Self-determination claims by EPR groups (EPRclaims)
Coding Instructions

Contact: Andreas Schädel, schaedel@icr.gess.ethz.ch

Project description

This project identifies ethnic groups in the EPR-ETH dataset (Vogt et al., 2015), which make claims for increased self-determination, covering the years 1946 through 2017. This project builds directly on the Self-Determination Movements (SDM) Dataset by Sambanis, Germann and Schädel (2018), which identifies 464 groups around the world that have made claims for increased self-determination between 1945 and 2012. This document provides coding instructions and lines out how information is mapped from the SDM dataset to the EPR dataset.

Coding of separatist claim

The claim variable, which indicates the type of a group's self-determination claim in a given year, is very closely based on the claim variable as coded by the SDM dataset (Sambanis et al. forthcoming). The SDM dataset distinguishes four types of claims: autonomy, independence, irredentism, and sub-state secession. Of these four claims, we only code the first three (see Table 1 below for a description of the three claim). Claims for sub-state secession, which the SDM dataset describes as a claim for the separation from an already autonomous entity to establish a separate autonomous unit or join another autonomous unit, are coded as claims for autonomy.

Table 1: Type of claims

Type of claim	claim	Description
Autonomy	1	Claim for establishment of territorial autonomy within an existing state or for an increase in the number of competencies of an already existing autonomous entity.
Independence	2	Claim for internationally recognized independence.
Irredentism	3	Claim for secession from the host state and subsequent merger with another state, typically the cultural motherland.

A self-determination group can be associated with multiple claims that vary in intensity. In such cases, we follow the coding rules of the SDM dataset and code the dominant claim. The SDM codebook mentions the example of the Corsican self-determination movement, which makes claims for both increased autonomy and independence. Since advocates of independence have always been in a minority position in Corsica, an autonomy claim is coded throughout. If the evidence is insufficient to arbitrate what is the dominant claim in a given

year or if multiple claims are being made by factions that have about equal strength, the most radical claim is coded (with autonomy < independence < irredentism).

Note: The claim variable relates to the situation on January 1, except if a country was not independent on January 1, if an ethnic group was not part of the given country on January 1, or in the first year of a movement's existence (first of January rule). In the latter cases it relates to the date of independence, the date the territory was merged with the state in question, or the date of the first organized self-determination claim, respectively.

Mapping SDM on EPR: SDM2EPR

To identify EPR groups that make claims for increased self-determination, we rely on the Self-Determination Dataset (SDM) by Sambanis, Germann and Schädel (forthcoming). The SDM dataset identifies 464 groups around the world that have made claims for increased self-determination between 1945 and 2012[1]. In order to map the information from the SDM dataset to EPR groups, Germann and Sambanis (2017) have developed SDM2EPR. SDM2EPR indicates for each group in the SDM dataset if there is an equivalent group in EPR. Mapping SDM to EPR allows us to identify those EPR groups that make self-determination claims and for each of these groups provides the start and end year of separatist activity. The exact documentation and coding rules for SDM2EPR are not publicly available, but the three mapping scenarios can be summarized as follows:

- **1:1 cases**: In many cases, linking SDM groups to EPR groups is straightforward as the SDM groups have a perfect or near-perfect equivalent in EPR (227 groups).
- **1:n/n:1 cases**: In an additional 61 cases, SDM and EPR groups have different levels of aggregation, but a match can still be established. SDM2EPR hereby distinguishes between n:1 and 1:n cases. In the former (n:1), EPR codes an umbrella group while the SDM dataset codes one or several smaller groups that form part of EPR's larger umbrella group. An example are the several indigenous groups in the United States coded by SDM (e.g. Cherokee, Iroquois, Pueblos), which are all matched with the EPR umbrella group 'American Indians'. In such cases, the EPR group is coded as separatist if the constituent SDM group(s) make(s) up more than 50% of the total population of the EPR umbrella group.[2] The start and end dates of separatist activity are determined by whatever is the first and the last year of activity coded for any of the constituent groups. In the latter (1:n), the SDM dataset has a higher level of aggregation than EPR. An example are the two EPR groups 'Northwestern Anglophones' and 'Southwestern Anglophones' in Cameroon, which the SDM dataset subsumes in one group ('Westerners'). In such cases, all constituent EPR groups are coded as separatist, but the start and end dates of separatist activity are determined individually based on case evidence.
- **No match**: Finally, 176 SDM groups have no equivalent in EPR, due to systematic reasons or due to coding practice. Their claims are not reflected in EPR.

For a significant share of the SDM groups, the type of claim is already provided by the SDM dataset. SDM2EPR hence also allows us to map the claims coding to the EPR groups. In 1:1 and 1:n cases, the claim coded by the SDM dataset is directly adopted. For n:1 cases where

---

[1] The SDM dataset codes a self-determination movement if at least one significant political organization makes a claim for increased self-determination within the nation-state context (see SDM codebook for additional information).
[2] Exceptions to the 50% rule are made for some indigenous groups because the SDM dataset, which seeks to disaggregate whenever possible, is susceptible to miss small indigenous communities.

several smaller SDM groups form part of a larger EPR umbrella group, the dominant claim is identified individually based on the claims and the population size of the constituent SDM groups. If an SDM group is coded to claim sub-state secession, this coding is changed into a claim for autonomy, as the claim for sub-state secession does not exist in EPR (see Table 1). For all SDM groups that are matched to an EPR group but for which the SDM dataset does not provide a claims coding, we code the claim ourselves.

Updating 2013-2017

Since the publication of the SDM dataset, the EPR dataset has been updated to 2017. Since the SDM dataset only codes self-determination claims until 2012, we code the period from 2013 until 2017 ourselves. This includes the identification of new movements that emerged between 2013 and 2017, the identification of movements that had been inactive before and reemerged in that period, or the identification of movements that terminated activity between 2013 and 2017. Furthermore, for all SDM groups that have a match in EPR and for which the SDM dataset provides information on the type of claims, we update this information for the period 2013-2017.

In some cases, the update of the EPR dataset led to revisions of previous codings or the SDM2EPR mapping. Furthermore, the EPR update also led to the inclusion of new countries, which the previous EPR coding had considered as too small (population <500,000). For SDM groups in these countries, the mapping could not rely on SDM2EPR but had to be done case-by-case. The following EPR groups were identified as being related to an SDM group:

country	SDM group	EPR group	SDM2EPR scenario
Belize	Mayans	Maya	1:1
Comoros	Anjouanese	Nzwani Comorans	1:1
Equatorial Guinea	Bubis	Bubi	1:1
Suriname	Indigenous Peoples	Indigenous Peoples (Arawaks and Caribs)	1:1

Documentation

For each EPR group that can be linked to one or more SDM groups, a documentation is provided. The documentation justifies the coding decision, lists coding sources and the following additional information:

- *Relevance period (EPR)*: Indicates the years in which EPR codes the group politically relevant.
- *SDM2EPR scenario*: Indicates the mapping scenario (1:1, n:1, 1:n) that matches the SDM group to the EPR group.
- *SDM group(s)*: Lists the group(s) that SDM2EPR matches with the EPR group.
- *SDM activity*: Indicates the years in which the corresponding SDM group makes self-determination claims.
- *SDM claim*: Indicates the type of claim as coded by the SDM dataset.

3

References

Sambanis, Nicholas, Micha Germann and Andreas Schädel (2018). "SDM: A new dataset on self-determination movements with an application to the reputational theory of conflict." *Journal of Conflict Research* 62(3): 656-686.

Vogt, Manuel, Nils-Christian Bormann, Seraina Rüegger, Lars-Erik Cederman, Philipp Hunziker and Luc Girardin (2015). "Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Dataset Family." *Journal of Conflict Resolution* 59(7):1327-1342.