Data Dictionary: Statistical Risk Assessment
Early Warning Project

Variable Name	Description	Transformation and Sourcing Notes from EWP	Source
mk_ongoing_count_log	Count of ongoing mass killings in a particular country (log base 2).	This variable is the count of ongoing mass killings, logged (base 2).	Early Warning Project
mean_mk_onset	The average annual rate of mass killing onsets (from the EWP dataset) in a given country (log base 2).	The number of mass killing onsets the country had up to the last year of included data, divided by the number of years the country has been included in our dataset. That number is then logged (base 2).	Early Warning Project
mk_onset_prev_year	Was there a mass killing onset in this country in the prior year (for example, for forecasts made based on 2022 data, this variable specifies if an onset occurred in 2021)?	● 0: No ● 1: Yes	Early Warning Project
newcountry	Indicates country is in its first, second, or third year in the EWP dataset (Note: In order to enter the dataset, countries must be independent, have a population over 500k, and have sufficient data available. Countries that were sufficiently large and independent prior to 1945 enter the dataset in 1945).	● 0: No ● 1: Yes	Early Warning Project, based on Kristian Skrede Gleditsch's list of independent states
mk_ever	Has the country (or its predecessor) experienced a mass killing onset in any year since 1945?	● 0: No ● 1: Yes	Early Warning Project

year	The year for the corresponding data in the model (for example, for forecasts made based on 2022 data, the year is 2022).	No transformations on the variable.	Early Warning Project
year_sq	This year (see row above) squared.	This variable represents the year (see row above) squared.	Early Warning Project
efindex	A measure of ethnic heterogeneity. The index "corresponds to the probability that two randomly drawn individuals within a country are not from the same ethnic group" (Drazanova 2019).	No transformations on the variable from the Historical Index of Ethnic Fractionalization (HIEF) Dataset. Where values are missing, we filled in with the UCLA Anderson School of Management's Ethnic Fractionalization Data.	Historical Index of Ethnic Fractionalization (HIEF) Dataset (efindex); UCLA Anderson School of Management's Ethnic Fractionalization Data (Alesina et al. 2003)
discrimpop	The percentage of the population that is coded as being discriminated against.	No transformations on the variable. Where values are missing, we fill in by carrying forward the most recent value.	Ethnic Power Relations (EPR) Dataset
v2csgender_binary	Restrictions on participation of women in civil society organizations.	Transformed to binary variable: ● 0: Women are rarely or never prevented from participating in civil society organizations. ● 1: Women are often or always prevented from participating in civil society organizations. Coded as 1 when v2csgender equals 0, 1, or 2. Coded as 0 when v2csgender equals 3 or 4. Meaning: Are women prevented from participating in civil society organizations (CSOs)? 0: Almost always. 1: Frequently. 2: About half the time. 3: Rarely. 4: Almost never	Varieties of Democracy (V-Dem) (v2csgender)

v2mecenefm_binary	Does the government directly or indirectly attempt to censor the print or broadcast media?	Transformed to binary variable: ● 0: Attempts at censorship are rare and/or limited. ● 1: Attempts at censorship are direct and/or routine. Coded as 1 when v2mecenefm equals 0, 1, or 2. Coded as 0 when v2mecenefm equals 3 or 4. Meaning: Does the government directly or indirectly attempt to censor the print or broadcast media? 0: Attempts to censor are direct and routine. 1: Attempts to censor are indirect but nevertheless routine. 2: Attempts to censor are direct but limited to especially sensitive issues. 3: Attempts to censor are indirect and limited to especially sensitive issues. 4: The government rarely attempts to censor major media in any way, and when such exceptional attempts are discovered, the responsible officials are usually punished.	Varieties of Democracy (V-Dem) (v2mecenefm)
widetargeting	Does any ongoing mass killing in the state target multiple groups (ethnic, racial, national, or religious) or a broadly defined group (e.g., political opponents)?	● 0: No ● 1: Yes	Early Warning Project
narrowtargeting	Do all ongoing mass killings in the state each only target exactly one ethnic, racial, national, or religious group?	● 0: No ● 1: Yes	Early Warning Project
partyban.new.0, partyban.new.1, and partyban.new.2	Are any parties banned?	Transformed to three binary variables: ● if partyban.new.0 is coded as 1: All parties except the state-sponsored party (and closely allied parties) are banned or elections are non-partisan or there are no officially recognized parties. ● if partyban.new.1 is coded as 1: Some parties are banned. ● if partyban.new.2 is coded as 1: No parties are officially banned. Coded partyban.new.0 coded as 1 when v2psparban_ord equals 0 or 1; 0 otherwise; partyban.new.1 coded as 1 when v2psparban_ord equals 2 or 3; 0 otherwise; partyban.new.2 coded	Varieties of Democracy (V-Dem) (v2psparban_ord)

		as 1 when v2psparban_ord equals 4; 0 otherwise. Meaning: Are any parties banned? 0: Yes. All parties except the state-sponsored party (and closely allied parties) are banned. 1: Yes. Elections are non-partisan or there are no officially recognized parties. 2: Yes. Many parties are banned. 3: Yes. But only a few parties are banned. 4: No. No parties are officially banned	
countryage_new_ln	Years (integer) country has existed in the EWP dataset (log base 2).	This variable is the count of years the country has existed in the EWP dataset, logged (base 2).	[Early Warning Project](#), based on [Kristian Skrede Gleditsch's list of independent states](#)
imr.fwd.fill.sqrt	Infant mortality (square root).	This variable is the square root of the infant mortality rate per 1,000 live births from the World Bank World Development Indicators. Where values are missing, we filled in with the CIA World Factbook and Mitchell's International Historical Statistics. The only remaining missingness we handle are cases where the data are not yet published by the sources; as is the case for 2022 data. For that, we fit log(imr) to the country's time series with a second order polynomial (up through 2021) and predict the value for 2022. This is then re-exponentiated and we take the square root to fill in the value of imr.fwd.fill.sqrt.	[World Bank World Development Indicators](#); [CIA World Factbook](#); [Brian Mitchell's International Historical Statistics](#)
coup.try.5yr	Did a successful or unsuccessful coup attempt occur in this year or the prior four years?	● 0: No ● 1: Yes	[Powell and Thyne Coup d'Etat Dataset](#)
reg.afr	Region: Africa	● 0: not in this region ● 1: in this region	[US Department of State](#)
reg.eap	Region: East Asia Pacific	● 0: not in this region ● 1: in this region	[US Department of State](#)
reg.eur	Region: Europe	● 0: not in this region ● 1: in this region	[US Department of State](#)
reg.sca	Region: South and Central Asia	● 0: not in this region ● 1: in this region	[US Department of State](#)
reg.mna	Region: Middle East and North Africa	● 0: not in this region ● 1: in this region	[US Department of State](#)
wdi.popsize.log2	Population size (log base 2).	This variable is the population size, logged (base 2) from the World Bank World Development Indicators (SP.POP.TOTL).	[World Bank World Development Indicators](#) (SP.POP.TOTL); [CIA World Factbook](#)

		Where values are missing, we filled in with the CIA World Factbook.	
gdppcgrowth.combined	Annual percentage change in GDP per capita.	No transformations on the variable from V-Dem (e_migdpgro). Where values are missing, we filled in with regression-adjusted values from the World Bank World Development Indicators (NY.GDP.MKTP.KD.ZG). Where still missing, we filled in with the CIA World Factbook.	Varieties of Democracy (V-Dem) (e_migdpgro); World Bank World Development Indicators (NY.GDP.MKTP.KD.ZG); CIA World Factbook
ios.iccpr1	State signatory of the First Optional Protocol to the International Covenant on Civil and Political Rights.	0: Not a signatory 1: Signatory No transformations on the variable from Ulfelder's dataset. Where values are missing, we filled in with values from the Office of the United Nations High Commissioner for Human Rights. All country years are coded as 0 prior to 1967 as the Protocol was not available for signature before that.	Jay Ulfelder's Country Memberships in Selected Intergovernmental Organizations and Accession to Selected Regional and Global Treaty Regimes; Office of the United Nations High Commissioner for Human Rights
includesnonstate	A dummy indicator for years in which the data also include non-state led mass killings in the outcome (1989 and after). Used to account for shift in overall probability of an onset.	0: Years before 1989. 1: Years 1989 and after.	Early Warning Project
minorityrule	Minority control: Does a minority control political rule?	Transformed to binary variable: 0: Majority rule 1: Minority rule Coded as 0 (Majority rule) if v2pepwrsoc_ord is greater than 1; Coded as 1 (Minority Rule) if v2pepwrsoc_ord is less than or equal to 1. Meaning: Is political power distributed according to social groups? 0: Political power is monopolized by one social group comprising a minority of the population. This monopoly is institutionalized, i.e., not subject to frequent change. 1: Political power is monopolized by several social groups comprising a minority of the population. This monopoly is institutionalized, i.e., not subject to frequent change. 2: Political power is monopolized by several social groups comprising a majority of the population. This monopoly is institutionalized, i.e., not subject to frequent change.	Varieties of Democracy (V-Dem) (v2pepwrsoc_ord)

		3: Either all social groups possess some political power, with some groups having more power than others; or different social groups alternate in power, with one group controlling much of the political power for a period of time, followed by another — but all significant groups have a turn at the seat of power. 4: All social groups have roughly equal political power or there are no strong ethnic, caste, linguistic, racial, religious, or regional differences to speak of. Social group characteristics are not relevant to politics.	
battledeaths.ln2	Battle-related deaths (log base 2 of battle-related deaths plus 1).	Battle-related deaths, logged (base 2 plus 1) from PRIO Battledeaths Dataset up to its conclusion in 2008. Then we use UCDP data up to the last year of included data for the assessment. Note: only battle deaths from conflicts type 3 (intrastate) and 4 (internationalized intrastate) are included, and battle deaths where there is more than one location (locationid) are removed.	Peace Research Institute Oslo (PRIO); Uppsala Conflict Data Program (UCDP)
judicialreform	Were the judiciary's formal powers altered this year in ways that affect its ability to control the arbitrary use of state authority?	0: The judiciary's ability to control arbitrary power was reduced via institutional reform. 1: There was no change to the judiciary's ability to control arbitrary power via institutional review. 2: The judiciary's ability to control arbitrary power was enhanced via institutional reform.	Varieties of Democracy (V-Dem) (v2jureform_ord)
religiousfreedom	Is there freedom of religion?	0: Not respected by public authorities. Hardly any freedom of religion exists. 1: Weakly respected by public authorities. Some elements of autonomous organized religious practices exist and are officially recognized. 2: Somewhat respected by public authorities. 3: Mostly respected by public authorities. There are minor restrictions on the freedom of religion, predominantly limited to a few isolated cases. 4: Fully respected by public authorities. The population enjoys the right to practice any religious belief they choose.	Varieties of Democracy (V-Dem) (v2clrelig_ord)

pol_killing_approved	Do political killings occur? Political killings are killings by the state or its agents without due process of law for the purpose of eliminating political opponents. These killings are the result of deliberate use of lethal force by the police, security forces, prison officials, or other agents of the state (including paramilitary groups).	Transformed to binary variable: ● 0: Political killings are not practiced systematically or not typically incited and approved by top leaders of government. ● 1: Political killings are practiced systematically and they are typically incited and approved by top leaders of government. Coded as 1 only if v2clkill_ord equals 0; 0: all other. Meaning: Is there freedom from political killings? 0: Not respected by public authorities. Political killings are practiced systematically and they are typically incited and approved by top leaders of government. 1: Weakly respected by public authorities. Political killings are practiced frequently and top leaders of government are not actively working to prevent them. 2: Somewhat respected by public authorities. Political killings are practiced occasionally but they are typically not incited and approved by top leaders of government. 3: Mostly respected by public authorities. Political killings are practiced in a few isolated cases but they are not incited or approved by top leaders of government. 4: Fully respected by public authorities. Political killings are non-existent.	[Varieties of Democracy (V-Dem)](#) (v2clkill_ord)
freediscussion	Are citizens able to openly discuss political issues in private homes and in public spaces?	No transformations on the variable. We used the continuous form as provided by V-Dem: "Ordinal, converted to interval by the measurement model, from low to high 0-1."	[Varieties of Democracy (V-Dem)](#) (v2xcl_disc)
social_inequality	Do some social groups—as distinguished by language, ethnicity, religion, race, region, or caste—have fewer civil liberties than others?	Transformed to binary variable: ● 0: Members of all salient social groups enjoy the same level of civil liberties. ● 1: Members of some social groups enjoy fewer civil liberties than the general population. Coded as 1 when v2clsocgrp_ord equals 0; 0: all other. Meaning: Do all social groups, as distinguished by language, ethnicity, religion, race, region, or caste, enjoy the same level of civil liberties, or are some groups generally in a more favorable	[Varieties of Democracy (V-Dem)](#) (v2clsocgrp_ord)

		position? 0: Members of some social groups enjoy much fewer civil liberties than the general population. 1: Members of some social groups enjoy substantially fewer civil liberties than the general population. 2: Members of some social groups enjoy moderately fewer civil liberties than the general population. 3: Members of some social groups enjoy slightly fewer civil liberties than the general population. 4: Members of all salient social groups enjoy the same level of civil liberties.	
even_civilrights	Is government respect (or lack of respect) for civil liberties equal across different geographic areas of the country?	Transformed to binary variable: ● 0: Not equally distributed. ● 1: Equally distributed. Coded as 1 when v2clrgunev_ord equals 2; 0: all other. Meaning: Does government respect for civil liberties vary across different areas of the country? 0: Yes. Government officials in some areas of the country respect civil liberties significantly more (or, alternatively, significantly less) than government officials in other areas of the country. 1: Somewhat. Government officials in some areas of the country respect civil liberties somewhat more (or, alternatively, somewhat less) than government officials in other areas of the country. 2: No. Government officials in most or all areas of the country equally respect (or, alternatively, equally do not respect) civil liberties.	[Varieties of Democracy (V-Dem)](#) (v2clrgunev_ord)
repress_civilsoc	Does the government attempt to repress civil society organizations (CSOs)?	Transformed to binary variable: ● 0: Not repressed ● 1: Repressed Coded as 1 when v2csreprss_ord equals 0. 0: all other. Meaning: Does the government attempt to repress civil society organizations (CSOs)? 0: Severely. The government violently and actively pursues all real and even some imagined members of CSOs. They seek not	[Varieties of Democracy (V-Dem)](#) (v2csreprss_ord)

| | | only to deter the activity of such groups but to effectively liquidate them.
1: Substantially. In addition to the kinds of harassment outlined in responses 2 and 3 below, the government also arrests, tries, and imprisons leaders of and participants in oppositional CSOs who have acted lawfully. Other sanctions include disruption of public gatherings and violent sanctions of activists (beatings, threats to families, destruction of valuable property).
2: Moderately. In addition to material sanctions outlined in response 3 below, the government also engages in minor legal harassment (detentions, short-term incarceration) to dissuade CSOs from acting or expressing themselves. The government may also restrict the scope of their actions through measures that restrict association of civil society organizations with each other or political parties, bar civil society organizations from taking certain actions, or block international contacts.
3: Weakly. The government uses material sanctions (fines, firings, denial of social services) to deter oppositional CSOs from acting or expressing themselves. They may also use burdensome registration or incorporation procedures to slow the formation of new civil society organizations and sidetrack them from engagement. The government may also organize Government Organized Movements or NGOs (GONGOs) to crowd out independent organizations.
4: No. Civil society organizations are free to organize, associate, strike, express them | |

| social_power_dist | Is political power distributed according to social groups? | Transformed to binary variable:

● 0: All social groups have roughly equal political power or there are no strong ethnic, caste, linguistic, racial, religious, or regional differences to speak of. Social group characteristics are not relevant to politics.
● 1: Political power is monopolized by one or a few social groups.

Coded as 1 when v2pepwrsoc_ord equals 0,1, or 2. 0: all other. Meaning: Is political power distributed according to social groups?
0: Political power is monopolized by one social group comprising a minority of the population. This monopoly is institutionalized, i.e., not subject to frequent change.
1: Political power is monopolized by several social groups comprising a minority of the population. This monopoly is institutionalized, i.e., not subject to frequent change.
2: Political power is monopolized by several social groups comprising a majority of the population. This monopoly is institutionalized, i.e., not subject to frequent change.
3: Either all social groups possess some political power, with some groups having more power
than others; or different social groups alternate in power, with one group controlling much of
the political power for a period of time, followed by another —but all significant groups have
a turn at the seat of power.
4: All social groups have roughly equal political power or there are no strong ethnic, caste,
linguistic, racial, religious, or regional differences to speak of. Social group characteristics are
not relevant to politics. | Varieties of Democracy (V-Dem) (v2pepwrsoc_ord) |

ses_power_dist	Is political power distributed according to socioeconomic position?	No transformations on the variable. We used the continuous form as provided by V-Dem: "Ordinal, converted to interval by the measurement model" with lower/smaller numbers indicating greater levels of power distributed by socioeconomic position, and higher/bigger numbers indicating lower levels of power distributed by socioeconomic position.	Varieties of Democracy (V-Dem) (v2pepwrses)
tradeshare.log2.combined	Trade as percentage ofGDP (log base 2).	This variable is trade as a percentage of gross domestic product, logged (base 2). V-Dem (multiple variables: e_gdppc, e_pop, e_cow_imports, and e_cow_exports) are the initial source. Where values are missing, we filled in with regression-adjusted values from the World Bank World Development Indicators. Where still missing, we filled in with the CIA World Factbook.	Varieties of Democracy (V-Dem) (e_gdppc, e_pop, e_cow_imports, and e_cow_exports); World Bank World Development Indicators; CIA World Factbook