SDM: A new dataset on self-determination movements with an application to the reputational theory of conflict

Nicholas Sambanis* Micha Germann* Andreas Schädel†‡

September 14, 2017

Abstract

This paper presents a new dataset on self-determination movements (SDMs) with universal coverage for the period from 1945–2012. The dataset corrects the selection bias that characterizes previous efforts to code SDMs and significantly expands coverage relative to the extant literature. For a random sample of cases, we add information on state-movement interactions and several attributes of SDM groups. The data can be used to study the causes of SDMs, the escalation of SD conflicts over time, and several other theoretical arguments concerning separatist conflict that have previously been tested with incomplete or inferior data. We demonstrate the usefulness of the new dataset by revisiting Barbara Walter's influential argument that governments will not accommodate SD challengers if they face several potential future challengers down the road because they want to build a reputation for strength. We do not find support for Walter's reputational theory of separatist conflict.

*Department of Political Science, University of Pennsylvania, Philadelphia, PA.

†Center for Comparative and International Studies (CIS), ETH Zurich, Zurich, Switzerland.

‡We thank the editor, the two anonymous reviewers, Lars-Erik Cederman, Simon Hug, Brian Philipps, Jonathan Wheatley, and conference participants at MPSA 2016, APSA 2015, and the 2015 European Network of Conflict Research (ENCoRe) Meeting in Barcelona, Spain, for their helpful comments. Germann and Schädel gratefully acknowledge funding through grants №135127, №143213, and №162220 by the Swiss National Science Foundation. We also want to thank Piia Bränfors, Nora Keller, Shelley Liu, and Rahel Zürrer for valuable research assistance, as well as Seraina Rüegger for her help with several Eastern European cases. Annalisa Zinn contributed significantly to an earlier version of the dataset presented in this article.

1 Introduction

Why do groups make claims for more self-determination? When do such claims lead to violence? When do states accommodate them? To address these important questions, a vast literature on ethno-nationalist conflict has drawn insights from both qualitative (e.g. Hechter 2000, Horowitz 1985) and quantitative data (e.g. Ayres & Saideman 2000, Hale 2000, Jenne, Saideman & Lowe 2007, Roeder 2007, Sambanis & Milanovic 2014). The approaches are varied: some focus on opportunity structures that favor the organization of violent challenges to the state (e.g. Cunningham 2013*b*, Walter 2006*b*), while others focus on institutional determinants of separatism (e.g. Beardsley, Cunningham & White 2017, Brancati 2006, Sambanis 2000), or the conditions leading to successful separatist challenges (e.g. Coggins 2011, Cunningham 2011, Walter 2003, Walter 2006*a*).

These studies make clear that separatist conflict is a persistent feature of world politics. Violent separatism spiked after the end of the Cold War; new states were created in East Timor and South Sudan; *de facto* independent regions emerged in Azerbaijan, Georgia, and Somalia; and separatist sentiment is fueling conflicts in Eastern Ukraine, Turkey, Catalonia, and Scotland, among other places. Peripheral demands for increased autonomy or secession have been a cause of more than a third of all civil wars since 1945 (Sambanis & Milanovic 2014).

To understand these cases, insights must be sought through the analysis of high-quality large-N data on self-determination movements (SDMs). Quantitative data on SDMs have been collected globally, identifying the claimants, targets, and dates of separatist conflict. The two main sources of data are the Minorities at Risk project (MAR) (Gurr 1993, Gurr 2000*b*, Minorities at Risk Project 2009) and the Peace and Conflict reports published biennially by the Center of International Development and Conflict Management (CIDCM) (e.g. Marshall & Gurr 2003, Marshall & Gurr 2005).[1] Many studies have drawn on these datasets (e.g. Ayres & Saideman 2000, Siroky & Cuffe 2015), and some have added variables to them (e.g. Cunningham 2014, Kelle 2017, Walter 2006*a*).[2]

[1] The CIDCM data are an offspring of MAR, but include groups that are not in MAR.

[2] Cunningham (2014), for example, added information on the number of factions within each movement and Walter (2006*a*) on whether governments accommodated separatist movements.

Unfortunately, coverage in both MAR and CIDCM is incomplete. All datasets building on MAR and CIDCM in the extant literature are biased in favor of including cases of highly mobilized, violent SDMs in less developed and to some extent less democratic countries. For many research questions, this skewed data coverage will translate into biased causal inferences about separatism.

We are not the first to recognize this problem, but the solutions offered to date have limitations. Sorens (2012, chapter 2) added around two dozen cases to MAR. But as we show later, MAR misses many more cases than that. Coggins (2011) and Griffiths (2015) created new datasets with somewhat broader coverage, including colonial cases that are excluded from MAR. However, they only include cases of secessionist claims and exclude movements making more limited claims for autonomy. Our dataset reflects an understanding of self-determination as a variable ranging from minor to major claims. The intensity of many SDMs changes between autonomism and secessionism depending on underlying conditions. If we coded exclusively secessionist cases, we would not be able to explain escalation processes. Our approach reduces the risk of selection effects in the study of separatist war onset as we include violent secessionist movements that started out as non-violent claims over more limited forms of autonomy.

Our research leads us to code a total of 464 movements in 120 countries from 1945-2012. This is more than double the number of SDMs in the MAR and CIDCM datasets. We distinguish between violent and non-violent claims and, for a randomly drawn sample of 106 movements, we also code details of government policy responses to SDMs as well as other variables that allow us to study escalation processes. We provide extensive documentation, including justifications for all coding decisions.

2 The SDM dataset

The SDM dataset is divided in two parts. The first part identifies SDMs around the world from 1945–2012, pinpoints the years during which those movements were active, and distinguishes between violent and non-violent periods of activity. The second part

provides more detailed information on state-movement interactions over time and includes information about attributes of claimants for a randomly drawn sample of cases.

2.1 Cross-country data on SDMs, 1945-2012

2.1.1 What is an SDM?

SDMs are movements by one or more political organizations that are connected to an ethnic group and make claims for increased self-determination from the state. We elaborate on elements of this definition below.

Although SD may be exercised in all spheres of social life, we refer strictly to a group-defined right to self-rule within the boundaries of a territory. SD thus 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; indigenous land rights; the power to define the territorial boundaries of the state within which the group exercises its right to self-determination; and, of course, national independence. This understanding of SD is therefore close to non-absolute definitions of sovereignty (see e.g. Krasner 1999).

Based on this broad conceptualization of SD, we code both movements demanding national independence (such as the Scots in the UK) and movements demanding their merger with another state (e.g. Serbs in Bosnia), but also movements that merely demand more internal autonomy (e.g. Mayans in Mexico). That said, we only code autonomy claims if they exceed a certain level of political significance: they must imply a significant redefinition of a state's institutional set-up. Federal states, in particular, are characterized by constant negotiations over the extent of policy centralization, so we do not code every claim made by regional governments in federal states. To give an example from the US context, we do not code claims that abortion or gay rights should be decided at the state level. By contrast, we code the Yakuts, a group in northeastern Siberia that has been claiming a radically higher level of autonomy, including greater control over natural resources in their region, since the early 1990s. We also include Switzerland's Jurassians, who pushed for the establishment of a new canton of their own. Both coding decisions

are consistent with MAR and CIDCM, though they miss many such cases.

SD claims as we understand them must be directed against a state. Thus, we do not code movements that demand increased policy autonomy from a supranational entity, such as the BREXIT movement. Also, because our definition of an SD claim entails the logic of political disintegration, we do not code movements to overthrow a government or gain increased representation at the center. Because these are aimed at political integration, we also do not include movements aimed at the merger of nation-states, such as the pan-African movement or the *enosis* movement in Cyprus. Finally, we do not include efforts to preserve the current level of SD (contrary e.g. to Hale 2008, p. 3) as the movement in the early 1990's in Moldova opposing unification with Romania.

For a case to be coded, there must be evidence of mobilization; of activity by a political organization. We must be able to identify a political party, cultural organization, armed group, or protest movement. The mobilization may take violent or non-violent forms, and may be both institutional (e.g. running for office) or extra-institutional (e.g. protests). Separatist sentiment alone is not enough. For example, although nationalist sentiment was present among Yugoslavia's Bosniaks long before the 1990s, we only code an SDM after Bosniak nationalism was expressed with the formation of the Party of Democratic Action (Stranka Demokratske Akcije (SDA)) in May 1990.[3]

The literature on autonomy and secession has focused almost exclusively on claims made by ethnic groups. We follow this practice, though we have a broader concept of ethnic identity that is defined not only by the usual descent-based attributes (race, religion, language), but also region of origin. Thus, several of our ethnic groups are regionally-defined groups. This conceptualization of ethnicity is consistent with Horowitz (1985) and allows us to capture more cases. Shared regional identity can be sufficient to define group differences that underlie autonomy movements, as was the case in the U.S. civil war. Our approach is consistent with MAR, which states that it includes groups defined by their "region of residence" (Gurr 2000*b*, p. 4) and codes a number of claims by groups that can be considered ethnic only by virtue of their shared regional identity,

[3]We do not include fringe organizations with very few followers.

such as the Crimean Russians in Ukraine or the Jurassians in Switzerland.[4]

Finally, we do not include colonial liberation movements.[5] This is not a definitional requirement—anti-colonial movements constitute SDMs according to our definition—but a choice made for practical reasons. The experience of anti-colonial movements is less relevant for contemporary separatism; and finding relevant information is harder for these cases. Moreover, MAR and CIDCM both exclude anti-colonial movements, as is customary practice.

2.1.2 Coding the global incidence of SDMs

The conceptualization discussed above leads us to code a total of 464 SDMs between 1945 and 2012. We relied upon a broad range of sources to identify SDMs. Existing datasets on SD disputes constituted a good starting point for coding, in particular MAR and CIDCM. We did not, however, take extant coding decisions for granted and dropped cases included in these datasets if we found no corroborative evidence of separatist mobilization. Examples of groups that were excluded based on this include the Bahais in Iran, Blacks in Ecuador, and Germans in Romania. Many cases that were missing from MAR and CIDCM were identified using three encyclopedic sources on separatism (Hewitt & Cheetham 2000, Minahan 1996, Minahan 2002). In addition, we consulted the World Directory of Minorities and Indigenous Peoples (Minority Rights Group International 2015), Degenhardt (1988), material provided by the Unrepresented Nations and Peoples Organization (UNPO), and the UCDP/PRIO database of armed conflicts (Gleditsch, Wallensteen, Eriksson, Sollenberg & Strand 2002, Themnér & Wallensteen 2014). Further, news archives were searched for keywords including self-determination, self-governance, self-rule, regionalism, separatism, and secessionism (in combination with country names). Finally, coders consulted a wealth of region- and country-specific academic literature (refer to the online supplement for additional information on sources and case histories).

[4]The CIDCM data also include regionally defined groups, such as the Cornish, the Trans-Dniester Slavs, the Jurassians, and the Crimean Russians.

[5]We define a colony as a dependent territory that is neither politically nor legally integrated with the metropole. Note that we include movements in overseas territories that are fully integrated with the metropole, such as Guadeloupe.

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). As is standard in the literature, we code separate SDMs if the same ethnic group organizes for SD in multiple countries. For example, we code separate SDMs for the Kurds in Iraq and Turkey. 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, a movement included in our dataset may consist of multiple organizations making claims on behalf of the same group, and the claims of the different organizations may vary in intensity. For example, we code a single Catalan SDM in Spain even though this movement has both organizations claiming increased autonomy and organizations claiming national independence.

For all 464 movements, we record the year in which the first organized SD claim was made (the movement start date). If we find evidence that a movement ceased to exist or abandoned its demands for increased SD, we code an end to the movement in the respective year (the end date). 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. Coding the date a movement ends is not always straightforward. Events that would qualify for movement termination are not reported systematically and in some cases movement activity stops very gradually. We therefore 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 cutoff rule is arbitrary, we chose a long time horizon so as not to omit non-violent groups with discontinuous activity, a low public profile due to fear of government repression, or groups in their early stages. If a movement ends and then restarts, we code

Table 1: Basic descriptives

SDMs	464
Countries	120
Average years of activity	30.15
Ongoing SDMs (2012)	327
SDMs with discontinuous activity	36
Violent SDMs	150
Average years until first violence	6.55

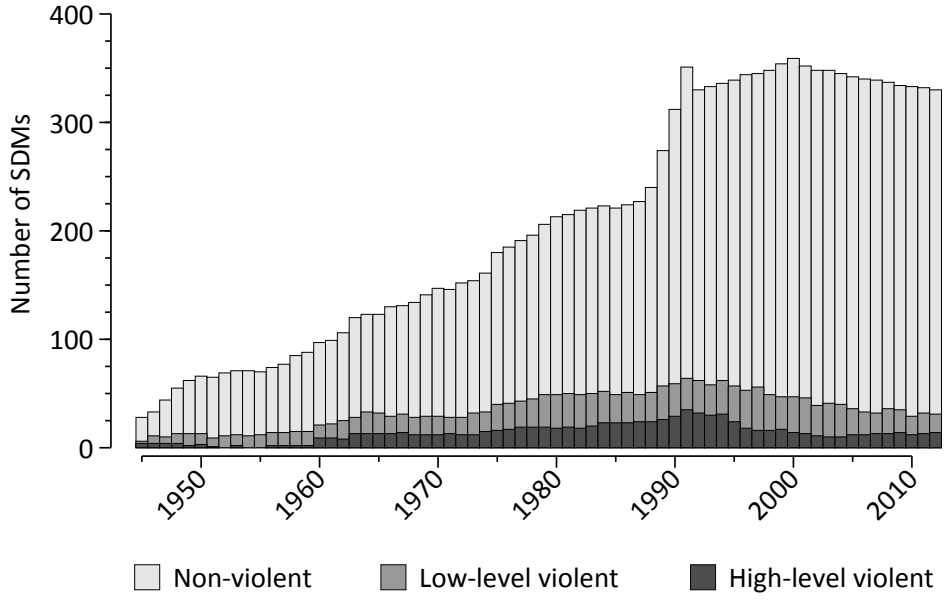
a second period of activity. For example, we code two phases of activity for the Estonians in the former USSR, with the first corresponding to the partisan resistance campaign in the immediate post-WWII phase (ending in 1956) and the second starting in 1987 in the context of glasnost and perestroika, ending with Estonia's independence (1991). No movement in our dataset had more than two phases of activity. Movements which were ongoing as of 2012, the last year we cover, are denoted as "ongoing." See Table 1 for a brief summary.

A first look at the data reveals a number of interesting facts. Figure 1 shows that the number of SDMs increased steadily and almost linearly until the late 1980s, with a spike occurring due to the breakdown of the USSR and of ethnofederal states in Eastern Europe (the number of movements in the former USSR increased from 5 in the early 1980s to as many as 68 in 1991). Since the early 1990s, the number of SDMs has stabilized and we even observe a small downward trend in recent years, though the overall number remains high (our 10-year rule probably contributes to the persistence of SDMs). The average duration for an SDM is over 30 years, revealing that autonomy concessions are hard-won or that new claims often escalate even after governments make concessions. More than 70% of all movements we cover are ongoing as of 2012.[6]

Looking at regional patterns (see Figure 2) the most striking observation is the explosion of separatist movements in the former USSR in the late 1980's/early 1990's. Many of the movements born under the crumbling Soviet regime have continued to operate in the USSR's successor states, including (as of 2012) 29 in Russia, 4 in Georgia, 4 in Ukraine, 2

[6]In addition, about 15% of the cases coded as having "ended" are due to country break-ups, while the respective groups continued to claim self-rule in a different host state. Compare the example of the Hungarians in Czechoslovakia/Slovakia mentioned earlier.

Figure 1: SDMs by calendar years distinguishing between violent and non-violent claims

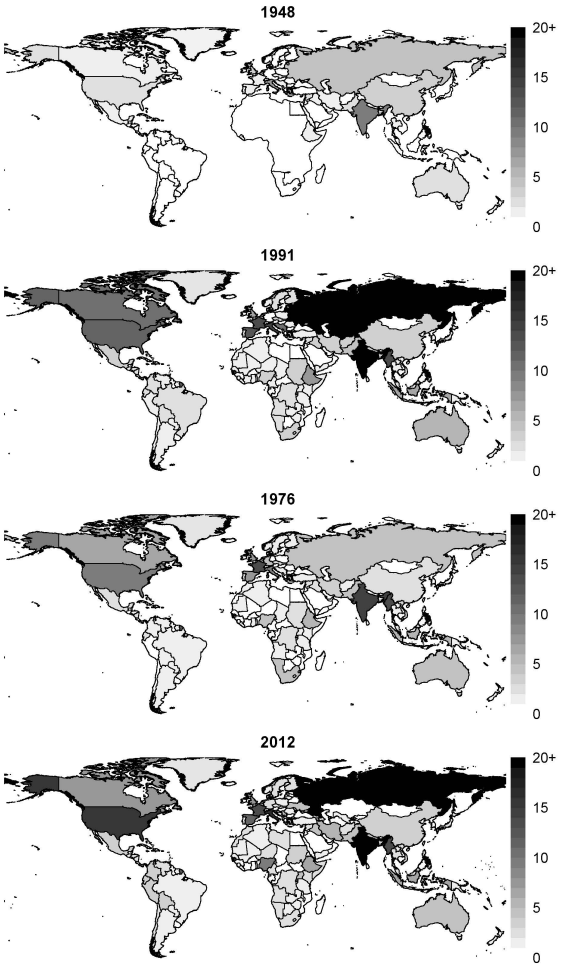


in Moldova, and 1 in Azerbaijan. Southern and Eastern Asia is another region with high separatist activity, in particular India (22 SDMs in 2012, up from 9 at independence), Myanmar (13 in 2012, up from 7), and Indonesia (6, up from 4).

Separatism has increased in a number of other regions. Western Europe harbored only 15 movements in the 1940s, but more than 50 in 2012. Italy has the highest number (16) but Spain saw an even more spectacular increase in SDMs, up from a single movement in the 1940s to 11 in 2012. France also has its share of separatism, though almost half of its 13 movements are in overseas territories, such as Guadeloupe. In the Americas, we code 60 active movements in 2012, up from 6 movements in the late 1940's. To a large extent, this increase is due to a wave of mobilization among indigenous groups starting in the 1970's. Meanwhile, SD activity remains absent from many African and Middle Eastern countries, though with some notable exceptions (Nigeria, Ethiopia, South Africa, and Uganda in Africa; Iran and Iraq in the Middle East).

Our data allow us to distinguish between violent and non-violent movements as well as between spells of low and high-intensity violence. Our coding of high-intensity violent SDMs follows Sambanis's (2004) definition of a civil war and identifies periods of high violence over SD demands based on data from Doyle & Sambanis (2006), as updated in

Figure 2: Number of SDMs by countries



Sambanis & Schulhofer-Wohl (2015). Our preferred measure of low-intensity separatist violence is an inclusive category that incorporates any of the definitions of low-level violence currently in use, drawing data from UCDP/PRIO (Gleditsch et al. 2002, Themnér & Wallensteen 2014), MAR (Minorities at Risk Project 2009), and the Peace and Conflict reports (in particular Marshall & Gurr 2003, Hewitt, Wilkenfeld & Gurr 2008), but also news reports and case studies.[7] Again, interesting patterns emerge. Figure 1 shows clearly that at any given point in time, violent SDMs are a small share of the total number of cases. Much of the quantitative literature on separatism has focused on violent movements because they are more destructive and possibly because they are easier to code. A broader perspective on separatism can emerge from our data by virtue of including non-violent cases; and the transition from nonviolence to violence can now be studied.

[7]We have coded a second set of violence indicators based exclusively on UCDP/PRIO data, which allows for quick robustness checks.

We note that most violent movements start non-violent. On average, it takes more than 6 years for violence to erupt (see Table 1). Finally, it is noteworthy that violent SD claims have declined since the early 1990's, consistent with predictions made by Gurr (2000*a*) about the decline of ethnic war.

2.2 Coding conflict dynamics for a random sample of movements

Part two of the SDM dataset includes more detailed information on state-movement interactions and attributes of SDM groups for a randomly selected subset of SDMs. New research was required to code many of these variables. Coding state-movement interactions over time for all 464 SDMs on our list was not practical, so we selected cases randomly to provide more in-depth data on conflict escalation. We over-sample SDMs that switched from non-violence to violence (we call these "switched" movements) since these cases contain more information about escalation than movements that started out as violent or non-violent cases. The sample consists of 106 SDMs, 59 of which were violent (we include sampling weights to correct for over-sampling switched movements).[8]

Our main focus is on how states react to groups making SD claims. We code new data on several types of government concessions and restrictions: policies by the state that increase or decrease a group's level of self-determination, including both autonomy and national independence, as well as policies that increase or decrease a group's cultural rights, such as language or religious rights. While the data on non-violent restrictions constitutes an original contribution to the literature, there exist other data on government concessions to SDMs (Cunningham 2014, Walter 2006*a*). However, this data is only available for the biased sample of SDMs contained in CIDCM (see below). We provide broader coverage and, by reviewing a wider set of sources than any previous study, we identify many previously overlooked government actions and correct other coding errors in the extant literature. A detailed discussion of our approach to coding government

[8]We exclude countries with a population of less than 500,000, so 6 movements in 6 micro-states were dropped before sampling: Antigua and Barbuda, St. Kitts and Nevis, Belize, Vanuatu, Kiribati, and Micronesia.

concessions to SDMs is explained in section 5.

For the random sample of cases, we also provide information on a number of other variables, including:

- *Type of SD claim:* Do SDMs demand autonomy or secession (national independence or the merger with a different state)?
- *Sovereignty declarations:* Do SDM leaders unilaterally declare autonomy or secession?
- *Central state access:* Do SDM groups have meaningful representation in the central state's executive?
- *Regional autonomy:* Do SDM groups have a meaningful level of regional autonomy?
- *De facto independence:* Are SDM groups *de facto* separated from their host states, but not *de jure*?
- *Relative group size:* What is the population of SDM groups relative to the country's total population?
- *Spatial concentration:* Are SDM groups territorially concentrated?
- *Ethnic kin:* Do SDM groups have transborder ethnic kin?
- *Hydrocarbons:* Does the territory claimed by SDMs contain hydrocarbon reserves (oil or gas)?
- *Land border:* Does the territory claimed by SDMs contain an international land border?
- *Seashore:* Does the territory claimed by SDMs have a sea outlet?

All variables are coded at the group-year level. Sources consulted to code these variables include the previously mentioned encyclopedic sources (Hewitt & Cheetham 2000, Minahan 1996, Minahan 2002), as well as the World Directory of Minorities and Indigenous Peoples (Minority Rights Group International 2015). However, in most cases we also consulted case-specific sources listed in our supplement. Data on hydrocarbon

Table 2: Descriptives for random sample variables

Concessions:	303	**Sovereignty declarations:**	56
Cultural rights concessions	51	Autonomy declarations	18
Autonomy concessions	242	Secession declarations	38
Independence concessions	10	SDMs with at least one declaration (%)	38.64
SDMs with at least one concession (%)	80.07	SDM groups with central state access (%)	44.26
Restrictions:	194	SDM groups with regional autonomy (%)	50.94
Cultural rights restrictions	44	SDM groups with de-facto independence (%)	9.06
Autonomy restrictions	143	SDM groups' relative share of country population (%)	7.33
Independence restrictions	7	Spatially concentrated SDMs groups (%)	75.23
SDMs with at least one restriction (%)	52.47	SDM groups with transborder ethnic kin (%)	67.6
Type of SD claim:		**Claimed territory:**	
SDMs with secession claim (%)	46.2	Contains hydrocarbon reserves (%)	51.45
SDMs with secession claim in first year (%)	31.45	Contains international land border (%)	64.29
		Contains seashore (%)	47.19

Note: Absolute figures represent counts whereas all relative figures are aggregated to the group-level and weighted to account for the oversampling of switched movements. The same group can receive more than one concession and/or restriction over time.

reserves draws on Lujala, Rød & Thieme (2007), whereas data on group sizes and inclusion in central government was culled from the Ethnic Power Relations (EPR) dataset (Vogt, Bormann, Rüegger, Cederman, Hunziker & Girardin 2015) for all cases where the SDM group corresponded to an EPR group; for all other cases we conducted original research with details relegated to our codebook and coding notes.

Table 2 presents summary statistics for the random sample. Most SDM groups are spatially concentrated and relatively small. The average SDM group makes up only around 7% of its host state's population. More than half of all SDMs occupy territory that contains hydrocarbons; and again about half are represented in the central government for at least part of the time of an active claim. Finally, whereas almost half of all SDMs claim secession at some point, only a third of them start out as secessionist. Thus, group claims escalate over time, probably in response to state-group interactions.

Based on our research, we code a large number of concessions (303) and restrictions (194), mostly over autonomy. Figure 3 shows how long it takes until the state makes the first concession or restriction after a group makes an SD claim. Clear differences emerge with respect to country GDP: the higher a country's income, the more likely an SDM is to get a concession. A movement's chances to get a concession after five years of activity are almost 50% in rich countries, but only 30% in medium-income countries and much lower in low-income countries. The risk of a restriction is lower overall, but high-income countries are less likely to restrict group rights. Autocracies are more likely to restrict

group rights than democracies and less likely to accommodate SD claims. Interestingly, anocracies are more likely than democracies to make concessions in the first five years of a challenge; but the pattern reverses and after about 10 years democracies are more likely to make concessions. At the same time, however, anocracies are also more likely to restrict group rights—even compared to autocracies. The data suggest that anocracies have an incoherent institutional commitment to pluralism. Finally, the third panel of plots shows differences related to violence. Movements that had violently challenged the state in the past are somewhat less likely to be granted a concession. But prior violence does not seem to affect the likelihood of the first restriction.[9]

3 The problem of incomplete case coverage

Next, we consider the issue of incomplete case coverage in existing lists of separatist movements. We include three such lists: the MAR dataset as well as the 2003 (Marshall & Gurr 2003) and the 2008 (Hewitt, Wilkenfeld & Gurr 2008) versions of the CIDCM data (CIDCM stopped publishing data on SDMs in 2008). The former two have been used widely in academic research (e.g. Ayres & Saideman 2000, Cunningham 2011, Cunningham 2013*a*, Siroky & Cuffe 2015, Walter 2006*a*).[10]

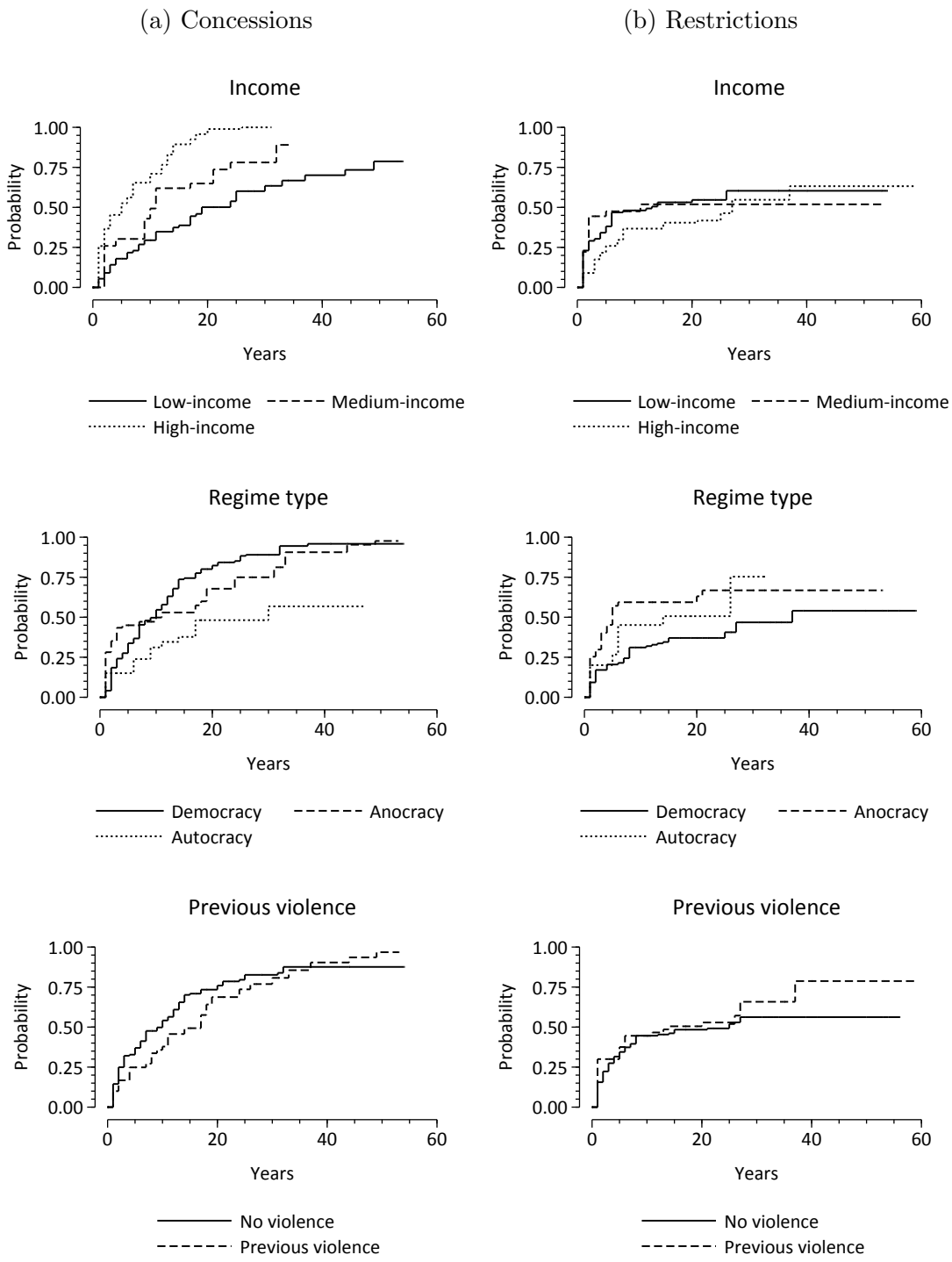
Simple counts are sufficient to highlight important differences between these datasets and ours (see Table 3). The SDM dataset has much broader coverage. While MAR identifies a total of 177 separatist groups in 81 countries, the SDM dataset in the same time frame (1945–2006) identifies more than double that number of separatist groups (459) in 118 countries. Moreover, MAR contains many cases (14) that our research suggests should not be counted (see supplement). The comparison to the other two datasets looks similar. CIDCM (2008), for instance, list only 175 movements in 82 countries.

What accounts for these differences? We argue that coding error is the main reason,

[9]The conclusions remain similar when looking at all concessions and restrictions and not only the first (see supplement).

[10]We do not include the more recently compiled datasets by Coggins (2011) and Griffiths (2015) due to their narrower focus on secession claims. We also do not include Sorens's (2012) extension of MAR because while Sorens includes autonomist groups, he excludes SDMs that explicitly rule out independence, which makes cross-dataset comparisons less meaningful.

Figure 3: Kaplan-Meier estimates of time until first concession/restriction by regime type, income, and previous violence



Note: Weighted estimates. A movement is considered to be located in a democracy if the host state's polity2 score is 6 or higher, in an anocracy if the host state's democracy score ranges between -5 and 5, and in an autocracy if the country's polity2 score is -6 or lower. In turn, a movement is considered to be located in a high-income country if the host state's GDP per capita is at least 10,000 constant US dollars (base: 2005 prices), in a medium income country if the GDP per capita is in-between 4,000 and 9,999 constant US dollars, and a low income country if GDP per capita is below 4,000 constant US dollars.

Table 3: Incomplete coverage in other datasets

	MAR		**CIDCM (03)**		**CIDCM (08)**		**All combined**	
	MAR	SDM	CIDCM	SDM	CIDCM	SDM	Combined	SDM
Period	1945–2006		1955–2002		1955–2006		1945–2006	
SDMs	177	459	148	450	175	456	237	459
Countries	81	118	78	117	82	118	98	118
Regional distribution:								
MENA	20	22	8	19	10	20	20	22
SS Africa	36	73	33	72	34	73	46	73
Central Asia	21	55	19	55	20	55	27	55
SE Asia	34	92	37	91	47	91	52	92
Europe	41	127	34	126	41	127	56	127
North America	9	27	5	26	7	27	11	27
Latin America	12	41	9	39	10	41	18	41
Oceania	4	22	3	22	6	22	7	22
Violent SDMs (%)	57.63	32.24	54.73	30.44	54.86	30.48	52.32	32.24
SDMs in democracies (%)	27.27	39.82	31.97	39.19	39.08	40.09	34.47	39.82
SDMs in high-income countries (%)	10.73	26.8	15.54	27.33	18.86	27.85	15.61	26.8

Note: A movement is considered violent if there was an incidence of low- or high-level violence over SD in at least one year. A movement is considered to be located in a democracy if the host state's polity2 score is 6 or higher in 75% or more of all years it was active. In turn, a movement is considered to be located in a high-income country if the host state's GDP per capita is at least 10,000 constant US dollars (base: 2005 prices) in 75% or more of all active years. MENA stands for Middle East and Northern Africa, SS Africa for Sub-Saharan Africa, and SE Asia for Southern and Eastern Asia. Refer to the supplement for more details on how we compared the different datasets.

though differences in coding rules also matter. First, there are differences due to the aggregation of some groups to larger entities. This applies mainly to indigenous groups. For example, while both MAR and CIDCM code a single Native American movement in the US, we code 5 separate groups (Cherokee, Dine (Navajo), Iroquois, Lakota (Sioux), and Pueblos). In addition, the other datasets systematically exclude some groups that we include. MAR only includes groups with a population of more than 100,000 or more than 1% of a country's population; and it only covers countries with a population of more than 500,000. By contrast, we code small groups, such as the Toubou in Niger, the Sumos in Nicaragua, the Iroquois in Canada, and the Cherkess in Russia. We also code a total of 6 movements in micro-states, including the Nevisians in St. Kitts and Nevis. These cases can be easily dropped to better approximate the coding rules in other datasets.

The starkest definitional difference emerges with the CIDCM data. According to Quinn & Gurr (2003, p. 26) only territorially concentrated groups are included and non-violent movements are included only if they are *currently* active. The same coding rules apply to the 2008 report. However, while most SDMs have at least a low level of spatial concentration[11]), one can find SD claims by groups that are not spatially concentrated by any coding rule, such as the Turkish Cypriots before the island's partition in 1974. Dropping inactive non-violent movements introduces another selection problem since cases where the government allows movements to fester are then over-represented.

In sum, definitional differences account for minor disparities across datasets. Coding error is the main source of difference. Critically, our research suggests that measurement error is non-random. A look at the regional distributions offers a valuable first insight as to what types of SDMs are missing from the other datasets. While MAR and CIDCM miss cases across all world regions, the region with the highest number of missed cases is Europe. Perhaps unsurprisingly, given these regional distributions, we find that MAR and CIDCM systematically under-represent non-violent movements (see Table 3). MAR and

[11]We code a group as concentrated if at least half of all group members reside in a contiguous territory where they make up an absolute majority. According to this rule, around 75% of all SDMs can be considered spatially concentrated (see above). MAR has a lower threshold, requiring only that at least one quarter of a group's population reside in a contiguous area where the group makes up a relative majority. According to this rule, the share of concentrated SDMs is even higher. The Peace and Conflict reports do not define what they mean by territorial concentration.

CIDCM may deliberately omit small non-violent movements, but this is not mentioned in any coding rule and it is not easy to say a priori which movements should be taken seriously. MAR and CIDCM, too, include some weakly mobilized non-violent groups, such as the Cornish in the UK or the Saamis in Norway, Sweden, and Finland. No dataset makes use of an explicit rule on the extent of public support for SD a condition for inclusion, so we err on the side of including small groups in most cases (the codebook explains why we choose not to include some fringe groups). In addition, our research suggests that MAR and CIDCM systematically omit groups from richer countries and, to a lesser extent, democracies.[12] Such systematic measurement error is likely to bias inferences in studies of separatism drawing on MAR and CIDCM. To illustrate this point, we used a standard civil war model to compare the country-level characteristics of the number of SDMs in a country across datasets (the results are reported in the supplement). We find that MAR and CIDCM systematically under-estimate the effects of GDP per capita and democracy, especially in more recent years.[13] We also find (smaller) differences regarding other variables, including whether states are newly independent or non-contiguous. Similar biases are likely to emerge for many other research questions (for a general discussion see e.g. Hug 2013).

4 Uses of the SDM data

The SDM dataset can be used to address several important and unanswered questions related to separatist mobilization. In particular, any question that involves the onset, duration, or escalation of separatist conflict can be studied using our data. One of the key advantages of our data is that they do not suffer from the case selection bias that is inherent in MAR and CIDCM and they can be used to study conflict escalation. The analysis of separatist conflict can now start much earlier than the onset of civil war. For example, the Tamil insurgency in Sri Lanka had its beginnings in moderate claims for

[12]However, additional analyses suggest that measurement error in MAR and CIDCM is primarily a function of whether movements are violent or not. The differences regarding democracy and wealth are to a large extent a result of the fact that violent movements tend to be in poorer and less democratic countries (see supplement).

[13]Coverage differences between MAR/CIDCM and SDM become more pronounced over time.

federalization long before the conflict escalated to secessionist war in 1983. Using our data, the sources of conflict escalation can be explored in detail.

For many questions, such as for analyses of the onset of separatist mobilization or transitions from non-violent to violent mobilization, our data must be supplemented with additional variables and merged, depending on the researchers' preferences, with country-level or group-level data (such as EPR (Vogt et al. 2015) or AMAR (Birnir, Wilkenfeld, Fearon, Laitin, Gurr, Brancati, Saideman, Pate & Hultquist 2015, Birnir, Laitin, Wilkenfeld, Waguespack, Hultquist & Gurr forthcoming)). In this respect, it is important to note that no extant group-level dataset provides information on all SDM groups. EPR, for example, does not include regionally defined groups. There may also be aggregation differences. It is important that researchers consider the implications of dropping SDM groups due to missing data and, if necessary, collect data on missing cases.

While our dataset can be used to study separatist mobilization at the country-level, disaggregation offers more leverage. Many of the variables of interest are defined at the group (or regional) level, such as the existence of hydrocarbon reserves in the separatists' territory. Furthermore, many states have multiple SDMs, so group-based analyses allow one to trace individual conflicts over time and explore the effects of state-movement interactions using the data provided in the random sample component of our dataset. For example, several studies have asked if government concessions on autonomy are likely to satisfy groups seeking SD or if accommodation will have the opposite effect, fanning the flames of separatist conflict (e.g. Cederman, Hug, Schädel & Wucherpfennig 2015, Chapman & Roeder 2007). How do autonomy restrictions affect conflict escalation? These questions could not be addressed fully in the extant literature, but answers are now within reach with our new data on government concessions and restrictions.

Our data can be used to study several other questions about the dynamics of SD conflicts. For example, the claims variable can be used to compare the conditions under which SDMs make more limited claims for autonomy as compared to maximalist claims for outright secession. The data on concessions and restrictions can be used to study why states accommodate some groups while ratcheting up restrictions against others,

a question we explore next by re-analyzing a prominent study of the determinants of government accommodation of SDMs.

5 Application

Next, we demonstrate the usefulness of our dataset by reconsidering Barbara Walter's (2006*a*) influential argument that government accommodation of SD claims is less likely to occur in countries with a large number of ethnic groups. Missing cases in Walter's data are likely to influence her conclusions, as are errors in the coding of government accommodation of SDMs. Correcting these data errors, we find that Walter underestimates governments' willingness to accommodate SDMs. Re-estimating her model using our data, we find no support for her reputational theory of separatist conflict.

5.1 Walter's reputation argument

Do governments facing many potential challengers refuse to accommodate SD claims so as to build a reputation for strength? Drawing on the theory of reputation in bargaining models, Barbara Walter answers this question affirmatively because a government's concessions to one group is likely to be seen as signaling weakness, which will invite future challenges and more demands for concessions. By contrast, refusing to accommodate SD claims should signal strength, which should discourage future challenges. Thus, Walter argues that the high costs of future disputes imply that even conciliatory governments will resist accommodation to deter future challengers.[14]

This reputation logic leads Walter to make two testable hypotheses. The first and arguably the main hypothesis is that "[a] government's decision to accommodate demands for self-determination will be negatively related to the number of challengers it expects to face in the future" (p. 314). The second hypothesis is that the higher the economic, strategic, or psychological value of land occupied by potential challengers, the less likely is

[14]Walter's (2006*a*) argument has antecedents in earlier versions of the 'ethnic domino theory' see Horowitz (1985), Saideman (1998), Gurr (2000*b*), Hale (2000), Toft (2003) and Walter (2003). For a recent review and analysis see Forsberg (2013).

a government to accommodate any challenges (p. 315). We focus on the first hypothesis though we also show that Walter's conclusions with respect to the second hypothesis are not robust using her own data.

Walter (2006*a*) tests her predictions using data on separatist movements culled from the 2003 CIDCM report (Marshall & Gurr 2003). Each movement resents one unit of observation. The dependent variable is a four-point ordinal variable coding the maximum level of accommodation of an SDM throughout the period of activity: 0 indicates no accommodation; 1 some reform but not over territory (e.g. inclusion in central government, linguistic or religious rights); 2 territorial autonomy; and 3 independence. Walter's main independent variable—the number of potential future challengers—is proxied by the number of ethnic groups in each country. The premise is that any ethnic group might at some point make a separatist claim. The value of land of potential future challengers is proxied via the cumulative economic, strategic, psychological, and human (share of population) value of land that is occupied by all minorities at risk (MARs) in each country. Walter controls for the economic, strategic, and psychological value of the land currently occupied by groups challenging the state; the relative capabilities of the contestants (using proxies such as government military personnel); a country's level of democracy and the total duration of the SD challenge in calendar years.

The first column in Table 4 replicates Walter's main results (corresponding to column 1 in Table 1 in her article). In line with her theory, Walter's results suggest that the higher the number of ethnic groups in a country, the lower the chances of accommodation. Furthermore, two of the measures of the value of land occupied by potential future challengers (the combined strategic value and the combined proportion of the population) are also negatively associated with accommodation.

5.2 Data issues

There are two main problems with the data used in Walter's analysis. First, she draws on an incomplete list of SDMs based on Marshall & Gurr (2003). As discussed previously, there are systematic omissions of cases, a problem that is exacerbated due to missing

Table 4: A first assessment of Walter's reputation argument

		SDM replacements	
	(1) Walter	**(2) Standard**	**(3) Strict**
Reputation model			
Number of ethnic groups	−0.10**	−0.04	−0.01
	(0.05)	(0.04)	(0.04)
Combined economic value	0.05	−0.04*	−0.00
	(0.03)	(0.03)	(0.03)
Combined strategic value	−0.31**	0.05	−0.12
	(0.14)	(0.11)	(0.11)
Combined psychological value	0.28	0.17	0.32
	(0.24)	(0.20)	(0.20)
Proportion of population (all groups)	−2.03**	−0.97	−1.42*
	(1.03)	(0.71)	(0.73)
Value of land under dispute			
Economic value	−0.04	0.10	0.02
	(0.09)	(0.08)	(0.08)
Strategic value	0.05	−0.14	0.05
	(0.19)	(0.17)	(0.17)
Length of residence	0.01	0.39	0.17
	(0.41)	(0.38)	(0.37)
History of autonomy	−0.39	−0.54*	−0.65**
	(0.30)	(0.32)	(0.30)
Proportion of population (current group)	1.37	−1.53	0.79
	(2.03)	(1.59)	(1.63)
Relative capabilities			
Government military personnel	−13.57	33.01	38.35
	(28.55)	(25.78)	(24.42)
Government instability	0.41	0.01	0.13
	(0.36)	(0.31)	(0.31)
Group concentration	0.02	0.31	0.14
	(0.21)	(0.20)	(0.20)
Neighboring ethnic groups	−0.14	−0.12	−0.13
	(0.14)	(0.12)	(0.12)
Other controls			
Level of democracy	0.11***	0.01	0.03
	(0.03)	(0.02)	(0.02)
Duration of conflict	0.00	0.01	0.02**
	(0.01)	(0.01)	(0.01)
Pseudo R^2	0.25	0.13	0.13
N	86	86	86

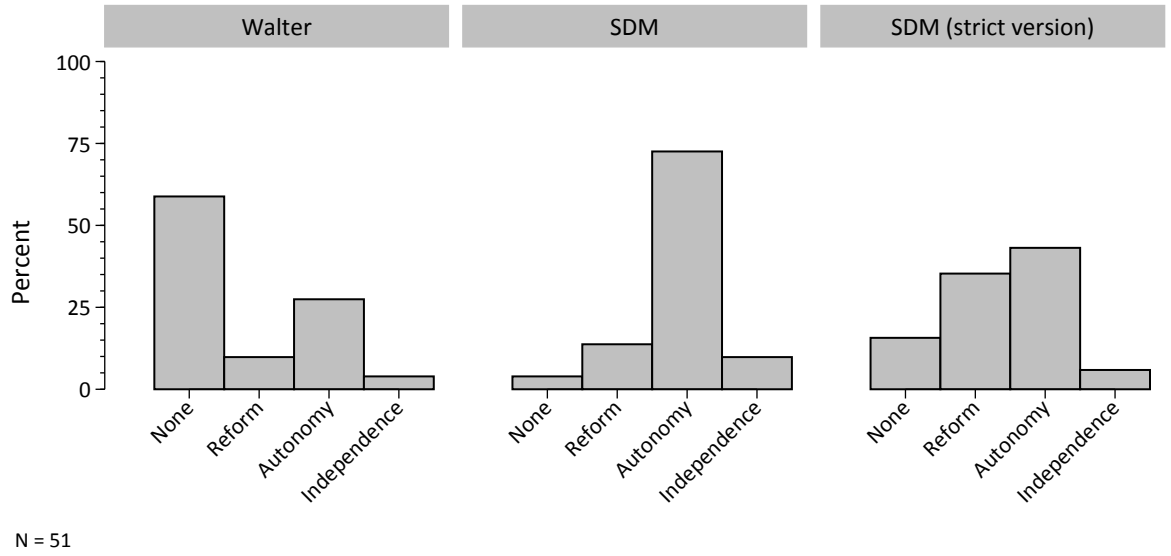
Note: Ordered probit regressions explaining the maximum level of accommodation ever granted to an SDM. Model 1 replicates Walter's (2006) main model. Model 2 replaces Walter's accommodation data with data based on the SDM dataset for the 51 cases we have in common (16 of the 51 cases are dropped due to missing data on other variables). Model 3 does the same, but uses a stricter version of SDM's accommodation variable. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

data in right-hand-side variables, which leaves Walter with only 86 SDMs in 51 countries, a small fraction of the total number of 450 movements we have coded in the period analyzed by Walter (see the supplement). Violent movements are over-represented in her data and movements in affluent democracies are under-counted. The correlates of government accommodation of SDMs are likely different in violent vs. non-violent cases or in democracies vs. dictatorships, so these omissions constitute a likely source of bias.

Second, there are problems in how she codes government accommodation. Walter's data severely under-count state policies to accommodate challengers. Take the Nagas and the Tripuris in India. In both cases, Walter codes no government accommodation of those groups, yet both were granted their own states, Nagaland and Tripura, in 1963 and 1972, respectively. The case of the Chechens in Russia is another example. According to Walter (2006*a*, p. 313), Russia did not make any concessions to the Chechens. Independence for Chechnya was not an option that the Russian government would allow, possibly consistent with the reputation argument (see Toft 2003). However, during much of the 1990's the Russian government pursued a policy of accommodating autonomy demands to co-opt ethnic/regional elites and this strategy included Chechnya. In 1992, prior to the escalation of the conflict, the Yeltsin government adopted the Federal Treaty, which created an asymmetrical federation in which ethnic republics (including Chechnya) were granted far-reaching autonomy. In 1997, after the first Chechen war, Yeltsin offered Chechnya a bilateral power-sharing arrangement in return for Chechnya's reintegration into Russia. The offer was rejected, but it cannot be argued that Russia showed no willingness to accommodate any Chechen demands (George 2009, Ross 2002). Miscoding countries like Russia and India—two large multi-ethnic countries with many potential challengers—is likely to bias the results in Walter's analysis.

Closer scrutiny of Walter's accommodation data reveals a number of other questionable codings. For a more systematic comparison, we coded our own measure of accommodation drawing on data from the SDM dataset for the period analyzed by Walter (1956–2002). The variable is coded as in Walter, measuring the maximum level of accommodation ever granted to a movement on a four-point scale (with 0 denoting no

Figure 4: Comparing the accommodation coding across datasets



accommodation; 1 cultural rights concessions or central government accommodation; 2 autonomy concessions; and 3 independence). There are 51 cases in common between our random sample and Walter's data. For those cases, Figure 4 reveals big coding differences in the two datasets. While Walter codes "no accommodation" in 59% of cases (see the first panel), we have identified at least a minimal form of accommodation in all but 4% of cases (see the second panel). In the overwhelming majority of cases, we have coded an "autonomy" concession (73%), while the corresponding figure in Walter is just 27%.

Conceptual differences in how "accommodation" is defined likely contribute to these coding differences (we discuss this below). However, most discrepancies are due to coding errors in Walter (2006*a*). Errors similar to those described above with reference to India and Russia are found in many other cases. Indonesia's Acehnese and Papuans, for instance, are coded with "no accommodation" even though Indonesia embarked on major decentralization reforms in 1999 that included these groups. The Eritreans are also coded as having received "no accommodation" despite being granted independence after Mengistu's overthrow in 1991. Walter also misses lesser forms of accommodation, as with the Native Hawaiians for whom in 1978 the Office of Hawaiian Affairs was established, an

institution with limited self-government capacity in support of native Hawaiian culture. Some of these discrepancies might be explained by differences in the coding of start dates for the conflicts. For example, Walter codes "no accommodation" for the Croats in Yugoslavia, but she appears to code a narrow period of the conflict in the late 1980's/early 1990's before the outbreak of the civil war in 1991. By contrast, we code another period of SD activity among the Croats in Yugoslavia from 1967 until 1972 in the context of the Croatian spring. While we agree with Walter that there was no accommodation in the late 1980s and early 1990s, there were significant concessions in the 1967–1972 period, including the 1971 constitutional amendment that greatly increased Croatia's autonomy.

As one might expect, these differences in the coding of accommodation influence the results. If we replace Walter's accommodation data with ours for the cases that we have in common and re-estimate her model without making any other changes, none of the variables related to the reputation logic remains significant at the 5% level (see column 2 in Table 4).

As mentioned above, coding errors alone are not the only source of coding differences, which also reflect a different conceptualization of accommodation, with our data reflecting a more inclusive concept. There are gradations of autonomy solutions and while we have a relatively broad view of what counts as an autonomy concession, Walter may have tried to code cases according to a stricter rule. For example, unlike Walter, we code an autonomy concession in case of the Bretons, given France's stepwise movements towards more devolved government (Cole 2006). We believe that such cases of modest decentralization reforms are consistent with the concept of accommodation. Yet it is true that Brittany's level of autonomy remained limited as compared to other cases.

Moreover, although she does not discuss this, Walter may have coded only policies that were fully implemented. We include also partially implemented concessions as long as there were meaningful steps toward reform. An example is Mexico's Mayans, whom we code as having received an "autonomy" concession due to the 1996 San Andrés Accord. Among other things, the Accord promised the Mayans increased autonomy over land, including control over natural resources. Yet it took Mexico several years to come up

with enabling legislation and the final version went significantly below what had originally been promised. Nevertheless, we code accommodation in this case because there was some improvement with regard to the status quo ex-ante. Such forward steps likely influence conflict dynamics and might also influence how other groups' assess their chances in making SD claims.[15] Groups do not know if a concession will be fully implemented but they often respond to negotiated settlements or initial steps in implementing agreements. Our approach can capture such cases.

Finally, other coding differences are noted in a small number of cases where the government's position indicates it is open to negotiation but is not able to implement concessions without the agreement of SD groups. This scenario concerns groups that have attained *de facto* independence, as in the case of the Turkish Cypriots after 1974. Whereas Walter codes "no accommodation" in this case, we code an autonomy concession due to a 1977 agreement to establish a bicommunal federal republic in return for the reintegration of the northern Turkish part—a significant concession from the perspective of the Greek Cypriot government, even though the agreement was not implemented due to the island's divided sovereignty (Solsten 1991).[16]

To make sure that the differences between Walter's and our results are not driven by conceptual differences in the coding of concessions, we coded a second, stricter version of the accommodation variable. The strict version is intended to address the possible conceptual sources of coding differences with Walter. Specifically, the strict version counts only concessions that were implemented to a very high degree and only codes a concession on autonomy when there was a major act of devolution (less significant concessions on self-rule are now coded as "reform"). The third panel in Figure 4 gives the breakdown of government accommodations according to this stricter coding rule. We now have many more "reforms" (35%) and considerably fewer "autonomy" concessions (43%). But even with this stricter coding, we have "no accommodation" in only 16% of cases (compared

[15]The Sahrawis in Morocco constitute another example. We code an independence concession for the Saharwis in 1988 because Morocco and POLISARIO agreed to hold an independence referendum. This referendum was never held. Nevertheless, there were steps toward implementation before Morocco postponed the referendum in 1991 (which we code as a restriction).

[16]Chechnya's rejection of a bilateral power-sharing arrangement constitutes another example (see above).

to 59% in Walter). Moreover, coding accommodation this way does not change the results of the analysis: we again reject Walter's hypotheses regarding the centrality of the reputation model as none of the variables in her model retain statistical significance at the 5% level (see column 3 in Table 4).[17]

5.3 Reanalysis based on SDM

Thus far we have established that Walter's results are not robust to substituting her measure of accommodation for ours for the cases we have in common. Next, we use our full dataset to estimate a new version of Walter's model.

While we try to remain as close as possible to Walter's original approach, both the model and the data now differ from her original article. Regarding data, the most important differences are that we cover a longer period (1945–2012), analyze a sample of randomly drawn cases from the population of SDMs, and that data for the accommodation variable now comes from SDM for all cases. There are also changes to the operationalization of some of the control variables. The strategic value of claimed land is proxied by whether it contains an international land border or seashore (we choose not to code ambiguous concepts such as "attack routes"). The economic value of the land under dispute is proxied by the presence of hydrocarbons, whereas Walter also included other resources such as coal, salt, and gold. Moreover, we employ slight variations of Walter's measures for the strength of claimants (see the supplement for details).

Regarding model specification, the most important difference is that we choose to drop all measures of the combined value of land under the control of potential challengers. The main reason is that Walter's data are confined to groups included in MAR. As argued by Walter (2006*a*, p. 318) herself, adding these variables to the model may introduce endogeneity since the number of minorities at risk in a country is a function of regime

[17] In the supplement we provide additional evidence that Walter's main result concerning the number of ethnic groups is not robust. First, it depends critically on India. India is influential because it constitutes the most ethnically diverse country in Walter's sample. Furthermore, almost 10% of the cases in Walter's sample are from India, and Walter erroneously codes several Indian cases with "no accommodation." If India is dropped from her analysis, the coefficient for the number of ethnic groups is halved and no longer statistically significant. Second, Walter's main result regarding the negative effect of the number of ethnic groups on accommodation loses statistical significance if standard errors are clustered by country. See the online appendix for more robustness tests.

behavior. For example, unresponsive government behavior may increase the number of mobilized or discriminated groups in a country, and thus the number of minorities at risk would go up. As the number of minorities at risk goes up, so will the value of the land occupied by minorities at risk. The endogeneity problem is compounded by MAR's tendency to over-represent violent, strongly mobilized or discriminated groups, as demonstrated above for SD groups (though this likely also applies to non-SD groups).[18]

We also choose not to control for the psychological value of the land that is under dispute because we do not believe that it can be measured without significant error. Generally speaking, the more contested a piece of land, the more likely the population will feel attached to it; and cross-country comparisons of emotional attachment are hard to make: is the psychological value of Kosovo the same for successive Serbian governments and does Kosovo mean more or less to Serbs than Belfast does to Irish Republicans? Walter makes a number of implausible assumptions here, such as that governments are less attached to territories that have previously enjoyed autonomy. We see no evidence of this and it is easy to find counter-examples where the government fights hard to keep a territory it is not supposed to be attached to (such as the Karen territory in Burma, where Burmese government attachment should be low according to Walter yet a separatist war has been fought since 1947).

We start our analysis with a cross-sectional regression to compare our results directly to Walter's. Each SDM constitutes one observation. Table 5 reports the results of ordinal probit regressions with standard errors clustered by country. As "switched" movements were over-sampled, sampling weights are employed. The results unambiguously go against the reputation model, irrespective of whether we use our preferred measure of government accommodation (column 1) or the strict version (column 2). The coefficient is positive in both models and even statistically significant in the strict accommodation model, directly contradicting Walter's reputation theory.

Next, we turn to a time-series cross-sectional (TSCS) evaluation of the model. In this case, the dependent variable records the level of accommodation in a given year (again

[18]The number of groups in the country that is used to test hypothesis 1 is not affected by the endogeneity concern, as it is not based on MAR.

Table 5: Reevaluating the reputation model based on SDM

	CS		TSCS	
	(1) Standard	**(2) Strict**	**(3) Standard**	**(4) Strict**
Reputation model				
Number of ethnic groups	0.02	0.08***	0.02	0.03*
	(0.03)	(0.03)	(0.02)	(0.01)
Value of land under dispute				
Oil/gas reserves	0.51*	0.35	0.32***	0.31***
	(0.27)	(0.28)	(0.08)	(0.09)
Strategic value	−0.03	0.06	−0.07	−0.06
	(0.24)	(0.24)	(0.07)	(0.08)
Proportion of population	0.74	2.71**	0.76**	0.82**
	(0.96)	(1.09)	(0.31)	(0.32)
Relative capabilities				
Government military personnel	71.89*	47.94	31.62*	32.69**
	(40.49)	(33.27)	(17.08)	(15.39)
Government instability	−0.08	0.35	0.33**	0.33**
	(0.32)	(0.32)	(0.15)	(0.15)
Group concentration	0.19	−0.25	0.07	−0.01
	(0.29)	(0.31)	(0.10)	(0.10)
Kin groups	−0.11	−0.04	−0.03	0.00
	(0.15)	(0.14)	(0.05)	(0.05)
Other controls				
Level of democracy	0.00	0.01	0.01*	0.01
	(0.03)	(0.03)	(0.01)	(0.01)
Duration of conflict	0.02**	0.01	−0.00	−0.00
	(0.01)	(0.01)	(0.00)	(0.01)
Pseudo R^2	0.07	0.09	0.05	0.05
N	101	101	3347	3347

Note: Cross-sectional (models 1 and 2) and time-series cross-sectional (models 3 and 4) weighted ordered probit regressions explaining the level of accommodation granted to an SDM. Models 2 and 4 use a stricter coding of the accommodation variable. Models 3 and 4 include a counter of the number of previous instances of accommodation as well as cubic polynomials of the time elapsed since the last instance of accommodation (not shown). Standard errors clustered at the country level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

ranging from 0–3) rather than the maximum level of accommodation for the conflict. Annual time-series variants for all our right-hand-side variables are used except for the number of ethnic groups, which does not vary over time.[19] To avoid simultaneity bias, government instability and the level of democracy are lagged one year. All other right-hand-side variables do not need to be lagged because they reflect conditions prevailing on January 1 of each year.

[19]We do not include country or group fixed effects because our main independent variable is time-invariant.

With TSCS estimation we can explore over-time changes in accommodation. Walter codes the maximum level of accommodation over the duration of the conflict, which is problematic since concessions are the outcome of dynamic interactions between states and SDMs. Regimes can change over the life of a conflict as can movement tactics and this can shape the state's willingness to make concessions. Moreover, governments may accommodate the same movement more than once. TSCS estimation allows us to go beyond an analysis of the maximum level of accommodation ever granted to a movement.

Columns 3 and 4 in Table 5 report the results of TSCS ordered probit regressions. To correct for the oversampling of "switched" movements, we again employ sampling weights. To account for possible time dependence, we include a counter of the number of previous concessions as well as cubic time polynomials (Carter & Signorino 2010), and cluster standard errors by country. Again, the results do not support Walter's hypothesis. For both our preferred coding of accommodation (column 3) and the strict coding of accommodation (column 4), the coefficient for the number of ethnic groups is positive and in the latter case it is even weakly significant. We conducted a large number of robustness tests. Among other things, we re-estimated all models using binary versions of our accommodation variables and coded a version where only autonomy and independence concessions are counted as accommodation; we used alternative measures for the number of ethnic groups in a country; used the number of first-tier administrative subdivisions in a country as an alternative proxy for the number of potential future challengers; and added a number of additional controls to the specification. Our substantive conclusion remains unchanged (see the supplement for details).

In sum, our reanalysis suggests that Walter's reputation theory cannot be substantiated. Her results were likely an artifact of data problems. However, our re-analysis does not constitute an outright rejection of the reputational theory of separatist conflict. Governments may still care to set a precedent, but only under a set of conditions that are not specified in Walter's theoretical model. We contend that a state's concern for reputation is unlikely to be just a function of the number of ethnic groups. Fighting a challenger can be costly for the state and may decrease its ability to fight a future challenger. Thus,

when deciding on whether to accommodate or invest in reputation, states should take into account their own resource constraints as well as the strength of both the current and potential future challengers (Treisman 2004). Furthermore, some governments may distinguish between violent and non-violent claims: democracies or federal states, for example, may feel less threatened by limited, non-violent claims for increased autonomy and reputational concerns need not arise in such contexts.

Furthermore, reputation is likely to be issue-specific. Economists have long argued that entry deterrence works only if markets are sufficiently similar (Milgrom & Roberts 1982). And in the context of separatism, states usually play on multiple boards. Several potentially relevant distinctions emerge. First, assuming that reputation-building works only if subsequent challenges resemble previous ones, fighting a weak challenger is unlikely to deter a strong challenger in the future. Thus, governments should only be concerned with reputation if the current and future challenger(s) they have in mind are similarly strong. Second, states may make a distinction depending on the administrative status of separatist territories (Griffiths 2015). Granting a concession to an entity with a relatively unique status does not necessarily set a precedent for other regions. For example, granting Puerto Rico (an unincorporated US territory) independence would not necessarily set a precedent for Alaska, Hawaii, or Texas (all fully incorporated US states). Finally, limited concessions, say on cultural rights, need not affect the state's reputation for strength as far as concessions on autonomy or independence are concerned. Overall, concern with reputation is likely to be much more context-dependent than Walter assumes.

6 Conclusion

Coding data is not glamorous, but it is important. Several studies in the civil war literature have found that empirical results with big theoretical and policy implications are sensitive to small coding changes. This paper improves the coding of self-determination movements and offers data that are likely to lead to broad revisions in what we think we know about separatist conflict. We discuss which cases should be counted as examples

of self-determination movements and explore a richer set of historical and other sources, finding that existing datasets have under-counted non-violent movements, particularly in affluent democratic countries. We code a total of 464 SDMs worldwide from 1945–2012, which significantly expands the set of cases used in all prior studies of separatist conflict. We also offer new data on patterns of escalation, separating violent from non-violent phases of SDMs.

We demonstrate the value of the new dataset by replicating Walter's influential study of government accommodation of SDMs. We code instances of government concessions and restrictions that significantly improve the quality of the data analyzed by Walter as well as other variables that contribute to the study of conflict escalation. Using our new data, we find no evidence that governments in countries with several potential challengers are unlikely to accommodate SDMs, thereby challenging Walter's (2006*a*) conclusion regarding the pervasive importance of the logic of reputation in separatist conflict.

The logic of reputation is likely to be context-dependent and cannot be tested with indirect static proxies such as the number of ethnic groups. Prior conflict with a particular group; the group's strength relative to the state; the strength of potential future challengers; an entity's administrative status; the sensitivity of territorial concessions in light of politically salient cleavages in the country—all these are factors that likely condition the centrality of reputational considerations in the state's decision-making calculus.

While our dataset can be expanded with more group-specific information, the range of substantively important questions that our dataset can help address is already large: What are the causes of violent escalation of SDMs? Why do some groups demand autonomy rather than independence? Do autonomy concessions help de-escalate violent SDMs? And finally, why do governments accommodate some groups while repressing others? The answers to all these questions are still open.

References

Ayres, William R. & Stephen M. Saideman. 2000. "Is Separatism as Contagious as the Common Cold or as Cancer? Testing International and Domestic Explanations." *Nationalism and Ethnic Politics* 6(3):91–113.

Beardsley, Kyle, David E. Cunningham & Peter B. White. 2017. "Resolving Civil Wars before They Start: The UN Security Council and Conflict Prevention in Self-Determination Disputes." *British Journal of Political Science* 47(3):675–697.

Birnir, Jóhanna K., David D. Laitin, Jonathan Wilkenfeld, David M. Waguespack, Agatha Hultquist & Ted R. Gurr. forthcoming. "Introducing the AMAR (All Minorities at Risk) Data." *Journal of Conflict Resolution* .

Birnir, Jóhanna K., Jonathan Wilkenfeld, James D. Fearon, David D. Laitin, Ted R. Gurr, Dawn Brancati, Stephen M. Saideman, Amy Pate & Agatha S. Hultquist. 2015. "Socially Relevant Ethnic Groups, Ethnic Structure, and AMAR." *Journal of Peace Research* 52(1):110–115.

Brancati, Dawn. 2006. "Decentralization: Fueling the Fire or Dampening the Flames of Ethnic Conflict and Secessionism?" *International Organization* 60(3):651–685.

Carter, David B. & Curtis S. Signorino. 2010. "Back to the Future: Modeling Time Dependence in Binary Data." *Political Analysis* 18(3):271–292.

Cederman, Lars-Erik, Simon Hug, Andreas Schädel & Julian Wucherpfennig. 2015. "Territorial Autonomy in the Shadow of Future Conflict: Too Little, Too Late?" *American Political Science Review* 109(2):354–370.

Chapman, Thomas & Philip G. Roeder. 2007. "Partition as a Solution to Wars of Nationalism: The Importance of Institutions." *American Political Science Review* 101(4):677–691.

Coggins, Bridget. 2011. "Friends in High Places: International Politics and the Emergence of States from Secessionism." *International Organization* 65(3):433–467.

Cole, Alistair. 2006. "Decentralization in France: Central Steering, Capacity Building and Identity Construction." *French Politics* 4(1):31–57.

Cunningham, Kathleen Gallagher. 2011. "Divide and Conquer or Divide and Concede: How Do States Respond to Internally Divided Separatists?" *American Political Science Review* 105(2):275–297.

Cunningham, Kathleen Gallagher. 2013*a*. "Actor Fragmentation and Civil War Bargaining: How Internal Divisions Generate Civil Conflict." *American Journal of Political Science* 57(3):659–672.

Cunningham, Kathleen Gallagher. 2013*b*. "Understanding Strategic Choice: The Determinants of Civil War and Nonviolent Campaign in Self-Determination Disputes." *Journal of Peace Research* 50(3):291–304.

Cunningham, Kathleen Gallagher. 2014. *Inside the Politics of Self-Determination*. Oxford: Oxford University Press.

Degenhardt, Henry W. 1988. *Revolutionary and Dissident Movements: An International Guide*. London: Longman Publishing Group.

Doyle, Michael W. & Nicholas Sambanis. 2006. *Making War and Building Peace: United Nations Peace Operations*. Princeton, NJ: Princeton University Press.

Forsberg, Erika. 2013. "Do Ethnic Dominoes Fall? Evaluating Domino Effects of Granting Territorial Concessions to Separatist Groups." *International Studies Quarterly* 57(2):329–340.

George, Julie A. 2009. *The Politics of Ethnic Separatism in Russia and Georgia*. New York, NY: Palgrave Macmillan.

Gleditsch, Nils P., Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg & Havard Strand. 2002. "Armed Conflict 1946-2001: A New Dataset." *Journal of Peace Research* 39(5):615–637.

Griffiths, Ryan D. 2015. "Between Dissolution and Blood: How Administrative Lines and Categories Shape Secessionist Outcomes." *International Organization* 69(3):731–751.

Gurr, Ted R. 1993. *Minorities at Risk: A Global View of Ethnopolitical Conflicts*. Washington, DC: United States Institute of Peace Press.

Gurr, Ted R. 2000*a*. "Ethnic Warfare on the Wane." *Foreign Affairs* 79(May/June):52–64.

Gurr, Ted R. 2000*b*. *Peoples versus States: Minorities at Risk in the New Century*. Washington, DC: United States Institute of Peace Press.

Hale, Henry E. 2000. "The Parade of Sovereignties: Testing Theories of Secession in the Soviet Setting." *British Journal of Political Science* 30(1):31–56.

Hale, Henry E. 2008. *The Foundations of Ethnic Politics: Separatism of States and Nations in Eurasia and the World*. Cambridge: Cambridge University Press.

Hechter, Michael. 2000. *Containing Nationalism*. Oxford: Oxford University Press.

Hewitt, Christopher & Tom Cheetham. 2000. *Encyclopedia of Modern Separatist Movements*. Santa Barbara, CA: ABC-CLIO.

Hewitt, Joseph J., Jonathan Wilkenfeld & 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.

Hug, Simon. 2013. "The Use and Misuse of the "Minorities at Risk" Project." *Annual Review of Political Science* 16(1):191–208.

Jenne, Erin K., Stephen M. Saideman & Will Lowe. 2007. "Separatism as a Bargaining Posture: The Role of Leverage in Minority Radicalization." *Journal of Peace Research* 44(5):539–558.

Kelle, Friederike L. 2017. "To Claim or Not to Claim? How Territorial Value Shapes Demands for Self-Determination." *Comparative Political Studies* 50(7):992–1020.

Krasner, Stephen D. 1999. *Sovereignty: Organized Hypocrisy*. Princeton, NJ: Princeton University Press.

Lujala, Päivi, Jan Ketil Rød & Nadja Thieme. 2007. "Fighting Over Oil: Introducing a New Dataset." *Conflict Management and Peace Science* 24(3):239–256.

Marshall, Monty G. & 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. & 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.

Milgrom, Paul & John Roberts. 1982. "Predation, Reputation, and Entry Deterrence." *Journal of Economic Theory* 27(2):280–312.

Minahan, James. 1996. *Nations without States: A Historical Dictionary of Contemporary National Movements*. London: Greenwood Press.

Minahan, James. 2002. *Encyclopedia of the Stateless Nations: Ethnic and National Groups around the World*. Westport, CT: Greenwood Press.

Minorities at Risk Project. 2009. "Minorities at Risk Dataset." `http://www.cidcm.umd.edu/mar/` [5/27/2014].

Minority Rights Group International. 2015. "World Directory of Minorities and Indigenous Peoples." `http://www.minorityrights.org/directory/` [8/28/2015].

Quinn, David & Ted R. Gurr. 2003. Self-Determination Movements: Origins, Strategic Choices, and Outcomes. In *Peace and Conflict 2003: A Global Survey of Armed Conflicts, Self-Determination Movements and Democracy*, ed. Monty G. Marshall &

Ted R. Gurr. College Park, MD: Center for International Development and Conflict Management pp. 26–38.

Roeder, Philip G. 2007. *Where Nation-States Come from: Institutional Change in the Age of Nationalism*. Princeton, NJ: Princeton University Press.

Ross, Cameron. 2002. *Federalism and Democratisation in Russia*. Manchester and New York, NY: Manchester University Press.

Saideman, Stephen M. 1998. Is Pandora's Box Half Empty or Half Full? The Limited Virulence of Secessioism and the Domestic Sources of Disintegration. In *The International Spread of Ethnic Conflict*, ed. David A. Lake & Donald Rothchild. Princeton, NJ: Princeton University Press.

Sambanis, Nicholas. 2000. "Partition as a Solution to Ethnic War: An Empirical Critique of the Theoretical Literature." *World Politics* 52(4):437–483.

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 & Branko Milanovic. 2014. "Explaining Regional Autonomy Differences in Decentralized Countries." *Comparative Political Studies* 47(13):1830–1855.

Sambanis, Nicholas & Jonah Schulhofer-Wohl. 2015. "Civil War as Sovereignty Rupture: Coding Intra-State Conflict, 1945–2012." Unpublished manuscript.

Siroky, David S. & John Cuffe. 2015. "Lost Autonomy, Nationalism and Separatism." *Comparative Political Studies* 48(1):3–34.

Solsten, Eric. 1991. *Cyprus: A Country Study*. Washington, DC: GPO for the Library of Congress.

Sorens, Jason. 2012. *Secessionism: Identity, Interest, and Strategy*. Montreal: McGill-Queen's University Press.

Themnér, Lotta & Peter Wallensteen. 2014. "Armed Conflicts, 1946-2013." *Journal of Peace Research* 51(4):541–554.

Toft, Monica Duffy. 2003. *The Geography of Ethnic Violence: Identity, Interests, and the Indivisibility of Territory*. Princeton, NJ: Princeton University Press.

Treisman, Daniel. 2004. "Rational Appeasement." *International Organization* 58(2):345–373.

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

Walter, Barbara F. 2003. "Explaining the Intractability of Territorial Conflict." *International Studies Review* 5(4):137–153.

Walter, Barbara F. 2006*a*. "Building Reputation: Why Governments Fight Some Separatists but Not Others." *American Journal of Political Science* 50(2):313–330.

Walter, Barbara F. 2006*b*. "Information, Uncertainty, and the Decision to Secede." *International Organization* 60(1):105–135.