

The EPR-Organizations (EPR-O) Dataset: Documentation and Codebook

Content

1. Citation
2. Introduction
3. Definitions
4. Relevant organizations
5. Ethnic group links
6. Variables
7. Time variation in variables
8. References

Acknowledgements

The EPR-Organizations dataset was developed with funding from the Swiss National Science Foundation (SNF) within the “Research for Development” (r4d) project on “Ethnic Power Relations and Conflict in Fragile States” (SNSF, grant no. 400240-147210). The original coding instructions were prepared together with Kristian Skrede Gleditsch, during a short-term scientific mission at the University of Essex in January 2015, funded by the European Network for Conflict Research.

I would like to thank Kathleen Cunningham for sharing her data on self-determination organizations, as well as Mohammad Aghdam, Baris Ari, Felicitas Fischer, Andreas Juon, Nora Keller, Fabian Morgenthaler, Radhika Nagesh, Kathrin Oestringer, Francisco Villamil, and Katherine Woolbright for excellent research assistance, and Luc Girardin for managing the coding interface and database.

1. Citation

When using the EPR-Organizations dataset in your research, please cite the following publication:

Vogt, Manuel, Kristian Skrede Gleditsch, and Lars-Erik Cederman. Forthcoming. “From Claims to Violence: Signaling, Outbidding, and Escalation in Ethnic Conflict.” *Journal of Conflict Resolution*.

2. Introduction

The EPR-Organizations (EPR-O) dataset is a database on ethnically based political organizations in a globally representative sample of 40 countries. Based on the Ethnic Power Relations (EPR) dataset (Cederman, Wimmer and Min 2010; Vogt et al. 2015), it identifies political organizations representing the interests of specific ethnic groups listed in EPR at the national level. These include political parties and civil society organizations, as well as self-determination organizations. The stratified random sample was drawn from the universe of countries in the EPR dataset. After being assigned random numbers, countries were split into two groups: those that have at some point experienced ethnic civil conflict and those that have not. From each group, the twenty countries with the highest numbers were selected. Table 1 lists all countries currently included in EPR-O.

Country coders then inspected a predefined “universe” of relevant political organizations, identified those that represent(ed) the interests of specific EPR groups, and coded a number of variables for each ethnic organization-group link, drawing on both primary and secondary sources. Most countries were evaluated by three different coders across various coding phases. EPR-O includes five general types of variables:

- the ethnic groups represented;
- the type of the organization-group link (that is, the way ethnicity is invoked by the organization);
- foundation and, if applicable, dissolution year;
- organization agenda/claims;
- and means of group representation.

Table 1: EPR-Organizations sample

Ethnic civil conflict countries	Countries without ethnic civil conflicts
Angola	Algeria
Azerbaijan	Australia
Bangladesh	Belgium
Bosnia and Herzegovina	Botswana
Burundi	Brazil
China	El Salvador
Iraq	Guinea-Bissau
Israel	Lithuania
Macedonia	Madagascar
Myanmar	Malawi
Pakistan	Malaysia
Russia	Mongolia
South Sudan	Mozambique
Spain	Namibia
Sri Lanka	Paraguay
Tajikistan	Peru
Trinidad and Tobago	Serbia (2006-)
Turkey	Taiwan
Yemen	Tanzania
Zimbabwe	Turkmenistan

This document details the structure and content of the dataset, including definitions, the types of organizations covered, the coding of ethnic group links, and a description of all variables.

3. Definitions

In line with the EPR dataset, EPR-O defines **ethnicity** as “any subjectively experienced sense of commonality based on the belief in common ancestry and shared culture”, indicated through markers such as “common language, similar phenotypical features, adherence to the same faith, and so on” (Cederman, Wimmer and Min 2010: 98-9). Hence, this definition includes ethnolinguistic, ethnosomatic (or “racial”), and ethnoreligious groups, but *not* tribes and clans (as, for example, the different clans in Somalia), which conceive of ancestry in genealogical terms, nor regions, which do not define commonality on the basis of shared ancestry.

Organizations are defined as any named non-state entities that recruit members and make political claims. Non-state entities exclude internationally recognized states, administrative

units of such states, and entities that claim to be a state.¹ Within-organization groups, such as factions, committees, or party bureaus, were also excluded.

According to their diverse purposes, political organizations can invoke ethnicity through different mechanisms. For example, the ethnic identity of a political party will often manifest itself through its electoral support base, whereas an ethnic mass movement may make explicit ethno-political claims vis-à-vis the state government. Moreover, the way in which ethnicity is politically expressed also depends on the specific local context. For instance, organizations making explicit ethnic claims are constitutionally banned almost everywhere in Africa (Moroff 2010), whereas this is not the case in other world regions. Acknowledging these different contexts, EPR-O defines an **ethnic organization** as:

an organization that represents the interests of one or more specific ethnic group(s), in opposition to other ethnic groups in the country, in one of the following ways: a) through explicit ethnic claims; b) through recruitment along ethnic lines; or c) through electoral support along ethnic group lines.

This definition implies that:

- 1) A given organization can be linked to more than one ethnic group. The crucial condition, however, is that interest representation occurs in opposition to *other* ethnic groups in the country.
- 2) There are three possible links between an ethnic organization and the group it represents:
 - a. Ethnic claims: public claims are made by the organization on behalf of (a) specific ethnic group(s). An ethnic claim is defined here as *consistent public demands in favor of the rights, benefits, or well-being of (a) specific ethnic group(s)*. This may include claims for self-determination, more influence for the group(s) in the national government, autonomy, language and other cultural rights, etc. An *ethnic denominator* in the name of an organization is considered proof of a public ethnic claim.
 - b. Membership recruitment: the organization recruits its members based on their ethnic identity, that is, members overwhelmingly join, or are admitted to, the organization *because they are from (a) specific ethnic group(s)*. An *ethnic denominator* in the name

¹ The only exception to this definition in the dataset is the “Republic of Nagorno-Karabakh” in Azerbaijan. Because no individual organization forming part of this self-proclaimed, but unrecognized state could be identified in the predefined universe of organizations, the “Republic of Nagorno-Karabakh” was included in EPR-O as a “placeholder” organization.

of an organization is considered proof of a strategy of ethnically based recruitment on the part of the organization.

- c. Electoral support: the electoral support (defined here as votes) of an organization/party is ethnically based, that is, voters overwhelmingly choose this party over other parties *because* they are from (a) specific ethnic group(s).

Following Horowitz (1985: 296-301), EPR-O focuses on the distribution of organization support/recruitment, rather than group support/recruitment. This means that the relevant criterion to determine b) and c) is the degree to which a given organization is composed of/supported by (a) specific ethnic group(s), rather than the degree/share of the ethnic group joining/supporting the organization.

The term “because” is included in both b) and c) to ensure that organization support/recruitment is not accidental (for example, a mere function of demographic conditions), but based on a conscious decision on the part of the individuals supporting/joining the organization. This means that ethnicity has to be the driving force for individuals to join, or vote for, a given organization.²

4. Relevant organizations

In the data collection process, EPR-O coders screened a predefined set of relevant political organizations to decide whether they could be associated with one or more ethnic groups, according to the above definition. This universe of relevant organizations was composed of entries in the following four sources:

- election archives of all national-level (parliamentary or presidential) elections in the included countries between 1946 and 2013;
- the Political Handbooks of the World (for the years 1984, 1989, 1991, 1995, 1998, 2001, 2005, and 2013) (Banks et al. 2014);
- the World Directory of Minorities and Indigenous Peoples (Minority Rights Group International 2007);

² For example, the fact that most members of the April 6 Youth Movement in Egypt happened to be Muslims would not make it an ethnic organization in the EPR-O dataset, as Muslims make up the vast majority of Egypt’s population and most of its members simply joined the organization to overthrow President Mubarak (a fellow Muslim) and put an end to authoritarian rule. In contrast, the Egyptian Muslim Brotherhood would meet the definition of an ethnic organization in EPR-O as its leaders and members follow a clear religious political agenda that excludes the country’s Christian minority.

- the Cunningham (2013) self-determination movement (SDM) dataset.

The universe of relevant organizations was limited to nationally active organizations. This criterion refers to the target of mobilization/political activism, rather than organizations' structural setup. Hence, an organization does not need to have a national institutional reach (e.g., branches in all regions of the country) to be included in the dataset; it may be locally based as long as it makes political demands on the national state or has a direct influence on national-level politics (e.g., a locally rooted organization mobilizing against nationally defined mining policies or a regionally confined rebel organization that fights national government troops). By contrast, local/regional organizations targeting local/regional government entities over local/regional concerns are not included in EPR-O.

Political parties were only included if they obtained at least 1% of the national vote in at least one parliamentary or presidential election *or* won at least one seat in parliament in at least one election. In cases of party coalitions, the “component” parties of the coalition were coded as separate political parties if they met either of the following two criteria: i) they participated as a self-standing party in at least one national (parliamentary or presidential) election, obtaining at least 1% of the national vote or winning any number of seats in the parliament (even if this occurred through reserved seats); or ii) they participated in national elections in different coalitions over time. If neither of these two criteria was met, only the coalition itself was included in the universe of relevant organizations (given that it obtained at least 1% of the national vote or won any number of seats in the parliament in at least one election).

With respect to non-electoral organizations, EPR-O distinguishes between two or more formally or informally associated organizations (such as Batasuna and ETA in Spain) based on the leadership structure. Two organizations were considered separate organizations if they were/are headed by different people. Regional divisions (or branches) of national organizations were not included as separate organizations even if they were/are connected to specific regions (inhabited by specific ethnic groups) of a country.³ The exceptions to this rule are independently established regional organizations that are politically active at the national level and form part of a national alliance of civil society organizations while maintaining their own leaderships.⁴

³ For example, the Socialist Party of the Basque Country in Spain, which was founded and acts as the regional affiliate of the national Spanish Socialist Workers' Party (PSOE), would *not* be considered a separate (ethnic) organization in EPR-O.

⁴ For example, in Ecuador, the regionally based indigenous organizations ECUARUNARI and CONFENIAE, which were established in 1972 and 1980, respectively, and which became part of the national umbrella

Finally, extra-systemic ethnic organizations – that is, ethnic organizations whose organizational bases lie outside the territory of the country of the ethnic group they represent – were only included if i) they are country-specific, that is, if they were established to represent the interests of (a) specific ethnic group(s) in that particular country only; ii) their political activism takes place at least partly within the territory of that country; *and* iii) their operations affect the national-level politics of the country. Hence, EPR-O does not include transnational organizations representing ethnic minority rights more broadly (such as local chapters of Human Rights Watch, supra-national indigenous organizations, etc.) because they are not necessarily an expression of ethnic mobilization in that country.

5. Ethnic group links

Organization-group linkages in EPR-O follow a many-to-many principle: each organization could be linked to more than one ethnic group, while an ethnic group may be represented by a single or several different organizations. Among the universe of relevant organizations defined in section 4, coders identified ethnically based organizations by coding each of the three “ethnic link variables” (ethnic claim, ethnic recruitment, and ethnic electoral support) according to the definitions provided in section 3, and linking the organizations to the corresponding EPR groups. To this end, coders consulted a large number of sources, including original documents and websites of organizations, as well as scholarly texts, to decide whether a given organization could be associated with one or more ethnic groups. An organization was only linked to an ethnic group if at least one trustworthy source provided convincing evidence. This generated one or more organization-group links, depending on the number of ethnic groups matched with a given organization.

6. Variables

In a first step, all variables were coded in a time-invariant fashion. This means that a variable was coded as true if it applied at any point in an organization’s lifetime. Time variation was subsequently introduced as described in section 7. The final dataset is structured in a panel

organization CONAIE in 1986 while continuing to exist under their own leaderships and to advance their own ethno-political demands, would be counted as separate ethnic organizations in EPR-O.

(organization-group year) format. All variables were coded based on the same types of sources as specified in section 5.

Organization-specific variables:

Country (string): Name of the country. From Gleditsch and Ward (1999).

Gwid (numeric): Unique country identifier. From Gleditsch and Ward (1999).

Orgid (numeric): Unique organization identifier. Since the same organization can be linked to multiple ethnic groups, duplicate orgid-year observations exist in the dataset.

Orgname (string): Complete name of the organization. Taken from the sources specified in section 4.

Variables specific to organization-group link:

Linkid: Unique organization-group link identifier. The linkid and year variables together uniquely identify observations in the dataset.

EPR group (string): Name of the ethnic group that the organization represents, as listed in the EPR dataset.

EPR groupid (numeric): Ethnic group ID used in EPR.

Year (numeric): Calendar year for which the variable codings apply.

Ethnic link variables: In what ways is the organization linked to the ethnic group? More than one variable may apply for each organization-group link.

Ethnic claim (dummy variable): Does the organization make explicit ethno-political claims for the group? Yes (1) or no (0).

Ethnic electoral support (dummy variable): If the organization participates in national electoral politics, do the party's voters overwhelmingly make their vote choice based on their ethnic identity? Yes (1) or no (0).

Ethnic recruitment (dummy variable): Do members overwhelmingly join, or are admitted to, the organization based on their ethnic identity? Yes (1) or no (0).

Founding year (numeric): Year in which the organization was founded. This may be before 1946. In cases, in which an organization emerged from another, already existing organization (for instance, a political party emerging from a social movement organization), we code the founding year of the newly emerging organization *if* it can be considered a *separate* political

organization, according to the criteria outlined in section 4. This rule also applies to organizations that were dissolved at one point and founded again later under the same name.

Dissolution year (numeric, if applicable): Year in which the organization was dissolved. Coded missing if organization still exists.

Organization agenda/claims variables: What goals has the organization pursued, or what has been its political agenda, with respect to a given ethnic group? Coding of this variable may be based either on explicit organization claims or on more implicitly signaled aims. More than one variable may be true for each organization-group link.

Governmental power (dummy variable): Has the organization attempted to increase the ethnic group's power or influence in the national-level government? Yes (1) or no (0).

Regional autonomy (dummy variable): Has the organization advocated regional autonomy for the ethnic group? Yes (1) or no (0).

Separatism/irredentism (dummy variable): Has the organization promoted the ethnic group's secession from the home country and/or the unification with an ethnic/national group living in another state? Yes (1) or no (0).

Economic redistribution (dummy variable): Has the organization aimed for a greater share of economic resources for the ethnic group that are distributed by or stem from the national government? Yes (1) or no (0).

Religious rights (dummy variable): Has the organization aimed at increasing the religious rights of the ethnic group (e.g., official recognition of the group's religion)? Yes (1) or no (0).

Linguistic rights (dummy variable): Has the organization aimed at increasing the linguistic rights of the ethnic group (e.g., making the group's language official language of instruction in schools)? Yes (1) or no (0).

Other goals (string): Other political goals not covered by the aforementioned dummy variables, including non-ethnic goals (e.g., ideological goals, such as the creation of a socialist state).

Means of group representation: What means has the organization used to advance the interests of the ethnic group? More than one variable may apply for each organization-group link.

Electoral participation (dummy variable): Has the organization participated in national (parliamentary or presidential) elections? Yes (1) or no (0).

Non-electoral, non-violent strategies (dummy variable): Has the organization used non-electoral, but non-violent means of ethnic group representation, such as lobbying, filing petitions, demonstrations, strikes, boycotts, etc.? Yes (1) or no (0).

Violence against government (dummy variable): Has the organization intentionally used violence against the state or state agents? Violence is defined here as actions intentionally leading to the loss of life or consciously accepting the possibility of the loss of life. Yes (1) or no (0).

Violence against civilians (dummy variable): Has the organization intentionally employed violence against civilians? Civilians are defined here as individuals that are not representing the state or any armed organization. This variable was only coded as true if intention could be assumed. Harm afflicted on civilians as an unintended side effect of violence against the government was not coded as violence against civilians. Yes (1) or no (0).

Other mobilization strategies (string): Other means of group representation not covered by the aforementioned dummy variables.

Sources (string): List of sources that coders relied on in their coding.

Comments (string): Detailed explanations of ambiguous coding decisions.

Confidence (ordinal): Coders' confidence in their assessment of a given organization-group link, using a scale of ordinal categories ranging from "certain" (1) to "not at all confident" (5).

7. Time variation in variables

EPR-O is structured in a panel (organization-group year) format. This final section of the codebook details how variation in variable codings over time was determined.

Time variation in EPR-O features a nested structure, consisting of two levels: the ethnic group level and the variables level, with the latter nested in the former. The ethnic group level refers to the "EPR groupid" variable (see section 6) and specifies the time period during which a given organization-group link is valid, that is, an organization is linked to a particular ethnic group. This "macro" time period defines the limits within which the other variables may be

true. Put differently, an organization's goals and strategies identified in the time-invariant coding may apply to the entire macro time period or to sub-periods *within* (but *not* outside) this macro period. Such sub-periods are indicated by time period changes at the variables level. Importantly, *time period breaks for different variables need not correspond to each other*, that is, they may coincide or overlap with, or completely differ from, each other.

At the ethnic group level, time variation was introduced in either of three scenarios:

- An organization emerged as a non-ethnic political entity and became “ethnicized” (i.e. started to represent a particular ethnic group) at a later point in time.
- An organization was formed as an ethnically based organization, but became “de-ethnicized” (i.e. aborted/lost its ethno-political base/agenda) at some point in time.
- An organization represented distinct ethnic groups in different periods of time. This may be the case, for example, if an organization's political base/agenda is closely tied to its leader. If the ethnic identity of the leader changes, so might the organization's ethnic identity.

Within the ethnic group-level macro period, time variation for other variables was introduced in case of clear “turning points” in the goals or strategies of an organization (i.e. organizations either taking up or abandoning certain goals or strategies) – for example, a new organization manifesto that calls for the ethnic group's secession from its current state, rather than just regional autonomy within that state, or a militant organization signing a peace agreement with the government and henceforth abstaining from violence.

With respect to organizations' use of violence against the government or civilians, which often occurs intermittently, that is, in a non-constant way, EPR-O distinguishes between two general scenarios. If violence is the main reason for an organization's existence, that is, if an organization was *established as a means to carry out acts of violence* (e.g., in the case of a terrorist organization), the corresponding violence variable(s) was/were coded as true for the entire macro period. By contrast, if an organization at times engaged in acts of violence, but violence was/is not its main *raison d'être*, sub-periods were coded during which the corresponding violence variable(s) applied: ideally, all individual years in which the organization carried out acts of violence, or, otherwise, the time span between the year of the organization's first act of violence against the government and/or civilians and the year of the last such act (if applicable).

With respect to the electoral participation variable, sub-periods correspond to one or more consecutive electoral cycles. The start year of such a sub-period is defined as the year, in which the (first) election took place; the end year as the last year of the corresponding legislature/presidential term (whatever occurs later).⁵ Thus, a given sub-period may be limited to a single electoral cycle of, for example, four years or could cover, for instance, four cycles of five years each, totaling twenty years. In the case of interruptions in an organization's electoral participation (caused either by the organization itself or the political circumstances), various sub-periods were coded. If electoral participation started in the same year as the macro time period and lasted uninterrupted until its end, no sub-periods were coded.

As in the case of the time-invariant coding, time period changes were coded for each organization-group link separately. Coders relied on the same types of sources as specified in section 5, and time variation was only coded if at least one trustworthy source provided convincing evidence for changes in a given variable.

8. References

- Banks, Arthur S., Muller, Thomas C., Overstreet, William R., and Isacoff, Judith F. Eds.. 2014. *Political Handbook of the World*. Thousand Oaks, CA: CQ.
- Cederman, Lars-Erik, Andreas Wimmer and Brian Min. 2010. "Why Do Ethnic Groups Rebel? New Data and Analysis." *World Politics* 62(1):87-119.
- Cunningham, Kathleen Gallagher. 2013. "Actor Fragmentation and Civil War Bargaining: How Internal Divisions Generate Civil Conflict." *American Journal of Political Science* 57(3): 659–672.
- Gleditsch, Kristian Skrede and Michael D. Ward. 1999. "Interstate System Membership: A Revised List of the Independent States since 1816." *International Interactions* 25(4):393-413.
- Horowitz, Donald L. 1985. *Ethnic Groups in Conflict*. Berkeley, CA: University of California Press.
- Minority Rights Group International. 2007. World Directory of Minorities and Indigenous Peoples. Retrieved April 10, 2016, from <http://minorityrights.org/directory/>
- Moroff, Anika. 2010. "Party Bans in Africa – an Empirical Overview." *Democratization* 17(4):618-41. doi: 10.1080/13510347.2010.491184.

⁵ In the case of *coup d'états* or other irregular endings of a legislative period, the year of the dissolution of the legislative body constitutes the end year.

Vogt, Manuel, Nils-Christian Bormann, Seraina Rügger, Lars-Erik Cederman, Philipp Hunziker and Luc Girardin. 2015. "Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Data Set Family." *Journal of Conflict Resolution* 59(7):1327-42.