase of the conflict who are more likely to be Acholi. More generally, this could also indicate that the messages are less successful with Acholi people. Below we argue that this need not be the case.

Fifth, we distinguish between radio stations according to their (ethnic) alignment to the Acholi. We exploit the general broadcast language of the radio stations. This dimension might indicate its ethnic identity or that of the population to which it usually caters. We calculate each stations ethnic distance by averaging the distance of each language used to broadcast content (including broadcasting of alternative content) from Acholi.[5] We distinguish between radio stations with linguistic distance equal or lower to 0.75 (the median in the sample) and stations with linguistic distance larger than 0.75. We then compute intensity of messaging separately for each type of radio station. We do not find any significant difference between defection messaging broadcast by the two types of radios. This suggests that messaging is not necessarily less successful with the Acholi people as the result in the previous point might have suggested. Ethnic identity of the radio station or its distance from Acholi does not seem to matter, while the actual broadcast language matters more.

To further analyze the heterogeneity of the effect by radio station, the left panel in figure B19 plots marginal effects of messaging intensity on LRA-related fatalities by each radio station. Marginal effects are estimated using equation (2) separately for each radio station's intensity of messaging. The specification includes only the radio-specific intensity of messaging, defined by equation (1) using information relative to the specific radio station (share of cell covered and daily minutes of broadcasting). The right panel in figure B19 plots instead marginal effects of intensity of messaging on LRA-related fatalities by excluding one radio station at a time. The estimate is stable, providing evidence that the effect is not driven by a specific radio station.

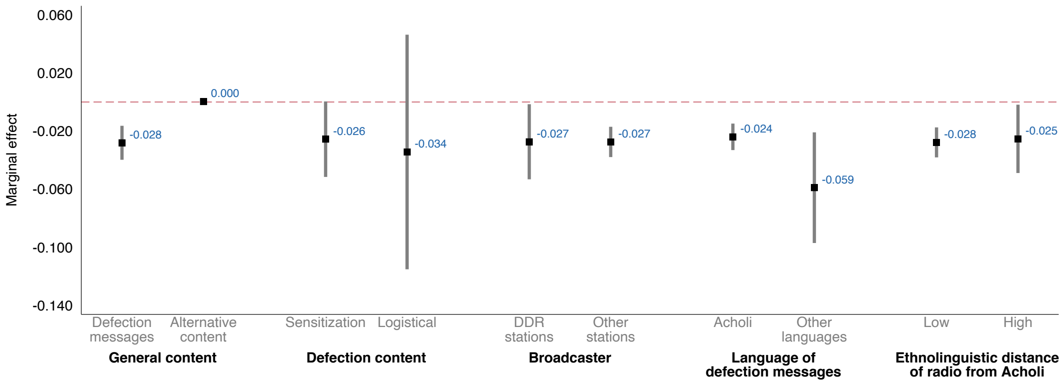
[5]The distance metric is based on language trees. See Section B.16 for a definition of linguistic distance.

Figure B17: Intensity of messaging, by content or type of broadcast content



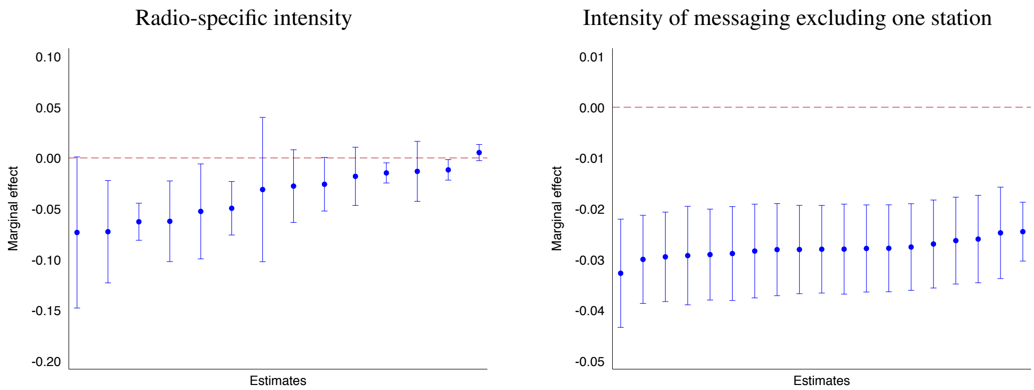
Notes. Figure A graphs the average intensity of messaging for different contents, as defined by equation (1) using information relative to the specific content (share of cell covered and daily broadcasting minutes). Figure B graphs the standard deviations of the same variables. Ethnolinguistic distance is computed using equation (4) and assuming a decay factor $\delta = 0.05$ as in Desmet et al. (2009). Data on language trees come from the Ethnologue dataset. See Appendix A for variable definitions. The time period is restricted to 2008–2015.

Figure B18: Effect of intensity on the number of fatalities, by content or type of broadcast



Notes. The figure presents marginal effects of the intensity of messaging estimated using equation (2) in which intensity of messaging is split into different measures of intensity or coverage. Ethnolinguistic distance is computed using equation (4) and assuming a decay factor $\delta = 0.05$ as in Desmet et al. (2009). Data on language trees come from the Ethnologue dataset. Confidence intervals are reported at 90% of confidence allowing standard errors to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). The dependent variable is the number of fatalities linked to the LRA. The number of fatalities is reported in logarithms (adding one unit before taking logarithms to accommodate 0 values). All specifications include cell and year FE, propagation controls, additional controls, and interactions between year dummies and macro-region dummies. Additional controls include commodity price and weather shocks (see Section III). See Appendix A for variable definitions. The time period is restricted to 2008–2015. The p-values of a joint-test of equality of the effects within each category are: 0.001 (general content); 0.890 (defection content); 0.995 (broadcaster); 0.152 (language of defection messages); 0.892 (ethnolinguistic distance of radio from Acholi).

Figure B19: Radio-specific effect of intensity of messaging on LRA-related fatalities



Notes. The left figure presents marginal effects of the intensity of messaging estimated using equation (2) separately for each radio station. Five stations with the smallest intensity have been merged to the nearest stations in order to allow for sufficient variation. The specification includes the intensity of messaging for the specific radio station under consideration. The right panel in figure B19 plots instead marginal effects of intensity of messaging on LRA-related fatalities by excluding one radio station at a time. Coefficients are ordered by point estimate, from the smallest to the largest. Confidence intervals are reported at 90% of confidence allowing standard errors to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). The dependent variable is the number of fatalities linked to the LRA. The number of fatalities is reported in logarithms (adding one unit before taking logarithms to accommodate 0 values). All specifications include cell and year FE, propagation controls, additional controls, and interactions between year dummies and macro-region dummies. Additional controls include commodity price and weather shocks (see Section III). See Appendix A for variable definitions. The time period is restricted to 2008–2015.

B.20 Measurement error in the dependent variables

The conflict data come from three event-based datasets: primarily LRACT, but also ACLED and UCDP. These datasets rely on the accurate reporting of events from different sources, such as, reports from news agencies, NGOs, and governments. Given the reliance on media, violence might be under-reported due to low media coverage in remote areas where much of the violence is concentrated (see, for instance, Eck, 2012). The main source of information, LRACT limits this possibility by drawing on a widespread network of field sources, some linked by high frequency (HF) radio. This allows LRACT verifiers to find corroborating accounts of events sourced from other channels, as well as report events that alternative event-based datasets fail to capture. Despite this, reporting errors in remote areas might bias estimates if, conditional on observable characteristics, the measurement error is correlated with the intensity of messaging. The longitudinal dimension of the data allows controlling for cell-level fixed effects, which capture time-invariant characteristics related to measurement error, such as remoteness. The potential issue of correlation between intensity of messaging and measurement error is therefore reduced to the idiosyncratic component of the measurement error.

One potential issue related to the idiosyncratic component of measurement error might arise if an increase in intensity of messaging induces a reduction in killings by the LRA, thus affecting the number of individuals who could observe and report a violent event. In this case, a higher intensity of messaging would be positively correlated with the error term, leading to an underestimation of the true effect of defection messaging. Alternatively, higher exposure to defection messaging might also shift attitudes towards the reporting of violent events. For instance, higher intensity might raise salience of reporting, therefore increasing the probability of reporting an event, and leading to an underestimation of its true effect. On the other hand, higher exposure to defection messaging might also translate into fear of retaliation for reporting an event, leading to the lower probability of recording one, and therefore an overestimation of the true effect.

To further understand the role of measurement error in the three main sources of information, we focus on two potential sources of error: one that is common to all the three datasets, and one that is dataset-specific. To understand the relative importance of these sources, we estimate a single-factor measurement model for the recording of fatalities in these datasets. We construct three indicator variables indicating whether each dataset records at least one fatality in a cell, and we label the variables as m_i, where $i = \{\text{LRACT, ACLED, UCDP}\}$. We assume that these measures are affected by a variety of factors, which load differently on to the observed variables. Standard for a measurement model, we allow for observable factors, X; pure random noise specific to each measure and assumed to be i.i.d. normal, ϵ_i; and one common unobservable factor, V, which reflects unobservable characteristics that load on to observed measures and that are common across the three sources. For instance, this could be capturing underlying true violence or the attitudes of the local population towards reporting. Each equation in the single-factor model is

Table B22: Single-factor measurement model for LRA-related positive fatalities

	Measure for LRA-related fatalities ($\mathbf{m_i}$)		
	m_{LRACT}	m_{ACLED}	m_{UCDP}
	(1)	(2)	(3)
Common unobservable factor (V)	0.784	1.000	1.106
	(0.148)	(constrained)	(0.211)
Raw correlations:			
m_{LRACT}	1.000	.	.
m_{ACLED}	0.282	1.000	.
m_{UCDP}	0.374	0.377	1.000
Observations	60600	60600	60600

Notes. The table presents estimates for λ_i in the single-factor model described by equation (5). For identification, we normalize $\lambda_{ACLED} = 1$. We assume a Bernoulli process and a probit link and estimate the model using maximum likelihood. m_{LRACT}, m_{ACLED}, and m_{UCDP} are indicator variables equal to 1 if at least one fatality is recorded by the corresponding dataset in a specific cell, and 0 otherwise. Errors are assumed to be i.i.d. and clustered at the cell-level. Observable factors (X) are composed by the distance polynomial and by cell characteristics (ruggedness, forest cover, and country indicators). The time period is restricted to 2008–2015. See Appendix A for further information on the variables.

represented by the following specification:

$$m_i = \alpha_i + \lambda_i V + \beta_i X + \varepsilon_i \tag{5}$$

where λ_i is the loading on the unobservable factor. β_i are the coefficients on observable regressors. Table B22 presents the estimates for λ_i and the pairwise correlation coefficients of the three measures. The common factor has comparable loadings across the three datasets (coefficients are not statistically different), and more importantly have the same sign. This indicates that all three datasets capture a common (unobservable) process in a similar way. In addition, this suggests that differences across datasets are more likely to be resulting from dataset-specific differences (such as the definition of a violent event, as discussed in section II.B) or to dataset-specific measurement error.

To address potential issues related to dataset-specific measurement error, we combine information from the three sources into a single dependent variable measuring the number of fatalities. Combining information from multiple sources affected by independent measurement error allows us to reduce the error. In Table B23 we estimate the effect of intensity of messaging using combined information from the LRACT, ACLED, and UCDP as dependent variables. Column 1 uses the average number of LRA-related fatalities across the three datasets, column 2 uses the median, while columns 3 and 4 use the minimum and maximum number of fatalities. Column 5 presents results using principal component analysis (PCA) applied to LRA-related fatalities across the three datasets, while column 6 makes use of PCA, but includes as measures not only LRA-related fatalities, but also non-LRA-related fatalities from the ACLED and UCDP datasets. All dependent variables are log-transformed (adding one unit before taking logarithms to accommodate 0 values). Results are robust to using combined information from the three sources.

Table B23: Intensity of messaging and combined information from LRACT, ACLED and UCDP

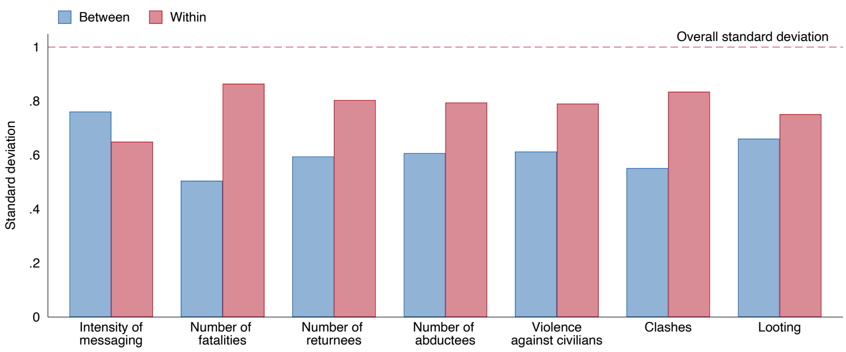
	Combination of number of fatalities in LRACT, ACLED, and UCDP					
Dependent variable:	Mean	Median	Min	Max	Principal Component Analysis	
	(1)	(2)	(3)	(4)	(5)	(6)
Intensity of messaging	-0.028	-0.020	-0.011	-0.038	-0.028	-0.028
	(0.005)	(0.003)	(0.002)	(0.007)	(0.005)	(0.005)
Observations	60600	60600	60600	60600	60600	60600
Number of cells	7575	7575	7575	7575	7575	7575
Includes non-LRA fatalities	No	No	No	No	No	Yes

Notes. Standard errors in parentheses are allowed to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). The dependent variables are the average number of LRA-related fatalities in LRACT, ACLED, and UCDP datasets (column 1); the median number of LRA-related fatalities in LRACT, ACLED, and UCDP datasets (column 2); the minimum number of LRA-related fatalities in LRACT, ACLED, and UCDP datasets (column 3); the maximum number of LRA-related fatalities in LRACT, ACLED, and UCDP datasets (column 4); the predicted number of fatalities computed using the first component of PCA and using the number of LRA-related fatalities in LRACT, ACLED, and UCDP datasets as measures (column 5); the predicted number of fatalities computed using the first component of PCA and using the number of LRA-related fatalities in LRACT, ACLED, and UCDP datasets, and the number of non-LRA-related fatalities in ACLED, and UCDP as measures (column 6). Dependent variables are measured in logarithms (adding one unit before taking logarithms to accommodate 0 values). All specifications include cell and year fixed effects, propagation controls, additional controls and interactions between year dummies and macro-region dummies. Additional controls include commodity price and weather shocks (see Section III). The time period is restricted to 2008–2015. See Appendix A for further information on the variables.

B.21 Topographical versus time variation

The identification strategy presented in the main text combines two sources of variation in the intensity of messaging: topographical and time variation. Figure B20 presents the between and within decomposition of the overall variation in intensity of messaging and in the outcome variables related to violence (number of fatalities, number of returnees, number of abductees, number of events characterized by violence against civilians, clashes with security forces and looting). The between-cell component displays the cross-sectional variation, while the within-cell component displays the temporal variation. All variables present significant variations in both components, suggesting both messaging and violence varies both over time and space. Outcome variables related to violence have a larger share of variation from the within component.

Figure B20: Between and within variation in the variables of interest



Notes. The figure plots the between and within decomposition of the overall standard deviation of the corresponding variables. All outcomes are computed from the LRACT dataset and measured in logarithms (adding one unit before taking logarithms to accommodate 0 values). The dotted line indicates the overall standard deviation for all variables. It is normalized to 1 as all variables are standardized before the decomposition. The spatial unit is the cell, the time unit is the year. The time period is restricted to 2008–2015.

To identify the effect attributed to topographical variation, we begin by focusing on free-space circular coverage of radio signals in addition to topography-corrected coverage. Free-space circular coverage does not correct signal reception for topography and is computed using a free-space propagation model, which captures the attenuation and maximum reach of an audible signal in the absence of obstacles in a flat terrain. Hence, it captures the coverage of the radio signal in the absence of (random) topographic variation. Panel A in Table B24 shows estimates of equation (2) when controlling for circular coverage. The correlation between intensity of messaging and circular coverage is equal to 0.43 in the whole study area (0.36 if restricted to be within 100 km from an antenna). Similarly, panel D in Table B24 shows estimates of equation (2) when controlling for intensity of messaging if there were no topographic obstacles. Estimates of the effect of intensity of messaging are robust to the robustness checks in panel A and D. Note however that free-space coverage captures the area targeted by messaging and is therefore potentially endogenous to idiosyncratic shocks related to violence.

Restricting variation to areas targeted by messaging and exploiting topographical variation within targeted areas, panels B and C in Table B24 show estimates of the effect of intensity of messaging using different sub-samples according to free-space coverage. Restricting the sample to these areas allows for the identification of the effect owing to topographical variation within the targeted area. Panel B restricts the sample to cells covered at a specific point in time (contemporaneous to the measurement of violence), while panel C restricts the sample to cells covered at any point in time over the period 2008–2015. The effect of defection messaging estimated according to this procedure is in line with the main results but tend to be slightly smaller, though not statistically different. This suggests that both topographical and time variation contribute to the identification of the effect of intensity of messaging.

An alternative strategy to separate the contributions of topographical and time variation to the effect of messaging intensity is to estimate equation (2) using different sets of area fixed effects.[6] First, we estimate equation (2) by pooling all the data over time without including any area fixed effects. On the other hand, to capture time variation, we estimate equation (2) using area fixed effects of different sizes. We split the study area in equally sized cells, gradually moving from larger to smaller cells, by dividing the latitude and longitude in the study area into equally sized groups. We focus on the following decomposition of the study area: 100 cells (10 groups per latitude by 10 groups per longitude), 400 cells (20 by 20), 900 cells (30 by 30), 1,600 cells (40 by 40), and 2,500 cells (50 by 50). Finally, we include cells at the $0.125° \times 0.125°$ resolution, as in the benchmark specification (7,575 cells). For comparability with Table B24, we restrict the sample to areas within circular coverage.

Figure B21 plots the marginal effects of intensity of messaging for the different specifications. For most outcomes, results correspond to the conclusions obtained from Table B24: both topographical and time variation contribute to the results. For fatalities, the effect from cross-sectional

[6]We do not control for macro-region by year fixed effects to avoid confounding the comparison of estimates with different area fixed effects.

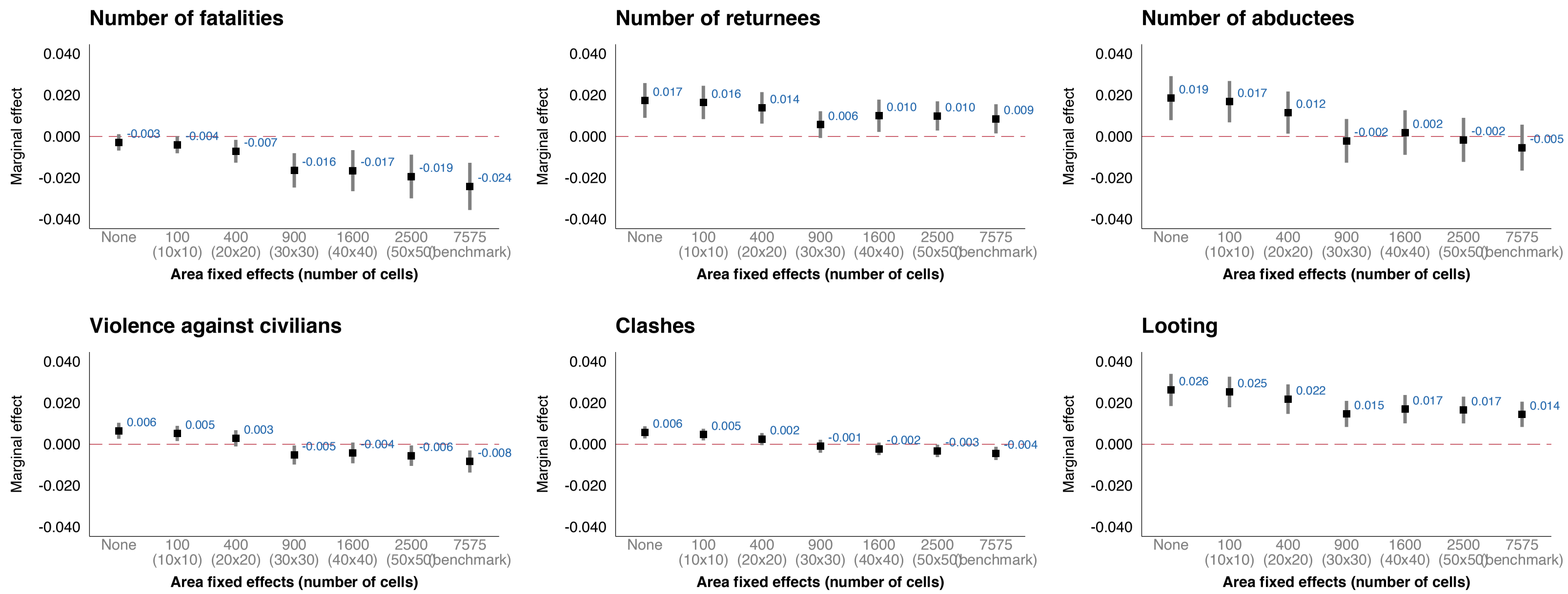
Table B24: Defection messaging and circular coverage as a control variable

Dependent variable:	**Number of fatalities**	**Number of individuals...** Returning	Being abducted	**Number of events involving...** Violence against civilians	Clashes	Looting
	(1)	(2)	(3)	(4)	(5)	(6)
A. Control for circular coverage of radio signal						
Intensity of messaging	-0.028	0.009	-0.005	-0.010	-0.006	0.016
	(0.005)	(0.004)	(0.005)	(0.004)	(0.003)	(0.004)
Cell covered (circular coverage)	-0.008	0.006	-0.010	-0.008	0.000	-0.002
	(0.006)	(0.005)	(0.009)	(0.004)	(0.003)	(0.004)
Observations	60600	60600	60600	60600	60600	60600
Number of cells	7575	7575	7575	7575	7575	7575
B. Sample restricted to cells within free-space circular coverage at a specific point in time						
Intensity of messaging	-0.023	0.008	-0.005	-0.008	-0.004	0.014
	(0.006)	(0.004)	(0.006)	(0.005)	(0.003)	(0.004)
Observations	11584	11584	11584	11584	11584	11584
Number of cells	2208	2208	2208	2208	2208	2208
C. Sample restricted to cells within free-space circular coverage at any point in time						
Intensity of messaging	-0.024	0.008	-0.004	-0.008	-0.005	0.015
	(0.005)	(0.004)	(0.005)	(0.005)	(0.003)	(0.004)
Observations	17664	17664	17664	17664	17664	17664
Number of cells	2208	2208	2208	2208	2208	2208
D. Control for intensity of messaging built using free-space circular coverage						
Intensity of messaging	-0.025	0.004	-0.011	-0.011	-0.006	0.012
	(0.006)	(0.004)	(0.006)	(0.005)	(0.003)	(0.004)
Intensity of messaging (free-space)	-0.004	0.012	0.011	0.002	-0.000	0.008
	(0.003)	(0.004)	(0.005)	(0.003)	(0.002)	(0.003)
Observations	60600	60600	60600	60600	60600	60600
Number of cells	7575	7575	7575	7575	7575	7575

Notes. The table reports marginal effects estimated using a fixed effects model. Standard errors in parentheses are allowed to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). The dependent variables are the number of fatalities linked to LRA (column 1), the number of individuals returning (column 2), the number of abductees (column 3) and the number of other LRA-related violent events (columns 4–6). The dependent variables are measured in logarithms (adding one unit before taking logarithms to accommodate 0 values). In panel A, coefficients are estimated using equation (2) and controlling for free-space circular coverage. Cell covered (circular coverage) is a dummy variable equal to 1 if the cell is covered by free-space circular coverage of radio signal, and 0 otherwise. In panel B, coefficients are estimated using equation (2) and restricting the sample to cells within free-space circular coverage at a specific point in time (contemporaneous to the measurement of violence). In panel C, coefficients are estimated using equation (2) and restricting the sample to cells within free-space circular coverage at any point in time during the period 2008–2015. In panel D, coefficients are estimated using equation (2) controlling for intensity of messaging if there were no topographic obstacles. All specifications include cell and year fixed effects, propagation controls, additional controls and interactions between year dummies and macro-region dummies. Additional controls include commodity price and weather shocks (see Section III). The time period is restricted to 2008–2015. See Appendix A for further information on the variables.

variation continues to be negative, but is smaller in magnitude compared to the overall effect. Including additional area fixed effects leads to larger estimates, which continue to be statistically different from zero. For the number of abductees, violence against civilians and clashes with security forces, pure cross-sectional variation leads to an opposite effect of intensity of messaging. This suggests that in the cross-section there could be adaptation of the amount of messaging that could be related to specific characteristics of cells, which are captured by area fixed effects. Though often the pure cross-section effects are statistically indistinguishable from zero.

Figure B21: Cross-sectional versus panel variation



Notes. The figure plots marginal effects of intensity of messaging estimated using equation (2) and either no area fixed effects (left-most specification) or different sets of area fixed effects (reported in the x-axis). Area indicator variables are generated by splitting the study area into equally sized cells. The dependent variables are the number of fatalities (upper left panel), the number of returnees (upper right panel), the number of abductees (middle left panel), the number of events characterized by violence against civilians (middle right panel), clashes with security forces (bottom left panel), and looting (bottom right panel). All outcomes are computed from the LRACT dataset and are measured in logarithms (adding one unit before taking logarithms to accommodate 0 values). For comparison, we allow the vertical axis to vary at the same scale for all outcome variables. Confidence intervals are computed at 95% of confidence, standard errors are clustered at the 0.125° × 0.125° resolution cell level. All specifications include propagation controls and additional controls. Additional controls include commodity price and weather shocks (see Section III). In the specification not using the 0.125° × 0.125° resolution cell fixed effects, we include time-invariant cell-level characteristics as controls. These include ruggedness, population (measured in 2005), nightlight (measured in 2007), country indicator variables, urban and forest cover, and baseline violence (measured as average number of fatalities in the pre-study period). The time period is restricted to 2008–2015. See Appendix A for further information on the variables.

B.22 Test for linearity in the effect of intensity of messaging

To test for linearity in the effect of intensity of messaging on the different outcomes analyzed in figure 6, we estimate the following specification:

$$y_{irt} = \gamma_i + \alpha\, f(dm_{it}) + \mathbf{X}'_{it}\beta + \delta_t M_r + u_{it} \tag{6}$$

in which we allow y_{irt} to depend non-linearly on dm_{it}. We estimate three versions of equation (6) allowing $f(dm_{it})$ to be a polynomial of second, third and fourth degree in dm_{it}. We test the null hypothesis of y_{irt} depending linearly on dm_{it} by testing for (joint-)significance of the terms in the polynomial $f(dm_{it})$ with power larger than one. For instance, when $f(dm_{it})$ is a polynomial of second degree, we estimate equation (6) by introducing dm_{it} and dm_{it}^2, and then we test for $dm_{it}^2 = 0$. A rejection of the null hypothesis indicates a non-linear relationship between y_{irt} and dm_{it}, conditional on the other controls. Table B25 presents F statistics and the corresponding p-values for these test.

Table B25: Test for linearity in the effect of intensity of messaging

	Polynomial degree for $f(dm_{it})$					
	Second		Third		Fourth	
	F	p-value	F	p-value	F	p-value
Dependent variables:	(1)	(2)	(3)	(4)	(5)	(6)
Number of fatalities	4.526	0.033	2.418	0.089	1.708	0.163
Number of returnees	4.097	0.043	6.836	0.001	4.878	0.002
Number of abductees	4.269	0.039	5.478	0.004	3.883	0.009
Events: violence against civilians	3.541	0.060	2.473	0.084	1.691	0.167
Events: clashes	4.286	0.038	2.852	0.058	1.828	0.140
Events: looting	1.137	0.286	2.906	0.055	1.973	0.116

Notes. F statistics and corresponding p-values are from a joint-test of significance of the terms in the polynomial $f(dm_{it})$ with power larger than one, estimated in equation (6). The dependent variables, listed in each row, are measured in logarithms (adding one unit before taking logarithms to accommodate 0 values). All specifications include cell and year FE, propagation controls, additional controls and interactions between year dummies and macro-region dummies. Additional controls include commodity price and weather shocks (see Section III). Standard errors are allowed to be correlated over time and space (Conley, 1999, 2008; Hsiang, 2010). The time period is restricted to 2008–2015. See Appendix A for further information on the variables.

B.23 Aggregate effect of defection messaging on the LRA insurgency (2008–2015)

The aggregate impact of defection messaging is calculated as the difference between the actual estimate (provided by event-based datasets) and the counterfactual estimate, which imposes an alternative level of intensity. We compute aggregate impacts from equation (2) by predicting the dependent variable under different scenarios of messaging, holding all other control variables constant. We focus on two scenarios: absence of defection messaging (zero intensity of messaging) and a reduction of 50% in messaging intensity. We estimate confidence intervals using 2000 bootstrap iterations.

We estimate equation (2) using three alternative specifications for the dependent variable. First, we focus on the logarithm, adding one unit before taking logarithms to accommodate 0 values. However, when estimating the model using a log-transformed dependent variable, predicting outcomes in levels exponentiating predicted values from equation (2) results in biased predictions.

This issue, known as the re-transformation problem (Duan, 1983; Wooldridge, 2010), requires imposing distributional assumptions relative to the error term, but in the presence of heteroskedasticity in the data this requires stronger assumptions (Manning and Mullahy, 2001; Manning et al., 1998). We therefore re-transform predicted values to levels assuming heteroscedastic idiosyncratic error terms, and we follow a non-parametric procedure as in Duan (1983). Assume that the main estimates are produced using the following specification:

$$log(y_{it} + a) = \mathbf{X'}_{it}\beta + \gamma_i + u_{it} \tag{7}$$

where y_{it} is the outcome of interest for predictions, a is a positive constant, $\mathbf{X}_{it}$ is a vector of time-varying characteristics, γ_i are cell fixed effects and u_{it} are idiosyncratic error terms. Predicting the outcome of interest requires the following re-transformation:

$$E[y_{it}|\mathbf{X}_{it}, \gamma_i] = e^{\{\mathbf{X'}_{it}\beta+\gamma_i\}}\, E[e^{u_{it}}|\mathbf{X}_{it}, \gamma_i] - a \tag{8}$$

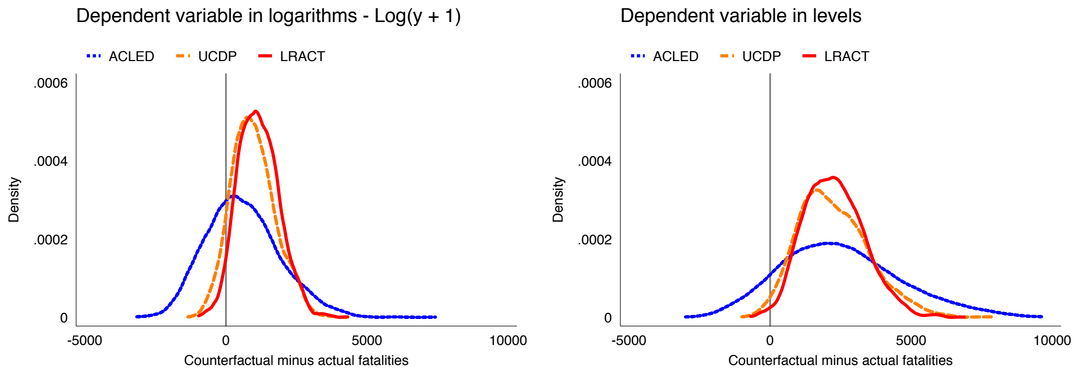
where $E[e^{u_{it}}|\mathbf{X}_{it}, \gamma_i]$ is known as the *smearing factor*. We assume heteroskedasticity and estimate the smearing factor using a fixed effect model similar to equation (7), where the dependent variable is the exponentiated predicted u_{it} from equation (7). Second, we focus on an alternative log-transformation by using the logarithm of the outcome of interest, adding 0.5 before taking logarithms. We follow a similar re-transformation to report predictions in levels. Finally, to avoid the re-transformation problem, we also present estimates using the dependent variable in levels.

Given the evidence of non-linearity in the effect of intensity of messaging (Appendix B.22), we maximize predictive power by using a more flexible functional form to estimate its effect. We include both intensity of messaging and its square as independent variables in equation (2). Table B26 presents the results. For comparison, Table B27 presents counterfactual estimates based on the benchmark specification of equation (2). Figure B22 shows the distribution of actual minus counterfactual estimates for fatalities under both specifications.

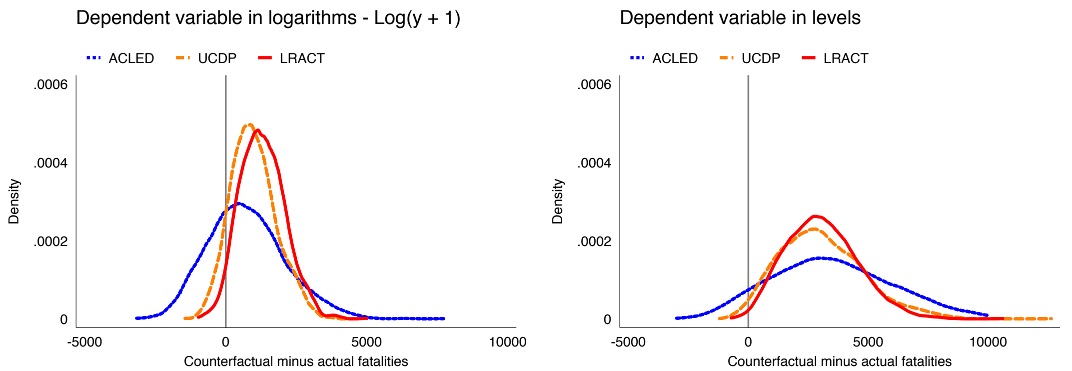
The difference between standardized effects presented in tables 2 and 3 on the one hand and aggregate effects on the other, relates to the intent-to-treat nature of any radio-based intervention. While fatalities and violent events are dispersed over the geographical space under consideration, they tend to occur in specific cells, rather than clusters of surrounding cells. On the contrary, intensity of messaging covers larger areas due to the propagation of radio waves. Thus, several cells in which no fatalities or events ever take place during the intervention period (possibly because there is no presence of armed groups) are also treated. As a result, treating larger numbers of these cells drives cell-level estimates of the impact of (standarized) intensity of messaging down, while the aggregate number of fatalities remains unchanged. To reinforce this statement, note that increasing the cell size raises the standard deviation of intensity of messaging, but estimates of its standardized effect on the number of fatalities also increase in magnitude (see Appendix B.3).

Figure B22: Reduction in fatalities attributed to defection messaging

A. Quadratic specification



B. Benchmark specification



Notes. Distribution of bootstrap-predicted values of the reduction in fatalities attributed to defection messaging using ACLED, UCDP and LRACT datasets. For each bootstrap iteration, this is computed as the difference between the actual estimate, as reported in the datasets, and the counterfactual estimate, computed imposing zero intensity of messaging and predicting outcomes using estimates from equation (2). In panel A, both intensity of messaging and its square are included as independent variables in equation (2). In panel B, intensity of messaging is included linearly as the independent variable in equation (2). 2000 iterations are performed. The left figure presents estimates using as dependent variable the logarithm of fatalities, adding one unit before taking logarithms to accommodate 0 values. The right figure presents estimates using as dependent variable the number of fatalities in levels. Re-transformation of predicted values in levels is performed assuming heteroskedasticity and using the procedure illustrated in Duan (1983). See Appendix A for further information on the variables.

Table B26: Aggregate effects of defection messaging: quadratic specification

	Actual	Counterfactual scenario				
	Estimate	Estimate	C.I. bounds		Impact	
			Lower	Upper	Δ	% change
	(1)	(2)	(3)	(4)	(5)	(6)
SCENARIO A: zero intensity of messaging						
Dependent variable in logs - Log(y + 1)						
Number of fatalities	3137	4288	3195	5541	-1151	-26.8%
Number of fatalities (UCDP)	3348	4317	3126	5744	-969	-22.4%
Number of fatalities (ACLED)	5263	5872	3906	8203	-609	-10.4%
Number of returnees	2073	1776	1451	2135	297	16.8%
Events: violence against civilians	989	1099	922	1284	-110	-10.0%
Events: clashes	501	553	458	657	-52	-9.4%
Events: looting	1158	837	702	976	321	38.3%
Dependent variable in logs - Log(y + 0.5)						
Number of fatalities	3137	4172	3115	5391	-1035	-24.8%
Number of fatalities (UCDP)	3348	4217	3053	5605	-869	-20.6%
Number of fatalities (ACLED)	5263	5737	3829	7999	-474	-8.3%
Number of returnees	2073	1799	1474	2155	274	15.3%
Events: violence against civilians	989	1090	916	1275	-101	-9.3%
Events: clashes	501	554	461	656	-53	-9.5%
Events: looting	1158	866	730	1009	292	33.7%
Dependent variable in levels						
Number of fatalities	3137	5407	3832	7260	-2270	-42.0%
Number of fatalities (UCDP)	3348	5608	3798	7872	-2260	-40.3%
Number of fatalities (ACLED)	5263	7798	4571	11659	-2535	-32.5%
Number of returnees	2073	1550	1214	1928	523	33.7%
Events: violence against civilians	989	1133	943	1339	-144	-12.7%
Events: clashes	501	553	447	672	-52	-9.4%
Events: looting	1158	625	509	744	533	85.4%
SCENARIO B: -50% in intensity of messaging						
Dependent variable in logs - Log(y + 1)						
Number of fatalities	3137	3594	2759	4506	-457	-12.7%
Number of fatalities (UCDP)	3348	3726	2753	4873	-378	-10.1%
Number of fatalities (ACLED)	5263	5469	3734	7483	-206	-3.8%
Number of returnees	2073	1916	1595	2257	157	8.2%
Events: violence against civilians	989	1035	875	1203	-46	-4.5%
Events: clashes	501	525	444	616	-24	-4.5%
Events: looting	1158	988	840	1151	170	17.2%
Dependent variable in logs - Log(y + 0.5)						
Number of fatalities	3137	3505	2692	4386	-368	-10.5%
Number of fatalities (UCDP)	3348	3650	2695	4777	-302	-8.3%
Number of fatalities (ACLED)	5263	5381	3680	7326	-118	-2.2%
Number of returnees	2073	1924	1602	2266	149	7.8%
Events: violence against civilians	989	1027	869	1192	-38	-3.7%
Events: clashes	501	524	443	614	-23	-4.4%
Events: looting	1158	998	850	1161	160	16.0%
Dependent variable in levels						
Number of fatalities	3137	4275	3153	5569	-1138	-26.6%
Number of fatalities (UCDP)	3348	4484	3146	6126	-1136	-25.3%
Number of fatalities (ACLED)	5263	6547	4201	9382	-1284	-19.6%
Number of returnees	2073	1813	1498	2156	260	14.3%
Events: violence against civilians	989	1060	893	1238	-71	-6.7%
Events: clashes	501	526	441	622	-25	-4.8%
Events: looting	1158	891	768	1020	267	29.9%

Notes. Estimates refer to the period 2008–2015. Actual estimate corresponds to estimates computed using the LRACT (if not otherwise indicated), UCDP and ACLED datasets. The counterfactual estimate is computed imposing alternative scenarios about intensity of messaging and predicting the outcome variables using estimates from equation (2). Intensity of messaging is included as an independent variable in equation (2) with a linear and a quadratic term. Confidence intervals are estimated iterating the procedure 2000 times using bootstrap. The lower bound corresponds to the 5^{th} percentile of the empirical distribution of predicted values, the upper bound corresponds instead to the 95^{th} percentile. Δ is computed as the difference between the actual estimate from column 1, and the counterfactual estimate from column 2. For log-transformed dependent variables, predictions are reported in levels assuming heteroskedasticity and using a procedure illustrated in Duan (1983). See Appendix A for further information on the variables.

Table B27: Aggregate effects of defection messaging: linear specification

	Actual	**Counterfactual scenario**				
	Estimate	Estimate	C.I. bounds		Impact	
			Lower	Upper	Δ	% change
	(1)	(2)	(3)	(4)	(5)	(6)
SCENARIO A: zero intensity of messaging						
Dependent variable in logs - Log(y + 1)						
Number of fatalities	3137	4437	3239	5816	-1300	-29.3%
Number of fatalities (UCDP)	3348	4328	3087	5804	-980	-22.6%
Number of fatalities (ACLED)	5263	6003	3931	8453	-740	-12.3%
Number of returnees	2073	1881	1538	2258	192	10.2%
Events: violence against civilians	989	1173	975	1395	-184	-15.7%
Events: clashes	501	606	502	719	-105	-17.3%
Events: looting	1158	875	731	1024	283	32.4%
Dependent variable in logs - Log(y + 0.5)						
Number of fatalities	3137	4262	3135	5561	-1125	-26.4%
Number of fatalities (UCDP)	3348	4204	3007	5625	-856	-20.4%
Number of fatalities (ACLED)	5263	5816	3857	8155	-553	-9.5%
Number of returnees	2073	1888	1550	2261	185	9.8%
Events: violence against civilians	989	1142	953	1351	-153	-13.4%
Events: clashes	501	597	496	707	-96	-16.1%
Events: looting	1158	903	757	1054	255	28.3%
Dependent variable in levels						
Number of fatalities	3137	6199	3922	8782	-3062	-49.4%
Number of fatalities (UCDP)	3348	6399	3834	9587	-3051	-47.7%
Number of fatalities (ACLED)	5263	8928	4883	13801	-3665	-41.0%
Number of returnees	2073	1786	1409	2211	287	16.1%
Events: violence against civilians	989	1332	1063	1651	-343	-25.7%
Events: clashes	501	636	521	761	-135	-21.2%
Events: looting	1158	668	527	806	490	73.3%
SCENARIO B: -50% in intensity of messaging						
Dependent variable in logs - Log(y + 1)						
Number of fatalities	3137	3674	2794	4649	-537	-14.6%
Number of fatalities (UCDP)	3348	3740	2745	4928	-392	-10.5%
Number of fatalities (ACLED)	5263	5543	3762	7620	-280	-5.0%
Number of returnees	2073	1972	1643	2332	101	5.1%
Events: violence against civilians	989	1074	905	1253	-85	-7.9%
Events: clashes	501	552	464	646	-51	-9.2%
Events: looting	1158	1007	851	1174	151	15.0%
Dependent variable in logs - Log(y + 0.5)						
Number of fatalities	3137	3557	2725	4493	-420	-11.8%
Number of fatalities (UCDP)	3348	3653	2688	4806	-305	-8.4%
Number of fatalities (ACLED)	5263	5429	3705	7414	-166	-3.1%
Number of returnees	2073	1972	1645	2332	101	5.1%
Events: violence against civilians	989	1055	890	1229	-66	-6.3%
Events: clashes	501	546	459	639	-45	-8.3%
Events: looting	1158	1017	861	1183	141	13.8%
Dependent variable in levels						
Number of fatalities	3137	4671	3244	6296	-1534	-32.8%
Number of fatalities (UCDP)	3348	4880	3187	6991	-1531	-31.4%
Number of fatalities (ACLED)	5263	7112	4320	10345	-1849	-26.0%
Number of returnees	2073	1931	1592	2303	142	7.4%
Events: violence against civilians	989	1159	958	1391	-170	-14.7%
Events: clashes	501	568	475	670	-67	-11.8%
Events: looting	1158	913	775	1055	245	26.8%

Notes. Estimates refer to the period 2008–2015. Actual estimate corresponds to estimates computed using the LRACT (if not otherwise indicated), UCDP and ACLED datasets. The counterfactual estimate is computed imposing alternative scenarios about intensity of messaging and predicting the outcome variables using estimates from equation (2). Intensity of messaging is included linearly as an independent variable in equation (2). Confidence intervals are estimated iterating the procedure 2000 times using bootstrap. The lower bound corresponds to the 5th percentile of the empirical distribution of predicted values, the upper bound corresponds instead to the 95th percentile. Δ is computed as the difference between the actual estimate from column 1, and the counterfactual estimate from column 2. For log-transformed dependent variables, predictions are reported in levels assuming heteroskedasticity and using a procedure illustrated in Duan (1983). See Appendix A for further information on the variables.

C Survey of radio stations

We collected defection messaging data by conducting original surveys of an exhaustive set of radio stations that have broadcast programs encouraging LRA defections at any time. In addition to the surveys, we cross-referenced and supplemented the data with publicly and privately available digital documentation of the stations and their broadcasts. To start with, we generated a complete list of participating radio stations by cross-referencing policy reports, direct exchanges with several international actors, a thorough review of local news reports and social media, and referrals from local radio operators. Collectively this produced a roster of all stations that broadcast defection messaging from the beginning of the insurgency to the time data collection began. With this list of radio stations in hand, we set out to survey each one and capture all defection broadcasts targeting the LRA since the beginning of the conflict. The data we sought included time-varying broadcast characteristics and frequency, as well as technical parameters of station equipment and the precise geographic locations of the broadcasting antennas.

We conducted two survey waves. In the first round, during 2016 and early 2017, we conducted surveys with logistical support from an international NGO called Invisible Children (IC). IC was directly involved in encouraging defection messaging and provided varying degrees of support to a majority of the radio stations involved in transmitting the "Come Home" messages. In this round we were able to survey 16 radio stations.[7] The second round was conducted by two local survey firms during late 2017 and early 2018: Gaplink (Uganda) and Innovative Hub for Research in Africa / Marakuja Kivu Research (DRC).[8] In total, the two firms covered 22 antennas from 19 stations, including the resurvey of seven stations from the first round. In total we collected information on 26 radio stations (30 antennas) of which 19 radio stations (21 antennas) have broadcast defection messaging targeted at the LRA during 2008–2015.

We conducted a second round of surveys for three reasons. First, we wanted to independently verify the information we gathered in the first round. To this end, we did not share any data from the first round with the firms conducting the second in order to be able to cross-reference two independent sources.[9] Second, we wanted to gather the information on the radio stations in the DRC that were missing from the first round. Third, we wanted to verify that in Uganda there was indeed no messaging during the period of the study (2008–2015), as initial reports had suggested.

The survey was designed for radio station managers at the participating stations and administered by field research assistants. Station managers responded based either on written records (if available) or from recall. In the majority of cases the managers had been personally involved

[7] IC collected data on 14 of 16 stations. For security reasons, we directly obtained data on one station from the management of the radio station (Catholic Radio Network) and on another one using IC's internal reports.

[8] These firms had previous experience in data collection in the region. See, for instance, Bauer et al. (2018); Sanchez de la Sierra (2019).

[9] Security reasons prevented us from revisiting radio stations in Central African Republic and South Sudan. In addition, one of the radio stations interviewed in phase 1 in DRC was closed by the time of phase 2. In phase 2, we also surveyed five additional stations that ultimately did not participate in defection broadcasts during 2008–2015. We initially included the five stations due to encouragement by survey coordinators to follow the referrals of station operators (to further ensure all stations appeared in the sample).

with the station throughout its broadcasting of defection messages and in some cases they were also founding members, both of which we believe to lead to accurate recall. As an additional check we chose to cross-reference stations across survey rounds against internal reports acquired from international actors, and against a host of sources available on the internet. This last category included articles and public documentation from the United Nations, press and NGO reports, and social media. For example, in a number of cases we were able to verify survey responses based on reports from smaller NGOs that provided assistance or equipment to stations and subsequently documented this with photos and descriptions on Facebook or Twitter. With the second round of surveys and the cross-referencing, we have a high degree of confidence that all radio stations involved in defection messaging are represented in the data and that the resultant data are accurate.

We collected data on all 19 radio stations (21 antennas in total) that have broadcast defection messaging targeted at the LRA since 2008. While we were able to identify and locate the two radio stations with four antennas broadcasting defection messages during the pre-2008 period in Uganda, the data on actual defection messaging during this period were ultimately unreliable. First, data for these stations are based uniquely on self-reported recalls of two stations of faraway in time events (up to 20 to 15 years before the interview), which could lead to problematic recall bias, especially as respondents were not directly involved in the programs. Second, we did not find any additional source that could verify the gathered information. Moreover we encountered contrasting information from sources that appeared equally credible. Finally, during the Ugandan phase of the conflict messaging was generally ad hoc. At some point the government radio network was emitting messages and at other points this shifted to private networks coupled with frequent changes in broadcasting equipment. We were unable to get data on these changes. In addition, messaging stopped endogenously in 2006 as part of the terms of the peace agreement. This implied that we could not reliably obtain information for this phase.

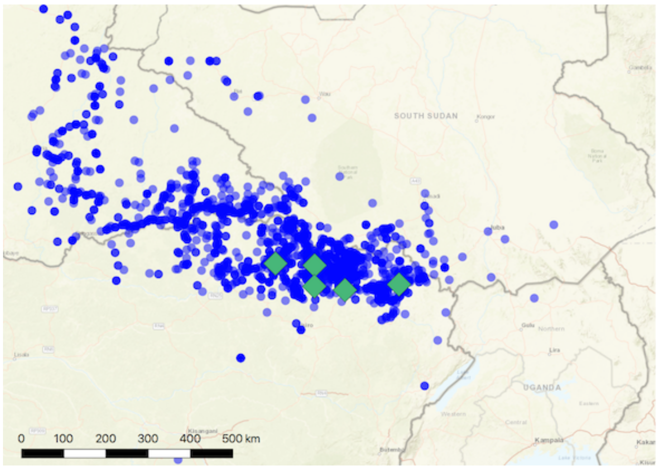
D Survey of returnees

This section provides a brief overview of the individual-level survey of returnees who had spent at least one month with the LRA.[10] We located and identified returnees based on lists in each community made available by civil society organizations and local associations (APRUs). Since the lists were numbered, respondents were randomly selected on the basis of these numbers. If the person corresponding to one of the numbers drawn was not available for the survey (either because they could not be located or because they refused to take the interview), the number was replaced by the next number on the list (non-response rate was 18%). The final sample comprises 89 returnees drawn from the four northern DRC provinces of Ango, Dungu, Faradje and Niangara. For security reasons, we did not extend the survey to South Sudan and CAR. In total, we interviewed returnees in more than 15 villages, organized in five clusters (see Figure

[10]The survey was conducted by Innovative Hub for Research in Africa / Marakuja Research Kivu, a research organization based in DRC.

D1). Sample size allows identifying (under unknown population and unknown standard deviation of outcome variables) a standardized effect of 0.3 in a one-sample mean comparison (assuming 80% of power and 5% of confidence).

Figure D1: Location of respondents



Notes. The distribution of respondents from the survey of returnees. To avoid the possibility of identification of respondents, each rhombus represents a cluster of respondents. Each dot represents an LRA-related violent event in the period 2008–2015. The geographic extent of the figure is restricted to the study area. Basemap source: Esri (see Appendix A for details and attributions).

Table D1 provides the summary statistics of the main variables. 58% (52) of the 89 respondents were male and 42% (37) were female. The average age of the respondents was approximately 27 years (25 for women and 29 for men). The main purpose of the survey was to understand whether in practice LRA members had access to radios and had heard defection messaging during their stint with the LRA. From Table D1 we see that 73% of respondents listened to the radio while with the LRA and 65.5% of the respondents had heard defection messages during their stint at the LRA. Furthermore, 94% of the respondents had heard other group members discussing the broadcasts asking rebels to return. Combining direct and indirect exposure, 95.5% of the sample was exposed to defection messaging. Finally, more than 67% of the respondents say that the broadcasts had influenced their decision to return. Hence, while the LRA were highly mobile and operated across vast expanses, the combatants were still exposed to radio defection messages either directly or indirectly quite frequently. In addition, such broadcasts had a direct influence on their decision to lay down arms and return home.

Table D2 tabulates the frequency of messaging exposure. The left panel shows the frequency of direct exposure and the right panel shows the frequency of indirect exposure. Approximately 72% of the 57 respondents who had heard defection messages on the radio were exposed to the messages at least once a week and among the 84 respondents who had heard other group members discuss messages 44% did so at least once a week. This is despite the fact that on average they spent less than 2 nights in the same place.

Table D1: Summary Statistics for survey of returnees

Variable	Mean	Std. Dev.	Min.	Max.	N
	(1)	(2)	(3)	(4)	(5)
Male	0.584	0.496	0	1	89
Age	27.045	9.060	17	53	89
Birth Year	1990.865	9.099	1965	2001	89
Year joined LRA	2010.101	2.927	2007	2018	89
Age when joined the LRA	19.247	9.645	6	49	89
Approximate length of stay with the LRA (months)	27.066	28.387	1	111.6	89
Listened to radio while with LRA	0.73	0.446	0	1	89
Heard defection messaging while with LRA	0.655	0.478	0	1	87
Heard members discuss def. broadcasts while in the LRA	0.944	0.232	0	1	89
Exposure to radio messages	0.955	0.208	0	1	89
Broadcasts influenced decision to return	0.674	0.471	0	1	89
Nights spent in same place while with LRA	1.553	1.139	1	5	85

Notes. For "Heard defection messaging while with LRA" and "Nights spent in same place while with LRA", the answer "I don't know" is set to missing.

Table D2: Frequency of hearing messages

	Defection broadcasts		**Group members discussing broadcasts**	
	Frequency	Percent	Frequency	Percent
	(1)	(2)	(3)	(4)
Once a month	3	5.26	32	38.1
Every two weeks	13	22.81	15	17.86
Every week	33	57.89	30	35.71
Everyday	8	14.04	7	8.33
Total	57	100	84	100

Notes. The table reports frequencies of respondents for each category. In columns 1 and 2, samples are restricted to respondents that report having heard defection broadcasts. In columns 3 and 4, samples are restricted to respondents who report to have heard group members discussing defection broadcasts.

E Defection messaging content

Following up on the heterogeneity-by-content analysis of Section B.19, we use machine learning techniques to categorize 86 digitized messages broadcast during the defection messaging campaign into two broad categories: sensitization and logistical content. The messages were digitized and made available by The Voice Project, one of the producers of content. These messages represent only a share of the total content broadcast during the period. In addition, information linking messages to specific radio stations or to specific periods is not available.

Examples of defection messages

First, we present some examples of messages broadcast during the defection messaging campaign. The following examples are drawn from a repository of broadcasts containing both audio files and transcripts hosted on The Voice Project. The first example is a message recorded by the chairperson of a village and addressed to children in LRA and to Joseph Kony:

> *My name is Pauline Achan; chairperson LCI of Akoyo village. As a mother, I will not talk much but I do appeal to you my children who are still in the bush that today if you hear my voice, you should not have any doubt. Some people used to say the people whose voices are played on radio are all dead, but today I am speaking from home in Odek and for you, who are still alive, you should hear me. Moses the son of Jackson stayed in the bush for eight years, but he is now farming together with us here at home without any problem. Now Lucore, I used to call you Lucore, if you are still alive please come home. Joseph Kony, you know me very well, I am the daughter of Obonyo Sione and I am your cousin. If you can hear me now please come back home because home is very good, girls have become tailors, others are builders and others are doing different useful work. Come back home because I am sure even you were abducted against your will. Thank you.*

The second example is a message recorded by a former LRA combatant:

> *My name is Opio but most of you from the bush know me by the name Aditi from Copil, I want to appeal to you my brothers to come out of the bush because whatever takes place there are not proper. For instance forcing people to kill is something I never wanted to do but I was forced into doing it. It is really not proper (sic) to be beaten like I was beaten while I was still there. So I appeal to everybody in the bush to come out. I still remember people like Owila we lived together with in Gilber battalion. I request you to come home because life at home is very good, there is now total peace. For me when I came back home, I was first taken to [Gulu Support the Children Organization] and later I was handed over to my parents who welcomed me with maximum happiness.*
>
> *You know very well that in the bush there is no proper medication should you get wounded but at home you have access to health services whenever you are sick and you can be treated. Human beings are not supposed to be treated like animals where your wounds are tied with banana leaves instead of receiving proper medication. So make up your mind and come out now.*

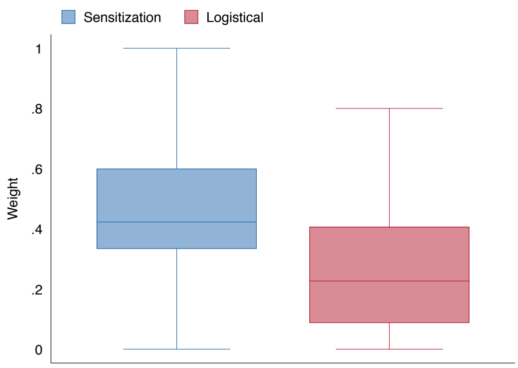
I have stayed with you people like Atingo nyim, Oyoo and Olwere and all other people. My primary interest is that you should come back home and live a good life instead of suffering in the bush. Come and stay with us the rest of people who returned. For us we are enjoying peace and we meet with returnees from other places from time to time. Thank you.

Content analysis

We use Natural Language Understanding (NLU) and Knowledge Studio services from IBM Watson to extract topics of interest along with the corresponding keywords from the raw text. Compared to traditional unsupervised topic models (see, for instance, Blei et al., 2003), supervised learning extracts relevant topics only, and thus helps us focusing on targeted contents. We initialize and load the model with three types of assets: entity types, documents (for training) and a dictionary that links entity types and documents. Similar to Section B.19, we specify two major entity types. *Sensitization* includes terms related to family members, friends, communities or tribes (i.e. relationship bonds), and content related to forgiveness and acceptance, along with expressions indicating opportunities outside the conflict sector. *Logistical* contains information about who returnees can turn to upon return, locations to which they can return, and the names of specific organizations.

We associate each of them with a dictionary rule to facilitate annotation, and we augment the dictionaries with the Watson NLU dictionary. We select and annotate 18 out of the 86 raw texts in the model, of which we use 14 for model training and four in the test set. The 18 chosen documents are a representative sample such that we have enough information for each entity type. A simple dictionary-based method might lead to significant bias, for instance by double counting the same word used in the same sentence multiple times. To avoid such bias, a human annotator corrected the documents. In the training step, we first randomly split the documents into a training set and test set. We then train the model and update it by looking at the tagging performance in the test set. If the model fails to tag labels correctly, we include more examples to enhance the performance of a specific entity type. In the last step, we deploy the model with NLU services to all the raw texts. We present a short summary of the results in the form of a box plot of the two alternative types of content (Figure E1). Results are in line with survey data, suggesting that defection messages do not focus on a specific type of content. Rather they broadcast both types of content, broadcasting a higher proportion of sensitization content within each message.

Figure E1: Content analysis



Notes. This figure plots the weights assigned to Sensitization and Logistical content based on 86 digitized messages made available by The Voice Project. Weights are produced by Natural Language Understanding (NLU) and Knowledge Studio services from IBM Watson. *Sensitization* content includes terms related to family members, friends, communities, or tribes (i.e. relationship bonds), forgiveness, acceptance, and the opportunity to start a new life outside the conflict sector. *Logistical* content refers to information about who returnees can turn to upon return, locations to which they can return, and the names of specific organizations.

F Paths of returning, forgiveness and reintegration

This section discusses the processes and institutions involved in the return and reintegration of LRA rebels into civil society. In addition to describing formal institutions, we outline a host of informal institutions that have been further facilitating reintegration. We also explain why communities and families might have been willing to accept returnees regardless of amnesty or what other national laws mandate. Finally, we provide a summarized list of the services and processes offered to returnees. We focus on the three countries affected by the LRA in the post-2008 period: Democratic Republic of Congo (DRC), South Sudan, and Central African Republic (CAR).

F.1 Institutions

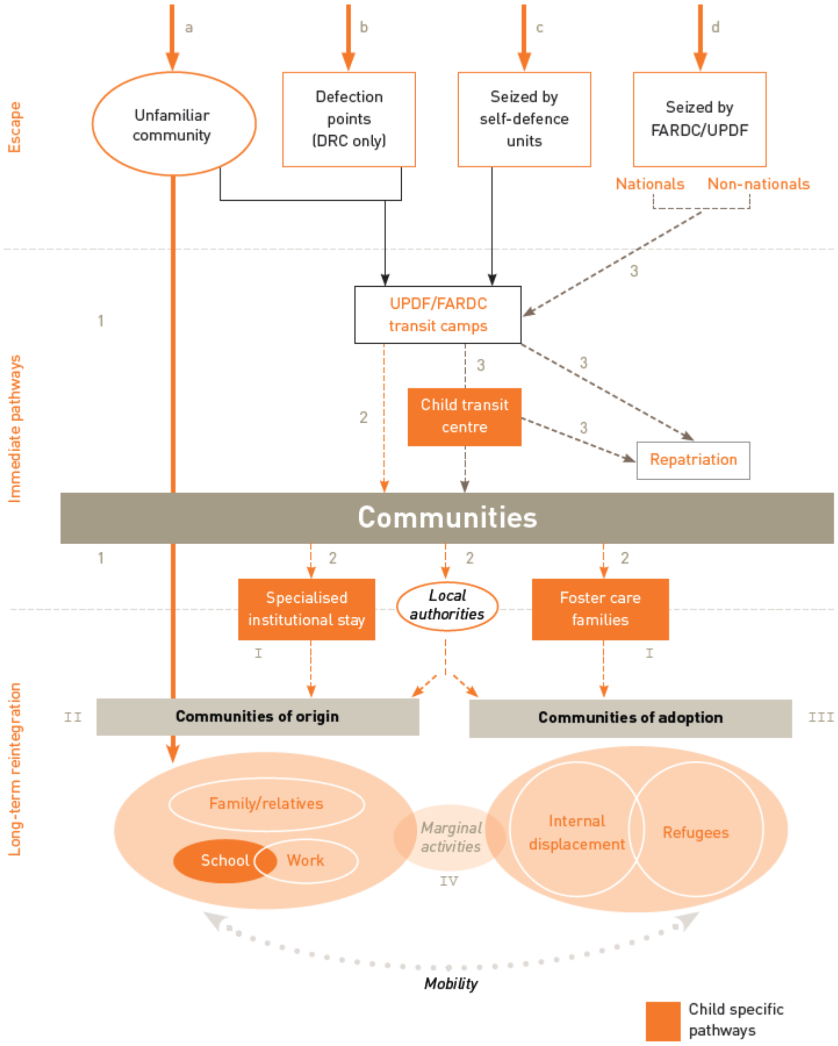
Formal institutions

Figure F1 provides a flowchart describing the different pathways available to returnees. The first step involves leaving the LRA to return to civilian life: they could voluntarily surrender at safe defection points; escape to unfamiliar communities; be freed, or be seized by either community defense groups or formal armed forces.[11] Once returnees leave the LRA, the next challenge is returning to their home community. When villages have been destroyed or are still under threat from the LRA, returnees have also made their way to new communities and urban centers, where they are attracted by potential employment opportunities. Refer to Medeiros (2014) for a detailed

[11]Community defense groups refer to groups such as the Arrow Boys, a civilian self-defense group in South Sudan. Formal armed forces refers to troops such as the Uganda Peoples' Defense Force (UPDF) or the Armed Forces of the Democratic Republic of the Congo (FARDC).

overview of the reintegration process.

Figure F1: Reintegration process



Notes. The flowchart maps the pathways available to both child and adult returnees while they make their way out of the LRA and into civilian life. It underscores both the short-term and long-term options available to the returnees in DRC and South Sudan. Stages supported by formal national institutions or international actors are framed in rectangles. Circles highlight informal pathways. Source: Medeiros (2014).

There are various layers of institutional support available to LRA returnees. At the top end of the formal institutional structure lies the Regional Cooperation Initiative for the elimination of the Lord's Resistance Army (RCI-LRA or simply RCI). This initiative was formed in late 2011 under the aegis of the African Union (AU) and was endorsed by the governments of the four countries affected by the LRA conflict (Uganda, DRC, CAR and South Sudan). One of the main outcomes of this initiative was the formation of the Regional Task Force (RTF), composed of troops from all four affected countries, which, advised by the US, has been facilitating the defection and return of LRA combatants. The RTF has maintained safe defection sites, accepting and referring defecting

LRA members (Shepard et al., 2015).

From a legal standpoint, the Ugandan Amnesty Act of 2000 offered a blanket amnesty to all Ugandan LRA combatants willing to abandon the group. While the US-led Operation Lightening Thunder in 2008 dispersed the LRA into the bordering regions of CAR, DRC and South Sudan, there were still a number of Ugandan nationals in the LRA, particularly in positions of leadership. Regardless of where they surrendered (or where they had been fighting), all Ugandan nationals were eligible for amnesty upon return except for a very small minority of high ranking LRA officials who were indicted by the International Criminal Court (ICC) as war criminals.[12] The logistics of the return were facilitated by Ugandan troops (under the RTF initiative) in the three countries where the LRA was operating during the post-2008 period. Ugandan returnees had the option of returning to the UPDF, which would process and repatriate them to Uganda. This is particularly true in the areas where the LRA operated due to the absence of state structures.[13] While there are isolated incidents of mistreatment by foreign armed forces upon surrender, most of the armies seem to have treated surrendering individuals humanely, particularly in the later stages of the conflict (Cakaj, 2011).

While legal amnesty was not available for non-Ugandan nationals, they were *de facto* receiving amnesty by returning through the DDRRR (disarmament, demobilization, repatriation, reintegration, and resettlement) programs of the UN (Medeiros, 2014). The UN in particular developed a series of standard operating procedures to deal with returnees from the LRA for all its missions across the region. The missions included MONUSCO in the DRC, UNMISS in South Sudan, and MINUSCA in Central African Republic.[14] For instance, in the specific case of the DRC, while LRA returnees did not receive any formal reintegration support from state authorities, they received *de facto* reintegration services through MONUSCO. MONUSCO headed the territory-wide DDR coordination for individuals who went through official channels of demobilization. Congolese authorities were not involved in the process.

Child returnees received special attention and treatment. Protection training provided by international agencies and NGOs such as the United Nations Children's Fund (UNICEF), the International Committee of the Red Cross (ICRC), the office of the United Nations High Commissioner for Refugees (UNHCR) and Save the Children led to improved practices regarding the treatment of children escaping the LRA (Shepard et al., 2015). For instance, the UPDF is reported to have improved its practices and followed a standard operating procedure (SOP) in dealing with children, which it agreed to with UNICEF. In South Sudan, this implied that once a child returnee arrived, the UPDF had to notify civilian agencies within 24 hours and then transfer the child to the Child Transit Center in Yambio (South Sudan) run by the Ministry of Gender, Child, and Social

[12]The radio messages made it clear that except for Joseph Kony and other top-ranking officials who were indicted by the ICC, everyone else would be forgiven upon return. This is not surprising as most LRA combatants were abducted and coerced to engage in violence (Le Sage, 2011).

[13]In CAR for example, the UPDF was the only *de facto* authority in the absence of a state (Shepard et al., 2015).

[14]Individuals could also return to the FARDC (Forces Armes de la Rpublique Dmocratique du Congo) and SPLA (Sudan Peoples Liberation Army), operating under the RTF initiative.

Development. Apart from these centralized systems, there were several alternatives. The network of foster families used by the Italian NGO COOPI was one such decentralized alternative. Under this initiative, custodians (of both children and adults) received training and continuous supervision. Furthermore, there were several initiatives arising out of local communities including local-volunteer-based committees. For instance, under the Diocesan Committee for Justice and Peace (CDJP) in Aru (DRC), local committees sometimes cared for returnees in their homes to ensure their well-being (Medeiros, 2014). Section F.3 provides a more detailed list of the organizations involved and services offered.

Informal institutions

One crucial aspect of reintegration is the willingness of families and communities to accept returnees back into the fold. This is where informal institutions come into play. Several (ad-hoc) homegrown initiatives in the DRC and South Sudan have been helpful for the reintegration of returnees. These include spontaneous welcome ceremonies, visits by chiefs, Christian prayer services, and local healing rituals. There have also been instances of group healing sessions by local non-governmental organizations (NGOs). These activities helped creating bonds within the community, contributing to the success of reintegration (Medeiros, 2014). Furthermore, such activities have been found to contribute to the psychological well-being and reconciliation of returnees. Cultural leaders have also played an important role by providing traditional cleansing ceremonies (for instance, rituals meant to remove bad spirits), for those who experienced or perpetrated violence. Such traditional belief systems enabled both return and reintegration (HHI, 2016).

A survey by the Harvard Humanitarian Initiative (HHI, 2016) identified the role of leadership committees in reintegration. Supported by NGOs, leadership committees worked within local communities to sensitize “community members around issues with returnees, focusing on welcoming them home and advising on how to interact peacefully with them” (HHI, 2016). They organized awareness campaigns, helped address disputes, offered counseling, and provided peer-to-peer support to family members of returnees (HHI, 2016). For instance, in 2013 two NGOs, namely Discover the Journey (DTJ) and Invisible Children (IC), teamed up with community leaders of the Mboki community in CAR to raise awareness among community members about the importance of accepting LRA returnees (HHI, 2016).

F.2 Forgiveness and reintegration

There are several reasons why communities are willing to forgive returnees and accept them back despite all the heinous crimes they might have committed. First, communities often perceived returnees as victims rather than perpetrators of crimes (Medeiros, 2014). This is particularly true because many LRA combatants were abducted against their own free will. Second, traditional belief systems played a substantial role in facilitating the willingness of communities to welcome individuals home. Combatants often undergo traditional rituals during their time in the LRA.

These rituals have the ostensible goal of protecting "members against the enemy's spells and bullets, to increase their endurance of physical pain and to inspire the aggression needed to perpetrate violence on civilians" (Medeiros, 2014). These ceremonies often involve the use of herbs and traditional medicines and are supposed to lead to states of altered consciousness among the fighters who are believed to have been under spells or intoxicated when they committed their crimes. Thus, both returnees and the communities who receive them could attribute their misdeeds to these rituals and absolve themselves of the crimes committed during their service to the LRA.

Third, upon return individuals often have the option to undergo tribal cleansing rituals to purify themselves. These rituals, conducted by traditional healers, can make forgiveness and reintegration easier. Consider, for instance, the case of the Zande ethnic group. Their territory has been the operating ground of the LRA in the DRC, South Sudan, and CAR since 2008. "The Zande ethnic group's customary belief-systems offer explanations for environmental adversity, insecurity, illness and conflicts between individuals, families or clans, and prescribe ways to address these difficulties" (Medeiros, 2014). There are specific rituals to overcome difficult situations. For instance, leaves from local plants, such as okpo, gundu and tande, can be used to cleanse a person of elements harmful to their mind such as anger. Feasting and jumping on animal blood (known as the *vuga* ritual) can act as a sign of purification. Another ritual, called *vonimi*, can also be effective in enabling forgiveness and reconciliation. Among the Acholi, which provided the majority of LRA members during the early stages of the conflict, most returnees underwent a public forgiveness ceremony called *Nyono Tong Gweno*, which includes stepping on an egg before declaring their wrongs (Clark, 2010). Another ceremony, called *mato oput*, includes drinking the juice from the oput tree, which has a similar theme of atonement. Such rituals offer returnees the possibility of purification (as understood in a traditional sense) and can enable forgiveness and reintegration (Medeiros, 2014).

Finally, Christian practices of confessing sins also facilitated this process. Church leaders are reported to have provided pastoral support even in remote areas. This includes prayer sessions that returnees believed to be an opportunity to communicate with God directly. These sessions involve the public acknowledgment and forgiveness of sins, which also facilitated forgiveness and acceptance within their communities (Medeiros, 2014).

F.3 Services and processes

This section provides a non-exhaustive list of organizations involved in the return and reintegration of returnees, as well as the processes and services available to them. For a more detailed overview, refer to Medeiros (2014), Shepard et al. (2015), HHI (2016), and Cakaj (2011).

- Defection and Debriefing:
  - Safe Defection Sites: the UN regional missions such as MONUSCO and the African Union-led Regional Task Force (RTF) maintained safe defection sites. The RTF was composed of troops from the armed forces of the four LRA affected countries. Hence,

in practice many of the sites were maintained by organizations such as the UPDF (Uganda Peoples' Defense Force), FARDC (Forces Armes de la Rpublique Dmocratique du Congo), and the SPLA (Sudan People's Liberation Army). Returning LRA members could turn themselves in at these sites without risk of prosecution.

  - Debriefing: several organizations such as the UPDF, FARDC, Joint Intelligence Operations Centre (JIOC) and US military advisers were involved in debriefing returnees. In practice, Ugandan nationals were debriefed by the UPDF. Most non-Ugandans were debriefed by UN missions.

- Repatriation:
  - Ugandan nationals were repatriated by the UPDF, while non-Ugandans were mostly repatriated by the UN missions.
  - The International Committee of the Red Cross (ICRC) has been primarily responsible for repatriating children and vulnerable mothers. The RTF and US advisers also assisted in the return of adult males to their home countries.

- Medical attention (targeting both mental and physical health):
  - The RTF provided immediate medical attention to returnees before handing them over to the appropriate agencies (Shepard et al., 2015) including, in the case of Ugandan nationals, the UPDF and, in the case of non-Ugandans, UN missions.
  - Intersos-DRC has funded mental health support for LRA-affected individuals (in combination with income-generating activities). This initiative primarily targeted adults with the most severe disturbances.
  - Specialized units such as the Catholic Medical Mission Board (CMMB), the Combini sisters, and Mdcins Sans Frontires (MSF) received investments from international donors at the height of the LRA conflict. They were responsible for supplying affordable or free medicine to state structures, refurbishing local facilities and providing technical expertise to hospitals and primary health care centers.
  - Some organizations provided specific services such as intensive group therapy sessions. This included organizations such as the Inter-Church Committee (ICC) and the Combini sisters (in South Sudan). The aim was to address post-traumatic effects, forgiveness, and reconciliation.

- Development initiatives (primarily sponsored by international development organizations) targeted at LRA returnees:
  - The Deutsche Gesellschaft fr Internationale Zusammenarbeit (GIZ), Germany's primary development organization, has funded some agricultural initiatives such as tool distribution in the West Equatorial State in South Sudan.

- Intersos-DRC has funded income-generating activities to LRA-affected individuals (in combination with mental health support).

- Children-specific services:
  - With support from international organizations such as Discover the Journey (DTJ), some NGOs, such as Solidarity and Integral Assistance to Destitute People (SAIPED), provided intensive group workshops focusing on parenting skills and family relationships with child returnees.
  - Save the Children and UNICEF collaboratively developed a training schedule for the RTF in 2013. They aim to improve practices to deal with child returnees.
  - The Elikya Center and Associazione per la Cooperazione Internazionale (COOPI) based in Dungu, in the Haut-Ul district in the DRC, provided specialized post-traumatic care for child returnees.
  - Some additional organizations have supported child returnees at various points in time including the Child Transit Centre in Yambio (South Sudan), and Catholic-run orphanages in the Ul region in DRC. International NGOs and the UN have also sometimes provided some form of family tracing service.

- Informal community-based initiatives providing support at various points in time:
  - Families and communities threw small parties for returnees, while local authorities such as local chiefs used their own means to help returnees in their basic needs.
  - Local priests held prayer meetings (together with neighbors and friends of returnees) and provided pastoral support. The prayer meetings often involved the public confession of sins, and offered what certain returnees believed to be the opportunity to communicate with God directly.
  - On return, individuals often have the option of undergoing local rituals (conducted by traditional healers) to purify themselves, such as those described in the above section.

- Awareness campaigns and support groups:
  - There have been community-based awareness campaigns (promoted by NGOs) to reduce stigmatization and highlight the importance of accepting LRA returnees, with the aim of facilitating the reintegration of returnees. The awareness campaign organized by Discover the Journey (DTJ) and Invisible Children (IC) in collaboration with community leaders in the Mboki community in Central African Republic, is one example.
  - In South Sudan, the Self-Help Women's Development Association (SHWDA) has been instrumental in forming support groups to enable the reintegration of returnees.

G Generalizability of results: radio for peace building

We evaluate a FM radio-based defection messaging program encouraging rebels to give up arms in the absence of a stable peace agreement. The success of such programs in persuading combatants to leave conflict behind and return to civilian life crucially depends on formal or informal reintegration processes that welcome rather than punish them. Such processes are relatively common, with or without stable peace agreements, in the form of DDR programs. Of the 40 DDR programs implemented between 1989 and 2010, 14 were instituted during ongoing conflicts. These included diverse cases such as Afghanistan, Cambodia, and several African countries (Banholzer, 2014). We expect radio defection messaging programs could be effective in many of these contexts as a means of informing rebels about possibilities and conditions of returning even during active conflict.[15]

Of the 216 peace agreements between 1975 and 2011, 80 had provisions for DDR (Sundberg and Melander, 2013). Moreover, 297 conflict amnesties were enacted between 1946 and 2010 (Dancy, 2018). Sending messages over the radio could also be effective in these contexts, by spreading information on the possibilities of return and/or amnesty for the intended recipient with whom it might be otherwise difficult to establish communication. Hence, findings are relevant to policy makers who focus on ending entrenched conflicts and violence, and more generally counteracting citizen violence through media-based peace building (Larreguy and Marshall, 2019).

Radio-based defection messaging can work across different demographics of combatants. We expect these interventions to be successful in the presence of conscripts and propaganda, as well as volunteers. In particular, defection messaging can be effective in the presence of child soldiers. Since it formed, the LRA has relied largely on the abduction and indoctrination of child soldiers (Beber and Blattman, 2013). However, the use of child soldiers is not limited to the LRA Insurgency. There are more than 250,000 child soldiers around the world spread across more than 20 countries as diverse as Afghanistan, Myanmar, Somalia, and Yemen.[16] While they represent individuals with little initial voluntary interest to be and remain combatants, successful indoctrination through the experiences of trauma and violence can make child conscripts highly resolute in their identity as a combatants (Gates, 2017). The intervention under question is likely to be successful in all the cases in which young people were coerced into taking up arms against their will and might be willing to return if adequate opportunities and incentives are created.

Furthermore, we believe that the applicability of radio-based programs is more general. For instance, radio programming can be used to discourage people from choosing violence in the first place. One such instance is the case of the radio-based campaign to counter Boko Haram. While rigorous quantitative evaluations are scant, anecdotal evidence suggests that radio has been effective in the fight against Boko Haram in northeastern Nigeria by changing people's beliefs

[15] Humphreys and Weinstein (2007) provide evidence of sub-groups least likely to participate (and participate successfully) in DDR are ideologues and youth, one of whom places high implicit value on continuing the conflict and the other who often expects low returns on demobilization. Both of these frictions are at play in the LRA insurgency.

[16] See, for instance, Their World, *Explainer - Child soldiers*.

about violence as a “means to an end”.[17] A media-based demobilization campaign is also supposed to have been successful in encouraging FARC members to defect (Ricks, 2016). Radio broadcasts were also part of the Chieu Hoi (Open Arms) campaign that supposedly led to almost 200,000 defections of Viet Cong and North Vietnamese troops (Daddis, 2017).

Again, recent evidence shows that by affecting perceptions of social norms and behaviors, radio programs can reduce inter-ethnic prejudice and willingness to defer to authority (Paluck and Green, 2009). Further evidence from Rwanda suggests that radio messaging can reduce ethnic salience and increase inter-ethnic trust and the willingness to interact with other ethnic groups, thus aiding in nation building (Blouin and Mukand, 2018). In the context of Northern Ireland, a media-based campaign is understood to have been instrumental in garnering acceptance of the Good Friday Agreement that ended three decades of conflict (Bratic, 2013). While these are not illustrations of defection promotion, they have obvious implications for reducing violence. While this paper provides causal evidence on their success, similar programs have indeed taken a notable role in multiple conflicts in Africa and globally. Below we provide examples of instances of conflicts and settings in which radio was used for peacekeeping:

- the Colombian conflict against the FARC;[18]
- the Boko Haram insurgency in Nigeria;[19]
- the Vietnam War, during which a South Vietnamese campaign, the *Chieu Hoi* (or Open Arms), encouraged defections by the Viet Cong, and have supposedly led to almost 200k defections of Viet Cong and North Vietnamese troops (Daddis, 2017);
- a variety of settings in which radio has been used more generally for peacekeeping operations and peace building (Rao, 2014; Bratic, 2013):
  - UNAMSIL Radio, set up in the year 2000 by the UN peacekeeping mission in Sierra Leone during and following the Sierra Leonean civil war;
  - Radio Kaoural and a host of community radio stations in central Mali;[20]
  - Radio Agatashya in the Great Lakes region of Africa;
  - Radio Amani in Nakuru county, Kenya;
  - Radio FERN in Bosnia;
  - STAR Radio in Liberia;
  - Radio UNTAC, managed by the United Nations Transitional Authority in Cambodia following the CambodianVietnamese War.

[17] See, for instance, PRI (*How a shortwave radio network is helping to counter Boko Haram*, 19/05/2017).
[18] Source: Time (*Why Colombia's leftist guerrillas are defecting*, 30/10/2009).
[19] Source: PRI (*How a shortwave radio network is helping to counter Boko Haram*, 19/05/2017).
[20] Source: UN Peacekeeping (*Community radios build bridge to peace in Central Mali*, 14/01/2019).

H Asymmetric effects of media on violence and peace

A number of studies have demonstrated the propensity of mass media to incite hatred and violence (Yanagizawa-Drott, 2014; Adena et al., 2015; DellaVigna et al., 2014). This paper represents a valuable addition to recent findings on the influence of media on conflict and does so while expanding the range of outcome variables to include peaceful outcomes. In addition, this paper also adds evidence on a substantively distinct question in the literature by providing one of the first rigorous empirical studies demonstrating that mass media can be instrumental in deescalating conflict, and that peaceful outcomes are far from the inverse of conflict outcomes. Below we present the context-specific case that recruitment and defection are distinct processes in theory and practice. In terms of the underlying mechanisms and processes, we draw on a number of scholars who maintain that media for peace is very different from media for violence.

H.1 Recruitment versus defection

Recruitment for and defection from armed groups are two separate behavioral processes (Gates and Nordås, 2015). Recruitment relates to a participation problem, in which individuals from outside the rebel group must be encouraged to join. Defection, on the other hand, relates to a retention problem in which group leaders devise incentive schemes to ensure continued allegiance from members. The literature has primarily relied on a principal-agent framework to highlight the incentive schemes that rebel leaders devise for recruitment and allegiance (Gates and Nordås, 2015; Gates, 2002).

In the case of the Rwandan genocide, in which radio messages persuaded listeners to commit violence against Tutsis, broadcasts not only incited hatred and encouraged violence, but made clear that the perpetrators would not be punished (and would even be supported by the government). This last aspect made it clear that there were few structural deterrents to participating in violence. Furthermore, much of the violence was the result of spillover effects through social ties which involved people partaking in a rising tide of violent behavior against neighbors (Yanagizawa-Drott, 2014).

On the other hand, once individuals become part of an organized armed group, they face powerful deterrents to defection. First, the threat of harsh punishment and retribution (often extending beyond an individual to their families and friends) can be a powerful deterrent. This is particularly true for the LRA, who meted out severe punishments to defectors. Second, the LRA forced recruits to commit heinous crimes against their families and communities. This broke the bonds they had, generated feelings of guilt, and resulted in high (perceived) risks of retaliation upon return. Third, cult-like rebel groups often invoke fringe ideologies and beliefs, such as the religious and spiritual frames that Joseph Kony invoked. This generates fear of even supernatural retribution among rebels, under the influence of which they might continue to comply with the armed group. Finally, and perhaps most importantly, in organizations such as the LRA, indoctrination of rebels plays an important role in maintaining allegiance. As rebels start internalizing the rules and ideologies of the group, they perceive the ideas and actions of the groups as "intrinsically rewarding". In such a context, encouraging defection may require redefining alternatives, as well as shifting perceived benefits (Gates and Nordås, 2015).

Moreover, groups such as the LRA, which rely on forced recruitment, generally do better at retaining members than groups that rely on voluntary recruits. Evidence suggests that on average groups with forced recruits are able to retain them for seven years, while voluntary recruits stay for two years (Gates, 2017). While they recruit individuals by force, these groups cannot solely rely on coercive measures for continued allegiance. They instead rely on the internalization of norms by recruits through "indoctrination, psychologically transformational experiences, or altering beliefs" to ensure high retention rates (Gates, 2017). For instance, indoctrination might ensure that rebels perceive certain activities as the only possible way to fulfill their moral or religious duties. Furthermore, rebel groups can also provide security, self-esteem, and a sense of community and camaraderie (Gates and Nordås, 2015). While propaganda plays a role during recruitment, the scope of indoctrinating recruits is much higher once they are already part of the group.

The presence of child soldiers leads to further high rates of cohesion and retention. Evidence from child psychology suggests that children demonstrate higher levels of altruism and group bonding (Harbaugh and Krause, 2000). Once recruited, children have a lower tendency to leave than adults, even if they were forced to join in the first place (Beber and Blattman, 2013). Part of the reason forced conscripts become loyal followers can be attributed to Stockholm syndrome (Gates, 2017). Hence, convincing someone to leave presents additional challenges in comparison to compelling someone to join a group.

Bénabou and Tirole (2011) provide an alternative framework to understand retention. They model retention as the product of a process of escalating commitments and radicalization. Escalating commitment occurs once an individual builds up a considerable stock of some asset such as social capital within the group. Someone who has already committed significantly to a group is more likely to continue investing in it, maintaining allegiance and committing even more resources to it, rather than defecting. This also follows from the literature on self-justification and psychology (Gates, 2017).

H.2 Media for peace versus media for violence

The distinction between media that foments conflict and media that deescalates conflict is not firmly established. This primarily owes to the limited empirical evidence on media for peace, noted by several scholars of media and conflict (see, for instance, Wolfsfeld, 2004; Hoffmann, 2014; Bratic, 2013). Existing research on this link "remain[s] concerned mainly, or exclusively, with the absence of war or direct violence, without explicitly or thoroughly conceptualizing their place in the processes of peace in a broader sense" (Hoffmann, 2014). We make some headway on this issue.

While we are in conversation with earlier papers on the incitement of conflict, the intervention we study and its outcome are fundamentally distinct from the escalation of conflict. These two processes speak to different pressures, incentives, and opportunity costs. Wolfsfeld (2004) describes this distinction, "...a case for media leading to peace is a much more difficult hypothesis to prove than the one that media can lead to war. In other words, the idea of lowering the level of hate is clearly in any case a much more difficult challenge than raising the level of hate. Like anything else, it is easier to start the fire and burn the building than to build one."

At a basic level, the decision to exit a conflict relates to a different set of choices than the decision to mobilize and join a conflict. In the case we study, this is particularly evident. For instance, observational accounts underscore that the decision to remain with the LRA rests primarily on ideological returns and fear of prosecution (even if recruitment might have depended mostly on force). The returns for demobilization, on the other hand, are rooted in forgiveness, and the social, economic, and welfare benefits to returning home. A media intervention fostering conflict would therefore need to target a very different set of motivations.

The path to peace is very different from the simple reversal of the path to conflict (Wolfsfeld, 2004). Peace, particularly at the individual level, involves confronting the legal, social, and psychological penalties of having participated in war. Social and state institutions (and processes) that allow combatants to return to society frequently underpin it. The existence and applicability of those institutions must be perceptible to combatants in an entrenched conflict. In contrast, convincing an individual to join a conflict is often ideologically motivated, with nebulous returns coming in a later period, or through the use of force.

In summary, while the "treatment vector" may be the same across attempts to incite violence and foster peace, we side with voices calling for theories that distinguish the effects of media on peace from the effects of media on conflict. Peace is ill-defined (and empirically researched) as the absence of conflict.

Appendix Bibliography

Adena, M., R. Enikolopov, M. Petrova, V. Santarosa, and E. Zhuravskaya (2015). Radio and the rise of the nazis in prewar germany. *The Quarterly Journal of Economics 130*(4), 1885–1939.

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2003). Fractionalization. *Journal of Economic growth 8*(2), 155–194.

Anselin, L. (2013). *Spatial econometrics: methods and models*, Volume 4. Springer Science & Business Media.

Athey, S. and G. W. Imbens (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives 31*(2), 3–32.

Banholzer, L. (2014). When do disarmament, demobilization and reintegration programs succeed? *German Development Institute Discussion Paper 8/2014*.

Bauer, M., N. Fiala, and I. Levely (2018). Trusting former rebels: An experimental approach to understanding reintegration after civil war. *The Economic Journal 128*(613), 1786–1819.

Beber, B. and C. Blattman (2013). The logic of child soldiering and coercion. *International Organization 67*(01), 65–104.

Beguería, S., S. M. Vicente-Serrano, F. Reig, and B. Latorre (2014). Standardized precipitation evapotranspiration index (spei) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *International Journal of Climatology, 34*(10).

Bénabou, R. and J. Tirole (2011). Identity, morals, and taboos: Beliefs as assets. *The Quarterly Journal of Economics 126*(2), 805–855.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003). Latent dirichlet allocation. *Journal of machine Learning research 3*(Jan), 993–1022.

Blouin, A. T. and S. W. Mukand (2018). Erasing ethnicity? propaganda, nation building and identity in rwanda. *Journal of Political Economy forthcoming*.

Bontemps, S., P. Defourny, and E. Van Bogaert (2009). Globcover 2009. Products Description and Validation Report. European Space Agency. http://due.esrin.esa.int/files/GLOBCOVER2009_Validation_Report_2.2.pdf.

Boots, B. N. and A. Getis (1988). *Point pattern analysis*, Volume 8. Sage Publications, Inc.

Bratic, V. (2013). Twenty years of peacebuilding media in conflict. *open knowledge network occasional working paper* (2).

Cakaj, L. (2011). Too far from home: demobilizing the lord's resistance army. Technical report.

Cameron, A. C. and P. K. Trivedi (2013). *Regression analysis of count data*, Volume 53. Cambridge university press.

CIA (2019). The world factbook 2019. Washington, DC: Central Intelligence Agency, 2019. https://www.cia.gov/library/publications/resources/the-world-factbook/index.html.

Ciccone, A. (2011). Economic shocks and civil conflict: A comment. *American Economic Journal: Applied Economics 3*(4), 215–227.

CIESIN-CIAT (2005). Gridded population of the world, version 3 (GPWv3) - population count grid. Center for International Earth Science Information Network and Centro Internacional de Agricultura Tropical. Palisades, NY.

Clark, J. N. (2010). The ICC, Uganda and the LRA: re-framing the debate. *African Studies 69*(1), 141–160.

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

Conley, T. G. (2008). Spatial econometrics. In *New Palgrave Dictionary of Economics - 2nd Edition*. Springer.

Cox, N. J. (2006). Speaking stata: In praise of trigonometric predictors. *The stata journal 6*(4), 561–579.

Daddis, G. A. (2017). *Withdrawal: Reassessing America's Final Years in Vietnam*. Oxford University Press.

Dancy, G. (2018). Deals with the devil? conflict amnesties, civil war, and sustainable peace. *International Organization*, 1–35.

DellaVigna, S., R. Enikolopov, V. Mironova, M. Petrova, and E. Zhuravskaya (2014). Cross-border media and nationalism: Evidence from serbian radio in croatia. *American Economic Journal: Applied Economics 6*(3), 103–132.

Desmet, K., J. Gomes, and I. Ortuño-Ortín (2018, June). The geography of linguistic diversity and the provision of public goods. Working Paper 24694, National Bureau of Economic Research.

Desmet, K., I. Ortuño-Ortín, and R. Wacziarg (2012). The political economy of linguistic cleavages. *Journal of development Economics 97*(2), 322–338.

Desmet, K., S. Weber, and I. Ortuño-Ortín (2009). Linguistic diversity and redistribution. *Journal of the European Economic Association 7*(6), 1291–1318.

Duan, N. (1983). Smearing estimate: a nonparametric retransformation method. *Journal of the American Statistical Association 78*(383), 605–610.

Easterly, W. and R. Levine (1997). Africa's growth tragedy: policies and ethnic divisions. *The quarterly journal of economics 112*(4), 1203–1250.

Eck, K. (2012). In data we trust? a comparison of ucdp ged and acled conflict events datasets. *Cooperation and Conflict 47*(1), 124–141.

Esteban, J.-M. and D. Ray (1994). On the measurement of polarization. *Econometrica: Journal of the Econometric Society*, 819–851.

Fan, Y. and H. van den Dool (2008). A global monthly land surface air temperature analysis for 1948-present. *Journal of Geophysical Research 113*(D01103).

Fearon, J. D. (2003). Ethnic and cultural diversity by country. *Journal of economic growth 8*(2), 195–222.

Fotheringham, A. S. and D. W. Wong (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A 23*(7), 1025–1044.

Friedman, J., T. Hastie, and R. Tibshirani (2001). *The elements of statistical learning*, Volume 1. Springer series in statistics New York.

Fujiwara, T., K. Meng, and T. Vogl (2016). Habit formation in voting: Evidence from rainy elections. *American Economic Journal: Applied Economics 8*(4), 160–188.

Funk, C., P. Peterson, M. Landsfeld, D. Pedreros, J. Verdin, S. Shukla, G. Husak, J. Rowland, L. Harrison, A. Hoell, et al. (2015). The climate hazards infrared precipitation with stations–a new environmental record for monitoring extremes. *Scientific data 2*, 150066.

Gates, S. (2002). Recruitment and allegiance: The microfoundations of rebellion. *Journal of Conflict resolution 46*(1), 111–130.

Gates, S. (2017). Membership matters: Coerced recruits and rebel allegiance. *Journal of Peace Research 54*(5), 674–686.

Gates, S. and R. Nordås (2015). Recruitment, retention and religion in rebel groups. In *Workshop 'Peace, Conflict, and Development', Centre for the Study of Equality, Social Organization (ESOP), University of Oslo*, Volume 24.

Gomes, J. F. (2014). The health costs of ethnic distance: Evidence from sub-saharan africa. Technical report, ISER Working Paper Series.

GSMA (2012). *GSMA Network Coverage Maps - Submission Guide January 2012*.

Harari, M. and E. La Ferrara (2013). Conflict, climate and cells: A disaggregated analysis.

Harbaugh, W. T. and K. Krause (2000). Children's contributions in public good experiments: The development of altruistic and free-riding behaviors. *Economic Inquiry 38*(1), 95–109.

Hausman, J. A., B. H. Hall, and Z. Griliches (1984). Econometric models for count data with an application to the patents-r&d relationship.

Hengl, T. (2006). Finding the right pixel size. *Computers & Geosciences 32*(9), 1283–1298.

HHI (2016). We mobilized ourselves: Community resilience in areas impacted by the lord's resistance army. Technical report, Case Study by Harvard Humanitarian Initiative, Harvard University.

Hidalgo, F. D., S. Naidu, S. Nichter, and N. Richardson (2010). Economic determinants of land invasions. *The Review of Economics and Statistics 92*(3), 505–523.

Hoffmann, J. (2014). Conceptualising 'communication for peace'. *Peacebuilding 2*(1), 100–117.

Hsiang, S. M. (2010). Temperatures and cyclones strongly associated with economic production in the caribbean and central america. *Proceedings of the National Academy of sciences 107*(35), 15367–15372.

Humphreys, M. and J. M. Weinstein (2007). Demobilization and reintegration. *Journal of Conflict Resolution 51*(4), 531–567.

Hurvich, C. M. and C.-L. Tsai (1989). Regression and time series model selection in small samples. *Biometrika 76*(2), 297–307.

Lancaster, P., G. Lacaille, and L. Cakaj (2011). *Diagnostic study of the Lord's Resistance Army*. Washington, DC: World Bank.

Larreguy, H. and J. Marshall (2019). The incentives and effects of independent and government-controlled media in the developing world. In B. Grofman, E. Suhay, and A. Trechsel (Eds.), *Oxford Handbook of Electoral Persuasion*. Oxford University Press.

Le Sage, A. (2011). Countering the lord's resistance army in central africa. Technical report, National Defense University, Washington DC Institute for National Strategic Studies.

Manning, W. G. et al. (1998). The logged dependent variable, heteroscedasticity, and the retransformation problem. *Journal of health economics 17*(3), 283–295.

Manning, W. G. and J. Mullahy (2001). Estimating log models: to transform or not to transform? *Journal of health economics 20*(4), 461–494.

Medeiros, E. (2014). Back but not home: supporting the reintegration of former LRA abductees into civilian life in congo and south sudan. Technical report.

Miguel, E., S. Satyanath, and E. Sergenti (2004). Economic shocks and civil conflict: An instrumental variables approach. *Journal of Political Economy 112*(4), 725–753.

Monfreda, C., N. Ramankutty, and J. A. Foley (2008). Farming the planet. part 2: Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles 22*.

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

National Oceanic and Atmospheric Administration (2014). DMSP-OLS nighttime lights time series version 4. National Geophysical Data Center, United States Department of Commerce. https://www.ngdc.noaa.gov.

Nunn, N. and D. Puga (2012). Ruggedness: The blessing of bad geography in africa. *The Review of Economics and Statistics 94*(1), 20–36.

Paluck, E. L. and D. P. Green (2009). Deference, dissent, and dispute resolution: An experimental intervention using mass media to change norms and behavior in rwanda. *American Political Science Review 103*(4), 622–644.

Raleigh, C., A. Linke, H. Hegre, and J. Karlsen (2010). Introducing ACLED: An armed conflict location and event dataset special data feature. *Journal of Peace Research 47*(5), 651–660.

Rao, S. (2014). Sending messages to reduce violent conflict. Technical report, GSDRC Helpdesk Research Report 1050. Birmingham, UK: GSDRC, University of Birmingham.

Ricks, T. E. (2016). The lesson of colombia's demobilization of farc can help us work against isis. *Foreign Policy January 28*.

Sanchez de la Sierra, R. (2019). On the origins of the state: Stationary bandits and taxation in eastern congo. *Journal of Political Economy*.

Shepard, B., L. Davis, and M. Jowell (2015). In new light: protection of civilians, the lord's resistance army and the african union regional task force. Technical report.

Silva, J. S. and S. Tenreyro (2010). On the existence of the maximum likelihood estimates in poisson regression. *Economics Letters 107*(2), 310–312.

Stolwijk, A., H. Straatman, and G. Zielhuis (1999). Studying seasonality by using sine and cosine functions in regression analysis. *Journal of Epidemiology & Community Health 53*(4), 235–238.

Storeygard, A., D. Balk, M. Levy, and G. Deane (2008). The global distribution of infant mortality: A subnational spatial view. *Population, Space and Place 14*(3), 209–229.

Sugiura, N. (1978). Further analysts of the data by akaike's information criterion and the finite corrections: Further analysts of the data by akaike's. *Communications in Statistics - Theory and Methods 7*(1), 13–26.

Sundberg, R. and E. Melander (2013). Introducing the UCDP georeferenced event dataset. *Journal of Peace Research 50*(4), 523–532.

The Resolve (2015). LRA crisis tracker map methodology & database codebook v2. Technical report.

The World Bank (2019). Global economic monitor (GEM). https://datacatalog.worldbank.org/dataset/global-economic-monitor.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267–288.

Tibshirani, R., M. Wainwright, and T. Hastie (2015). *Statistical learning with sparsity: the lasso and generalizations*. Chapman and Hall/CRC.

Tollefsen, A. F., H. Strand, and H. Buhaug (2012). PRIO-GRID: A unified spatial data structure. *Journal of Peace Research 49*(2), 363–374. https://grid.prio.org/.

UKW/TV-Arbeitskreis (2018). FMLIST - database of worldwide FM radio, DAB and TV. https://www.fmlist.org (accessed March 2, 2018).

UN-DDR (2014). *Operational Guide to the Integrated Disarmament, Demobilization, and Reintegration Standards*. United Nations Inter-Agency Working Group on DDR.

USGS (2016a). Historical statistics for mineral and material commodities in the united states, data series 140. *Data Source 2016*.

USGS (2016b). Mineral resources data system (mrds). United States Geological Survey. https://mrdata.usgs.gov/mrds/.

Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives 28*(2), 3–28.

Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences*, Volume 100. Academic press.

Wolfsfeld, G. (2004). *Media and the Path to Peace*. Cambridge University Press.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT press.

Yanagizawa-Drott, D. (2014). Propaganda and conflict: Evidence from the rwandan genocide. *The Quarterly Journal of Economics 129*(4), 1947–1994.