

Self-Determination Movements (SDM) Dataset

Codebook

Nicholas Sambanis

Department of Political Science
University of Pennsylvania

Micha Germann

Department of Political Science
University of Pennsylvania

Andreas Schädel

Center for Comparative and International Studies
ETH Zurich

May 2017

Table of Contents

Introduction.....	5
1 Cross-Country Data on SDMs (1945-2012)	5
1.1. Coding the Global Incidence of SDMs.....	5
1.1.1. Definition of an SDM.....	5
1.1.2. Identification of SDMs in the dataset	7
1.1.3. Sources	8
1.2. Separatist Violence	10
2 Random Sample Data	11
2.1. Sampling Procedure.....	11
2.2. Sources.....	11
2.3. EPR2SDM.....	12
3 Variable Descriptions.....	13
country	13
ccode.....	13
group.....	13
groupid.....	13
year	13
region	13
rcode	13
active.....	14
phase1	14
phase2	14
fyphase.....	14
fyphase1.....	14
fyphase2.....	15
fygroup.....	15
startdate1	15
enddate1	15
tyrule1	15
startdate2.....	15
enddate2.....	15
tyrule2.....	15
perstart	15
perend	15
violsd	16
lviolsd	16
hviolsd	16
ambiguous.....	16
nviolsd	16
prioract.....	16
priorconflict	16
priornonviol	16
priorviol	17
violsd_onset.....	17
hviolsd_onset.....	17
shortpeace_violsd	17
shortpeace_hviolsd	17
py_violsd_phase	17
py_violsd_country	18

py_hviolsd_phase	18
py_hviolsd_country	18
violsd_onset_woambig	18
hviolsd_onset_woambig	18
shortpeace_violsd_woambig	19
shortpeace_hviolsd_woambig	19
py_violsd_woambig_phase	19
py_violsd_woambig_country	19
py_hviolsd_woambig_phase	19
py_hviolsd_woambig_country	19
doublecount	19
doublecount_oldhost	19
doublecount_newhost	20
sample	20
weight	20
below500000	20
violent	20
switched	20
con	20
cultcon	21
autcon	21
indcon	21
res	21
cultres	21
autres	21
indres	21
priorcon1	23
priorres1	23
priorcon2	24
priorres2	24
majterr1	24
majterr2	24
majterr3	24
claim	25
landborder	26
seashore	26
hydrocarbon	27
groupcon	27
kin	27
regaut	27
defacto	28
pwrstat	28
excluded	29
groupsize	29
epr2sdm_scenario	30
epr_group	30
epr_gwgroupid	30
sovdec	30
autdec	30
subdec	30
inddec	31
irrdec	31
ucdp_conflictid	31
ucdp_territoryname	31
ucdp_intensitylevel	31
ucdp_cumulativeintensity	31
ucdp_typeofconflict	31
ucdp_terrconflict	31
ucdp_hterrconflict	32
ucdp_terrconflict_onset	32
ucdp_hterrconflict_onset	32
shortpeace_ucdp_terrconflict	32

shortpeace_ucdp_hterrconflict.....	32
py_ucdp_terrconflict_phase.....	32
py_ucdp_terrconflict_country.....	32
py_ucdp_hterrconflict_phase.....	33
py_ucdp_hterrconflict_country.....	33
accomm.....	33
strictaccomm.....	33
References.....	34

Introduction

This document describes the coding rules and sources used for constructing the Self-Determination Movements (SDM) dataset. The SDM dataset identifies a total of 464 self-determination movements (SDMs) around the world from 1945 to 2012. The dataset is divided in two parts. The first part pinpoints the years in which these movements were active and distinguishes between periods of violent and non-violent conflict over self-rule. The second part provides more detailed information on state-movement interactions and several attributes of SDM groups for a random sample of 106 movements.

The SDM data comes with extensive coding documentation. These are available in two separate files. The “Coding Notes I” file includes justifications for the global cross-country data on SDMs. The “Coding Notes II” file includes justifications for the variables coded only for the random sample of movements.

The first section of the codebook gives an overview of our approach to compiling the global data on violent and non-violent SDMs. The second section introduces the random sample data. Finally, the third section describes all variables included in the SDM dataset.

1 Cross-Country Data on SDMs (1945-2012)

1.1. Coding the Global Incidence of SDMs

1.1.1. Definition of an SDM

We define SDMs as constituted by one or more political organizations that are connected to an ethnic group and make claims for increased self-determination (SD) from the state.

More specifically, we understand SD as a group-defined right to self-rule within the boundaries of a territory. SD as defined here includes the authority to raise and spend tax revenues; the ability to maintain a defense or police force; autonomy over policies governing cultural or linguistic practices within a given territory; the power to define the territorial boundaries of the state within which a group exercises its right to SD; some forms of indigenous land rights; and, of course, national independence.

Given our broad conception of SD, we include both movements claiming secession (national independence or the merger with a different state) and movements demanding increased internal autonomy short of independence. However, we require that SD claims cross a minimal level of political significance. Claims for increased internal autonomy must imply a significant re-definition of a state's institutional set-up. This is especially relevant for federal states, which are characterized by constant negotiations between the center and its regions over policy centralization. For example, because these do not imply a significant institutional re-definition, we do not include claims by US states that gay or abortion rights should be decided at the state level. An example of an autonomy claim in a federal state that we do include is the Jurassians in Switzerland, who mobilized for their own separate canton. Another is the Yakuts, a group in northeastern Siberia that has claimed increased control over their natural resources, among other things.

SDMs as defined here must be directed against the state. Thus, we do not include the BREXIT movement because it was directed against a supranational entity. However, while SDMs must be directed against the state, they need not be primarily directed against the national government. The claims of some SDMs are primarily targeted at regional governments. For example, the Jurassians' claim for their own canton was primarily directed against the canton

which harbored the Jura region, Bern.¹

Furthermore, our definition of an SD claim entails the logic of political disintegration (see Wood 1981). Thus, we do not include center-seeking opposition movements that want to overthrow the government or gain increased representation. Because these are aimed at political integration rather than disintegration, we also do not include movements that are aimed at the merger of nation-states. For example, we do not include the Greek Cypriot *enosis* movement that aimed at the unification of Cyprus with Greece. Nor do we include the pan-African movement or movements related to the re-unifications of Germany, Vietnam, or South Korea. (However, as stated already, we do include movements that want to secede from an existing state and merge with another state, such as the Serbs in Bosnia.) Finally, because these do not aim at *increased* SD, we also do not include groups that merely want to retain their current level of SD. For example, we do not include the Moldovan movement in the early 1990s that was opposed to unification with Romania and aimed to preserve national independence. (However, we do include the earlier Moldovan movement that sought independence from the USSR).

Separatist sentiment alone is insufficient for us to code an SDM. We require evidence of political mobilization, for example by a political party, a cultural organization, or a rebel group. SDMs need not have majority support among the group they represent. However, we require a minimum level of political significance and exclude fringe movements. Examples of fringe movements that were excluded include the Bornholm movement (Bornholm is a Danish island), which is represented by a political party that, however, secured at most a couple of hundred votes at a time, and the movement that claims independence for Seborga, a small town in Liguria, Italy, with approximately 300 inhabitants. Similar examples include the movement that demands self-rule for Wessex, a region in England, or the dozens (or even hundreds) of self-proclaimed “micro-nations,” such as the Principality of Sealand, a supposedly independent state located on an offshore platform off the coast of England.

While we follow much of the existing literature and focus on SD claims associated with ethnic groups, we employ an explicitly broad understanding of ethnicity that includes region of origin in addition to standard ethnic markers (in particular language, religion, and race). This broad definition of ethnicity is consistent with Horowitz (1985) and allows us to include SDMs associated with groups that are primarily regionally defined. Many prominent SDMs are strongly associated with regional identities, including the Texans or the Southerners in the US. The inclusion of regional identities thus allows us to cover the full spectrum of identities that may give rise to separatist claims.

Finally, we do not include colonial liberation movements. This is not a definitional requirement – movements that want to sever colonial ties and become independent or at least gain increased autonomy do constitute SDMs according to our definition. But we made the choice to exclude anti-colonial movements for both theoretical (colonial cases are less relevant to understand contemporary separatism) and pragmatic reasons (finding relevant information is much harder for colonial cases). We define a colony as a dependent territory that is neither politically nor legally integrated with the metropole. Thus, we do not include self-determination movements in the British Crown Fiefdoms and Crown Colonies or the Dutch dependencies (i.e. Aruba and the Netherlands Antilles) because the latter represent colonies according to our definition (e.g. the British dependencies do not have representation in the British Parliament while the Dutch dependencies, self-governing in all internal affairs, neither have representation in the Dutch parliament nor abide by the Dutch constitution; see Banks et al. 1997). Note that not all overseas entities are colonies according to this definition. For example, the French Overseas Departments and Territories no longer fall under this definition as they are now represented in the French Parliament.

¹ However, it is only possible for SD claims to be targeted against regional governments if these have significant autonomy that they could devolve to a separatist region.

1.1.2. Identification of SDMs in the dataset

The SDM dataset includes a total of 464 (non-colonial) SDMs in the period between 1945 and 2012. In the dataset, SDMs are identified by the ethnic group on whose behalf they make claims for increased self-rule and their host state (e.g. Scots-UK).² We code separate movements if the same ethnic group makes SD claims in different states. For example, we code separate Kurdish movements in Azerbaijan, Iraq, Iran, Syria, and Turkey. We also code separate SDMs if a group changes its host state affiliation. For example, we code separate Hungarian movements in Czechoslovakia and Slovakia (see below). We also code separate SDMs if identities overlap. For example, several of Dagestan's ethnic groups have agitated for increased self-determination for themselves, but there has also been a movement demanding increased self-determination for Dagestan as a whole. Thus, we code separate SDMs for several Dagestani groups, including the Avars and the Lezgins, but also an umbrella Dagestani movement. Finally, the SDMs identified in the dataset may consist of one or multiple organizations making claims on behalf of the same group. We do not identify individual organizations and, importantly, the nature of the claims made by different organizations associated with the same movement may vary, too. For example, while some Catalan organizations in Spain claim independence, others merely claim increased autonomy. In the dataset, these are combined to a single Catalan SDM.

The dataset indicates the start date of each movement, defined as the first year in which a political organization made public claims for increased SD, including declarations of independence or autonomy. The organization in question needs not be a new organization; it can also be a pre-existing one that shifted its demands to SD. We code a movement as active as long as there is at least one political organization that publicly claims increased SD. If we find evidence that a movement ceased to exist or abandoned its demands for SD, we code an end to the SDM in the year in which either of those events occurred. A successful secession also ends an SDM, as does a country break-up that leads to a change in a group's host state affiliation. For example, we code an end to the Hungarian movement in Czechoslovakia in 1992 because Czechoslovakia ceased to exist. If the same group continues to claim self-rule under the new host, this is counted as a new challenge. For example, the Hungarians in the former Czechoslovakia continued to claim self-rule in independent Slovakia. Movements that were ongoing as of 2012, the last year we cover, are coded as “ongoing.” It is possible that a movement ends and resurges after a while. In such cases, we code a second period of activity, including a second start and end date. While possible in theory, in practice no movement in the dataset has more than two phases of activity.

Coding end dates is not always straightforward. Events that allow us to code an end to a movement are not reported systematically in our sources, and in some cases movement activity stops gradually. Thus, we established a time threshold beyond which a lack of detectable activity in our sources is considered a termination. We applied a “ten years of inactivity rule,” which means that we coded an end to a movement if we did not find evidence of organized separatist activity for ten years. While any cut-off rule is arbitrary, we err on the side of a long time threshold so as not to omit nonviolent groups with low public profiles that might exist in authoritarian states or groups in their early stages of formation. Moreover, a large proportion of non-violent SDMs function as political parties, which means that participation in elections is a key indicator of their activity. This suggests that some nonviolent SDMs may only be publicly observable (i.e. covered by the media) during election campaigns and we should therefore have a threshold that is at least as long as the longest electoral cycle since a group might only make public claims during elections. According to the CIA World Factbook (2002), the longest electoral cycle held in any country in the world is that of the French Senate, whose members are elected for nine-year terms. Our ten year rule allows us to include groups in such long electoral cycles if they are still active.

² When possible, we avoided coding umbrella ethnic groups, such as all indigenous groups in a country, and instead coded the constituent groups that made claims for SD. For example, we include six indigenous groups in Canada: the Cree, the Dene, the Haida, the Innu, the Inuit, and the Iroquois.

A special case emerges if SDMs were active before to the first year they are covered in our dataset. This may occur under two scenarios. First, when SDMs were active before 1945, the first year we cover. The Scottish SDM, for example, has been continuously active since the late 19th century. Second, when SDMs change their host state affiliation. For example, the Katanga SDM in Congo-Zaire emerged when Congo-Zaire was still under Belgian colonial rule. Similarly, the Hungarian movement in Slovakia emerged while Slovakia was still part of Czechoslovakia (see above). If SDMs were active before the first year they are covered in our dataset, the start date reflects the year in which the ongoing phase of separatist activity began. Thus, some start dates are before 1945 or before a country's date of independence. For example, the start date of the Scottish movement is coded with 1886.³

1.1.3. Sources

We used four main sources for the identification of SDMs. The first source is the Minorities at Risk (MAR) dataset (Gurr 1993, 2000; MAR 2009). MAR contains two concepts that can be used to identify SDMs. On the one hand, MAR includes a group typology that differentiates between ethno-nationalist groups, indigenous groups, national minorities, and religious sects, among others. MAR's definition of ethno-nationalism is compatible with our definition of an SDM (though narrower due to its exclusive focus on spatially concentrated groups with a history of autonomy).⁴ Accordingly, we included most of the groups MAR has identified as "ethno-nationalist."⁵ The second concept in MAR that allows for the identification of SDMs is the separatism index (SEPX), which identifies separatist movements, including autonomy movements. We considered two versions of the SEPX variable: the one in the combined phase I-IV release (version 2.05) and the one in the phase V release. They have slightly varied but similar definitions. In particular, both the phase I-IV and the phase V SEPX variables distinguish between "active" movements (meaning they were active in the 1980s and/or after that) and "historical" movements (meaning they were active before the 1980s but not thereafter). We checked both "active" and "historical" cases and included them if we found corroborating evidence on organized separatist activity.⁶ However, MAR's list of SDMs is not exhaustive. Some omissions are systematic. MAR only surveys countries with a population of more than half a million and only includes ethnic groups with a population of more than 100,000 or more than 1% of a country's population. Thus, MAR systematically excludes ethnic groups that are numerically small or reside in small countries. A larger number of omissions are due to coding error. To capture these cases, we supplemented MAR data with data from other sources.

Our second main source is the Peace and Conflict reports published by the Center for International Development and Conflict Management (CIDCM) at the University of Maryland. The 2001, 2003, 2005, and 2008 reports all include chapters on SDMs (see Gurr et al. 2001; Quinn & Gurr 2003; Khosla 2005; Quinn 2008). The respective appendices list all SDMs they found, separating violent from non-violent movements. Their definition of an SDM is relatively loose but appears compatible with ours (though it is narrower, see below).⁷ We included

³ Note that start dates do not necessarily reflect the first year an ethnic group has claimed self-rule. Start dates only indicate the first year of the present spell of separatist activity and do not consider possible separatist activity before this. For example, while the present spell of Scottish separatism originated in 1886, there may have been another phase of separatist mobilization before that.

⁴ Gurr (2000: 17) defines ethno-nationalist groups as "regionally concentrated peoples with a history of organized political autonomy with their own states, traditional ruler, or regional government who have supported political movements for autonomy since 1945."

⁵ For some groups MAR labelled "ethno-nationalist" we found no corroborating evidence of organized claims for self-determination. This includes the Palestinians in Lebanon and Jordan and the Germans in Romania.

⁶ Note that the TYPE and SEPX variables, apart from the rough distinction between "active" and "historical" movements, do not contain information on start and end dates of movements. When available, MAR's coding notes often contain much more detailed (and helpful) information. In addition, a non-zero MAR protest scores can point to an active separatist movement, though it has to be verified that the protest score relates to a self-determination claim.

⁷ For example, Gurr & Quinn (2003: 26) refer to SDMs as "the quest of national and indigenous peoples for self-governance" and "movements seeking greater self-determination."

movements listed in the CIDCM reports if we found corroborating evidence.⁸ While the CIDCM reports constitute an offspring of MAR, they still include several cases not listed in MAR, including the Flemings in Belgium and the Cornish in the UK. Still, the CIDCM reports miss many SDMs. They systematically exclude SDM groups if they are not spatially concentrated as well as non-violent movements that were no longer active in the 2000s and, similarly to MAR, the CIDCM reports are affected by significant coding error.

Our third main source is James Minahan's (2002) *Encyclopedia of Stateless Nations* as well as its predecessor, *Nations without States: A Historical Dictionary of Contemporary National Movements* (Minahan 1996). Minahan aims to cover what he labels "stateless nations," a concept that is by and large compatible with our definition of an SDM. According to Minahan (1996: xvi, 2002: xii), a stateless nation is defined over the following three criteria: i) there must be a body of people "with a national claim to a recognizable geographic area"; ii) this people must have "the display of the outward trappings of national consciousness," including in particular the adoption of a flag; and iii) "the formation of a specifically nationalist organization or political grouping that reflects the nation's claim to self-determination." The two encyclopedias by Minahan constitute the most comprehensive source of data on SDMs that we have found and we included most of the groups listed by Minahan.⁹ Furthermore, the narratives offered by Minahan are often relatively detailed and allow for the identification of start and end dates of SDMs. That said, we also found that the information provided by Minahan is not always fully accurate (especially when it comes to exact dates). Thus, we always sought to corroborate information from Minahan with other sources. Finally, despite the breadth of scope of that study, Minahan's data is not complete. In particular, Minahan systemically excludes groups that have already achieved independence (e.g. the Eritreans), but he also misses some other cases.

The fourth major source that we consulted is Christopher Hewitt and Tom Cheetham's (2000) *Encyclopedia of Modern Separatist Movements*, which covers "ethnic separatism and related topics" (Hewitt & Cheetham 2000: xi). Hewitt & Cheetham (2000: xi-xii) define a separatist movement as a movement for the separation of a particular territory or region from the state of which it is a part, explicitly including both movements for autonomy and independence though excluding anti-colonial movements. They require politically organized activity: "[a] separatist movement is considered to exist whenever there is violence in pursuit of separatist goals or whenever a separatist political party contests elections" (Hewitt & Cheetham 2000: xv). The encyclopedia also contains information on irredentist movements, including both movements for the separation from a state and merger with the ethnic kin state and irredentist claims made by states. According to the authors, the compendium contains information on nearly 300 cases of ethnic separatism and irredentism since 1945 (Hewitt & Cheetham 2000: xii). While not all of these cases fit our definition of an SDM (in particular, we do not include irredentist territorial claims by governments), most of them do so we included them in our dataset. This source is often relatively rich in detail and thus helpful in particular with regard to start and end dates of movements. However, it is not clear if the book is the result of a systematic survey of the universe of SDMs, which raises the possibility of a biased sample.

To overcome the weaknesses of our four main sources, we complemented them with a number of additional sources. This also allowed us to extend coverage of the dataset to 2012 (the last year covered in our main sources is 2006). The list of additional sources we consulted includes *Revolutionary and Dissident Movements* (1988) by Henry Degenhardt, the *World Directory of Minorities* (1997) by Minority Rights Group International, the online version of the *World Directory of Minorities and Indigenous Peoples* by Minority Rights Group International, the UCDP/PRIO database of internal armed conflicts (Gleditsch et al. 2002; Themnér &

⁸ For non-violent movements, the Peace and Conflict reports also give a rough indication as to the start date of SDMs.

⁹ There are three main reasons why we did not include all "stateless nations" identified by Minahan. First, in some cases we were not able to find supporting evidence of organized separatist activity (e.g. the Manchu and Nan Chao in China or the Pomaks in Bulgaria). Second, in other cases it seemed that proclamations of autonomy by "stateless nations" were the actions of fringe or maverick leaders. Finally, in other cases reported by Minahan separatist activity was limited to the colonial or pre-1945 era, which we do not cover.

Wallensteen 2014), UCDP's online Conflict Encyclopedia, the member list of the Unrepresented Nations and Peoples Organization (UNPO), and the civil war data by Doyle & Sambanis (2006), as updated in Sambanis & Schulhofer-Wohl (forthcoming). In addition, we canvassed news sources, including Keesing's and Lexis-Nexis, for keywords including self-determination, self-governance, self-rule, land rights, regionalism, separatism, secessionism, irredentism, and sovereignty (in combination with country names). Finally, we consulted a host of country- and case-specific academic literature. The coding notes list the sources used case-by-case.

1.2. Separatist Violence

The dataset identifies periods of separatist violent conflict between SDMs and states. It also provides information on the onset of separatist violence, dropping movement-years of ongoing separatist violent conflict. We distinguish between periods of high-level and low-level separatist violent conflict. High-level separatist violence (HVIOLSD) is defined as a civil war over regional autonomy or secession. The definition of a civil war follows Sambanis (2004) and the data is drawn from Doyle & Sambanis (2006), as updated in Sambanis & Schulhofer-Wohl (forthcoming).¹⁰

Our preferred indicator of low-level separatist violence (LVIOLSD) constitutes an inclusive category incorporating violent conflict over SD that fits *any* of the definitions for “low-violence” currently being used in the literature on civil war and ethnic conflict. We require only that the group in question must participate in the violence (thus excluding one-sided violence), that the violence is over self-rule, and that the violence concerns the movement-state dyad (thus excluding inter-ethnic violence and inter-factional clashes). We used three main sources for the identification of periods of LVIOLSD: the UCDP/PRIO list of internal armed conflicts over territory (Gleditsch et al. 2002; Themnér & Wallensteen 2014), the list of “armed self-determination conflicts” in the CIDCM reports (Gurr et al. 2001; Marshall & Gurr 2003, 2005; Hewitt et al. 2008), and MAR (Gurr 1993, 2000; MAR 2009). In the case of MAR, we required a score of three or higher on the anti-government rebellion scale, with the exception of a score of three that solely corresponds to sovereignty declarations.¹¹ Note that the MAR rebellion scale does not indicate whether a rebellion involved separatist aims; thus we manually established a rebellion's aims in those cases where we coded solely based on MAR.¹² In addition to these three main sources, we coded LVIOLSD if news reports or case study evidence suggest at least 25 deaths per year due to separatist violence.

¹⁰ When we refer to Doyle & Sambanis (2006) in the coding documentation, we always refer to the version as updated by Sambanis & Schulhofer-Wohl (forthcoming). Note: Doyle & Sambanis (2006)/Sambanis & Schulhofer-Wohl (2014) do not follow the conflict dyad approach. According to their conception, a civil war refers to the aggregation of violence in a given country. The SDM coding of HVIOLSD is disaggregated to the group-level. That is, we only code HVIOLSD if the state-movement dyad in question rose to the level of a civil war. If there is a civil war involving several groups with different or competing goals and the group in question was not a major participant in the war, we code low-level violence only.

¹¹ The MAR anti-government rebellion scale is defined as follows: 0 = None reported; 1 = Political banditry (often the last stage of guerilla wars), sporadic terrorism; 2 = Campaigns of terrorism; 3 = Local rebellions: armed attempts to seize power in a locale; declarations of independence by a minority-controlled regional government; 4 = Small-scale guerilla activity (fewer than 1,000 armed fighters; fewer than six attacks reported per year; and attacks in a small part of the area occupied by the group or in one or two other areas); 5 = Intermediate-scale guerilla activity; 6 = Large-scale guerilla activity (more than 1,000 armed fighters; frequent armed attacks and attacks affecting a large part of the area occupied by the group); 7 = Protracted civil war, fought by rebel military units with base areas.

¹² Note that the Peace and Conflict reports code ongoing armed conflict if episodes of armed conflict are separated by up to five years (see e.g. Marshall & Gurr 2005: 25). We do not apply a similar five-years rule. Thus, if the Peace and Conflict reports suggest ongoing armed conflict but we found evidence for conflict de-escalation and subsequent re-escalation within maximally five years (drawing on UCDP coding notes, MAR coding notes, case study evidence or newspaper reports), we do not follow the Peace and Conflict reports and code a de-escalation and a re-escalation. A similar problem applies to the MAR rebellion score for all years before 1985, when it is only available on a five-year basis. If MAR suggests a LVIOLSD code for a five-year phase until and including 1980-1984, we only followed MAR and coded LVIOLSD throughout the five-year phase if we found no better evidence suggesting clear start and end dates.

To check the robustness of results, the dataset allows for the exclusion of “ambiguous” cases of separatist violence. Ambiguity refers to mixed motives, that is, when violence was not only over SD but also about capturing the central government. Furthermore, we include short peace identifiers that allow for the exclusion of new conflict onsets after a short peace of no more than two calendar years.

Finally, the dataset also includes alternative indicators for separatist violence incidence and onset based solely on the UCDP/PRIO list of internal armed conflicts over territory (Gleditsch et al. 2002; Themnér & Wallensteen 2014).

2 Random Sample Data

The second part of the SDM dataset includes several more detailed variables for a randomly selected subset of 106 SDMs. Most variables are coded at the group-year level (see the variable descriptions for exceptions). Our focus was on how states react to SD challenges; specifically, whether they make concessions to SDM groups or whether they restrict their rights. We coded several additional variables related to state-movement interactions and attributes of SDM groups. The section below provides descriptions and coding rules for all variables. In this section, we summarize how the sample was drawn, the main sources we used, as well as the EPR2SDM mapping, which may be of use to researchers who want to merge the SDM data with the EPR dataset and/or merge EPR variables with the SDM data.

2.1. Sampling Procedure

Several of the variables included in the second part of the SDM dataset proved time-consuming to code. Instead of covering the universe of SDMs, we thus chose to pick a sample of SDMs at random. We did not include movements in countries with a population of less than 500,000 by the end of 2012 in the random draw.¹³ Importantly, we did not sample all types of SDMs with the same probability. One of the main research questions that can be addressed with our data are the conditions under which conflicts about SD escalate from non-violence to violence. However, transitions from non-violence to violence are relatively rare. Thus, we over-sampled “switched” SDMs – movements that switched from non-violence to violence at least once during their activities. The random sample includes 54 out of the totally 118 switched movements ($p \sim 46\%$). There are two other types of SDMs: non-violent movements and violent movements with no observed switch from non-violence to violence. Both were sampled with approximately the same probability: 47 out of 308 non-violent movements were sampled ($p \sim 15\%$) and 5 out of 32 violent-no-switch movements ($p \sim 16\%$). The dataset includes sampling weights to correct for the differential sampling probabilities.

2.2. Sources

We used a broad range of sources to code the random sample variables. Most of these sources are case-specific; and all sources are listed on a case-by-case basis in the coding notes. However, the main sources included Minahan (1996, 2002), Hewitt & Cheetham (2000), and the online version of the *World Directory of Minorities and Indigenous Peoples* by Minority Rights Group International. In addition, for our coding of whether the territory claimed by SDMs includes oil or gas, we drew on Lujala et al.'s (2007) PETRODATA. Finally, for data on group sizes and access to central state power we drew on data from the *Ethnic Power Relations* (EPR) dataset (Cederman et al. 2010; Vogt et al. 2015) for all cases where SDM groups correspond to an EPR group.

¹³ This concerns a total of six movements in the following states: Antigua & Barbuda, Belize, Kiribati, Micronesia, Saint Kitts and Nevis, and Vanuatu.

2.3. EPR2SDM

As just stated, we directly adopted data on group sizes and central state access from EPR for some of the cases. To this purpose, we “docked” all random sample cases to the EPR dataset. This docking may be useful to other researchers if they want to add additional data from EPR to the SDM dataset or merge the SDM data with the EPR dataset, so we included it in the data release.

The unit of analysis in both EPR and SDM is the ethnic group (or the country-group-year, to be more exact), and so matching EPR groups with SDM groups is often straightforward. However, not all SDM groups have a perfect equivalent in the EPR dataset. There are several reasons. First, EPR has a narrower understanding of ethnicity that does not include region of origin. EPR also excludes some tribal groups and clans; for example, contrary to SDM, EPR does not include the Isaaqs in Northern Somalia.

Second, groups in EPR and SDM may be differentially aggregated. For example, EPR codes a single umbrella group of Native American Indians in the US, whereas the SDM dataset several individual groups, including the Lakota (Sioux) and the Cherokee. Though cases where the SDM dataset employs a higher level of disaggregation are more prevalent, aggregation differences can run both ways. The SDM dataset codes an umbrella Southerners group in Sudan, for example, while EPR distinguishes between numerous South Sudanese groups, including the Dinka and the Nuer.

Third, EPR does not include overseas territories, whereas the SDM dataset includes them as long as they are integrated with the metropole. Fourth, EPR does not include 1945. And finally, there can be coding error in EPR. EPR aims to code all politically relevant ethnic groups around the world. While the existence of an organized claim to SD implies political relevance according to the EPR definition,¹⁴ EPR coders appear to have missed some of the groups included in the SDM dataset.

As a result, a total of five different scenarios emerge for the relationship between SDM and EPR groups:

- *1:1*. These are cases where an SDM group has a perfect equivalent in EPR.
- *1:n*. These are cases where a combination of EPR group corresponds to an SDM, as in the case of the Southerners in Sudan.
- *n:1*. These are cases where an SDM group constitutes a part of a larger EPR group, as in the example of the Native American Indians mentioned above. This scenario also includes cases of regionally defined SDM groups that form part of a larger EPR group, such as the Westerners in Canada (this is a regionally defined sub-group of English-speakers in Canada).¹⁵
- *1945*. SDM groups that were active in 1945 by definition have no match in EPR.
- *Irrelevant*. This is an umbrella category that includes all SDM cases for which there is no equivalent in EPR, including cases where EPR coders mistakenly did not consider a group politically relevant; SDM groups located in overseas entities; and SDM groups that are considered to be tribes or clans by EPR.

We culled data on group sizes and central state access directly from EPR for all 1:1 cases. In all other cases the data was hand-coded while relying on EPR data to the extent possible. For example, we combined group sizes of the constituent EPR groups for 1:n cases.

¹⁴ EPR considers a group politically relevant if at least one political organization claims to represent the group in national politics or if group members are subject to state-led political discrimination.

¹⁵ If cleavages cross-cut (e.g. in cases of regional identities), an SDM may form part of more than one EPR group (e.g. the Far-Eastern Slavs SDM relates to a subset of both EPR’s Russians and Ukrainians).

3 Variable Descriptions

country

The internationally-recognized sovereign state that includes the territory for which a given SDM is seeking more SD.

ccode

The host state's Correlates of War (COW) country code.

group

The name of the ethnic group that the SDM represents.

groupid

Numeric group identifier.

year

Calendar year.

region

The geographic location of the territory for which a movement is seeking greater SD. Note: this variable relates to the location of the territory that is claimed by SDMs. Thus, dependencies are not classified on the basis of the geographic location of the metropole, even though in the dataset such movements are listed under the metropole. The variable distinguishes between eight geographic regions (see Table 1).

rcode

Numeric region identifier (see Table 1).

Table 1: Regions

region	rcode	Description
North Africa and Middle East	1	Morocco, Algeria, Tunisia, Libya, Egypt, Turkey, Cyprus, Lebanon, Israel, Syria, Jordan, the Saudi Arabian peninsula, Iraq, Iran and Bahrain.
Sub-Saharan Africa	2	All countries on the African continent except for Morocco, Algeria, Tunisia, Libya and Egypt, in addition to all off-shore islands.
Central Asia	3	Armenia, Azerbaijan, Georgia, Kazakhstan, Turkmenistan, Uzbekistan, Tajikistan, Kyrgyzstan, Afghanistan and Pakistan. Russian SDMs in the Caucasus region or in regions bordering

		these Central Asian states have been listed under Central Asia.
Southern and Eastern Asia	4	All of the countries of mainland Asia except for the countries of Central Asia (i.e. from India eastwards), in addition to all of the islands of South-East Asia. Russian SDMs movements east of the Ural mountains and not in the Caucasus region or bordering one of the Central Asian states have been listed under this geographic area.
Europe	5	Countries of both Eastern and Western Europe. According to the National Geographic Atlas of the World, a commonly accepted division between Asia and Europe is formed by the Ural Mountains, the Ural River, Caspian Sea, the Caucasus Mountains, and the Black Sea with its outlets, the Bosphorus and Dardanelles. Any country to the West of this divide belongs to Europe; any country to the East of this divide is part of Asia. Russian SDMs have been listed under Europe or Asia according to this definition.
North America	6	Canada, the United States (including Hawaii) and Mexico.
Latin America	7	Central America, South America and the Caribbean Islands.
Oceania	8	Refers to Australia, New Zealand and the South Pacific Islands.

active

Binary variable indicating years in which an SDM was active. We track concessions and restrictions up to ten years before a movement's first year of activity recorded in the dataset. Thus not all group-years are coded as active. For example, if a movement starts in 1972, the dataset also includes 1962-1971, but these years are denoted as inactive.

phase1

Dummy variable flagging a movement's first phase of activity.

phase2

Dummy variable flagging a movement's second phase of activity.

fyphase

Dummy variable indicating the first years of a spell of separatist activity.

fyphase1

Dummy variable indicating the first year of a movement's observed activity. Note: this is not necessarily the start date as a movement may have been active already under a different host state (e.g. under a colonial regime) or before 1945.

fyphase2

Dummy variable indicating the first year of a movement's second phase of activity, where applicable.

fygroup

Dummy variable indicating the first year an SDM is included in the dataset. Since we track concessions and restrictions up to ten years before the first year we cover, the first year of inclusion is necessarily "inactive."

startdate1

Indicates the start date of a movement (first phase). The start date may precede the first year a movement is coded in the dataset if a movement has "prior activity," that is, if it was active under a different host state (colonial or non-colonial) or before 1945.

enddate1

Indicates the end date of a movement (first phase), defined as the last year in which a political organization made public demands for SD. If a movement ceased its activities in a given country due to a host state change, enddate1 is coded with 8888. If a movement was ongoing as of 2012, enddate1 is coded with 9999.

tyrule1

Dummy indicating whether the end date (first phase) was coded based on the "ten years of inactivity" rule.

startdate2

Indicates the start date of a movement (second phase), where applicable.

enddate2

Indicates the end date of a movement (second phase), where applicable. If a movement ceased its activities in a given country due to a host state change, enddate2 is coded with 8888. If a movement was ongoing as of 2012, enddate2 is coded with 9999.

tyrule2

Dummy indicating whether the end date (second phase) was coded based on the "ten years of inactivity" rule.

perstart

Indicates the start year of a given period. Periods are defined by a movement's involvement in separatist violence and host state changes.

perend

Indicates the end year of a given period. Periods are defined by a movement's involvement in separatist violence and host state changes.

violsd

Dummy indicating whether an SDM was involved in a violent separatist conflict in a given year, including both high- and low-level separatist violence.

lviolsd

Dummy indicating whether an active SDM was involved in low-level separatist violence in a given year.

hviolsd

Dummy indicating whether an active SDM was involved in high-level separatist violence in a given year.

ambiguous

Dummy identifying group-years when separatist violence involved mixed motives, that is, group-years when the violence was not only over SD, but also over capturing the central government.

nviolsd

Dummy indicating whether an SDM was non-violent in a given year.

prioract

Indicates whether a movement was involved in “prior activity,” that is, whether a movement was active already before the first year it is covered. Prioract is coded 1 if the movement was active before 1945 and continued to be active at least in 1945. It is coded 2 if the movement was active under colonial rule and continued to claim SD after the host state’s independence. It is coded 3 if a movement changes from one non-colonial host state to another non-colonial host state. Finally, it is coded 0 if there was no prior activity.

priorconflict

If a movement was involved in “prior activity,” this variable indicates the years (if any) the movement was involved in separatist violence before the first year it is covered. Only separatist violence during the ongoing phase of contention is considered. In addition to the sources for the separatist violence coding noted above, this variable draws on the UCDP/PRIO list of extrasystemic armed conflicts (Gleditsch et al. 2002; Themnér & Wallensteen 2014) and the Correlates of War dataset (Sarkees & Wayman 2010) for the period before 1946, the first year that is covered by UCDP/PRIO.

priornonviol

Binary variable indicating whether a movement involved in “prior activity” was non-violent in the year preceding the first year that it is covered in our dataset.¹⁶

¹⁶ A special case emerges if a movement splits up due to a country break-up. For instance, in 1991, the Armenian movement in the Soviet Union split up into an Armenian movement in Azerbaijan and one in Georgia (while Armenians in Armenia gained independence). In these cases, priornonviol relates exclusively to the territory within the new host state and not to the old movement as a whole. Thus, while the Armenians in the Soviet Union are coded as involved in separatist violence in 1991, priornonviol for the Armenians in Georgia is coded 0 because the separatist violence was over Nagorno-Karabakh. This rule applies analogously to the “priorviol” variable.

priorviol

Binary variable indicating whether a movement involved in “prior activity” was involved in separatist violence in the year preceding the first year that it is covered.

violsd_onset

Binary variable indicating onsets of violent separatist conflict, counting both cases of high-level and low-level separatist violence. Cases of ongoing separatist violence are dropped (i.e., they are coded missing). This includes cases of ongoing separatist violence in the first year of movement activity. That is, if there is separatist violence in the first year a movement is covered in our dataset, the onset variable is coded missing if the movement was active and violent in the year before that (i.e. if priorviol = 1). For instance, the Estonians were involved in separatist violence in 1945, but the violence had started in 1944. Thus the dataset indicates prior violent activity and we do not code an onset in 1945. The same logic applies in the context of host state changes. The Armenians in Azerbaijan, for instance, were involved in separatist violence in 1991, the year of Azerbaijan’s independence. However, we do not code an onset in this year because the violence started when Azerbaijan was still part of the Soviet Union.

hviolsd_onset

Binary variable indicating onsets of high-level separatist violence. Cases of ongoing high-level separatist violence are dropped, including cases of ongoing high-level separatist violence in the first year of movement coverage.¹⁷ The comments under violsd_onset apply analogously.

shortpeace_violsd

Binary variable indicating periods of short peace, defined as periods when episodes of separatist violence are separated by no more than two calendar years of peace (including episodes prior to the first year of activity recorded in the dataset). This version of the shortpeace variables considers both low-level and high-level incidences of separatist violence. To illustrate the logic, the Chechens were involved in separatist violence in 1994-1996 and 1999-2012. The two episodes of separatist violence are separated by a short peace of two calendar years (1997-1998). Thus the shortpeace_violsd variable flags the years 1997-1999 with 1. At analysis stage, shortpeace_violsd can be used to test the robustness of results to the exclusion of onsets of separatist violence after periods of short peace by excluding all observations that are coded 1.

shortpeace_hviolsd

Binary variable indicating periods of short peace, counting only high-level separatist violence (HVIOLSD). The comments under “shortpeace_violsd” apply analogously.

py_violsd_phase

Py stands for peace years. The dataset includes two versions of peace years. While both count the number of calendar years since the last episode of violent separatist conflict, they differ with regard to the date at which the peace year count starts. The first is what we call “phase-based.” We use the suffix “_phase” to denote phase-based peace years variables. Phase-based peace years count calendar years based on the ongoing phase of movement activity. With phase-based peace

¹⁷ In the following four cases, we do not code an onset of high-level separatist violence in the first year of movement coverage due to prior high-level violence: Estonians (1945), Latvians (1945), Lithuanians (1945), and Ukrainians (1945). In all other cases with high-level separatist violence in the first year of movement coverage and prior violent activity we do code an onset because prior activity had not risen to the war level: Bosnian Serbs (1992), Armenians in Azerbaijan (1991), Saharawis in Morocco (1975), Karens (1948), and East Timorese (1975).

years, the year of the start date is always coded as 0. It is 1 in the second year, 2 in the third, etc. The count reverts back to 0 if there was conflict in the previous year. Importantly, the count continues if a movement changes its host state affiliation. For example, we cover Myanmar's Shan movement from 1948 onwards, the year of Myanmar's independence. However, the movement's start date is coded with 1942, when Myanmar was still a British colony. Thus, the phase-based peace variable starts with 6 in 1948 (there was no separatist violence in 1942-1947). The analogous logic applies if a movement was active before 1945. The Scottish movement's start date, for instance, is coded with 1886. Accordingly, in 1945 the count stands at 59, given that there was no separatist violence in 1886-1944. If a movement has two phases of activity, the count restarts at 0 in the first year of the second phase.

The second version of peace years is what we call "country-based." We use the suffix "_country" to indicate country-based peace years. The country-based variant starts counting in 1945 or, if a country was not independent by then, in the year that the country attained independence.¹⁸ In the example of the Shans mentioned above, country-based peace years variables start with 0 in 1948. Importantly, the country-based peace years may not be 0 in a movement's first year of activity. For instance, the Assamese movement in India emerged in 1979 and any country-based peace years variable is thus pegged to 32, given that India had attained independence in 1947. A special case emerges if the territory including an ethnic group joined (or formally integrated with) a country after 1945 or after the country's date of independence. If so, we start counting in the year the respective territory merged with the state in question. Just like phase-based peace years, country-based peace year counts revert back to zero in the year following an episode of separatist violence.

For both phase-based and country-based peace years, we include alternatives for different levels of separatist violence. The present variable counts peace years based on VIOLSD, that is, based on both low-level and high-level incidences of separatist violence.

py_violsd_country

Country-based peace years variable based on VIOLSD. See "py_violsd_phase" for a more detailed discussion of the different versions peace years variables.

py_hviolsd_phase

Phase-based peace years variable based on HVIOLSD. See "py_violsd_phase" for a more detailed discussion of the different versions peace years variables.

py_hviolsd_country

Country-based peace years variable based on HVIOLSD. See "py_violsd_phase" for a more detailed discussion of the different versions peace years variables.

violsd_onset_woambig

Analogous to "violsd_onset," but dropping "ambiguous" onsets of separatist violence.

hviolsd_onset_woambig

Analogous to "hviolsd_onset," but dropping "ambiguous" onsets of high-intensity separatist violence.

¹⁸ With only few exceptions, the dates of independence were drawn from Gleditsch & Ward (1999).

shortpeace_violsd_woambig

Analogous to “shortpeace_violsd,” but excluding separatist violence denoted as “ambiguous.”

shortpeace_hviolsd_woambig

Analogous to “shortpeace_hviolsd,” but excluding high-intensity separatist violence denoted as “ambiguous.”

py_violsd_woambig_phase

Analogous to “py_violsd_phase_woambig,” but excluding separatist violence denoted as “ambiguous.”

py_violsd_woambig_country

Analogous to “py_violsd_country,” but excluding separatist violence denoted as “ambiguous.”

py_hviolsd_woambig_phase

Analogous to “py_hviolsd_phase,” but excluding high-level separatist violence denoted as “ambiguous.”

py_hviolsd_woambig_country

Analogous to “py_hviolsd_country,” but excluding high-level separatist violence denoted as “ambiguous.”

doublecount

Movements that change from one host state to another may be listed twice in the year of the host state change, once in the old and once in the new host state. This is fine for analysis, but can be misleading if one wants to know about the number of active self-determination movements in a given year. The double count variable identifies movement-years in which a movement is listed twice in the sense defined above. It is coded 1 (“listed twice (old host)”) in the last year of activity in the old host state and 2 (“listed twice (new host)”) in the first year of activity in the new host state. If a movement is involved in a host state change from one non-colonial host state to another but is not listed twice, the doublecount variable is coded 3 (“not listed twice (old host)”) and 4 (“not listed twice (new host)”) in the year of the host state change under the old and new host, respectively.¹⁹ The number of active groups in a given year can then be derived by excluding any observation with doublecount = 1 (or, alternatively, by excluding observations with doublecount = 2). Note: since we do not code anti-colonial movements, the double count problem does not apply to host state changes in the context of decolonization.

doublecount_oldhost

This gives the name of the old host state (rump state) if doublecount is non-missing.

¹⁹ There are three scenarios under which a movement can be involved in a non-colonial host state change but not counted twice. First, a small number of movements are not coded under the old host for specific reasons (see e.g. the Bosnian Serbs). Second, if the host state change occurred exactly at the turn of the year (December 31/January 1). Third, if the movement splits up in multiple sections in the context of a secession and they continue their activities in the old and the new state(s) (e.g. after Montenegro’s secession from Serbia and Montenegro in 2006, the Sandzak Muslims are coded as active in both countries).

doublecount_newhost

This gives the name of the new host state if doublecount is non-missing.

sample

Dummy that indicates whether a movement forms part of the random sample.

weight

Sampling weight (inverse probability weight) that can be used to correct for the over-sampling of “switched” movements in the random sample.

below500000

Binary variable indicating movements in countries with a population of less than 500,000 by the end of 2012. Countries with a population of less than 500,000 were not included in the random draw.

violent

Binary variable indicating movements that engaged in separatist violence, including both low- and high-level separatist violence. The variable exclusively considers the time a movement is observed in a given country. Violent episodes associated with the movement under a different country or before 1945 are not considered.

switched

Binary variable identifying movements that switched from non-violence to violence at least once. Since we are interested in tactic switches from non-violence to violence in response to government policy, the switched variable focuses on the period in which a movement is observed under the header of a given country. Thus, switches from non-violence to violence before a country’s independence or before 1945, the first year we cover, are not considered. However, we make an exception to this general rule if a group changes tactics in the first year of a country’s independence or in 1945. We consider this a tactic switch if the movement was both active and non-violent prior to this. The Gagauz, for instance, took up arms in 1991, the first year of Moldova’s independence, and had been both active and non-violent before that under the header of the Soviet Union. Thus the Gagauz are coded as switched under Moldova.²⁰

con

This binary variable tracks concessions by the state to SDM groups in the form of policies that increase SDM groups’ levels of self-rule (i.e., increase their autonomy or even give them independence) and/or increase their cultural rights status (in particular linguistic and religious rights). We include partially implemented policies as long as there were meaningful steps towards reform. However, we do not include token or very minimal concessions that do not effectively increase a group’s status. The con variable tracks concessions up to ten years before the first year of movement activity. Thus, if we cover a movement from 1972, the con variable tracks concessions from 1962. If there was a host state change in the ten years before the first year that is covered, we track actions by the former host state. Analogously, the con variable

²⁰ The following movements are coded as switched based on this rule: the Croatian Serbs (Croatia), the Gagauz and the Trans-Dniester Slavs (Moldova), the Somalis (Kenya), and the Chams, Khmer Krom, and Montagnards (Vietnam).

tracks concessions up to ten years before the first year of a movement's second phase of activity, where applicable. See below for more detailed comments on how we handled specific scenarios.

cultcon

Binary variable flagging concessions on cultural or minority rights, including concessions on language rights, religion, and education. See below for more details.

autcon

Binary variable flagging concessions on autonomy (territorial self-rule short of secession). See below for more details.

indcon

Binary variable flagging concessions on independence or secession. See below for more details.

res

Binary variable flagging years in which the state restricts an SDM group's level of self-rule (i.e., decreases its level of autonomy or revokes its independence) or reduces its cultural rights status (in particular linguistic and religious rights). Analogously to the concessions variable, we include restrictions that were only partially implemented if there were meaningful steps towards reform. Again, token or very minimal policies are not included. Finally, analogously to the concessions variable, res tracks restrictions up to ten years before movement activity. See below for more details.

cultres

Binary variable flagging restrictions related to cultural or minority rights, including language rights, religion, and education. See below for more details.

autres

Binary variable flagging restrictions related to territorial self-governance or self-rule short of secession or independence. These are instances of centralization, possibly to the extent where the center imposes direct rule. See below for more details.

indres

Binary that identifies restrictions on independence or secession. See below for more details.

Notes on coding concessions and restrictions:

The following provides more detailed information on the nature of policies coded as concessions and restrictions as well as on the time point at which we code concessions and restriction. Note that some of the issues we faced were more of a case-specific nature. Explanations of how we dealt with individual cases can be found in the coding notes.

- Unilateral actions by regional governments controlled by SDM groups, such as the forming of a de facto state or unilateral legislation, are not coded as concessions. We also do not code a concession if a group attains independence without the consent of the rump state. Analogously, the same applies to restrictions.
- We code concessions and restrictions irrespectively of the level of government that

adopts a policy. That is, we include concessions and restrictions irrespectively of whether they are made by the national government or by a regional government. However, note that we do not code “concessions” by regional governments if these are controlled by the group in question and the national government had no role in them (see the comments on unilateral power grabs above).

- A critical question relates to the point in time at which concessions and restrictions should be coded. Often, a significant amount of time elapses between the point at which a concession or restriction is announced, when the relevant legislation is passed, and when it is fully implemented. We generally code the earliest possible date, that is, when a policy was announced and there were some visible steps towards implementation (barring the special case of de facto independent groups, see below). This is done to allow us to capture the full range of possible reactions to concessions and restrictions. Furthermore, this allows us to code a restriction at a later point in time if a previously announced policy is later reversed and thus never fully implemented (see below).
- Federal systems are predicated on shared sovereignty and often feature a constant ebb and flow of competency transfers between the national center and the federal units. Day-to-day transfers of sovereignty in federal systems are not coded.
- Promises are not coded if there are no steps towards implementation (barring the special case of de facto independent groups, see below). If there are significant steps towards implementation, we code a concession. If the state begins to implement a promise, but then backtracks, we code a concession in the year implementation begins and a restriction in the year the center backtracks.
- We code an autonomy concession if the state grants a referendum on autonomy or independence (irrespective of the outcome). We do not code a restriction if the referendum comes out against the constitutional change, except under very specific circumstances. To give an example, we code a restriction for the Turkish Cypriots due to the 2004 referendum on the Annan plan, in which the Turkish Cypriots were clearly outvoted by the Greek Cypriots.
- Changes in the level of self-government must relate to the regional level; changes in municipal self-rule are not coded.
- Changes in non-territorial autonomy are included as autonomy concessions or restrictions, respectively. Non-territorial forms of autonomy are rare; the archetypical example is the Ottoman millet system. See Coakley (1994) for an extended discussion. An example that is coded in our dataset is the personal power-sharing system that was enshrined Cyprus’ 1960 constitution.
- The denial to live in a given territory constitutes the ultimate restriction of group autonomy. Implementation implies the deportation of a group. Thus, the denial to live in a given territory is coded as an autonomy restriction. The deportation of groups such as the Crimean Tatars under Stalin constitutes a good example. Conversely, if a group is granted the right to return, we code an autonomy concession.
- In contrast to the above, we do not include relocation policies wherein the state motivates members of an ethnic group (typically a country’s dominant group) to settle in the homeland of an SD group. While it is possible that relocation policies affect SD^M groups’ level of self-rule, it is very difficult to differentiate between state-initiated relocation policies from “normal” migration flows, which exist everywhere.
- The imposition of a blockade on an SDM group’s homeland is considered an autonomy restriction. Blockades may include the imposition of an economic blockade or the building of a wall around the group’s settlement area. If a blockade is lifted, we code an autonomy concession.
- The imposition of martial law, declarations of a state of emergency, government crack downs, and one-sided violence are not considered restrictions. These are outside the scope of our definition and, if considered, would induce tautology problems if restrictions are used to explain separatist violence. That said, declarations of emergency or the like may go hand in hand with the imposition of a restriction as defined above.
- Often, concessions on regional autonomy also include concessions on cultural rights. That is, if an autonomous region is set up or if additional powers are devolved to an

autonomous region, the group's status often increases with regards to both territorial self-rule and cultural rights. The decision whether a group gained only with regards to territorial self-rule or also with regard to cultural rights can be very difficult. To guarantee consistency, we never coded both a cultural rights and autonomy concession based on the same act. That is, if an act grants increased territorial self-rule, we coded only an autonomy concession – even if the act also affected a group's cultural rights. We only code both a cultural rights and an autonomy concession in the same year if there are two clearly distinguishable events, with one involving territorial autonomy and the other cultural rights. Analogously, the same logic is applied to restrictions.

- We do not code a restriction if some groups gain independence but not others. For example, we do not code the fact that Croatia gained independence but Serbian territories in Croatia had to remain with Croatia as a restriction for the Croatian Serbs. This follows the logic that there is no explicit policy by the state. However, this is clearly a borderline scenario; due to Croatia's secession, the Croatian Serbs, who had previously been part of the most powerful and numerous group, became an essentially powerless minority.
- A special case emerges if a group has achieved *de facto* independence. Given the state's lack of control over the separatist territory, the state cannot implement policies against the will of the SDM group's leaders. The following rules apply to cases with *de facto* independence.
 - First, we do not code a concession if a group attains *de facto* independence because this represents a unilateral power grab (see above). We do not code a concession even if there is evidence that the center implicitly accepts the separatists' *de facto* independent status. That is, we do not code a concession if the center stops its attempts to reintegrate the break-away region, but without granting formal independence.
 - Second, we code a concession or a restriction if the center changes the constitutional status of the *de facto* independent group, even if the legislation cannot be implemented.
 - Third, we code a concession if the center actively offers SDM groups significantly increased autonomy, even if the offer implies reintegration into the state. If the separatists reject the offer and the center renews its promise, for example after another round of negotiations, we code another concession. Analogously, we code a restriction if the center backtracks on an earlier promise.
 - Fourth, we continue to code policies or actions like the imposition of an economic blockade which directly affect the center-periphery relations despite the *de facto* independence.
 - Finally, we do not code a restriction if a group loses its *de facto* independent status. *De facto* independence, by definition, represents a unilateral power grab. The restoration of the constitutional status quo does not represent a restriction in the sense employed here.

priorcon1

Binary variable that is coded 1 if the last policy change affecting a group before the first year an SDM is coded as active constitutes a concession as defined above. This includes government actions beyond the ten years before movement activity that are covered by the concessions and restrictions variables. Where applicable, we included policies by former host states.

priorres1

Binary variable that is coded 1 if the last policy change affecting a group before the first year an SDM is coded as active constitutes a restriction as defined above. This includes government actions beyond the ten years before movement activity that are covered by the concessions and restrictions variables. Where applicable, we included policies by former host states.

priorcon2

Analogous to priorcon1, but relating to the second phase of activity (where applicable).

priorres2

Analogous to priorres1, but relating to the second phase of activity (where applicable).

majterr1

The majterr1 variable identifies major territorial changes affecting an SDM group in a given year. We distinguish between 10 different forms of major territorial changes (see Table 2). Major territorial changes may, but do not have to be, the consequence of a concession or a restriction. For example, establishment of de facto independence is by definition not the result of a government concession. Internationally recognized independence may but does not have to be the result of a concession – Kosovo, for example, attained independence without Serbia’s consent. Major territorial changes are always coded in the year that they are implemented. Since concessions and restrictions may be coded earlier (i.e., at the initiation of a policy process), the respective major territorial change may not be coded in the same year. See below for more detailed definitions of regional autonomy and de facto independence.

majterr2

It is possible that an SDM group is affected by multiple major territorial changes in the same year. If so, majterr2 gives the second major territorial change.

majterr3

Analogous to majterr2.

Table 2: Major Territorial Changes

Major territorial change	majterr	Description
Establishment of regional autonomy	1	An SDM group, or at least a significant part of the group, attains regional autonomy, excluding cases of de facto independence and sub-state secession.
Sub-state secession	2	An SDM group’s homeland, or at least a significant part thereof, is separated from an autonomous region to form an autonomous region of its own. This scenario can only arise in federal or quasi-federal states and states with autonomous regions. Partitions of a group’s homeland into multiple entities are not coded as long as the group retains meaningful autonomy. ²¹

²¹ A special case emerges if a group is granted an autonomous status within a larger federal unit. The sub-state secession category is reserved for cases of full separation from a federal unit. Thus such instances are coded with “Establishment of regional autonomy.” Analogously, if a group loses its autonomous status within a larger federal unit, we code “Revocation of regional autonomy.”

Establishment of de facto independence	3	An SDM group, or at least a significant part of the group, attains de facto independence.
Independence	4	An SDM group, or at least a significant part of the group, attains internationally recognized independence.
Revocation of regional autonomy	5	An SDM group's regional autonomy is revoked (or at least the autonomy of a significant part of the group), excluding cases of de facto independence and sub-state mergers.
Sub-state merger	6	An SDM group's autonomous homeland, or at least a significant part thereof, is merged with another autonomous entity, implying a loss of autonomy. We do not code a sub-state merger if two or more previously existing entities dominated by the same group are merged because the group retains autonomy.
Revocation of de facto independence	7	An SDM group, or at least a significant part of the group, loses its de facto independent status.
Revocation of independence	8	An SDM group, or at least a significant part of the group, loses internationally recognized independence. ²²
Host change (old)	9	An SDM group, or at least a significant part of the group, changes its host state, but is not granted independence itself. Note: MajtterrX is coded 9 in the last year of activity before a host state change and 10 in the first year of activity after a host state change.
Host change (new)	10	An SDM group, or at least a significant part of the group, changes its host state, but is not granted independence itself. Note: MajtterrX is coded 9 in the last year of activity before a host state change and 10 in the first year of activity after a host state change.

claim

The claim variable gives the type of SD claim made by SDMs. Broadly speaking, SDMs can either demand outright secession or increased autonomy. Each can be further divided into two sub-categories (see Table 3). Note that different actors belonging to the same movement may make claims of varying intensity. For example, some of the organizations associated with the Corsican movement have made claims for increased autonomy within France while others demanded outright independence. The claim variable reflects the dominant claim. In the Corsican example, we code an autonomy claim throughout because advocates of independence have always been in a minority position. If there is insufficient evidence as to what claim is dominant, we code the most radical claim (with autonomy < sub-state secession < independence < irredentism).

Note: The claim variable generally reflects the situation on January 1 of each year, with the following exceptions: i) if a country was not independent by January 1; ii) if an SDM group did not yet form part of a state by January 1; iii) in the first year of a movement's existence. In

²² The random sample does not contain any case of a group losing independence.

the latter three cases the claim variable reflects the situation at the date of independence, the date the territory was merged with the state in question, or the date of the first organized SD claim, respectively. We refer to this as the “first of January rule.”

Table 3: Claims

Type of SD claim	claim	Description
Autonomy	1	SDMs claim the establishment of territorial autonomy or an increase in the number of competencies if they are already autonomous, barring claims for sub-state secession.
Sub-state secession	2	SDMs make claims for the separation from a pre-existing autonomous entity and the establishment of an autonomous entity of their own or the merger with another autonomous unit. This scenario can only arise in federal or quasi-federal states and in states with autonomous regions. Switzerland’s Jurassians, who made claims for the separation from the canton of Bern, constitute a good example. ²³
Independence	3	SDMs claim internationally recognized independence.
Irredentism	4	SDMs make claims for secession and a subsequent merger with another state, typically their cultural motherland.

landborder

Dummy indicating whether the territory that is claimed by SDMs includes a land border with an internationally recognized country. In some cases different factions of the same movement make claims that vary in terms of their territorial scope. If so, the variable is coded on the basis of all land that is claimed. When the contours of the claimed territory are unclear, we relied on group settlement patterns, as indicated in GeoEPR (Wucherpfennig et al. 2011) or Minahan (2002), for example. The first of January rule applies.

seashore

Dummy indicating whether the territory that is claimed by SDMs has a sea outlet. While geographers use the term “sea” to refer to a “second rank body of saltwater,” we use the term more broadly to refer to all oceans (Atlantic, Pacific, Indian, Arctic, and Southern) as well as any large body of saltwater that is connected to the oceans (e.g. the Mediterranean Sea and the Black Sea). Access to the Caspian Sea is not coded because the Caspian Sea is not connected to the oceans. In some cases different factions of the same movement make claims that vary in terms of their territorial scope. If so, the variable is coded on the basis of all land that is claimed. When the contours of the claimed territory are unclear, we relied on group settlement patterns, as indicated in GeoEPR (Wucherpfennig et al. 2011) or Minahan (2002), for example. The first of January rule applies.

²³ A special case emerges for groups that make claims for autonomy within a larger autonomous unit. The sub-state secession category is reserved for cases of full separation from pre-existing autonomous units. Thus a claim for autonomy within a larger autonomous unit is coded as a claim for autonomy. Russia constitutes an example where such claims have been made.

hydrocarbon

Dummy indicating whether the territory that is claimed by SDMs includes hydrocarbon reserves (oil or gas) based on the PETRODATA by Lujala et al. (2007). Hydrocarbon reserves are considered from the year they were discovered. Both onshore and offshore reserves are included. Note that Lujala et al. use a 30 kilometers buffer around point estimates of hydrocarbon reserves and merge overlapping buffer polygons to a single polygon. Onshore reserves are coded if the claimed territory overlaps with one of the resulting polygons. Offshore reserves are considered if they would come to lie within the claimed territory's exclusive economic zone (EEZ) assuming it was independent. According to the United Nations Convention on the Law of the Sea, the EEZ "shall not extend beyond 200 nautical miles from the baselines from which the breadth of the territorial sea is measured." In some cases different factions of the same movement make claims that vary in terms of their territorial scope. If so, the variable is coded on the basis of all land that is claimed. When the contours of the claimed territory are unclear, we relied on group settlement patterns, as indicated in GeoEPR (Wucherpfennig et al. 2011) or Minahan (2002), for example. The first of January rule applies.

There are some caveats worth noting. First, the variable was hand-coded via visual inspection and we did not geo-reference the territory claimed by SDMs. Second, the 30 kilometer buffer around hydrocarbon reserves may lead to false positives, especially if groups are small. Third Lujala et al. record the date of discovery per polygon. Since polygons can represent combinations of many different hydrocarbon fields, it may be that the date of discovery does not apply to the portion of a polygon that overlaps with a movement's claimed territory. In other words, hydrocarbon may in some cases be coded 1 even if the reserve(s) in question were discovered only at a later point. Fourth, in some cases the date of discovery is missing in Lujala et al. If so, we relied on the date of discovery of a reasonably proximate reserve; if there is none, we assumed that oil/gas was discovered before SDMs' start dates. Finally, Lujala et al. cover the period until 2003. Thus, the variable does not account for oil/gas fields that were discovered after 2003.

groupcon

Dummy indicating whether an SDM group is spatially concentrated. A group is considered spatially concentrated if at least 50% of group members reside in a geographically contiguous territory where they make up at least 50% of the local population. The regional base must be larger than an urban area. The first of January rule applies.

kin

This variable indicates numerically significant trans-border ethnic kin.²⁴ Ethnic kinship is defined as a perception of ethnic bonds between ethnic groups (for our understanding of ethnicity see above). Kin groups have to be numerically significant, that is, they must have a population of at least 100,000. The variable is coded 0 if there is no transborder ethnic kin, 1 if there is at least one kin group in a non-adjacent country (but none in a neighboring country), and 2 if there is at least one kin group in a neighboring country. Countries are considered neighboring if they share a land border or if they are separated by a body of water less than 150 statute miles (241 kilometers) in width.

regaut

Dummy indicating whether an SDM group enjoyed regional autonomy in a given year. Federal states, such as Switzerland or India, are the most typical – but not the only – systems of regional

²⁴ We include kin groups in former colonies.

autonomy. Two conditions need to be jointly fulfilled for a regional autonomy code. First, there must be a regional government that operates below the national level but above the local level and commands over meaningful and significant autonomy. Meaningful autonomy requires that regional powers do not exist only on paper but that regional institutions can effectively exercise their powers. In some countries, the constitution grants far-reaching competencies to the regions but in practice political power remains highly centralized. This would not consider meaningful autonomy. The scope of autonomy is significant if it involves core competencies of the state. We would not consider this condition fulfilled if a regional government merely collects taxes or if it has competencies of an administrative nature (e.g. garbage collection). Significant regional power may involve decision-making as regards cultural rights (e.g. language and education) or the right to levy taxes.

The second condition requires that a group is significantly represented in the regional government and that group representation is not token, that is, group members must exert actual influence on the decisions of this entity. In operational terms, this means that an ethnic group must be significantly represented in the executive of a regional government and that its governance must not be directed by an outside authority. A token government that is appointed by and whose affairs are directed by the central government would not fulfill this condition, as there is no effective regional decision-making in line with the group's interests. The first of January rule applies. Note: Non-territorial forms of autonomy (see Coakley 1994) are not coded. Note: if a group is coded as de facto independent, this automatically implies regional autonomy.

defacto

Dummy indicating whether an ethnic group is de facto independent in a given year. De facto independence applies to SDM groups that have excluded themselves from the institutions of their host state and de facto exercise control over an ethnic group's homeland, or at least a significant share thereof. De facto control over small territories (e.g. rebel bases) is insufficient for us to code de facto independence. De facto independent states are not recognized by the international community. We do not require that there has been a formal declaration of independence, but typically de facto entities have declared independence. The first of January rule applies. Note: if a group is coded as de facto independent, this automatically implies regional autonomy.

pwrstat

This variable indicates the level of an ethnic group's access to central state power in a given year. The data is drawn from EPR for cases where they correspond to SDM groups (Cederman et al. 2010; Vogt et al. 2015).²⁵ All other cases were hand-coded relying on EPR coding rules. The EPR operationalization of access to central state power focuses on power access at the central state level of elite group members who claim to represent this given ethnic group. Critically, EPR looks only at access to executive power and does not take into account access to other branches of the central state, namely legislative and judicial institutions. Moreover, EPR focuses on access to those executive institutions where power is effectively exercised. Depending on the country, this may be the presidency, the cabinet, senior posts in the administration, and/or the army. Only major power shifts with substantial changes in the representation of a country's leadership are taken into account, disregarding day-to-day changes, such as cabinet reshuffles or the promotion of certain officer groups in the army. Importantly, the assessment of group access to executive power is absolute, meaning that it is independent of demographic over- or under-representation. Based on this definition, the pwrstat variable distinguishes between six levels of access to central state power. The typology (see Table 4), including definitions, follows EPR, except that groups that self-excluded themselves from central state power and erected a de facto independent state

²⁵ Note that EPR does not code ethnic groups' power access if the central state collapses (failed state). However, the state collapse rule concerns only very selected cases: Chad (1979), Liberia (1990-1996), and Sierra Leone (1993-1996, 1998-2001). None of these country-years form part of the random sample.

are coded as powerless. The de facto independence variable captures cases of self-exclusion. The first of January rule applies.

Table 4: Access to Central State

	Power status	Pwrstat	Description
Undivided power	Monopoly	1	Elite members hold monopoly power in the executive and all other ethnic groups are excluded (no SDM group has monopoly power).
	Dominance	2	Dominance is very close to the monopoly status; if a group is dominant it allows for some limited but ultimately politically irrelevant representation of “token” members of other groups (no SDM group has monopoly power).
Power-sharing	Senior partner	3	Power-sharing regimes occur if elite members of a given ethnic group share power with elite members of at least one other ethnic group. A group is classified as a senior partner if the representatives in question have at least as much executive influence as other powerful partners in the power-sharing arrangement.
	Junior partner	4	If elite members participate in a power-sharing arrangement but its influence and absolute representation is clearly lower compared to that of other participating groups.
Exclusion	Powerless	5	If group members do not hold executive power without being openly and systematically discriminated against.
	Discriminated	6	If group members are subject to explicit and targeted discrimination that effectively blocks their access to state power. Discrimination can be implemented through denial of political rights, including citizenship, or a systematic ban of ethnic parties.

excluded

Binary variable that identifies SDM groups that are excluded from central state power in a given year (that is, pwrstat = “Powerless” or pwrstat = “Discriminated”). The first of January rule applies.

groupsize

SDM groups' population relative to the country's total population. Data on group sizes is culled from EPR for SDM groups with a perfect equivalent in EPR (Cederman et al. 2010; Vogt et al.

2015). For other groups we generally drew on group size estimates provided by Minahan (2002). The first of January rule applies.²⁶

epr2sdm_scenario

This gives the scenario how EPR groups map onto SDM groups, distinguishing between “1:1”, “1:n”, “n:1”, “1945”, and “Irrelevant” (see above).

epr_group

Gives the name(s) of the EPR group(s) that are associated with an SDM group (if available).

epr_gwgroupid

Gives the EPR numeric group identifier(s) of EPR group(s) that are associated with an SDM (if available).

sovdec

Binary variable indicating whether the leaders of an SDM issued a unilateral declaration of sovereignty in a given year. Sovereignty declarations constitute unilateral acts that defy the existing institutional framework and in which SDM leaders publicly announce that the relationship with the state has been or is about to be unilaterally changed to their own advantage (Kahn 2000: 59). Sovereignty declarations as understood here occur without the consent of the state; typically the central government declares them invalid or illegal. Sovereignty declarations differ with regard to the claims that are made, ranging from autonomy over the separation from a pre-existing autonomous unit (sub-state secession) to independence and secession and subsequent merger with another state (irredentism). The form that the sovereignty declaration takes is irrelevant; sovereignty declarations may involve some sort of a written document, but it is also possible that sovereignty is declared in a speech in the context of a rally, for example. The originators of the declaration must represent a significant part of a movement. Declarations issued by marginal organizations that do not represent the SDM in question are not coded.

autdec

Binary variable indicating unilateral declarations of autonomy. Autonomy declarations are sovereignty declarations as defined above in which SDM leaders unilaterally announce the establishment of an autonomous sub-state within an existing state or an increase of the number of competencies of an already existing autonomous entity.

subdec

Binary variable indicating unilateral declarations of sub-state secession. Sub-state secession declarations are sovereignty declarations as defined above in which the SDM leaders unilaterally announce that a given entity is separated from an already autonomous entity to establish a separate autonomous unit. This scenario can only occur if a territory currently forms part of a larger autonomous unit.²⁷

²⁶ If group sizes are hand-coded, we report four digits after the comma as long as our sources allow, except if the group size is smaller than .00005 (then 5 digits after the comma).

²⁷ A special case emerges with declarations proclaiming an autonomous status within a larger federal unit. The sub-state secession category is reserved for cases of full separation from pre-existing federal units. Thus a declaration proclaiming autonomy within a larger federal unit is coded as an autonomy declaration.

inddec

Binary variable indicating unilateral declarations of independence. Independence declarations are sovereignty declarations as defined above in which SDM leaders unilaterally announce the formation of an independent state.

irrdec

Binary variable indicating unilateral irredentist declarations. Irredentist declarations are sovereignty declarations as defined above in which SDM leaders unilaterally announce their secession and subsequent merger with another state, typically the cultural motherland.

ucdp_conflictid

UCDP/PRIO armed conflict identifier. We merged all internal armed conflicts over territory recorded in v4_2014a of the UCDP/PRIO Armed Conflict Dataset to the SDM dataset, except for five conflicts that do not involve separatist aims as we define them.²⁸

ucdp_territoryname

The name of the territory over which the UCDP/PRIO internal armed conflict is fought.

ucdp_intensitylevel

The intensity level of an UCDP/PRIO internal armed conflict, with 1 = “Minor” (25-999 battle-related deaths in a given year) and 2 = “War” ($\geq 1,000$ battle-related deaths in a given year).²⁹

ucdp_cumulativeintensity

Binary variable indicating whether a UCDP/PRIO internal armed conflict has exceeded 1,000 battle-related deaths since the conflict onset.

ucdp_typeofconflict

The type of a UCDP/PRIO internal armed conflict, with 3 = “Internal armed conflict” and 4 = “Internationalized internal armed conflict.”³⁰

ucdp_terrconflict

Binary variable indicating a UCDP/PRIO internal armed conflict over territory, irrespective of the intensity level.

²⁸ The following five internal armed conflicts over territory identified by UCDP/PRIO could not be linked to an SDM: Israel-Hezbollah (over Southern Lebanon, id = “1-251”), China-Taiwanese insurgents (over Taiwan, id = “Jan 18”), South Vietnam – FNL (over South Vietnam, id = “Jan 52”), Oman – Free Oman (over Oman, id = “Jan 61”), and Malaysia – CCO (over North Borneo, id = “Jan 83”). Furthermore, UCDP codes an internal armed conflict over Palestine involving Israel and Palestinian insurgents from 1948 onwards (id = “1044”), while we code movement activity only from 1963, the year the PLO was formed, because there was no organized movement for self-determination before the formation of the PLO.

²⁹ In two cases (Chechens and Western Somalis), single group-years are associated with more than one UCDP/PRIO internal armed conflict over territory. If so, we used the highest intensity level, the highest cumulative intensity level, and the highest type of conflict (with internationalized internal armed conflict > internal armed conflict).

³⁰ By definition, extrasystemic armed conflict (1) and interstate armed conflict (2) do not apply to the SDM dataset.

ucdp_hterrconflict

Binary variable indicating a UCDP/PRIO high-level internal armed conflict (i.e. intensity level = 2).

ucdp_terrconflict_onset

Binary variable indicating the onset of a UCDP/PRIO internal armed conflict over territory, irrespective of the intensity level. Cases of ongoing armed conflict are dropped. In line with UCDP/PRIO practice (but contrary to our own practice, see “violsd_onset”), the onset coding is independent from SDMs' prior activities under a different host or before 1946. That is, we code onsets in the first year of movement coverage irrespective of whether there was a conflict in the last year. For example, we code an onset for the Estonians in 1946, the first year covered by UCDP/PRIO, even though the armed conflict started in 1944. Analogously, we code an onset for Azerbaijan's Armenians in 1991, the year of Azerbaijan's independence, even though there was violence already under the header of the Soviet Union.

ucdp_hterrconflict_onset

Binary variable indicating the onset of a UCDP/PRIO high-level internal armed conflict (i.e. intensity level = 2), dropping cases of ongoing high-level internal armed conflict. The comments under “ucdp_terrconflict_onset” apply analogously.

shortpeace_ucdp_terrconflict

Binary variable indicating periods of short peace based on the UCDP/PRIO list of internal armed conflicts over territory, irrespective of intensity level. The comments under “shortpeace_violsd” apply analogously.

shortpeace_ucdp_hterrconflict

Binary variable indicating periods of short peace based on the UCDP/PRIO list of high-intensity internal armed conflicts over territory (i.e. intensity level = 2). The comments under “shortpeace_violsd” apply analogously.

py_ucdp_terrconflict_phase

Phase-based peace years variable for UCDP/PRIO internal armed conflicts over territory, considering armed conflicts of both low and high intensity. The peace years accounts for prior internal armed conflicts over territory involving the same group under a different host and prior extrasystemic conflicts. Since UCDP/PRIO coverage starts only in 1946, we drew on alternative sources for the pre-1946 period. See “py_violsd_phase” for a more detailed discussion of the different versions of peace years variables.

py_ucdp_terrconflict_country

Country-based peace years variable for UCDP/PRIO internal armed conflicts over territory, considering armed conflicts irrespective of the intensity level. See “py_violsd_phase” for a more detailed discussion of the different versions of peace years variables.

py_ucdp_hterrconflict_phase

Phase-based peace years variable for UCDP/PRIO high-level intensity armed conflicts. See “py_violstd_phase” for a more detailed discussion of the different versions of peace years variables.

py_ucdp_hterrconflict_country

Country-based peace years variable for UCDP/PRIO high-level internal armed conflicts over territory. See “py_violstd_phase” for a more detailed discussion of the different versions of peace years variables.

accomm

Four-point ordinal variable indicating whether the state accommodated an SDM group in a given year, including concessions on cultural rights, autonomy, and independence, as well as concessions on access to central state power (see Table 5).

strictaccomm

An alternative, stricter variant of the accommodation variable. Concessions that were not implemented to a very high degree are not considered. Furthermore, the strict variant is coded with “Autonomy” only if an autonomy concession led to a very significant increase in a group’s level of autonomy; all other concessions on autonomy rights are coded as “Reform.” Concessions on independence are only considered if they led to internationally recognized independence or the merger with another state.

Table 5: Accommodation

Type of accommodation	accomm	Description
None	0	No accommodation.
Reform	1	Concession on cultural rights (cultcon = 1) or concession on central state access (change from pwrstat = discriminated/powerless to pwrstat = junior partner/senior partner or from pwrstat = junior partner to pwrstat = senior partner). ³¹
Autonomy	2	Autonomy concession (autcon = 1).
Independence	3	Independence concession (indcon = 1).

³¹ Note: Since the pwrstat variable relates to the situation on January 1 of a calendar year, it does not indicate concessions on access to central state power in the last year of movement activity. We manually added information for the last year of movement activity (for both phase I and phase II, where applicable).

References

- Banks, Arthur S., Alan J. Day, and Thomas C. Muller (1997). *Political Handbook of the World: 1997*. Binghamton, NY: CSA Publications.
- Cederman, Lars-Erik, Andreas Wimmer, and Brian Min (2010). "Why Do Ethnic Groups Rebel: New Data and Analysis." *World Politics* 62(1): 87-119.
- Coakley, John (1994). "Approaches to the Resolution of Ethnic Conflict: The Strategy of Non-Territorial Autonomy." *International Political Science Review* 15(3): 297-314.
- Correlates of War Project (2011). "State System Membership List, v2011." <http://correlatesofwar.org> [May 25, 2015].
- Degenhardt, Henry W. (ed.) (1988). *Revolutionary and Dissident Movements*. London: Longman Publishers.
- Doyle, Michael W., and Nicholas Sambanis (2006). *Making War and Building Peace: United Nations Peace Operations*. Princeton, NJ: Princeton University Press.
- Gleditsch, Kristian S., and Michael D. Ward. 1999. "A Revised List of Independent States Since the Congress of Vienna." *International Interactions* 25(4): 393-413.
- Gleditsch, Nils P., Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Havard Strand (2002). "Armed Conflict 1946-2001: A New Dataset." *Journal of Peace Research* 39(5): 615-637.
- Gurr, Ted R. (1993). *Minorities at Risk: A Global View of Ethnopolitical Conflict*. Washington, DC: U.S. Institute of Peace.
- Gurr, Ted R. (2000). *Peoples versus States*. Washington, DC: U.S. Institute of Peace.
- Gurr, Ted R., Monty G. Marshall, and Deepa Khosla (2001). *Peace and Conflict 2001: A Global Survey of Armed Conflicts, Self-Determination Movements and Democracy*. College Park, MD: Center for International Development and Conflict Management.
- Hewitt, Christopher, and Tom Cheetham (2000). *Encyclopedia of Modern Separatist Movements*. Santa Barbara, CA: ABC-CLIO.
- Hewitt, Joseph J., Jonathan Wilkenfeld, and Ted R. Gurr (eds.) (2008). *Peace and Conflict 2008*. Boulder, CO: Paradigm Publishers.
- Horowitz, Donald L. (1985). *Ethnic Groups in Conflict*. Berkeley, CA: University of California Press.
- Kahn, Jeffrey (2000). "The Parade of Sovereignities: Establishing the Vocabulary of the New Russian Federalism." *Post-Soviet Affairs* 16(1): 58-89.
- Keesing's Record of World Events. <http://www.keesings.com> [May 25, 2015].
- Khosla, Deepa (2005). "Self-Determination Movements and Their Outcomes." In: Monty G. Marshall and Ted R. Gurr (2005), *Peace and Conflict 2005: A Global Survey of Armed Conflicts, Self-Determination Movements, and Democracy*, 21-27. College Park, MD: Center for International Development and Conflict Management.
- Lexis Nexis. <http://www.lexis-nexis.com> [May 25, 2015].
- Lujala, Päivi, Jan Ketil Rød, and Nadia Thieme (2007). "Fighting over Oil: Introducing a New Dataset." *Conflict Management and Peace Science* 24(3): 239-256.
- Marshall, Monty G., and Ted R. Gurr (2003). *Peace and Conflict 2003: A Global Survey of Armed Conflicts, Self-Determination Movements and Democracy*. College Park, MD: Center for International Development and Conflict Management.

- Marshall, Monty G., and Ted R. Gurr (2005). *Peace and Conflict 2005: A Global Survey of Armed Conflicts, Self-Determination Movements, and Democracy*. College Park, MD: Center for International Development and Conflict Management.
- Minahan, James (1996). *Nations without States: A Historical Dictionary of Contemporary National Movements*. London: Greenwood Press.
- Minahan, James (2002). *Encyclopedia of the Stateless Nations*. Westport, CT: Greenwood Press.
- Minorities at Risk Project (2009). College Park, MD: University of Maryland.
- Minority Rights Group International (1997). *World Directory of Minorities*. London: Minority Rights Group International.
- Minority Rights Group International. *World Directory of Minorities and Indigenous Peoples*. <http://www.minorityrights.org/744/directory/world-directory-of-minorities-and-indigenous-peoples.html> [May 25, 2015].
- National Geographic Society (United States) (1999). *National Geographic Atlas of the World*. Seventh Edition. Washington, DC: National Geographic Society.
- Quinn, David (2008). "Self-Determination Movements and Their Outcomes." In: Joseph J. Hewitt, Jonathan Wilkenfeld, and Ted R. Gurr (eds.), *Peace and Conflict 2008*, 33-38. Boulder, CO: Paradigm Publishers.
- Quinn, David, and Ted R. Gurr (2003). "Self-Determination Movements: Origins, Strategic Choices, and Outcomes." In: Monty G. Marshall and Ted R. Gurr, *Peace and Conflict 2003: A Global Survey of Armed Conflicts, Self-Determination Movements and Democracy*, 26-38. College Park, MD: Center for International Development and Conflict Management.
- Sambanis, Nicholas (2004). "What Is Civil War? Conceptual and Empirical Complexities of an Operational Definition." *Journal of Conflict Resolution* 48(6): 814-858.
- Sambanis, Nicholas, and Jonah Schulhofer-Wohl (forthcoming). "Civil War as Sovereignty Rupture: Coding Intra-State Conflict, 1945-2012." Yale University: Unpublished manuscript.
- Sarkees, Meredith R., and Frank W. Wayman. *Resort to War. A Data Guide to Inter-State, Extra-State, Intra-State, and Non-State Wars, 1816-2007*. Washington, DC: CQ Press.
- Themnér, Lotta, and Peter Wallensteen (2014). "Armed Conflict, 1946-2013." *Journal of Peace Research* 47(2): 243-250.
- United Nations (1982). *United Nations Convention on the Law of the Sea*. http://www.un.org/depts/los/convention_agreements/texts/unclos/unclos_e.pdf [November 2, 2015].
- United States Central Intelligence Agency (2002). *The 2002 CIA World Factbook*. <https://www.cia.gov/library/publications/download/download-2002/> [May 25, 2015].
- Unrepresented Nations and Peoples Organization (UNPO) (2014). "Members." <http://unpo.org/members.php> [May 25, 2015].
- 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 Data Set Family." *Journal of Conflict Resolution* 59(7): 1327-1342.
- Wood, John R. (1981). "Secession: A Comparative Analytical Framework." *Canadian Journal of Political Science* 14(1): 107-134.
- Wucherpennig, Julian, Nils B. Weidmann, Luc Girardin, Lars-Erik Cederman, and Andreas Wimmer (2011). "Politically Relevant Ethnic Groups across Space and Time: Introducing the GeoEPR Dataset." *Conflict Management and Peace Science* 28(5): 423-437.