GROW^{up} Research Front-End Documentation
RFE Release 2.0

Nils-Christian Bormann, Luc Girardin, Philipp Hunziker, Manuel Vogt
ETH Zurich
http://growup.ethz.ch/rfe
growup@gess.ethz.ch

August 17, 2015

Contents

Chapter 1

Group-Level Data

1.1 The EPR Universe

The sample universe of ethnic groups in the RFE group-level data is adopted from the EPR (*Ethnic Power Relations*) Core dataset (Vogt et al. ming, for more detailed information concerning versions and sources see 3.2). Ethnic groups are included in the EPR Core dataset according to a two-step coding procedure:

1.1.1 Country Selection

Firstly, the EPR Core dataset defines a time-variant list of countries for which ethnic groups are coded. Specifically, the EPR Core dataset covers all countries in the period 1946 - 2013 that meet the following criteria:

(i) Administered by an intact sovereign state, i.e. overseas colonies and failed states are not included.[1]

(ii) Population in 1990 is greater than or equal to 500'000 inhabitants.

Newly independent states are included in the dataset beginning with the year of independence. For example, Macedonia (independent on 20/11/1991) is included from 1991 onwards. Given these criteria, the EPR Core dataset covers a total of 8752 country-years from 165 countries.

1.1.2 Group Selection

Secondly, ethnic groups are coded on the basis of this list of country-years. For this purpose, the EPR Core dataset defines ethnicity as any subjectively experienced sense of

[1] However, countries considered "administered" by a failed state *are* included in the dataset if the period during which they are coded as failed states falls in between periods when the country meets the EPR Core dataset inclusion criteria. During these periods, however, the coding on the ethnic group level (see below) is not continued.

commonality based on the belief in common ancestry and shared culture. Given this definition, an ethnic group (i.e. a group of individuals sharing a common ethnicity) is included in the EPR Core dataset if it is politically relevant at least once in the sample period. An ethnic group is classified as politically relevant if at least one political organization claims to represent it in national politics or if its members are subjected to state-led political discrimination.

With these two selection procedures in place, the EPR Core dataset identifies 817 politically relevant ethnic groups (before considering hierarchies, see 1.2.3) in 141 countries across the globe for the period 1946 - 2013. For the remaining 24 countries, no ethnic group is coded as politically relevant during the entire sample period. For these cases, national *placeholder groups* are defined, which are coded as politically irrelevant for the entire sample period (e.g. Germans in Germany).[2]

1.2 Group Activity, Relevancy, and Hierarchies

The unit of observation in the RFE group-level data is the ethnic-group-year. An EPR ethnic group is included in the RFE group-level data for all years during which the group's host state meets the EPR inclusion criteria (see 1.1). Consequently, researchers should note that ethnic groups are included in the RFE group-level data even if they are currently inactive or irrelevant, two concepts that are discussed subsequently.

1.2.1 Group Activity

An ethnic group is deemed active in a given group-year if it is currently physically present in a country *and* is not currently represented by an active ancestor or descendant (see below). Physical presence may be altered through border changes or large migration flows. In the RFE group-level data, whether a group is active during a given group-year is indicated by the *isactive* variable.

1.2.2 Group Relevance

An ethnic groups is deemed relevant in a given group-year if, in accordance to the EPR 2014 definition, at least one political organization claims to represent it in national politics or if its members are subjected to state-led political discrimination (for more information, see the EPR 2014 codebook at `http://www.icr.ethz.ch/data/epr`). Due to the EPR inclusion criteria (see above), all groups in the RFE group-level data except placeholder groups are coded as relevant at least once during the sample period. Group relevance and activity are nested concepts, i.e. a group can only be relevant if it is active. In the RFE group-level data, whether a group is relevant during a given group-year is indicated by the *isrelevant* variable.

[2]These national placeholder groups are relevant for the coding of transnational ethnic links in the EPR-TEK data.

1.2.3 Group Hierarchies

Building on the EPR dataset, the RFE group-level data tracks hierarchical group relationships over time. There are two kinds of hierarchical transformations that an ethnic group may experience:

Split An ethnic group may split into n ($n \geq 2$) smaller groups. This is the case if the politically relevant concept of ethnicity is redefined over time and individuals that have formerly been considered part of the same ethnic group are now considered members of distinct groups.

Unification Several ethnic groups may unify into one larger group. This is the case if individuals that have formerly been considered part of different ethnic groups are now considered members of the same politically relevant ethnic group.

If either of these hierarchical transformations occur, the RFE group-level data defines a hierarchical relationship between the groups involved. The hierarchically superior group in these transformations (i.e. the group defined by the more inclusive definition of ethnicity) is called the *ancestor group*, whereas the hierarchically inferior groups are called *descendant groups*. Consequently, a split is a transformation from one ancestor group into several descendant groups, and a unification is a transformation from several descendant groups into one ancestor groups.

Hierarchical transformations are mirrored in the RFE group-level data in three different ways:

- Hierarchical transformations affect **group activity**: In the case of a split, the descendant groups are coded as inactive in the group-years prior to the split, whereas the ancestor group is coded as inactive in the group-years after the split. Analogously, in the case of a unification, the ancestor group is coded as inactive prior to the unification, and the descendant groups as inactive after the unification. This is to avoid that the same individuals in a country are double-counted as members of several different hierarchically related ethnic groups at the same time.

- Hierarchical transformations are tracked by the following variables: **active_anc_groupid** and **active_desc_groupid_arr** are strings of groupids that are non-missing if the group in question has one or more currently active ancestors or descendants. Note that these variables are mutually exclusive in the sense that a group may only have an active ancestor *or* one or more active descendants. **relevant_anc_groupid** and **relevant_desc_groupid_arr** are analogously defined variables indicating currently relevant ancestors and descendants.

- Hierarchical relationships are reflected in variables of the **family type**: The *family* prefix indicates that a variable applies to this group, its active ancestor group (if any), or its active descendant groups (if any). For example, the *family_warhist* indicates

the number of conflict onsets the given group or any of its ancestors or descendants have experienced up until this year. *family* type variables are useful for tracking phenomena that are a function of time while considering that members of a given ethnic group may have been members of another hierarchically related ethnic group in the past.

1.3 Ethnic Conflict Data

The information on ethnic conflicts in the RFE group-level data is compiled from two different sources: The ACD2EPR dataset (Wucherpfennig et al. 2012), the UCDP Actor Dataset (Uppsala Conflict Data Program 2014), and the Uppsala/PRIO Armed Conflict Database (ACD) (Gleditsch et al. 2002). For more information on these datasets and the exact versions used, please see 3.2. These sources allow the identification of ethnic conflicts and their mapping onto the ethnic groups in the EPR Core Dataset 2014 (EPR) in three steps:

- The ACD2EPR dataset links UCDP rebel organizations to the EPR ethnic groups between 1946 and 2013.
- The UCDP Actor Dataset lists, among others, all rebel organizations involved in Internal or Internationalized Armed Conflicts.
- For each of these rebel–ethnic group dyads we code the occurrence, duration, and intensity of intrastate conflicts based on the information in the ACD dataset.

The remainder of this section briefly describes the structure of these data sources and establishes how the latter are used to code the conflict related variables in the RFE group-level data.

1.3.1 ACD2EPR

The units of observation in the ACD2EPR dataset are rebel-organization/ethnic-group dyads. For every rebel actor included in the UCDP Actor Dataset, ACD2EPR reports whether the organization can be linked to any EPR group in the same country through ethnic claims or significant ethnic recruitment.[3][4] More specifically, for the RFE group-level conflict data, a link between a rebel organization and an ethnic group is coded if the ACD2EPR *claim* variable indicates an explicit or implicit ethnic claim by the rebel organization ($claim >= 1$), and the *recruitment* variable indicates that the rebel organization has recruited from the respective ethnic group significantly ($recruitment = 1$). Note

[3]Note that previous versions of the ACD2EPR data linked ethnic groups to rebel organizations in the Non-State Actor dataset by Cunningham et al. (2009). These actors are now fully integrated into the UCDP Actor Dataset, which therefore provides backwards compatibility to the the NSA data.

[4]The original ACD2EPR dataset also establishes links between rebel organizations and ethnic groups among other dimensions, but these are not relevant for the RFE group-level data conflict coding.

that the ACD2EPR coding is time invariant. Thus, links between rebel organizations and ethnic groups exists for the entire lifespan of a rebel organization.[5] The RFE coding rule is that for a definite link between an ethnic group and a rebel organization to be established, the two must be connected through both *claim and recruitment*. With this coding rule in place, we are able to produce a time invariant one-to-many table connecting EPR groups one or multiple rebel organizations (if any).

1.3.2 UCDP Actor Dataset

The UCDP Actor Dataset dataset lists all actors and the UCDP IDs of the armed conflicts that these actors are involved in (see 1.3.3). By subsetting the list to only include non-state actors that are involved in Internal and Internationalized Armed Conflicts, we obtain the initial list of rebel organizations for which the ACD2EPR dataset codes links to EPR groups. Consequently, combining the EPR-group-to-rebel-organization table derived above with the Actor Dataset, we are able to create a new table indicating whether EPR groups are involved in one or more ACD conflicts through links with one or more rebel organizations.

1.3.3 ACD

Conflicts in the ACD dataset follow the UCDP conflict definition (please see 3.2 for links to more information and references) and are assigned a unique *UCDP ID*. ACD conflicts are assigned to one of four categories: Extrasystemic, Interstate, Internal, and Internationalized Internal Conflicts. Of these types, only the latter two are relevant for the RFE group-level conflict coding, since we focus exclusively on ethnic civil wars.

ACD conflicts are distinguished along the incompatibility they originate from, not their temporal dimension. Hence, a conflict that ends and reoccurs after a substantial period of time is assigned the same UCDP ID, regardless of the time period separating the actual conflict episodes, as long as the incompatibility remains the same. The ACD dataset provides information on the temporal dimension of conflict termination and recurrence with a variable indicating the end of a conflict episode, whereas according to the ACD coding rules, a conflict episode ends in a given year if there is no conflict-related activity in the next calendar year.[6] For the purpose of the RFE group-level data conflict coding, this definition is altered: For all conflict onset and incidence variables in the RFE group-level dataset, a conflict episode is only considered terminated if there is *no conflict-related activity in the following two calender years*. This recoding of ACD conflict episode is adopted in order to prevent an inflation of conflict episodes (and thus conflict onsets, see 1.3.4) in low-activity conflicts that reoccur repeatedly. With this 2-year episode coding rule in place, we are able to merge the table mapping EPR groups onto ACD conflicts discussed above (see 1.3.2) to the modified ACD conflict table, and thus create a time-variant dataset

[5] "Lifespan" refers to the period(s) during which a rebel organization is active in intrastate wars identified by the ACD.

[6] "Conflict-related activity" refers to the UCDP threshold of at least 25 battle deaths per annum.

indicating whether and when EPR groups are involved in ACD conflict episodes through links with one or several rebel organizations in the UCDP Actor Dataset.

1.3.4 Conflict Onset

The RFE group-level data follows a standard conflict onset coding based on the mapping of conflicts onto ethnic groups discussed above. Specifically, the coding rule imposed is that ethnic group G experiences conflict onset in year t if

(i) a rebel organization linked to G enters an ACD conflict active in year t,

and

(ii) no rebel organization linked to G has been active in the same ACD conflict in the two calendar years preceding t.

Rule (ii) ensures that conflicts linked to ethnic groups via multiple rebel organizations don't generate an excessive number of conflict onsets simply because individual rebel organizations start and stop fighting while the overall conflict continues to be violent.

1.3.5 KO and DO Options

The RFE group-level data conflict onset variables come in two variants: The KO and the DO option. KO stands for *Keep Ongoing* and is the default option. Conflict onset variables with the *KO* option take the value of 1 for group-years in which a group experiences conflict onset (see 1.3.4), and 0 in all other years.

DO stands for *Drop Ongoing* and represents a slightly more advanced coding. DO conflict onset variables indicate conflict onset in the same manner as the KO variables, but are censored if a given group has been involved in any conflict in the previous two calendar years. The reasoning underlying this definition is that in these years, positive conflict onset codings are highly unlikely due to the rules established in section 1.3.4, and thus in many statistical applications these observations will be excluded from the analysis.

1.3.6 Conflict Incidence

The RFE group-level data includes a number of conflict incidence variables, which indicate whether in a given year a group is involved in an ACD conflict episode through links with one or several rebel organizations. Conflict incidence variables assume the value of 1 in years when at least one rebel organization linked to an ethnic group is involved in an ACD conflict episode, and 0 otherwise.

Please note that because of the 2-year episode coding rule imposed on the ACD data, conflict episodes may terminate for a single calendar year and then continue without a new group-level conflict onset occuring (see 1.3.3 and 1.3.4).

1.4 Geographical Data

The group-level RFE offers two types of variables referring to geospatial information. The variables with the *geo* and *ggp* labels are derived directly from GeoEPR data set, version 2014. The actual spatial information (i.e., the group-level settlement polygons) is not distributed via the RFE, but can be downloaded at `http://www.icr.ethz.ch/data/geoepr`. You may use the groupid in conjuction with the *ggp_startdate* and *ggp_enddate* variables to merge the RFE group-year data with the spatial information contained in the GeoEPR 2014 shapefile.

The remaining geospatial variables are raster derived data, i.e. variables creating by overlaying the GeoEPR 2014 settlement polygons with various geospatial raster datasets. Raster derived variables with the with the *_total* label are derived by taking total sum of all raster values (e.g., population, nightlights) intersecting with a given GeoEPR settlement polygon. Raster derived variables with *_corr* label are derived by first classifying settlement areas depending on whether they intersect with the settlement polygons of other groups within the same country. In cases where settlement polygons overlap, the sum of the raster values are divided evenly among the present groups. Hence, this coding ensures that raster values (e.g., local population or nightlight emissions) are *not* double counted.

1.5 TEK Data

All variables with the label *tek* are derived from the Transborder Ethnic Kin dataset, version 2014. For more information, see `http://www.icr.ethz.ch/data/tek`.

1.6 Ethnic Dimensions Data

All variables with the label *ic* are derived from the EPR Ethnic Dimensions dataset, version 2014. For more information (including value label definitions), see `http://www.icr.ethz.ch/data/ic`.

1.7 Group-Level Variables

1.7.1 gwgroupid

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Unique ID of ethnic group
Sources:	EPR

1.7.2 year

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Year of observation (Group-Level)
Sources:	NA

1.7.3 groupname

Type:	String
Value Range:	NA
Description:	Name of ethnic group
Sources:	EPR

1.7.4 countries_gwid

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Unique GW ID of country
Sources:	GW

1.7.5 countryname

Type:	String
Value Range:	NA
Description:	Name of country
Sources:	GW

1.7.6 isactive

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether group is physically present in a country and not represented by active ancestor or descendant groups
Sources:	EPR

1.7.7 isrelevant

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether group is politically relevant
Sources:	EPR

1.7.8 groupsize

Type:	Double Precision
Value Range:	(0, 1]
Description:	This group's population size as a fraction of the country's total population
Sources:	EPR

1.7.9 statusid

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	ID indicating this group's political status
Sources:	EPR

1.7.10 statusname

Type:	String
Value Range:	NA
Description:	Name of the corresponding value on StatusID
Sources:	EPR

1.7.11 regaut

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating wether group has meaningful political autonomy.
Sources:	EPR

1.7.12 status_pwrrank

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Political status of this group ranked on a scale from 1 (discriminated) to 7 (monopoly); self-exclusion ranks at 3.
Sources:	EPR

1.7.13 status_monopoly

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group is assigned the status 'monopoly'
Sources:	EPR

1.7.14 status_dominant

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group is assigned the status 'dominant'
Sources:	EPR

1.7.15 status_senior

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group is assigned the status 'senior partner'
Sources:	EPR

1.7.16 status_junior

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group is assigned the status 'junior partner'
Sources:	EPR

1.7.17 status_selfexclusion

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group is assigned the status 'self-exclusion'
Sources:	EPR

1.7.18 status_powerless

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group is assigned the status 'powerless'
Sources:	EPR

1.7.19 status_discrim

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group is assigned the status 'discriminated'
Sources:	EPR

1.7.20 status_egip

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group is an EGIP (status 'junior partner' or higher)
Sources:	EPR

1.7.21 status_excl

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group is a MEG (status ' self-exclusion' or lower)
Sources:	EPR

1.7.22 egip_groups_count

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Count variable indicating the number of EGIPs in this group's country
Sources:	EPR

1.7.23 excl_groups_count

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Count variable indicating the number of MEGs in this group's country
Sources:	EPR

1.7.24 rlvt_groups_count

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Count variable indicating the number of relevant groups in this group's country
Sources:	EPR

1.7.25 actv_groups_count

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Count variable indicating the number of active groups in this group's country
Sources:	EPR

1.7.26 lpop

Type:	Double Precision
Value Range:	[0, 1]
Description:	Sum of the ethnically relevant population in this group's country (as a fraction of total population)
Sources:	EPR

1.7.27 lsize

Type:	Double Precision
Value Range:	[0, 1]
Description:	This group's population size as a fraction of the ethnically relevant population of this group's country (groupsize / lpop)
Sources:	EPR

1.7.28 egippop

Type:	Double Precision
Value Range:	[0, 1]
Description:	Sum of the population of all EGIP groups in this group's country (as a fraction of total population)
Sources:	EPR

1.7.29 legippop

Type:	Double Precision
Value Range:	[0, 1]
Description:	EGIP population as a fraction of ethnically relevant population in this group's country (egippop / lpop)
Sources:	EPR

1.7.30 exclpop

Type:	Double Precision
Value Range:	[0, 1]
Description:	Sum of the population of all MEG groups in this group's country (as a fraction of total population)
Sources:	EPR

1.7.31 lexclpop

Type:	Double Precision
Value Range:	[0, 1]
Description:	MEG population as a fraction of ethnically relevant population in this group's country (exclpo / lpop)
Sources:	EPR

1.7.32 discrimpop

Type:	Double Precision
Value Range:	[0, 1]
Description:	Sum of the population of all discriminated groups in this group's country (as a fraction of total population).
Sources:	EPR

1.7.33 ldiscrimpop

Type:	Double Precision
Value Range:	[0, 1]
Description:	Discriminated population as a fraction of ethnically relevant population in this group's country (discrimpop / lpop).
Sources:	EPR

1.7.34 rbal

Type:	Double Precision
Value Range:	(0, 1]
Description:	If this group is a MEG: groupsize / (egippop + groupsize); If this group is an EGIP: groupsize / egippop
Sources:	EPR

1.7.35 rlbal

Type:	Double Precision
Value Range:	(0, 1]
Description:	If this group is a MEG: lsize / (legippop + lsize); If this group is an EGIP: lsize / legippop
Sources:	EPR

1.7.36 downgraded1

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has lost power (according to the status_pwrrank variable) in the previous year
Sources:	EPR

1.7.37 downgraded_excl1

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has been downgraded from EGIP to MEG in the previous year
Sources:	EPR

1.7.38 downgraded_regaut1

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether group has lost political autnomy in the previous year
Sources:	EPR

1.7.39 upgraded1

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has gained power (according to the status_pwrrank variable) in the previous year
Sources:	EPR

1.7.40 upgraded_excl1

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has been upgraded from MEG to EGIP in the previous year
Sources:	EPR

1.7.41 upgraded_regaut1

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether group has gained political autnomy in the previous year
Sources:	EPR

1.7.42 downgraded2

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has lost power (according to the status_pwrrank variable) in the previous 2 years
Sources:	EPR

1.7.43 downgraded_excl2

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has been downgraded from EGIP to MEG in the previous 2 years
Sources:	EPR

1.7.44 downgraded_regaut2

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether group has lost political autnomy in the previous 2 years
Sources:	EPR

1.7.45 upgraded2

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has gained power (according to the status_pwrrank variable) in the previous 2 years
Sources:	EPR

1.7.46 upgraded_excl2

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has been upgraded from MEG to EGIP in the previous 2 years
Sources:	EPR

1.7.47 upgraded_regaut2

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether group has gained political autnomy in the previous 2 years
Sources:	EPR

1.7.48 downgraded5

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has lost power (according to the status_pwrrank variable) in the previous 5 years
Sources:	EPR

1.7.49 downgraded_excl5

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has been downgraded from EGIP to MEG in the previous 5 years
Sources:	EPR

1.7.50 downgraded_regaut5

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether group has lost political autnomy in the previous 5 years.
Sources:	EPR

1.7.51 upgraded5

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has gained power (according to the status_pwrrank variable) in the previous 5 years
Sources:	EPR

1.7.52 upgraded_excl5

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has been upgraded from MEG to EGIP in the previous 5 years
Sources:	EPR

1.7.53 upgraded_regaut5

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether group has gained political autnomy in the previous 5 years
Sources:	EPR

1.7.54 downgraded10

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has lost power (according to the status_pwrrank variable) in the previous 10 years
Sources:	EPR

1.7.55 downgraded_excl10

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has been downgraded from EGIP to MEG in the previous 10 years
Sources:	EPR

1.7.56 downgraded_regaut10

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether group has lost political autnomy in the previous 10 years.
Sources:	EPR

1.7.57 upgraded10

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has gained power (according to the status_pwrrank variable) in the previous 10 years
Sources:	EPR

1.7.58 upgraded_excl10

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has been upgraded from MEG to EGIP in the previous 10 years
Sources:	EPR

1.7.59 upgraded_regaut10

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether group has gained political autnomy in the previous 10 years
Sources:	EPR

1.7.60 downgraded_hist

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has ever lost power (according to the status_pwrrank variable) since its first appearance in the EPR dataset
Sources:	EPR

1.7.61 downgraded_excl_hist

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has ever been downgraded from EGIP to MEG since its first appearance in the EPR dataset
Sources:	EPR

1.7.62 downgraded_regaut_hist

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has ever lost autonomy since its first appearance in the EPR dataset
Sources:	EPR

1.7.63 upgraded_hist

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has ever gained power (according to the status_pwrrank variable) since its first appearance in the EPR dataset
Sources:	EPR

1.7.64 upgraded_excl_hist

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has ever been upgraded from MEG to EGIP since its first appearance in the EPR dataset
Sources:	EPR

1.7.65 upgraded_regaut_hist

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether this group has ever gained autonomy since its first appearance in the EPR dataset
Sources:	EPR

1.7.66 family_downgraded1

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.67 family_downgraded_excl1

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.68 family_downgraded_regaut1

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.69 family_upgraded1

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.70 family_upgraded_excl1

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.71 family_upgraded_regaut1

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.72 family_downgraded2

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.73 family_downgraded_excl2

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.74 family_downgraded_regaut2

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.75 family_upgraded2

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.76 family_upgraded_excl2

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.77 family_upgraded_regaut2

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.78 family_downgraded5

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.79 family_downgraded_excl5

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.80 family_downgraded_regaut5

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.81 family_upgraded5

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.82 family_upgraded_excl5

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.83 family_upgraded_regaut5

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.84 family_downgraded10

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.85 family_downgraded_excl10

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.86 family_downgraded_regaut10

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.87 family_upgraded10

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.88 family_upgraded_excl10

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.89 family_upgraded_regaut10

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.90 family_isrelevant

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	EPR

1.7.91 active_anc_groupid

Type:	String
Value Range:	NA
Description:	String indicating groupid of currently active ancestor of this group
Sources:	EPR

1.7.92 active_des_groupid_arr

Type:	String
Value Range:	NA
Description:	String indicating comma-separated groupids of currently active descendants of this group
Sources:	EPR

1.7.93 relevant_anc_groupid

Type:	String
Value Range:	NA
Description:	String indicating groupid of currently relevant ancestor of this group
Sources:	EPR

1.7.94 relevant_des_groupid_arr

Type:	String
Value Range:	NA
Description:	String indicating comma-separated groupids of currently relevant descendants of this group
Sources:	EPR

1.7.95 onset_ko_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating group-level conflict onset / ko option
Sources:	ACD2EPR, ACD

1.7.96 onset_ko_terr_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating group-level territorial conflict onset / ko option
Sources:	ACD, ACD2EPR

1.7.97 onset_ko_gov_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating group-level governmental conflict onset / ko option
Sources:	ACD2EPR, ACD

1.7.98 incidence_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating ongoing group-level conflict
Sources:	ACD, ACD2EPR

1.7.99 incidence_terr_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating ongoing territorial group-level conflict
Sources:	ACD2EPR, ACD

1.7.100 incidence_gov_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating ongoing governmental group-level conflict
Sources:	ACD, ACD2EPR

1.7.101 onset_do_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating group-level conflict onset / do option
Sources:	ACD2EPR, ACD

1.7.102 onset_do_terr_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating group-level territorial conflict onset / do option
Sources:	ACD, ACD2EPR

1.7.103 onset_do_gov_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating group-level governmental conflict onset / do option
Sources:	ACD, ACD2EPR

1.7.104 family_onset_ko_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	ACD, ACD2EPR

1.7.105 family_onset_ko_terr_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	ACD, ACD2EPR

1.7.106 family_onset_ko_gov_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	ACD, ACD2EPR

1.7.107 family_incidence_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	ACD, ACD2EPR

1.7.108 family_incidence_terr_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	ACD, ACD2EPR

1.7.109 family_incidence_gov_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Family version of the respective variable described above
Sources:	ACD, ACD2EPR

1.7.110 warhist

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Count variable indicating number of conflict onsets this group has previously experienced
Sources:	ACD, ACD2EPR

1.7.111 warhist_terr

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Count variable indicating number of territorial conflict onsets this group has previously experienced
Sources:	ACD, ACD2EPR

1.7.112 warhist_gov

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Count variable indicating number of governmental conflict onsets this group has previously experienced
Sources:	ACD, ACD2EPR

1.7.113 family_warhist

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Family version of the respective variable described above
Sources:	ACD, ACD2EPR

1.7.114 family_warhist_terr

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Family version of the respective variable described above
Sources:	ACD, ACD2EPR

1.7.115 family_warhist_gov

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Family version of the respective variable described above
Sources:	ACD, ACD2EPR

1.7.116 peaceyears

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Years since group first appears in dataset or since the end of the last ongoing conflict episode
Sources:	ACD, ACD2EPR

1.7.117 peaceyears_terr

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Years since group first appears in dataset or since the end of the last ongoing territorial conflict episode
Sources:	ACD2EPR, ACD

1.7.118 peaceyears_gov

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Years since group first appears in dataset or since the end of the last ongoing governmental conflict episode
Sources:	ACD, ACD2EPR

1.7.119 family_peaceyears

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Family version of the respective variable described above
Sources:	ACD, ACD2EPR

1.7.120 family_peaceyears_terr

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Family version of the respective variable described above
Sources:	ACD, ACD2EPR

1.7.121 family_peaceyears_gov

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Family version of the respective variable described above
Sources:	ACD, ACD2EPR

1.7.122 geo_typeid

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	GeoEPR 2014 settlement type ID.
Sources:	GEOEPR

1.7.123 geo_typename

Type:	String
Value Range:	NA
Description:	GeoEPR 2014 settlement type name.
Sources:	GEOEPR

1.7.124 geo_concentrated

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether group's settlement area is concentrated within its host country's territory.
Sources:	GEOEPR

1.7.125 geo_statewide

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether group's settlement area is statewide.
Sources:	GEOEPR

1.7.126 geo_urban

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether group's settlement area is urban.
Sources:	GEOEPR

1.7.127 geo_migrant

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether group's settlement area is migrant.
Sources:	GEOEPR

1.7.128 geo_dispersed

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether group's settlement area is dispersed.
Sources:	GEOEPR

1.7.129 geo_unknown

Type:	Integer
Value Range:	[0; 0]
Description:	Binary flag indicating whether group's settlement area is unknown.
Sources:	GEOEPR

1.7.130 ggp_startdate

Type:	Date
Value Range:	NA
Description:	Start-date of GeoEPR episode corresponding to this group-year.
Sources:	GEOEPR

1.7.131 ggp_enddate

Type:	Date
Value Range:	NA
Description:	End-date of GeoEPR episode corresponding to this group-year.
Sources:	GEOEPR

1.7.132 area_sqkm

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Group's settlement area in square kilometers (derived from GeoEPR).
Sources:	GEOEPR

1.7.133 pop90_total

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Group's population in 1990; derived by GeoEPR/CIESIN GRUMPv1 overlay. All population within group polygon is summed.
Sources:	GRUMP, GEOEPR

1.7.134 pop90_corr

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Group's population in 1990; derived by GeoEPR/CIESIN GRUMPv1 overlay. Population intersecting with multiple polygons is divided evenly.
Sources:	GRUMP, GEOEPR

1.7.135 pop00_total

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Group's population in 2000; derived by GeoEPR/CIESIN GRUMPv1 overlay. All population within group polygon is summed.
Sources:	GRUMP, GEOEPR

1.7.136 pop00_corr

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Group's population in 2000; derived by GeoEPR/CIESIN GRUMPv1 overlay. Population intersecting with multiple polygons is divided evenly.
Sources:	GRUMP, GEOEPR

1.7.137 pop10_total

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Group's population in 2010; derived by GeoEPR/CIESIN GRUMPv1 overlay. All population within group polygon is summed.
Sources:	GRUMP, GEOEPR

1.7.138 pop10_corr

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Group's population in 2010; derived by GeoEPR/CIESIN GRUMPv1 overlay. Population intersecting with multiple polygons is divided evenly.
Sources:	GRUMP, GEOEPR

1.7.139 nightlight_total

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Stable nightlight emissions in group polygon; derived by GeoEPR/DMSP-OLS overlay. All nightlights within group polygon are summed.
Sources:	GEOEPR, DMSP

1.7.140 nightlight_corr

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Stable nightlight emissions in group polygon; derived by GeoEPR/DMSP-OLS overlay. Nightlights intersecting with multiple polygons are divided evenly.
Sources:	DMSP, GEOEPR

1.7.141 gdp90_total

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Group's contribution to host country's GDP in 1990; derived by GeoEPR/G-Econ overlay. All gross cell product within group polygon is summed.
Sources:	GEOEPR, GECON

1.7.142 gdp90_corr

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Group's contribution to host country's GDP in 1990; derived by GeoEPR/G-Econ overlay. Gross cell product intersecting with multiple polygons is divided evenly.
Sources:	GECON, GEOEPR

1.7.143 gdp95_total

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Group's contribution to host country's GDP in 1995; derived by GeoEPR/G-Econ overlay. All gross cell product within group polygon is summed.
Sources:	GEOEPR, GECON

1.7.144 gdp95_corr

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Group's contribution to host country's GDP in 1995; derived by GeoEPR/G-Econ overlay. Gross cell product intersecting with multiple polygons is divided evenly.
Sources:	GECON, GEOEPR

1.7.145 gdp00_total

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Group's contribution to host country's GDP in 2000; derived by GeoEPR/G-Econ overlay. All gross cell product within group polygon is summed.
Sources:	GEOEPR, GECON

1.7.146 gdp00_corr

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Group's contribution to host country's GDP in 2000; derived by GeoEPR/G-Econ overlay. Gross cell product intersecting with multiple polygons is divided evenly.
Sources:	GECON, GEOEPR

1.7.147 gdp05_total

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Group's contribution to host country's GDP in 2005; derived by GeoEPR/G-Econ overlay. All gross cell product within group polygon is summed.
Sources:	GECON, GEOEPR

1.7.148 gdp05_corr

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Group's contribution to host country's GDP in 2005; derived by GeoEPR/G-Econ overlay. Gross cell product intersecting with multiple polygons is divided evenly.
Sources:	GECON, GEOEPR

1.7.149 elevsd

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Standard deviation of gridded elevation measurements (0.008330 decimal degree resolution) intersecting with group polygon.
Sources:	GTOPO30, GEOEPR

1.7.150 meanelev

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Mean elevation of territory intersecting with group polygon.
Sources:	GTOPO30, GEOEPR

1.7.151 tek_groupid_arr

Type:	String
Value Range:	NA
Description:	String indicating comma-separated groupids of TEK groups associated with this group
Sources:	TEK

1.7.152 tek_count

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Count variable indicating number of TEK groups associated with this group
Sources:	TEK

1.7.153 tek_isrelevant

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether at least one of this group's TEK groups is currently politically relevant
Sources:	TEK

1.7.154 tek_excl

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether at least one of this group's TEK groups is currently politically excluded (MEG)
Sources:	TEK

1.7.155 tek_egip

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating whether at least one of this group's TEK groups currently has EGIP status; note: for this coding; placeholder groups are considered EGIP
Sources:	TEK

1.7.156 ic_religion1

Type:	String
Value Range:	NA
Description:	String indicating largest religion for this group
Sources:	IC

1.7.157 ic_rel1_size

Type:	Double Precision
Value Range:	[0, 1]
Description:	Fraction of group associated with largest religion
Sources:	IC

1.7.158 ic_religion2

Type:	String
Value Range:	NA
Description:	String indicating second largest religion for this group
Sources:	IC

1.7.159 ic_rel2_size

Type:	Double Precision
Value Range:	[0, 1]
Description:	Fraction of group associated with second largest religion
Sources:	IC

1.7.160 ic_religion3

Type:	String
Value Range:	NA
Description:	String indicating third largest religion for this group
Sources:	IC

1.7.161 ic_rel3_size

Type:	Double Precision
Value Range:	[0, 1]
Description:	Fraction of group associated with third largest religion
Sources:	IC

1.7.162 ic_language1

Type:	String
Value Range:	NA
Description:	String indicating largest language for this group
Sources:	IC

1.7.163 ic_lang1_size

Type:	Double Precision
Value Range:	(0, 1]
Description:	Fraction of group associated with largest language
Sources:	IC

1.7.164 ic_language2

Type:	String
Value Range:	NA
Description:	String indicating second largest language for this group
Sources:	IC

1.7.165 ic_lang2_size

Type:	Double Precision
Value Range:	[0, 1]
Description:	Fraction of group associated with second largest language
Sources:	IC

1.7.166 ic_language3

Type:	String
Value Range:	NA
Description:	String indicating third largest language for this group
Sources:	IC

1.7.167 ic_lnag3_size

Type:	Double Precision
Value Range:	[0, 1]
Description:	Fraction of group associated with third largest language
Sources:	IC

1.7.168 ic_phenotype1

Type:	String
Value Range:	NA
Description:	String indicating largest phenotype class for this group
Sources:	IC

1.7.169 ic_pheno1_size

Type:	Double Precision
Value Range:	[0, 1]
Description:	Fraction of group associated with largest phenotype class
Sources:	IC

1.7.170 ic_phenotype2

Type:	String
Value Range:	NA
Description:	String indicating second largest phenotype class for this group
Sources:	IC

1.7.171 ic_pheno2_size

Type:	Double Precision
Value Range:	[0, 1]
Description:	Fraction of group associated with second largest phenotype class
Sources:	IC

1.7.172 ic_phenotype3

Type:	String
Value Range:	NA
Description:	String indicating third largest phenotype class for this group
Sources:	IC

1.7.173 ic_pheno3_size

Type:	Double Precision
Value Range:	[0, 1]
Description:	Fraction of group associated with third largest phenotype class
Sources:	IC

Chapter 2

Country-Level Data

2.1 EPR Countries

The unit of observation in the RFE country-level data is the country-year. The sample universe of country-years in the RFE country-level data is adopted from the EPR (*Ethnic Power Relations*) dataset (Vogt et al. ming. Consequently, the country-years included in the RFE country-level dataset are equivalent to the country-years for which the EPR dataset codes ethnic groups, and thus for which the RFE group-level dataset provides information on the ethnic-group-level (see 1.1.1).

2.2 Country-Level Conflict Data

The information on intrastate conflicts in the RFE country-level data originates primarily from the UCDP ACD dataset (Gleditsch et al. 2002). Information on whether intrastate conflicts are coded as ethnic is added using the RFE group-level conflict data, which is compiled from the ACD2EPR dataset (Wucherpfennig et al. 2012) and the ACD dataset (Gleditsch et al. 2002 (see 1.3). The remainder of this section briefly describes how the ACD data is aggregated into country-year format and elaborates the coding rules for country-level conflict onset and incidence.

2.2.1 ACD Aggregation

The ACD dataset codes inter- and intrastate armed conflicts on a yearly basis. Conflicts in the ACD dataset follow the UCDP conflict definition (please see 3.2 for links to more information and references) and are assigned a unique *UCDP ID*. ACD conflicts are assigned to one of four categories: Extrasystemic, Interstate, Internal, and Internationalized Internal Conflicts. Of these types, only the latter two are relevant for the RFE country-level conflict coding, since we focus exclusively on civil wars. ACD conflicts are distinguished along the incompatibility they originate from, not their temporal dimension. Hence, a conflict that ends and reoccurs after a substantial period of time is assigned the same UCDP ID,

regardless of the time period separating the actual conflict episodes, as long as the incompatibility remains the same. ACD conflicts are merged with the EPR country-level data by assigning each country-year observation one or several UCDP IDs if the ACD dataset reports an ongoing conflict in the respective country – according to ACD's conflict location variable – and year.

2.2.2 Conflict Onset

In the RFE country-level data a conflict onset occurs if a country experiences an intrastate conflict in a given year, and the respective conflict (as identified via its UCDP ID) has been inactive in the given country in the previous two calendar years.

The RFE country-level data also offeres conflict onset variables that distinguish between ethnic and non-ethnic conflicts. Conflict onsets are coded as ethnic if, according to the RFE group-level conflict data (see 1.3), at least one ethnic group is linked to the respective ACD conflict *in the onset year*. This implies that the number of ethnic conflict onsets defined in the RFE country-level data may differ from the aggregated number of onsets in the RFE group-level data, even though the two datasets are based on the same sources.

2.2.3 KO and DO Options

The RFE country-level data conflict onset variables come in two variants: The KO and the DO option. KO stands for *Keep Ongoing* and is the default option. Conflict onset variables with the *KO* option take the value of 1 for country-years in which a country experiences conflict onset (see 2.2.2), and 0 in all other years.

DO stands for *Drop Ongoing* and represents a slightly more advanced coding. DO conflict onset variables indicate conflict onset in the same manner as the KO variables, but are censored if a given country has experienced any conflict in the previous two calendar years. The reasoning underlying this definition is that in these years, positive conflict onset codings are highly unlikely due to the rules established in section 2.2.2, and thus in many statistical applications these observations will be excluded from the analysis.

2.2.4 Conflict Incidence

The RFE country-level data includes a number of conflict incidence variables, which indicate whether in a given year a country is involved in at least one ACD conflict. Conflict incidence variables assume the value of 1 in years when at least one ACD conflict episode is ongoing, and 0 in all other years.

2.3 Country-Level Information on Ethnic Groups

The RFE country-level data includes a number of variables referring to the composition of a country's ethnic groups in terms of size and power status. This information is aggregated

directly from the RFE group-level data, which is based on the EPR 2014 dataset; for detailed information please see 1.

Furthermore, the EPR dataset includes a set of countries for which no ethnic group is coded as politically relevant during the entire period covered by the dataset. As elaborated above (see 1.1.2), for these countries a *placeholder-group* is coded which remains politically irrelevant throughout the entire period covered by the dataset. The variable *cntr_Relevance* indicates for which countries this is the case. For these countries, all variables referring to aggregate group sizes or power status are coded as missing.

2.4 Country-Level Variables

2.4.1 countries_gwid

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Unique GW ID of country
Sources:	GW

2.4.2 countryname

Type:	String
Value Range:	NA
Description:	Name of country
Sources:	GW

2.4.3 year

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Year of observation (Country-Level)
Sources:	NA

2.4.4 egip_groups_count

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Count variable indicating the number of EGIP groups in this country
Sources:	EPR

2.4.5 excl_groups_count

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Count variable indicating the number of MEG groups in this country
Sources:	EPR

2.4.6 regaut_groups_count

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Count variable indicating number of groups with regional autonomy in this country.
Sources:	EPR

2.4.7 regaut_excl_groups_count

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Count variable indicating number of MEG groups with regional autonomy in this country.
Sources:	EPR

2.4.8 regaut_egip_groups_count

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Count variable indicating number of EGIP groups with regional autonomy in this country.
Sources:	EPR

2.4.9 rlvt_groups_count

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Count variable indicating the number of relevant groups in this country
Sources:	EPR

2.4.10 actv_groups_count

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Count variable indicating the number of active groups in this country
Sources:	EPR

2.4.11 lpop

Type:	Double Precision
Value Range:	$\mathbb{R}_{\geq 0}$
Description:	Sum of the ethnically relevant population in this country (as a fraction of total population)
Sources:	EPR

2.4.12 egippop

Type:	Double Precision
Value Range:	[0, 1]
Description:	Sum of the population of all EGIP groups in this country (as a fraction of total population)
Sources:	EPR

2.4.13 legippop

Type:	Double Precision
Value Range:	[0, 1]
Description:	EGIP population as a fraction of ethnically relevant population in this country (egippop / lpop)
Sources:	EPR

2.4.14 exclpop

Type:	Double Precision
Value Range:	[0, 1]
Description:	Sum of the population of all MEG groups in this country (as a fraction of total population)
Sources:	EPR

2.4.15 lexclpop

Type:	Double Precision
Value Range:	[0, 1]
Description:	MEG population as a fraction of ethnically relevant population in this country (exclpo / lpop)
Sources:	EPR

2.4.16 discrimpop

Type:	Double Precision
Value Range:	[0, 1]
Description:	Sum of discriminated population in this country (as a fraction of total population).
Sources:	EPR

2.4.17 ldiscrimpop

Type:	Double Precision
Value Range:	[0, 1]
Description:	Sum of discriminated population as a fraction of ethnically relevant population in this country (discrimpop / lpop)
Sources:	EPR

2.4.18 maxexclpop

Type:	Double Precision
Value Range:	[0, 1]
Description:	Size of the largest MEG group in this country (as a fraction of total population)
Sources:	EPR

2.4.19 lmaxexclpop

Type:	Double Precision
Value Range:	[0, 1]
Description:	Size of the largest MEG group in this country as a fraction of ethnically relevant population (maxexclpop / lpop)
Sources:	EPR

2.4.20 regautpop

Type:	Double Precision
Value Range:	[0, 1]
Description:	Sum of population with regional autonomy in this country (as a fraction of total population).
Sources:	EPR

2.4.21 regautexclpop

Type:	Double Precision
Value Range:	[0, 1]
Description:	Sum of population with regional autonomy and excluded (MEG) in this country (as a fraction of total population)
Sources:	EPR

2.4.22 regautegippop

Type:	Double Precision
Value Range:	[0, 1]
Description:	Sum of population with regional autonomy and included (EGIP) in this country (as a fraction of total population)
Sources:	EPR

2.4.23 cntr_relevance

Type:	String
Value Range:	NA
Description:	R indicates countries where ethnicity is coded as being relevant at least once in the sample period; P indicates countries where only a placeholder group is coded
Sources:	EPR

2.4.24 nstar

Type:	Double Precision
Value Range:	[0, 1]
Description:	N*(0.5; 5); see Cederman L.-E. and L. Girardin (2007). Beyond fractionalization: Mapping ethnicity onto nationalist insurgencies. The American Political Science Review 101(1): pp. 173-185.
Sources:	EPR

2.4.25 onset_ko_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating conflict onset / ko option
Sources:	ACD

2.4.26 onset_ko_terr_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating territorial conflict onset / ko option
Sources:	ACD

2.4.27 onset_ko_gov_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating governmental conflict onset / ko option
Sources:	ACD

2.4.28 onset_ko_eth_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating ethnic conflict onset / ko option
Sources:	ACD, ACD2EPR

2.4.29 onset_ko_noneth_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating nonethnic conflict onset / ko option
Sources:	ACD, ACD2EPR

2.4.30 onset_ko_terr_eth_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating territorial ethnic conflict onset / ko option
Sources:	ACD, ACD2EPR

2.4.31 onset_ko_gov_eth_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating governmental ethnic conflict onset / ko option
Sources:	ACD, ACD2EPR

2.4.32 onset_ko_terr_noneth_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating territorial nonethnic conflict onset / ko option
Sources:	ACD, ACD2EPR

2.4.33 onset_ko_gov_noneth_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating governmental nonethnic conflict onset / ko option
Sources:	ACD, ACD2EPR

2.4.34 incidence_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating ongoing conflict
Sources:	ACD

2.4.35 incidence_terr_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating ongoing territorial conflict
Sources:	ACD

2.4.36 incidence_gov_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating ongoing governmental conflict
Sources:	ACD

2.4.37 incidence_eth_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating ongoing ethnic conflict
Sources:	ACD, ACD2EPR

2.4.38 incidence_noneth_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating ongoing non-ethnic conflict
Sources:	ACD, ACD2EPR

2.4.39 incidence_terr_eth_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating ongoing territorial ethnic conflict
Sources:	ACD, ACD2EPR

2.4.40 incidence_gov_eth_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating ongoing governmental ethnic conflict
Sources:	ACD, ACD2EPR

2.4.41 incidence_terr_noneth_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating ongoing territorial nonethnic conflict
Sources:	ACD, ACD2EPR

2.4.42 incidence_gov_noneth_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating ongoing governmental nonethnic conflict
Sources:	ACD, ACD2EPR

2.4.43 onset_do_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating conflict onset / ko option
Sources:	ACD

2.4.44 onset_do_terr_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating territorial conflict onset / do option
Sources:	ACD

2.4.45 onset_do_gov_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating governmental conflict onset / do option
Sources:	ACD

2.4.46 onset_do_eth_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating ethnic conflict onset / do option
Sources:	ACD, ACD2EPR

2.4.47 onset_do_noneth_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating nonethnic conflict onset / do option
Sources:	ACD, ACD2EPR

2.4.48 onset_do_terr_eth_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating territorial ethnic conflict onset / do option
Sources:	ACD, ACD2EPR

2.4.49 onset_do_gov_eth_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating governmental ethnic conflict onset / do option
Sources:	ACD, ACD2EPR

2.4.50 onset_do_terr_noneth_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating territorial nonethnic conflict onset / do option
Sources:	ACD, ACD2EPR

2.4.51 onset_do_gov_noneth_flag

Type:	Integer
Value Range:	[0; 1]
Description:	Binary flag indicating governmental nonethnic conflict onset / do option
Sources:	ACD, ACD2EPR

2.4.52 warhist

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Count variable indicating number of conflict onsets this country has previously experienced
Sources:	ACD

2.4.53 warhist_terr

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Count variable indicating number of territorial conflict onsets this country has previously experienced
Sources:	ACD

2.4.54 warhist_gov

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Count variable indicating number of governmental conflict onsets this country has previously experienced
Sources:	ACD

2.4.55 peaceyears

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Years since country first appears in dataset or since the end of the last ongoing conflict episode
Sources:	ACD

2.4.56 peaceyears_terr

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Years since country first appears in dataset or since the end of the last ongoing territorial conflict episode
Sources:	ACD

2.4.57 peaceyears_gov

Type:	Integer
Value Range:	$\mathbb{N}$
Description:	Years since country first appears in dataset or since the end of the last ongoing governmental conflict episode
Sources:	ACD

2.4.58 nightlight_total

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Stable nightlight emissions in country polygon; derived by Cshapes/DMSP-OLS overlay.
Sources:	DMSP, CSHAPES

2.4.59 elevsd

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Standard deviation of gridded elevation measurements (0.008330 decimal degree resolution) intersecting with country polygon (from Cshapes).
Sources:	GTOPO30, CSHAPES

2.4.60 meanelev

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Mean elevation of territory intersecting with group polygon.
Sources:	GTOPO30, CSHAPES

2.4.61 pop90_total

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Country's population in 1990; derived by Cshapes/CIESIN GRUMPv1 overlay.
Sources:	GRUMP, CSHAPES

2.4.62 pop00_total

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Country's population in 2000; derived by Cshapes/CIESIN GRUMPv1 overlay.
Sources:	GRUMP, CSHAPES

2.4.63 pop10_total

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Country's population in 2010; derived by Cshapes/CIESIN GRUMPv1 overlay.
Sources:	GRUMP, CSHAPES

2.4.64 gdp90_total

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Country's GDP in 1990; derived by Cshapes/G-Econ overlay.
Sources:	GECON, CSHAPES

2.4.65 gdp95_total

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Country's GDP in 1995; derived by Cshapes/G-Econ overlay.
Sources:	GECON, CSHAPES

2.4.66 gdp00_total

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Country's GDP in 2000; derived by Cshapes/G-Econ overlay.
Sources:	GECON, CSHAPES

2.4.67 gdp05_total

Type:	Double Precision
Value Range:	$\mathbb{R}_{>0}$
Description:	Country's GDP in 2005; derived by Cshapes/G-Econ overlay.
Sources:	GECON, CSHAPES

Chapter 3

Sources

3.1 Terms and Conditions

By downloading data offered via the GROW^{up} Research Front-End (henceforth referred to as RFE), I agree to the following:

1. I agree that any books, articles, conference papers, theses, dissertations, reports, or other publications that I create using data distributed via the RFE reference the bibliographic citation accompanying the data. These citations include the data authors, data identifier, and other relevant information.

2. I understand that the data distributed via the RFE is in part compiled from external data sources and agree that the use of RFE data compiled from external data sources implies acceptance with the terms and conditions associated with these sources.

3. The distributor makes no warranties, expressed or implied, by operation of law or otherwise, regarding or relating to the data distributed via the RFE.

3.2 Sources Release 2.0

3.2.1 Gleditsch and Ward Interstate System Membership

Abbreviation:	GW
Version:	v5
Reference:	Gleditsch, K. S. And M. D. Ward (1999). Interstate System Membership: A Revised List of the Independent States since 1816. International Interactions 25.
URL:	`http://privatewww.essex.ac.uk/~ksg/statelist.html`

3.2.2 Ethnic Power-Relations Dataset

Abbreviation:	EPR
Version:	v2014
Reference:	Vogt, M., N.-C. Bormann, S. Rüegger, L.-E. Cederman, P. Hunziker, and L. Girardin (Forthcoming). Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Dataset Family. Journal of Conflict Resolution.
URL:	http://www.icr.ethz.ch/data

3.2.3 ACD2EPR Docking Dataset

Abbreviation:	ACD2EPR
Version:	v2014
Reference:	Wucherpfennig, J., N. Metternich, L.-E. Cederman, and K. S. Gleditsch (2012). Ethnicity, the state and the duration of civil war. World Politics 64 (1).
URL:	http://www.icr.ethz.ch/data

3.2.4 UCDP Armed Conflict Dataset

Abbreviation:	ACD
Version:	v4-2014
Reference:	Gleditsch, N. P., P. Wallensteen, M. Eriksson, M. Sollenberg, and H. Strand (2002). Armed conflict 1946-2001: A new dataset. Journal of Peace Research 39(5).
URL:	http://www.pcr.uu.se/research/ucdp/datasets/

3.2.5 GeoEPR Dataset

Abbreviation:	GEOEPR
Version:	v2014
Reference:	Wucherpfennig, J., N. B.Weidmann, L. Girardin, L.-E. Cederman, and A. Wimmer (2011). Politically relevant ethnic groups across space and time: Introducing the GeoEPR dataset. Conflict Management and Peace Science 28(5).
URL:	http://www.icr.ethz.ch/data

3.2.6 Not Applicable

Abbreviation:	NA
Version:	NA
Reference:	NA
URL:	NA

3.2.7 CShapes Dataset

Abbreviation:	CSHAPES
Version:	v0.4-2
Reference:	Weidmann, N. B., D. Kuse, and K. S. Gleditsch (2010). The Geography of the International System: The CShapes Dataset. International Interactions 36 (1).
URL:	http://nils.weidmann.ws/projects/cshapes

3.2.8 DMSP-OLS Nighttime Lights Time Series (Average Visible, Stable Lights, & Cloud Free Coverages)

Abbreviation:	DMSP
Version:	v4
Reference:	Image and data processing by NOAA's National Geophysical Data Center. DMSP data collected by US Air Force Weather Agency.
URL:	http://ngdc.noaa.gov/eog/dmsp.html

3.2.9 Global 30 Arc-Second Elevation (GTOPO30)

Abbreviation:	GTOPO30
Version:	v30
Reference:	Distributed by NASA's Land Processes Distributed Active Archive Center (LP DAAC) (1996).
URL:	https://lta.cr.usgs.gov/GTOPO30

3.2.10 Global Rural Urban Mapping Project

Abbreviation:	GRUMP
Version:	v1
Reference:	Center for International Earth Science Information Network - CIESIN - Columbia University, International Food Policy Research Institute - IFPRI, The World Bank, and Centro Internacional de Agricultura Tropical CIAT (2011). Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Population Count Grid. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC).
URL:	http://sedac.ciesin.columbia.edu/

3.2.11 G-Econ

Abbreviation:	GECON
Version:	v4
Reference:	Nordhaus, W., Azam, Q., Corderi, D., Hood, K., Victor, N. M., Mohammed, M., and Weiss, J. (2006). The G-Econ database on gridded output: methods and data. Yale University, New Haven.
URL:	http://gecon.yale.edu

3.2.12 Transborder Ethnic Kin (TEK) Dataset

Abbreviation:	TEK
Version:	v2014
Reference:	Cederman, L.-E., K. S. Gleditsch, I. Salehyan, and J. Wucherpfennig (2013). Transborder Ethnic Kin and Civil War. International Organization 67 (2).
URL:	http://www.icr.ethz.ch/data

3.2.13 Ethnic Dimensions Dataset

Abbreviation:	IC
Version:	v2014
Reference:	Bormann, N.-C., L.-E. Cederman, and M. Vogt (2013). Language, Religion, and Civil War. Paper Presented at the Annual Meeting of the American Political Science Association in Chicago, Il, August 29th-September 1st, 2013.
URL:	http://www.icr.ethz.ch/data

Bibliography

Cunningham, D. E., K. Skrede Gleditsch, and I. Salehyan (2009). It takes two. *Journal of Conflict Resolution 53*(4), pp. 570–597.

Gleditsch, N. P., P. Wallensteen, M. Eriksson, M. Sollenberg, and H. Strand (2002). Armed conflict 1946-2001: A new dataset. *Journal of Peace Research 39*(5), pp. 615–637.

Uppsala Conflict Data Program (2014). UCDP Actor Dataset 2.2-2014. *Uppsala University*.

Vogt, M., N.-C. Bormann, S. Rüegger, L.-E. Cederman, P. Hunziker, and L. Girardin (Forthcoming). Integrating data on ethnicity, geography, and conflict: The ethnic power relations dataset family. *Journal of Conflict Resolution*.

Wucherpfennig, J., N. Metternich, L.-E. Cederman, and K. S. Gleditsch (2012). Ethnicity, the state and the duration of civil war. *Wolrd Politics 64*(1), pp. 79–115.