

Election Violence (ELVI) Event Dataset

Clayton Besaw (cbesaw@oneearthfuture.org)

January 22, 2021

The Election Violence (ELVI) dataset builds upon the National Elections Across Democracy and Autocracy (NELDA) dataset and our own Rulers, Elections and Irregular Governance (REIGN) dataset. ELVI provides election related violence indicators for most national elections going back to January 1975 until the present. It also provides a number of country and region indicators related to stability and violence.

The key differences between ELVI and NELDA are as follows:

1. NELDA ranges from 1945 to 2015, while ELVI ranges from 1975 to 2021.
2. ELVI includes constitutional referendum election events alongside presidential and legislative election events.
3. ELVI only extends election violence outcome indicators and includes country-specific variables used in the forecasting of election violence risk.
4. ELVI does not yet include microstates.

Country-election-event is the unit of analysis and ELVI contains 2667 election observations as of January, 2021.

ELVI is updated monthly with public updates released typically on the first Tuesday of every new month. This codebook serves as a living document for navigating ELVI measurements and is subsequently updated each month alongside the new ELVI data updates.

Citation

When using REIGN, please use the following citation:

Besaw, Clayton., Frank, Matthew. 2021. Election Violence (ELVI) Indicators and Forecast Dataset. Broomfield, CO: One Earth Future. Available at <https://oefdatascience.github.io/REIGN.github.io/>

We also encourage users to cite NELDA as well, given that our data builds upon a portion of this amazing work. You can find the citation here:

Susan D. Hyde and Nikolay Marinov, 2012, Which Elections Can Be Lost?, *Political Analysis*, 20(2), 191-201.

CODEBOOK AND MEASURE DESCRIPTIONS

Country and Time Measures

- **ccode**: ccode is a one to three digit number that uniquely identifies each of the 201 countries included in the dataset. These numbers align with those used by other large cross-national datasets, including Correlates of War. The ccode is most useful for merging REIGN with other datasets that also share a cow country code.
- **country**: country provides the full colloquial name for the country observed in the row. Not all country names are full legally recognized country names.

- **year:** year is the calendar year, ranging from 1950 to 2021.
- **month:** month identifies the month number ranging from 1 (January) to 12 (December)
- **dates:** dates displays the exact date (MM/DD/YYYY) that an election event took place on. For upcoming elections, the date is most credible estimated date in which to expect the election to be held at the time of data publication.

Election Violence Indicators

- **elecViolence1:** elecViolence1 indicates whether there was any form of election related violence in the week leading up to an election, on election day or the week after an election. It is coded as “vio” if any election related violence was observed and “peace” otherwise. *Note: Descriptives below only correspond to elections that have been observed. Upcoming elections will have a missing indicator (NA) in the ELVI event data sheet.*

Election Violence? (Any)	N	Percent
peace	1821	68.28
vio	804	30.15
All	2667	100.00

- **l.elecViolence1:** l.elecViolence1 is a dummy indicator for whether the previous national election recorded for a country was violent. It is equal to 1 if the previous election had any election related violence and 0 otherwise.

Previous Election Violence? (Any)	N	Percent
0	1901	71.28
1	766	28.72
All	2667	100.00

- **elecViolence2:** elecViolence2 indicates whether there was any form of government conducted election related violence in the week leading up to an election, on election day or the week after an election. This includes kinetic violence alongside repression and arbitrary detention. It is coded as “vio” if any government related violence was observed and “peace otherwise”. *Note: Descriptives below only correspond to elections that have been observed. Upcoming elections will have a missing indicator (NA) in the ELVI event data sheet.*

Election Violence? (Government)	N	Percent
peace	2201	82.53
vio	424	15.90
All	2667	100.00

- **l.elecViolence2:** l.elecViolence2 is a dummy indicator for whether the previous national election recorded for a country was experienced government conducted violence. It is equal to 1 if the previous election had government conducted election related violence and 0 otherwise.

Previous Election Violence? (Government)	N	Percent
0	2238	83.91
1	429	16.09
All	2667	100.00

- **tae1:** tae1 (time after event 1) measures the number of peaceful election events that have taken place since the last election experiencing any form of election related violence. For countries that have had a previously violent election, the tae1 measure will equal 1. If a peaceful election takes place after a previously violent election, tae1 then increase to 2. For every peaceful election taking place afterwards, tae1 increases by a value equal to the number of peaceful election events that have taken place since the last violent election. Larger values of tae1 indicate greater distance from the last violent election. If a country experiences a new violent election, tae1 resets to 1.

	N	Mean	SD	Min	Max
Time After Event (Any)	2667	4.76	5.12	0	31

- **nvio1:** nvio1 (number of violent events 1) measures the number of cumulative elections experiencing any form of election related violence before the observed event. It is equal to 0 if a country has never experienced a violent election event and is equal to any positive integer that is appropriate to a country's election violence history. Larger values of nvio1 indicate a larger amount of previously violent elections.

	N	Mean	SD	Min	Max
Number of Previously Violent Elections (Any)	2667	2.19	3.04	0	17

- **tae2:** tae2 (time after event 2) measures the number of peaceful election events that have taken place since the last election experiencing government conducted election related violence. For countries that have had a previously violent election, the tae2 measure will equal 1. If a peaceful election takes place after a previously violent election, tae2 then increases to 2. For every peaceful election taking place afterwards, tae2 increases by a value equal to the number of peaceful election events that have taken place since the last violent election. Larger values of tae2 indicate greater distance from the last violent election. If a country experiences a new violent election, tae1 resets to 1.

	N	Mean	SD	Min	Max
Time After Event (Government)	2667	5.97	5.45	0	31

- **nvio2:** nvio2 (number of violent events 2) measures the number of cumulative elections experiencing government conducted election related violence before the observed event. It is equal to 0 if a country has never experienced a violent election event and is equal to any positive integer that is appropriate to a country's election violence history. Larger values of nvio2 indicate a larger amount of previously violent elections.

	N	Mean	SD	Min	Max
Number of Previously Violent Elections (Government)	2667	1.22	2.03	0	13

Political Indicators

- **gov_democracy:** gov_democracy is a dummy variable that is equal to 1 if a country is characterized as a presidential democracy or a parliamentary democracy. It is coded as 0 otherwise.

Democracy?	N	Percent
0	1105	41.43
1	1562	58.57
All	2667	100.00

- **gov_interim:** gov_interim is a dummy variable that is equal to 1 if a country is characterized as a civilian or military-led provisional government. This typically occurs after government collapse due to conflict or sudden regime change.

Provisional Regime?	N	Percent
0	2605	97.68
1	62	2.32
All	2667	100.00

- **dem_duration:** dem_duration is a numeric variable that measures the logged number of months that a country has had a democratic government. The value is 0 if a government is currently a non-democracy. Larger values correspond to longer periods of successive democratic governance.

	N	Mean	SD	Min	Max
Democratic Duration in Months (Logged)	2667	1.69	1.69	0.00	5.45

- **logtenure:** logtenure measures the logged number of months that a country's government type has been in place. logtenure is equal to 0 in the month that a government type (regime change) changes. Larger values indicate a longer successive period of the current form of governance.

	N	Mean	SD	Min	Max
Government Type Duration in Months (Logged)	2667	3.74	1.33	0.69	6.38

- **regav_dem:** regav_dem is the regional average of democratic duration (logged). Regions are split into the Americas, Europe, Middle East and North Africa, Sub-Saharan Africa and Asia. Larger values indicate a longer regional presence of democratic governance.

	N	Mean	SD	Min	Max
Average Regional Democratic Duration (Logged)	2667	1.68	1.03	0.04	3.50

- **del_regdem:** del_regdem (delta of regav_dem) is the monthly change in regav_dem value month to month. Positive values indicate an increase in the regav_dem from the previous month, while negative values indicate a decrease in regav_dem from the previous month. Decreases happen when countries within a region go from democratic to non-democratic governance. Greater negative values represents a stronger regional trend away from democratic governance.

	N	Mean	SD	Min	Max
Change in Regional Democratic Duration	2667	0.03	0.07	-0.37	0.24

- **del_regten:** del_regten (delta of average regional government duration) is the monthly change in regionally averaged governance duration from the previous month. Positive values indicate increasing stability regarding government continuity within the region. Conversely, negative values indicate increasing government instability at the regional level.

	N	Mean	SD	Min	Max
Change in Regional Political Stability	2667	0.01	0.12	-0.87	0.27

Socio-Economic Indicators

- **pcgdp**: pcgdp measures the estimated annual median GDP per Capita for a country. The measure is derived from multiple sources of GDP estimates. Sources include the World Bank, the International Monetary Fund's World Economic Outlook and the Gleditsch Trade and GDP indicators for historical economic data. GDP estimates are updated throughout the year as new forecasts are released. For missing historical data and states with weak source coverage, machine learning imputation is used to estimate values.

	N	Mean	SD	Min	Max
Median GDP per Capita	2667	7466.86	12298.47	0.00	102576.71

- **growth**: growth measures the level of annual economic growth based on GDP per Capita. Values greater than 1 indicate economic growth. Values of 1 indicate no change. Values less than 1 indicate a decline in economic indicators.

	N	Mean	SD	Min	Max
Economic Growth	2667	1.05	0.13	0.63	1.37

- **gdpdiff**: gdpdiff measures the relative difference between a country's GDP per Capita and the region it belongs to. Values greater than 1 indicate a higher GDP than the regional average, while values less than 1 indicate a lower GDP than the regional average.

	N	Mean	SD	Min	Max
Relative GDP to Region	2667	0.89	1.30	0.00	10.36

- **growthdiff**: growthdiff measures the relative difference between a country's annual economic growth and the region it belongs to. Values greater than 1 indicate a higher level of growth than the regional average, while values less than 1 indicate a lower level of growth than the regional average.

	N	Mean	SD	Min	Max
Relative Economic Growth	2667	1.00	0.11	0.56	1.50

- **regav_pcgdp**: regav_pcgdp measures regionally averaged annual median GDP per Capita. Larger values indicate a higher regional GDP, while smaller values indicate a lower regional GDP.

	N	Mean	SD	Min	Max
Regionally Averaged GDP per Capita	2667	9676.83	9637.86	598.97	36389.65

- **regav_growth**: regav_growth measures regionally averaged economic growth. Values greater than 1 indicate regional economic growth. Values of 1 indicate no change. Values less than 1 indicate a decline in regional economic indicators.

	N	Mean	SD	Min	Max
Regionally Averaged GDP per Capita	2667	9676.83	9637.86	598.97	36389.65

- **del_reggdp**: del_reggdp (delta of regionally averaged GDP) measures the annual change in the average regional GDP since the previous year. Positive values indicate larger changes to the regional GDP since the previous year.

	N	Mean	SD	Min	Max
Change in Regional GDP	2667	0.02	0.30	-1.65	1.65

- **del_reggrow**: del_reggrow (delta of regionally averaged economic growth) measures the annual change in the average regional economic growth since the previous year. Positive values indicate a higher regional average of economic growth. Negative values indicate a contraction in regional average economic growth.

	N	Mean	SD	Min	Max
Change in Regional Economic Growth	2667	0.00	0.08	-0.29	0.23

- **lnpop2**: lnpop2 measures the estimated logged population of a country for that year. The measure is derived from the World Bank and the International Monetary Fund’s World Economic Outlook. Higher values indicate a larger population.

	N	Mean	SD	Min	Max
Population (logged)	2667	87.93	27.34	31.48	199.29

- **logIMR**: logIMR measures the logged annual infant mortality rate. Data for infant mortality rate comes from the World Bank and is updated annually. Larger values indicate a higher infant mortality rate for that country-year. Because World Bank data is generally 1 to 2 years behind, **2020** and **2021** are estimated using country-specific ARIMA modeling.

	N	Mean	SD	Min	Max
Infant Mortality Rate (logged)	2667	3.18	1.11	0.21	5.17

- **regav_IMR**: regav_IMR measures the regionally averaged logged annual infant mortality rate. Larger values indicate a higher regional infant mortality rate on average.

	N	Mean	SD	Min	Max
Regionally Averaged Infant Mortality Rate (logged)	2667	3.15	0.93	1.20	4.79

Country Characteristics

- **SPI**: SPI measures the Standardized Precipitation Index (SPI) for each country month. You can interpret SPI like a Z-score. Values of 0 correspond to historically average levels of rainfall for that country-month. Positive values above 0 indicate more rainfall than historically expected, with values greater than 1.96 generally corresponding to a statistically significant increase in rainfall. Conversely, you can interpret negative values as indicating less than expected rainfall. The underlying data comes from NOAA’s PREC/L observation data, with future months forecasted by using country-specific optimized ARIMA models. As of January, 2021, ground-truth data ends at *Dec 2020*, with the forecast taking place from *January 2021* to *December 2021*. SPI is equivalent to REIGN’s precip measure.

	N	Mean	SD	Min	Max
Standardized Precipitation Index (SPI)	2667	-0.07	0.79	-3.01	4.72

- **logpredict**: logpredict is an estimated probability of the risk of a military coup attempt taking place in the country-month. couprisk is estimated by using machine-learning classification and a stacking

algorithm that combines the individual predictions of a (a) random forest algorithm, (b) logistic regression model and (c) linear probability regression model. A Generalized estimating equation is then used to combine the weighted probabilities of each previous model to take advantage of each model's strengths while minimizing poor predictions. Model training is performed through rolling-origin cross-validation using 1950 - 1974 as the initial training window. As such, coup risk estimates only begin in January, 1975 until January, 2021.

	N	Mean	SD	Min	Max
Structural Risk of Coup Event (Raw)	2667	0.00	0.01	0.00	0.13

- political_violence:** `political_violence` is a numeric variable that measures the relative level (z-score) of political violence experienced within the borders of a country for that year. Data is derived from the UCDP's georeferenced event dataset (up to 2019) and START's Global Terrorism Database (GTD) (up to 2018). Values of 0 correspond to historically average levels of political violence for that country-year. Positive values above 0 indicate more violence than historically expected, with values greater than 1.96 generally corresponding to a statistically significant increase in violence. Conversely, you can interpret negative values as indicating less than expected violence. As of January, 2021, ground-truth data ends at Dec 2019. Values in *2020* and *2021* are forecasted with country-specific ARIMA models.

	N	Mean	SD	Min	Max
Relative Level of Political Violence	2667	0.07	2.10	-2.14	30.70